

Research Article

Luca Grassetti*, Michele Lambardi di San Miniato and Valentina Mameli

Revisiting player contributions in regularized adjusted plus-minus models: an application on NBA historical data

<https://doi.org/10.1515/jqas-2024-0138>

Received September 13, 2024; accepted April 6, 2026;

published online April 27, 2026

Abstract: In basketball and other team sports, assessing players' and lineups' performances is crucial for team management. This assessment should address both individual players' contributions and the synergies among them. Here, we aim to analyze the importance of collaboration among NBA players from the 2002–2003 to the 2022–2023 season. To this end, we propose a modification of the extended Regularized Adjusted Plus-Minus model, where the efficacy of each lineup is determined by adding up the non-negative contributions of players on the court, adjusted for a positive or negative interaction term. The proposed specification improves the readability of the estimated values of players' and lineups' effects, which can be straightforwardly visualized and used more efficiently for actual decision-making. The main analyses based on this model specification revealed that, in the NBA, the importance of synergies among players decreases over the years, suggesting a shift from team to individual-oriented style of play. The only notable exception occurs during the COVID-19 lockdown period, when the relevance of lineups effects temporarily increases, indicating a change towards a team-oriented style of play. The proposed model can also be adopted to evaluate the difference between home and away-specific performances of players and lineups.

Keywords: basketball analytics; lineups and players performances; players synergies; NBA data; sports management

*Corresponding author: Luca Grassetti, Department of Economics and Statistics, University of Udine, Udine, Italy, E-mail: luca.grassetti@uniud.it. <https://orcid.org/0000-0003-1997-8001>

Michele Lambardi di San Miniato and Valentina Mameli, Department of Economics and Statistics, University of Udine, Udine, Italy, E-mail: michele.lambardi@uniud.it (M. L. D. S. Miniato), valentina.mameli@uniud.it (V. Mameli). <https://orcid.org/0000-0003-2423-4250> (M. L. D. S. Miniato). <https://orcid.org/0000-0002-0540-144X> (V. Mameli)

1 Introduction

Sports analytics has a relatively long history, and its relevance has grown with technological advancements. According to Marković et al. (2020), any data record contributing to the decision-making process can be considered part of these analytics. For this reason, the beginning of sports analytics can be dated back to the early days of organized sports. Nowadays, the availability of large quantities of complex data influences the interest of statisticians (see, for instance, Baumer et al. 2023) and machine learners (as in Aalbers and Van Haaren 2019; Assunção and Pelechrisinis 2018) in applying sophisticated methodologies to gain insights and enable data-driven decision-making in a variety of sports-related domains. Purely theoretical contributions in peer-reviewed journals are rare, making it challenging to provide a detailed overview of the current state of the art in this field. Notwithstanding, from an empirical point of view, the interest in analytics is fostered by the large availability of data from sports events. The analysis of these data can be diverse (see Zuccolotto and Manisera 2020). In this complex framework, the assessment of teams' or athletes' performance represents a marginal aspect in the statistical literature while being a crucial aspect of team management. Data-driven techniques can help in this respect, but they require extensive data processing, and the knowledge acquired through such methods is difficult to translate into everyday decision-making practice in sports.

We develop our study within the model-based framework; in this context, the Regularized Adjusted Plus-Minus (RAPM) model (Sill 2010) can be an effective and efficient solution for determining players' performances, allowing for direct comparisons among players. The present work focuses on assessing the performance of the National Basketball Association (NBA) players over a 21-season period from 2002–2003 to 2022–2023, with particular attention to the role of player synergies. Specifically, we propose an empirical study to explore how individual contributions and synergistic effects can influence team outcomes in the

NBA. To conduct this analysis, we build on the Extended RAPM (E-RAPM) model introduced by Grasseti et al. (2021), which allows for the joint estimation of lineup and player-specific effects. From the decision makers' point of view, this model offers a direct interpretation of total lineup performance, which, for each lineup, is determined by the sum of pure players' effects, adjusted by adding the pure lineup effect. On the one hand, the total lineup performance can be used to measure the efficacy of the different five-man units. On the other hand, the individual contributions of each player to this performance are still somewhat difficult to discern due to the unconstrained nature of the pure players' effect. Although not central to our contribution, we adopt a slightly modified version of the E-RAPM model obtained by constraining the players' effects to be non-negative. In this way, the scoring system in a match can be seen as a process in which the players contribute to the score in an additive fashion, and their synergistic behaviors impact the final result either positively or negatively. This simple modification improves the readability of the results and provides an intuitive and more accessible visualization tool that supports practical decision-making. Hereafter, we refer to the proposed model specification as the Modified E-RAPM model (ME-RAPM).

Previous research on the Euroleague revealed the importance of synergies in explaining team performance (Grasseti et al. 2021). However, the extent of cooperation and physicality of the NBA is actually different. According to Mandić et al. (2019), the most noticeable distinction between the two competitions is the game tempo. With a longer positional and planned offense, European basketball is more tactical than NBA basketball, which is characterized by shorter possession times, less focus on defense, and more turnovers and transition plays, resulting in a more appealing game for the general public. Motivated by these considerations and a preliminary analysis of the NBA season 2022–2023, we raised the following interesting research questions:

- Does synergy among players matter in NBA gameplay?
- Has the relevance of the individual players' contribution changed over time in the NBA?

To address these questions, we analyze NBA play-by-play data from the 2002–2003 season to the 2022–2023 season. Specifically, we use the historical data collected, cleaned, and improved by BigDataBall (<https://www.bigdataball.com/>).

The insights obtained from these analyses lead to further investigations into other relevant aspects of players' performances. By slightly modifying the model specification, we introduce home and away-specific

performance metrics that allow us to explore an additional research question:

- Is there any difference in the home and away-specific performances of NBA players?

The paper is organized as follows. Section 2 overviews the existing proposals in basketball players' performance analysis, focusing on the Adjusted Plus-Minus literature and RAPM estimation. In addition, it introduces the modification to the RAPM model specification adopted in the empirical analyses. Descriptive analyses in Section 3 prepare the reader for the application to NBA play-by-play data. Section 4 reports the discussion on model estimation results. Section 5 focuses on the evolution of players' teamwork in the NBA. Finally, Section 6 gives conclusions and draws a possible future development.

2 Literature review

Many are the possible applications of sports analytics in basketball, and an interesting overview can be found in Oliver (2004). The recent literature has focused on some important topics for team management, including analysis of team performance (see Metulini and Gnecco 2022), match and tournament outcome prediction (see Kvam and Sokol 2006; Manner 2016; Ruiz and Perez-Cruz 2015; Yang and Lu 2012), injury prevention (see Cohan et al. 2021), and player salary determination (see Özbalta et al. 2021; Sarlis and Tjortjjs 2024). In this very diversified framework, evaluating the performance of individuals and groups of individuals is arguably relevant to all of these strands (Terner and Franks 2021).

The classical players' performance analysis can be carried out by considering various metrics. In Sarlis and Tjortjjs (2020), the authors comprehensively review some useful metrics available at the match level using box-scores data, including the Plus-Minus (PM) and Adjusted PM metrics. Although box-score statistics can be used for performance assessments at the players' and teams' levels, play-by-play data offer more insight into players' contributions to the team's performance (as, for instance, in Deshpande and Jensen 2016; Gong and Chen 2024). Profiling players' performance based on their on-court statistics is a very relevant topic in the sports management framework (Tiedemann et al. 2010). These analyses can be invaluable in making well-informed managerial decisions and may provide a competitive edge in the world of sports. As reported in Hvattum (2019), coaches and team managers can make decisions based on players' ratings, which can be estimated

in different ways. Indeed, players' performance can also be assessed using suitable modeling techniques.

The origins of model-based approaches in sports can be traced back to the early 2000s and the significant work of Rosenbaum (2004), who introduced the Adjusted PM (APM), a regression-based version of the PM metric. The model-based approach is further discussed in Omidiran (2011) and is considered one of the most popular methods for evaluating player performance. This model specification entails sparse design matrices, multicollinearity, and severe overparameterization. For these reasons, regularization is crucial to improve accuracy. Regularized APM (RAPM) was formulated by Sill (2010) and typically employs ridge regression to estimate player efficiency, as summarized in Engelmann (2017), but other regularization methods can also be adopted, such as those reviewed by Efron and Hastie (2016). The RAPM model is arguably one of the best solutions because it allows for computing the efficiency of players while accounting for the opposing lineup and including some exogenous variables, such as the current time and score and the difference in teams' ranking, to mitigate confounding. Typically, the RAPM model is based on play-by-play data aggregated over shifts, defined as playing periods without substitutions. In a very recent work, Damoulaki et al. (2025) developed an RAPM-based indicator on a finer aggregation of play-by-play data at the level of possessions. The authors employed a model-based approach to analyze a single possession outcome. Their approach enables the study of offensive and defensive player performance by applying a lasso-regularized multinomial regression model. In this context, a by-product of the model estimation is the performance of lineups and teams, defined as groups of players.

However, considering the players' effects as the only measure of the group performance is simplistic since the team and lineup performances also depend on players' interactions, hereafter referred to as synergistic effects. In this respect, a peculiar use of the RAPM model is the one introduced in Grassetti et al. (2019), where the PM concept is applied to estimate the lineup effects only, encompassing the synergies among players on the court. As the number of observed lineups exceeds by far the number of players, this model specification involves more sparsity and, more relevantly, more complex results. From a technical perspective, it takes a step forward in terms of the goodness of fit. However, from a practical point of view, the estimated lineup effects are less interpretable than the classical RAPM players' contributions. More recently, in Grassetti et al. (2021), the original RAPM model is extended to include synergies among players, namely, the pure lineup

effects. Considering P players, L lineups, and T shifts, the E-RAPM model specification can be compactly expressed in a matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}^{(l)}\boldsymbol{\mu} + \mathbf{Z}^{(p)}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \quad (1)$$

where \mathbf{y} is the T -dimensional vector of differences between home and away scored points divided by the shift-specific number of possessions and $\mathbf{X}\boldsymbol{\beta}$ is the linear predictor for the effect of exogenous variables. With only two non-zero entries per row, $\mathbf{Z}^{(l)}$ is the $T \times L$ design matrix for lineups' effects, whereas $\mathbf{Z}^{(p)}$ is the $T \times P$ design matrix for the players' effects, presenting ten non-zero entries per row. These matrices are sparse, and their non-zero entries are either $+1$ or -1 when the lineups and players belong to home or away teams, respectively. The players' effects vector, denoted by $\boldsymbol{\gamma}$, has length P . The vector $\boldsymbol{\mu}$ represents the effects of the lineups and has length L . Finally, the model assumes the following normal distribution for the error term

$$\varepsilon_t \sim \mathcal{N}\left(0, \frac{\sigma_\varepsilon^2}{n_t}\right) \quad t = 1, \dots, T,$$

where n_t is the observed number of possessions during the generic shift t .

The classical RAPM literature employs *ridge regression* to estimate players' effects. As the model in equation (1) considers two sparse design matrices for the players and lineup effects, the empirical Bayes or full Bayesian approaches can be adopted to achieve a feasible and more flexible regularization. These solutions require the specification of priors for the lineup and player effects, which have been assumed to be independent Gaussian distributions, namely

$$\boldsymbol{\mu} \sim \mathcal{N}\left(\mathbf{0}, \sigma_\mu^2 \mathbb{I}_L\right) \text{ and } \boldsymbol{\gamma} \sim \mathcal{N}\left(\mathbf{0}, \sigma_\gamma^2 \mathbb{I}_P\right),$$

where \mathbb{I}_D is the $D \times D$ identity matrix. In the full Bayesian approach, to mimic the regularization obtained by ridge regression, one must introduce suitable priors for σ_μ^2 and σ_γ^2 , as shown in Grassetti et al. (2021). The full Bayesian solution is computationally and memory-intensive. Notwithstanding, its flexibility brings valuable advantages when considering further model generalizations, such as the one proposed in the next subsection.

2.1 The modification of the E-RAPM model

As just introduced, from a managerial point of view, reading the E-RAPM model in terms of players' contributions to the teams' performance is not straightforward. In fact, the cumulative performance of players on the court can be defined by summing up the typical RAPM measures, but this summarization does not represent a clear and easily

readable tool, as the resulting index cannot be directly disentangled to represent the individual contributions. This work proposes a model in which estimated effects still allow to assess player performance but are also easy to read and visualize for decision-makers.

In particular, we propose to treat the shifts' scores for the home and away teams as the outcome of players' efforts, borrowing the idea of input contribution in a production process. In this regard, an E-RAPM-like model specification can be applied to the non-negative home and away-specific scores. Specifically, let y_{ht} and y_{at} be the measures of the home and away teams' efforts observed at shift t . The two contributions to the outcome of shift t may be expressed as the total effect of the players on the court and an additional effect related to the specific lineups as

$$y_{ht} = Z_{ht}^{(l)} \boldsymbol{\mu} + Z_{ht}^{(p)} \boldsymbol{\gamma}, \quad \text{and} \quad y_{at} = Z_{at}^{(l)} \boldsymbol{\mu} + Z_{at}^{(p)} \boldsymbol{\gamma}, \quad (2)$$

where ht and at indicate the home and away teams playing at shift t . Regarding the lineups, the two design matrices are extremely sparse since each row, $(Z_{ht}^{(l)})_{t=1, \dots, T}$ and $(Z_{at}^{(l)})_{t=1, \dots, T}$ only has one non-zero element indicating the presence of the lineup at shift t . The two design matrices for players at home and away are also sparse. Indeed, $(Z_{ht}^{(p)})_{t=1, \dots, T}$ and $(Z_{at}^{(p)})_{t=1, \dots, T}$ show five non-zero elements indicating the five players on the court at shift t . The vectors $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ collect the lineup's and player's effects, respectively. As y_{ht} and y_{at} show non-negative and heavy-tailed distributions, it is natural to define the players' effects as non-negative contributions to the team scores and consider unconstrained synergistic effects (due to lineup composition) to adjust the lineup's total performance. The resulting model is called ME-RAPM hereafter. Dealing with non-negative player contributions, one can finally supply the team's management and coaching with an effective tool for building the team roster first and selecting the best lineup solutions during the games, being able to quantify the players' contributions and their synergies clearly and simultaneously. Moreover, the proposed solution also helps in shrinking the low-performing players' effects, thereby minimizing the potential influence of negative outliers in their practical use. The shrinkage effect of the proposed solution is visualized in Figure 2.1 in the Supplementary Material.

Following the model formulation in equation (1), the overall score at each shift t can then be modeled as a function of the difference between the home and the away teams' efforts, defined in (2) as

$$y_t = X_t \boldsymbol{\beta} + y_{ht} - y_{at} + \varepsilon_t, \quad t = 1, \dots, T, \quad (3)$$

where X_t is a matrix of additional covariates at shift t , with the associated vector of coefficients $\boldsymbol{\beta}$ and the idiosyncratic error term $\varepsilon_t \sim N\left(0, \frac{\sigma_\varepsilon^2}{n_t}\right)$.

Note that the model defined by combining equations (2) and (3) is a reformulation of the original E-RAPM model, which still fully accounts for the adjustment due to the simultaneous presence of players on the court. Modeling the two scores separately provides a more accurate representation of the data-generating process and represents a further justification for defining non-negative contributions from players.

Moreover, this formulation can be directly used to split the estimated effects into home and away-specific measures, defining two distinct vectors collecting the home and away performances of players ($\boldsymbol{\gamma}_h$ and $\boldsymbol{\gamma}_a$) and lineups ($\boldsymbol{\mu}_h$ and $\boldsymbol{\mu}_a$) as well. As a result, the equations defining the home and away-specific scores can be defined as follows:

$$y_{ht} = Z_{ht}^{(l)} \boldsymbol{\mu}_h + Z_{ht}^{(p)} \boldsymbol{\gamma}_h, \quad \text{and} \quad y_{at} = Z_{at}^{(l)} \boldsymbol{\mu}_a + Z_{at}^{(p)} \boldsymbol{\gamma}_a. \quad (4)$$

The resulting model (Generalized ME-RAPM - GME-RAPM hereafter) present a noticeable complexity due to the sparsity of the design matrices, in particular when considering the generalization of lineup effects.

In a Bayesian framework, the non-negative player effects can be modeled considering a half-normal distribution ($\mathcal{HN}(\mathbf{0}, \sigma_\gamma^2 \mathbb{1}_p)$). Such priors enable easier evaluation of the role of players and lineups in the team's performance. A Gaussian distribution ($\mathcal{N}(\mathbf{0}, \sigma_\mu^2 \mathbb{1}_L)$) can still be assumed for the lineup effects $\boldsymbol{\mu}$. Finally, the model specification assumes prior distributions for $\boldsymbol{\beta}$ and the standard deviations for the player effect (σ_γ), the lineup effect (σ_μ), and the error term (σ_ε). Similarly to Grasseti et al. (2021), we adopt weakly informative priors such as half-Cauchy distributions with a moderately large scale hyperparameter (equal to ten) for the standard deviation parameters, and an independent zero-mean normal distribution with a large standard deviation (equal to ten) for the components of $\boldsymbol{\beta}$, according to the response values' range; refer to the works of Gelman (2006) for further information on non-informative priors. The resulting model specification is a typical example of hierarchical Bayesian models. The same priors can also be adopted when considering the modification introduced in equation (4).

Hierarchical models often face estimation issues such as funneling. This issue may arise when some prior distributions are ruled by parameters that, in turn, present a given prior, leading to unstable chains. When the priors for players' effects are either Gaussian or half-normal, one may adopt hierarchically non-centered parameterizations

(NCPs, Papaspiliopoulos et al. 2007). This approach stabilizes sampling from posterior distributions in hierarchical models, making the full Bayesian solution computationally viable.

3 Data exploration

In this section, we present a few descriptive analyses that motivate our choices on data subsetting and stratification. These choices are crucial in drawing meaningful conclusions.

The NBA is a professional basketball league comprising 29 American clubs and one Canadian club, based in Toronto. The league is split up into the Eastern and Western conferences. The teams and their abbreviations, along with conference affiliations, are given in the Supplementary Material in Section S1 to improve the readability of the following results.

To describe the main characteristics of the NBA data, we first focus on the 2022–2023 season. The older seasons, starting from 2002 to 2003, are then analyzed to highlight relevant patterns with a special focus on lineups' and players' utilization.

The data cleansing process is as follows. First, we aggregate the play-by-play data and turn them into shift-by-shift data to match the RAPM model specification. Secondly, for each shift, we compute points scored and possessions played by the home and away teams, as well as the elapsed and

remaining time at the end of the shift. Lastly, we drop shifts when no possessions occurred or when the game is “gone”, that is, with a points difference larger than 20 and less than 6 min left, because it can be considered garbage time. The percentage of time passed in this condition varies between 2.37 % and 4.27 % in the 21 considered seasons.

3.1 2022–2023 NBA season

The first analysis we propose is a fine-tuning of the data-cleansing process. In particular, all the analyses hereafter will not consider the playoffs because games in this phase of the tournament are fewer than and significantly different from the regular season in many respects (Cabarkapa et al. 2022). For instance, Figure 1 displays the effective number of shifts, lineups, and players per game over months in 2022–2023, computed as described in Section 3.2. A pronounced shift in team dynamics is observed in April, coinciding with the start of the NBA playoffs.

The 2022–2023 regular season consists of 1,230 games with 467 plays per game on average. After aggregation and data cleansing, the dataset includes 35,128 shifts, with an average of 14.3 lineups and 10.3 players per team, and 28.4 shifts per game. Each shift consists of 3.4 possessions per team, on average.

Before starting the analysis of players' utilization, we describe the rationale behind the proposed model specification. By looking at the joint and marginal distributions of home and away points per possession scores in Figure 2,

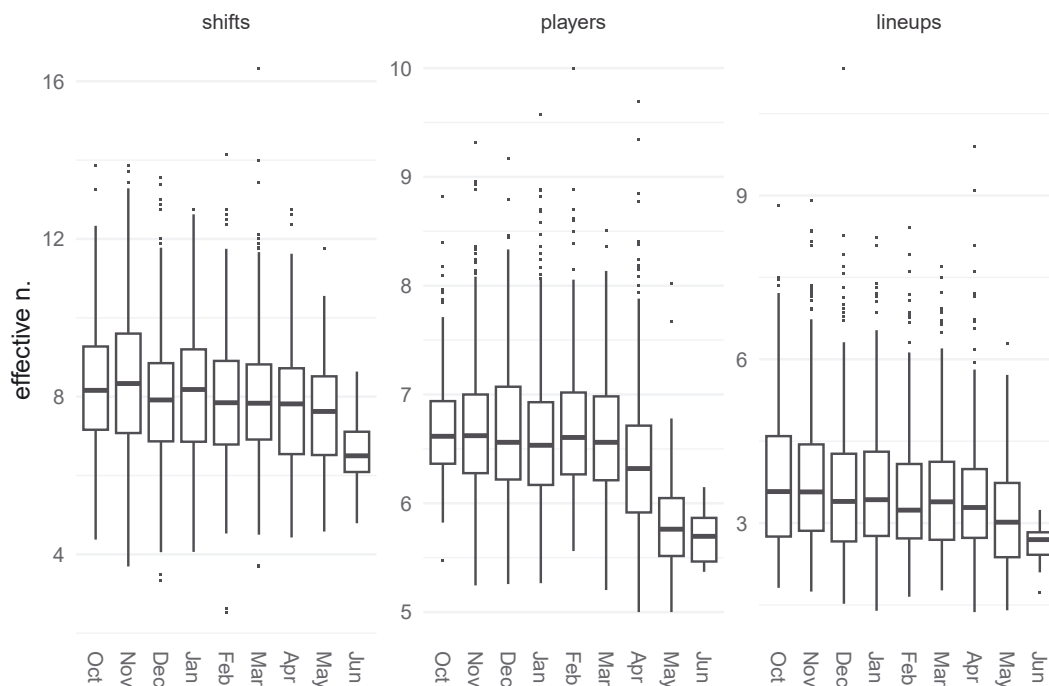


Figure 1: 2022–2023 NBA season. The players' usage over the season: the effective number of shifts, players, and lineups per game and month.

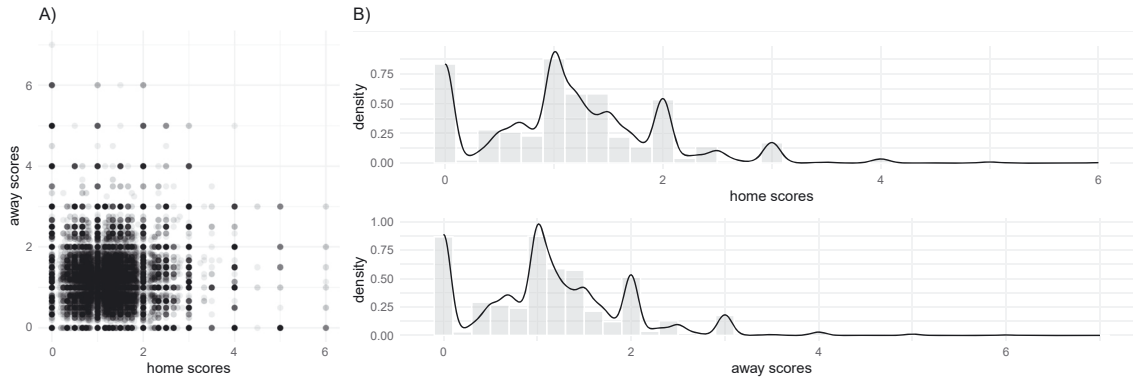


Figure 2: 2022–2023 NBA season. The scatterplot of points per possession at home and away (A) and the empirical marginal distributions of points per possession at home and away (B).

we note the scores exhibit a peculiar distribution with a remarkable positive skewness and zero inflation, supporting the idea that the player’s contributions can be considered non-negative.

To study the players’ utilization, we finally assess the concentration of possessions per lineup and player, which can help to evaluate the gameplay characterizing each team. Summary statistics describing the distribution of possessions per lineup and player are provided in the Supplementary Material, in Section S2.1.

To illustrate this distribution, Figure 3 presents a comparative analysis across four selected teams: on the left panels, a rescaled Lorentz curve for the concentration of possessions among lineups for four selected teams is displayed, while on the right panels, the distribution of possessions per player is reported. The chosen teams exhibit some rather distinct behaviors. For example, Boston Celtics’ and Denver Nuggets’ lineups play a significantly larger amount of possessions compared to San Antonio Spurs’ and Brooklyn Nets’. Boston Celtics and Denver Nuggets are smaller teams, both with only 18 players, and they put fewer lineups on the court during the season. The average difference between the home and away teams’ points per 100 possessions can also be expressed as a function of the possessions scored when specific lineups or players are on the court. In Section S2.1 of the Supplementary Material, we provide a representation of this relationship for the four previously selected teams.

All these analyses demonstrated that teams adopt very different strategies when choosing the players’ rosters and also when combining their players on the court. As a result, assessing players’ and lineups’ performances can be crucial for managing the player rotation, which in turn represents a fundamental aspect of teams’ performance, as clearly stated in Clay and Clay (2014).

3.2 NBA seasons from 2002–2003 to 2022–2023

The analysis in the previous subsection suggests that the teams may behave rather differently in terms of players’ and lineups’ usage. Here, we show that changes in terms of players’ usage have occurred over time, too, so it is better to stratify the analysis based on seasons in the following.

The number of games played in the regular season is essentially constant over the years, except in the 2011–2012 season for the lock-out and the 2019–2020 and 2020–2021 seasons during the COVID-19 pandemic. The number of plays per game varies from 444 to 476, with an average value of 455.6. Over time, there has been some variability in the number of shifts per game, ranging from a minimum of 27.1 to a maximum of 29.8, with an average of 28.3. The lineups’ and players’ average numbers per team and game are 13.8 and 10.2, respectively. The average number of possessions per team and shift is 3.2 and ranges between 3 and 3.4. Further details can be found in the Supplementary Material (Section S2.2).

To evaluate the differences among the seasons, we analyze the distribution of the effective number of players, lineups, or shifts per game over the years. These measures are computed as the ratios between the sum of the possessions recorded for players, or lineups, or shifts and their maximum within the team and game. In other words, the relevance of players, lineups, or shifts is measured by the number of played possessions. The analysis shows that the effective number of players and shifts increases over time, but the number of lineups remains mostly unchanged. This trend is illustrated in the Supplementary Material (Section S2.2).

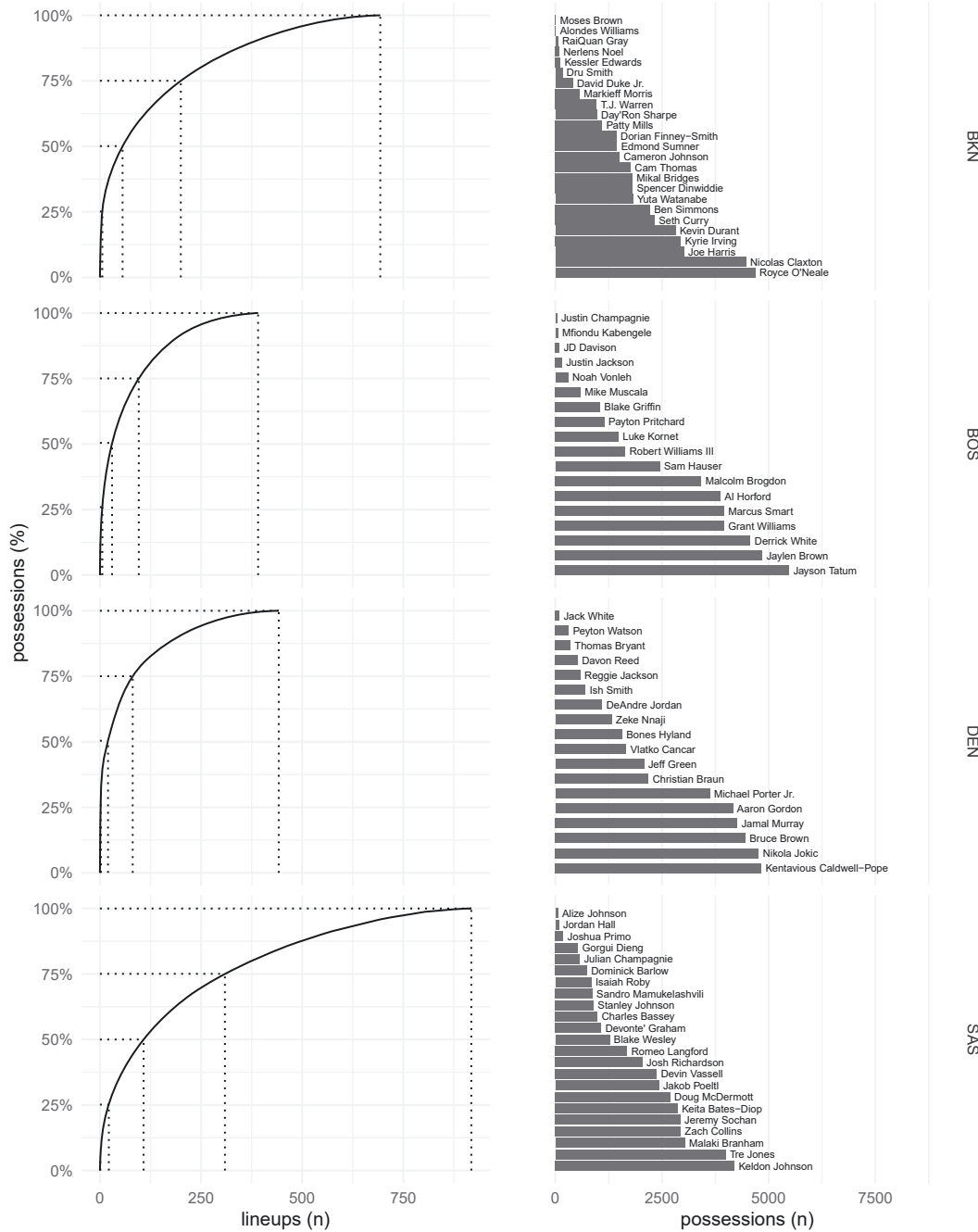


Figure 3: 2022–2023 NBA season. The concentration of possessions across lineups (left) and the distribution of possessions per player (right) for the four selected teams.

4 Model estimation results

This section collects the analysis of the model estimation results for the model specifications introduced in Section 2, using data from the 2022–2023 NBA season. This section is intended to facilitate the reading of the main results reported in Section 5 but it also gives general insights into the empirical analysis.

The first subsection highlights the similarities and differences between the E-RAPM and ME-RAPM models and compares their estimated performances with other established metrics. The second subsection introduces a graphical tool for visualizing player and lineup effects derived from the ME-RAPM model. The third subsection examines the estimates of the ME-RAPM model that accounts for home and away-specific performances.

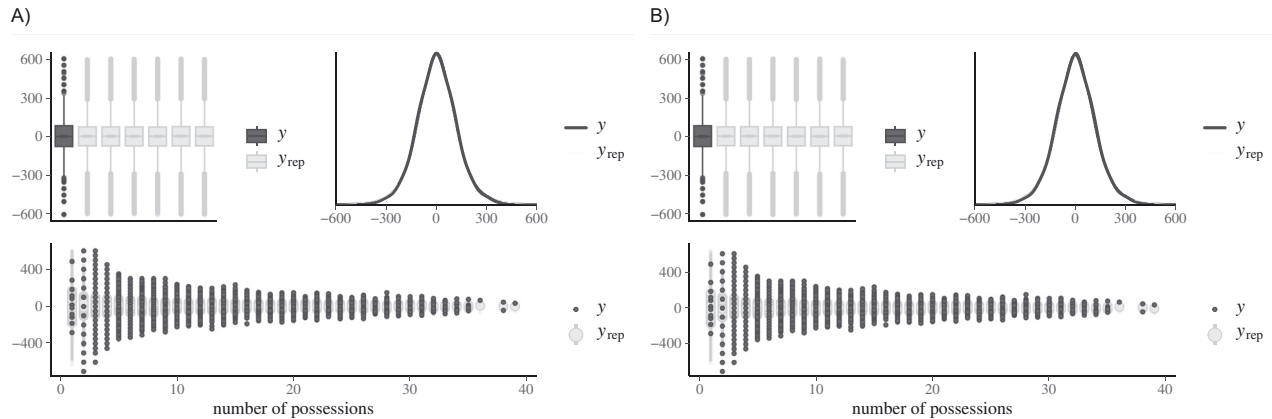


Figure 4: 2022–2023 NBA season. Bayesian predictive model checking: comparison of the standardized observed score with some samples simulated from the predictive distribution for the E-RAPM model (A) and for the ME-RAPM model (B). Top panel, left: boxplots of observed response and six simulated samples. Top panel, right: density plots of the observed variable and 50 simulated samples. Bottom panel: 50 % intervals computed using the predictive distribution, with the observed score superimposed, sorted by possessions.

All model estimates are obtained using R (R Core Team 2024), adopting the full Bayesian approach. The `rstan` library (Stan Development Team 2024) is used to estimate the model. To visualize model estimates and diagnostic outputs, we use the `bayesplot` library (Gabry and Mahr 2024), and some `ggplot2` tools (Wickham 2016) for more specialized plotting solutions.

4.1 Model fitting, estimation, and evaluation

Using 2022–2023 NBA season data, we analyze and contrast E-RAPM with the ME-RAPM model. In general, the estimation is based on four MCMC chains, and for each chain, 2,000 warmup iterations are discarded, while 2,000 post-warmup iterations are retained for inference. The \hat{R} (\hat{R}) statistic (Gelman et al. 1995) serves as a convergence diagnostic. When this measure exceeds 1.01, there are some concerns regarding the simulation’s stability, as suggested in Vehtari et al. (2021). In our simulations, all model parameters have \hat{R} statistics close to 1 and less than 1.01, thus not signaling any convergence problem. The Effective Sample Size (ESS) is sufficiently large to yield precise estimates. Moreover, the expected log-predictive density (ELPD) is used to assess the model’s goodness of fit (Vehtari et al. 2017). Higher ELPD values indicate greater posterior predictive accuracy. According to the ELPD, the E-RAPM and ME-RAPM models exhibit comparable performance, with scores of $-221,973.7$ and $-221,975.5$, respectively.

To further check the models’ fit, we examine their posterior predictive distributions. We also evaluate whether models can simulate data that approximately match the observed score, normalized by the number of possessions, by drawing 50 simulated datasets from the posterior

predictive distributions of the E-RAPM and the ME-RAPM. Figure 4 compares the observed score, normalized by the number of possessions, with data simulated from the two posterior predictive distributions. All the comparisons indicate that, despite a few outliers, both models provide adequate fits.

Model estimates are given in Table 1. The two models show similar estimates except for the players’ effects scale parameters, which differ due to the models’ specifications. The estimates show the typical home effect, which has been largely discussed in the sports analytics literature. For instance, in Harris and Roebber (2019) the authors evaluate the reasons for the decline of the home advantage over the NBA seasons using a neural network-based approach. In Ribeiro et al. (2016), the authors identify a general decline in this phenomenon over the years. Still, they also distinguish some team-specific positive patterns, identifying heterogeneity in the phenomenon. Following the idea of identifying a period-specific home effect as in Jones (2007), we consider a model based on the period dummies that capture this effect. In particular, β_h identifies the home advantage, while the parameters corresponding to the periods are typically not significant. The scale parameters also reflect a common characteristic of the RAPM models, where the residual error component is often prominent. The extent of the lineup and player effects can be inferred from the scale parameters estimates. As one can notice, the magnitude of γ parameters is superior to that of μ ones for all the considered models, even if not directly comparable. Further details on the distribution of the estimated effects for players and lineups can be found in the Supplementary Material (Section S3.1), where the previously mentioned

Table 1: 2022–2023 NBA season. E-RAPM and ME-RAPM model estimation results. Along with player and lineup effects, the models considered a simple dummy variables model matrix for the period of the game. For all the parameters, the mean of the posterior distribution is reported along with the 95 % credible interval, effective sample size (ESS), and \hat{R} fit measures.

Parameters	Mean	2.5 %	97.5 %	ESS	\hat{R}
E-RAPM model					
β_h	3.115	1.266	4.991	3,353.273	1.000
β_2	-1.655	-4.308	0.985	4,526.652	1.000
β_3	-1.092	-3.823	1.559	4,543.722	1.000
β_4	-0.791	-3.506	1.893	4,458.421	1.000
$\log(\sigma_\epsilon)$	5.496	5.488	5.503	10,453.077	1.000
$\log(\sigma_\gamma)$	1.061	0.864	1.237	2,628.683	1.001
$\log(\sigma_\mu)$	0.019	-2.881	1.336	708.208	1.002
ELPD	-221,973.665 (158.5)				
ME-RAPM model					
β_h	3.164	1.268	5.033	2,790.335	1.001
β_2	-1.706	-4.424	0.949	3,516.008	1.001
β_3	-1.119	-3.859	1.631	3,792.364	1.001
β_4	-0.825	-3.538	1.842	3,485.661	1.001
$\log(\sigma_\epsilon)$	5.496	5.488	5.503	10,633.934	1.000
$\log(\sigma_\gamma)$	1.523	1.319	1.707	2,312.410	1.002
$\log(\sigma_\mu)$	-0.042	-2.916	1.299	728.272	1.001
ELPD	-221,975.525 (158.493)				

shrinkage of the effects of the worst-performing players is clearly visible in the right-hand panel.

To test the validity of the estimated RAPM measures and equivalence between the original and the modified RAPM model specifications, we compare our estimated player performances with official rankings. Players considered in this analysis are only those who played at least half of the regular season. Figure 5 shows that the performance measures considered in the comparison are all positively correlated, but two main clusters of measures can be identified in the picture. One indicates that the RAPM results are closely related to the estimated Net Rating (E_NR), the Plus-Minus (PM), and the Net Ratings (NR). The other one identifies a more heterogeneous group, including Efficiency (EFF), Player Impact Estimate (PIE), and Player Efficiency Rating (PER), which are weakly related to model-based results. Moreover, these three measures are highly correlated and also strongly related to the Win Shares (WS), the Value Over Replacement of Players (VORP), and the Box score Plus-Minus (BPM). A lower correlation can be observed between the two model-based measures and WS, VORP, and BPM.

We can also notice that the two RAPM-based results present an almost perfect positive correlation, supporting the equivalence of the two model specifications. This analysis encompasses different “players stats” data sources,

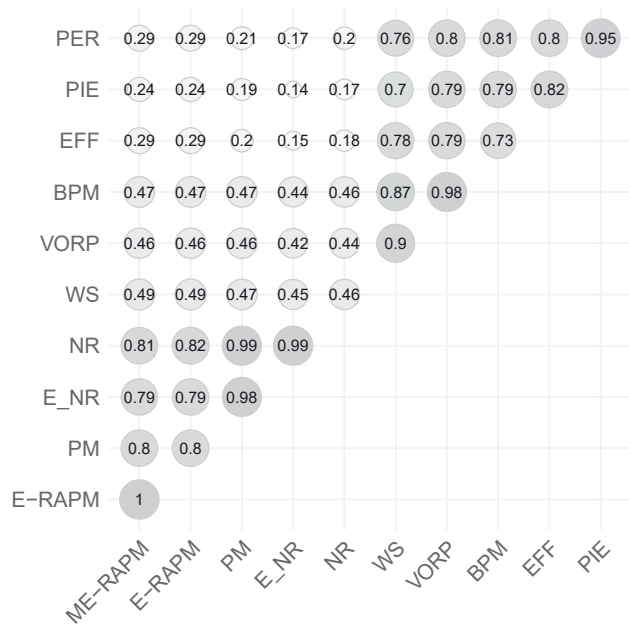


Figure 5: 2022–2023 NBA season. Spearman correlation matrix visualization compares some players’ efficiency measures, and the RAPM estimated players’ effects (E-RAPM and ME-RAPM).

such as the official NBA for EFF, E_NR, PM, PIE, and NR (<https://www.nba.com/stats/players>), the Basketball Reference for VORP, WS, and BPM (https://www.basketball-reference.com/leagues/NBA_2023_advanced.html), and the ESPN for PER (<https://insider.espn.com/nba/hollinger/statistics>). The complete set of performance measures can be found in the Supplementary Material (Section S3.2).

4.2 Visualizing player and lineup effects

In this section, we present a graphical representation to visualize the player and lineup effects estimated by the ME-RAPM model. These visualizations help to identify individual player contributions as well as the combined impact of specific lineups on team performance. This provides an example of the types of visual analysis that will be used in the next section.

The non-negative characterization of players’ effects introduced in the proposed specification enables a more intuitive visualization of the total lineup performance. By summing the estimated effects of the players, we can express a lineup’s overall impact in a clear and informative way.

To illustrate this, Figure 6 compares Denver Nuggets, who won the 2022–2023 championship, and the worst-performing team of the regular season, San Antonio Spurs. The figure shows each team’s best and worst lineups and displays the total lineup effect as a single white bar, allowing

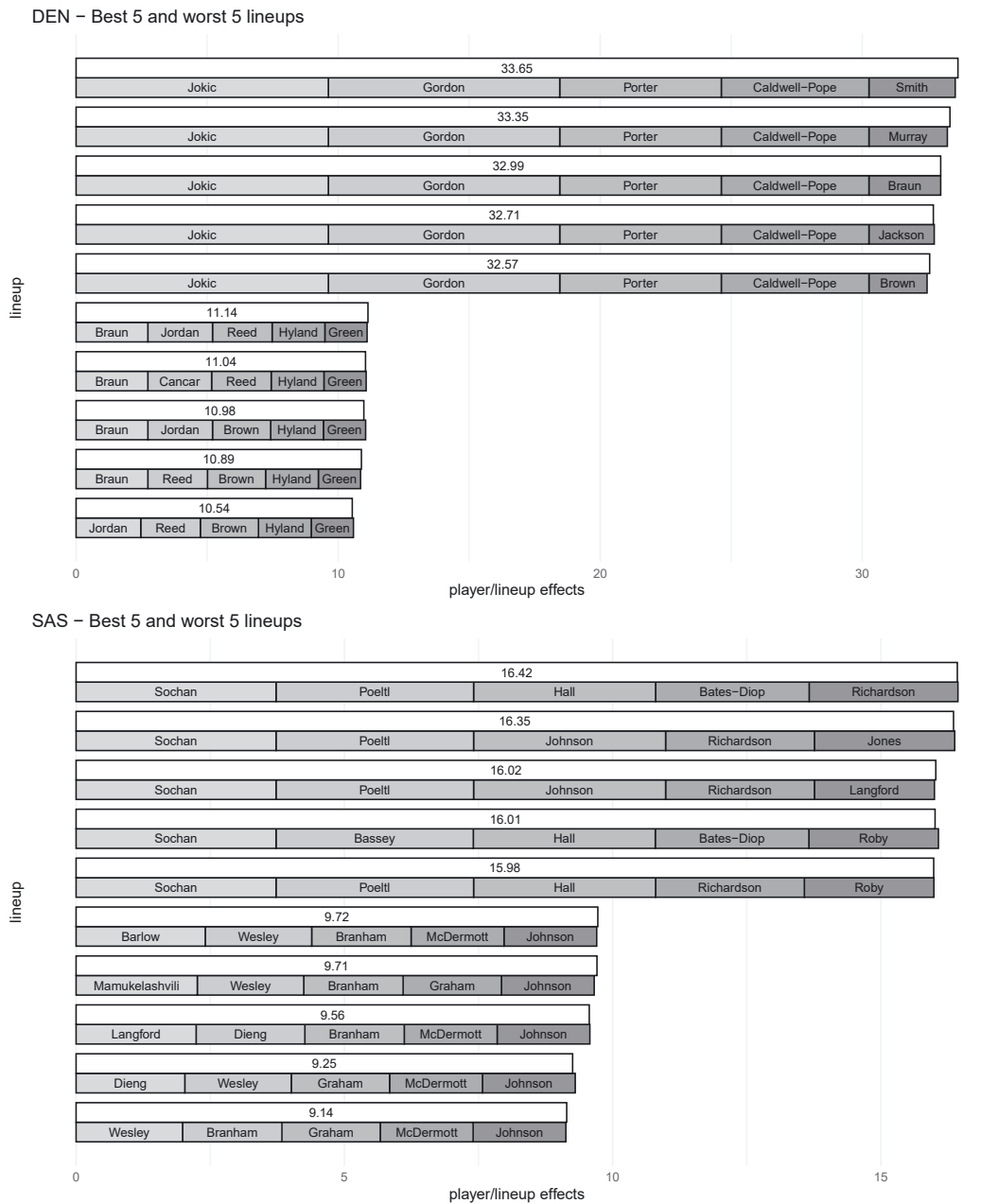


Figure 6: 2022–2023 NBA season. The five best and worst lineups total effects compositions. DEN (NBA champion) versus SAS (worst western conference team).

for immediate visual comparison of overall performance. The stacked grey bars identify the individual player effects, making it easy to see how each player contributes to the lineup’s total impact, since all the contributions are positive. The difference between the stack of grey bars and the white bar represents the synergistic effect. Notice that, comparing the two teams, we consider the x-axis scale to be free for better readability of the results. Regarding the best lineups, Denver Nuggets’ total effects are almost double those of San

Antonio Spurs. On the contrary, the estimates for worst lineups are similar. The lineup effects in the considered case are negligible, and consequently, the lineup total performance is mainly equivalent to the sum of pure players’ effects.

4.3 Splitting home and away effects

Looking at the results in Table 2, one can appreciate the effects of the modification introduced in Section 2.1. As a

Table 2: 2022–2023 NBA season. ME-RAPM model with home and away-specific effects estimation results. Along with player and lineup effects, the models considered a simple dummy variables model matrix for the period of the game. For all the parameters, the mean of the posterior distribution is reported along with the 95 % credible interval, effective sample size (ESS), and \hat{R} fit measures.

Parameters	Mean	2.5 %	97.5 %	ESS	\hat{R}
ME-RAPM model with home and away-specific effects					
β_h	-5.155	-12.296	1.821	3,659.614	1.001
β_2	-1.499	-4.165	1.150	10,674.630	1.000
β_3	-1.229	-3.925	1.446	11,890.102	1.000
β_4	-0.494	-3.261	2.259	11,217.020	1.000
$\log(\sigma_\varepsilon)$	5.496	5.489	5.503	14,721.758	1.000
$\log(\sigma_{\gamma,h})$	1.631	1.381	1.854	3,833.422	1.001
$\log(\sigma_{\mu,h})$	0.402	-2.334	1.736	1,348.340	1.000
$\log(\sigma_{\gamma,a})$	1.188	0.752	1.523	3,616.516	1.000
$\log(\sigma_{\mu,a})$	0.077	-2.680	1.468	1,976.188	1.002
ELPD	-222,001.404 (158.480)				

result of the half-normal specification of players’ performances, splitting players and lineups effects into home and away-specific terms, the home advantage is captured by the difference in the estimated size of the players’ performances. In fact, this difference also determines an average difference in favor of the home-specific performances as illustrated in Figure 7, where most points lie above the bisector line. The difference between the estimated regression model and the bisector line measures the average home advantage. One can also notice that this advantage is relatively more relevant for the worst-performing players. Additional evidence of how estimated players’

effects capture the home effect is given in the Supplementary Material Section S4 (Figure 4.1). Finally, the lineup effects also show larger variability when playing at home; however, these terms are less relevant, and their credible intervals exhibit a substantial overlap.

5 Analysis of historical NBA data

This section collects the results obtained using historical NBA datasets to study the evolution of teamwork relevance over the years. We estimate the models separately for each available season to evaluate the changes highlighted in the preliminary analyses in more detail. The estimation results show that the strength of synergies among players, characterizing the 1980s and 1990s, is weakening over time. The early 2000s represent, from this point of view, a transition period in which the synergistic effects are still present, but their relevance slowly decreases. Today, the lineup in the NBA can be merely considered a “sum of individuals”, as clearly arises from the following data analysis. In the case of the 2018–2019 season, we compare the results of the NBA analysis with those based on Euroleague data. The model estimation results will include the scale parameters of players and lineup effects as the estimation procedure treats them as parameters driving the sampling process. As the number of estimated effects is huge, the model estimation results will be mainly summarized graphically to make their inspection feasible. As observed in the analyses of the previous section, synergistic effects in the NBA courts in recent years are mostly negligible, but some interesting patterns have emerged over the years.

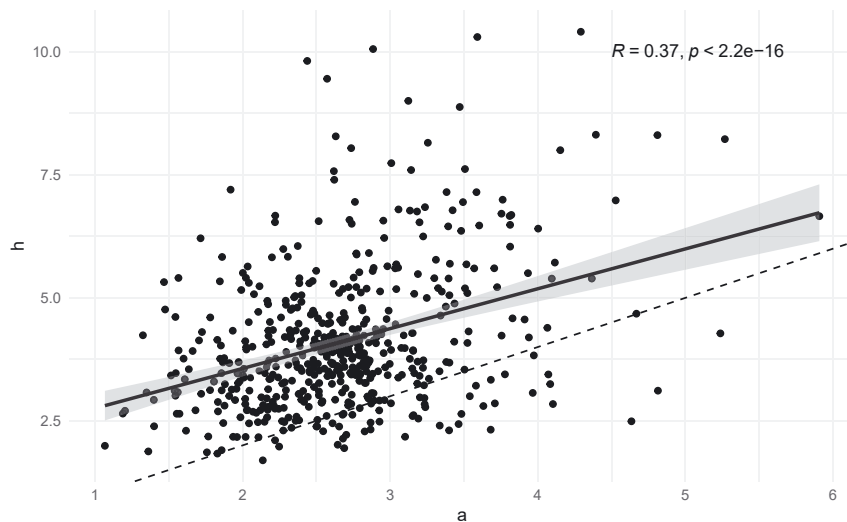


Figure 7: 2022–2023 NBA season. A direct comparison of home and away-specific estimated players’ performances.

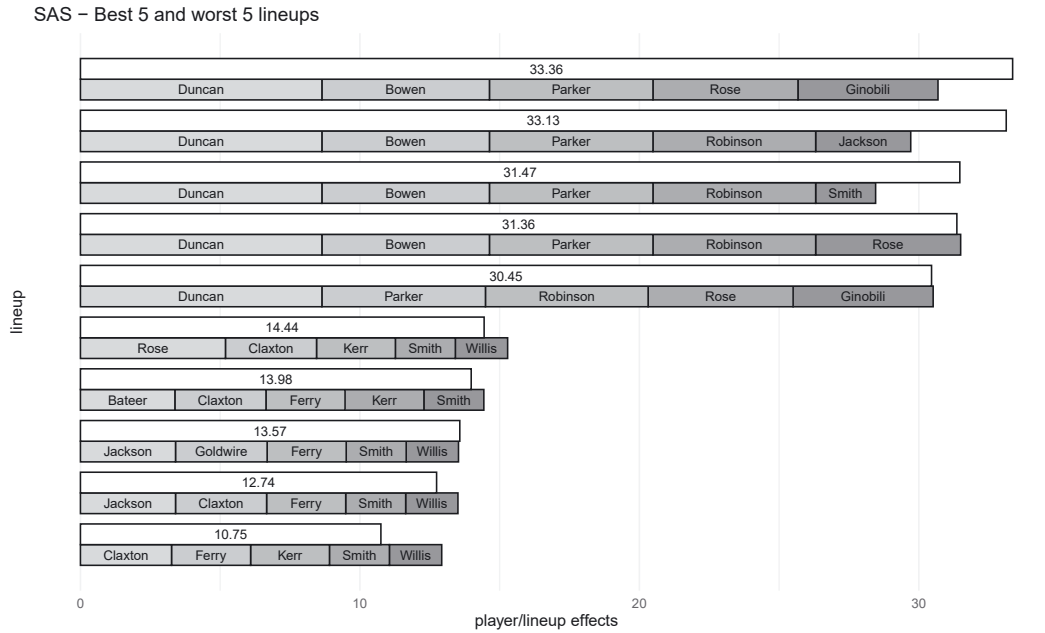


Figure 8: 2002–2003 NBA season. The five best and worst lineups total effects compositions. NBA champion SAS.

5.1 The evolution of teamwork in NBA games

The following analyses are meant to answer the first two research questions introduced in Section 1. In particular, the first analysis developed on the 2002–2003 NBA season dataset shows that, in the early 2000s, the player interactions (synergies) played a relevant role. To give an example, Figure 8 shows the composition of the total lineup effects for the five best and worst lineups of the NBA title winner, San Antonio Spurs. On the one hand, all the first five lineups include Duncan, Bowen, and Parker. Ginobili or Robinson's presence seems to generate a similar total effect for the four-man units. As the second and fourth lineups differ for only one player, namely Rose or Jackson, the estimation results suggest that Robinson and Rose's simultaneous presence on the court brings a negative synergistic effect. All the other solutions present a total lineup effect that is more prominent than the mere sum of players' effects. Conversely, all the worst five lineups show adverse synergistic effects. Moreover, one of the five best lineups includes Smith, who presents a low individual contribution, demonstrating that a positive synergistic players' behavior may play a relevant role in the evaluation of lineups. These players' synergistic effects have a more prominent role than in the 2022–2023 season results reported for the same team in Figure 6. All these results are in line with the general assessment made in Lefko (2016), where mostly qualitative evaluations are made

on the importance of players' synergies and the superstars' role.

More generally, the ME-RAPM model is used to study the evolution of collaborative behavior among players on the court. In particular, to answer the second research question, one can consider the scale parameter of the lineup effects as a proxy of the strength of collaboration among players. Figure 9 shows that this parameter declines over time. For space reasons, the figure only reports six equally spaced NBA seasons from 2002–2003 to 2022–2023. The plot also shows that the players' effects variance parameter presents substantial stability over time. This phenomenon is possibly connected with the prevalence of superstars' performances over teamwork. In other words, the effects of the players are far more relevant in the modern NBA than the synergies among them.

Interestingly, this behavior did not occur in the Euroleague 2018–2019 championship (see Grassetti et al. 2021), supporting the idea that European basketball still emphasizes teamwork. Moreover, the NBA championship is characterized by players' superior athleticism, whereas the Euroleague puts more emphasis on basketball tactical aspects, as highlighted in Mandić et al. (2019). To support these last considerations, we use the Euroleague dataset to estimate the ME-RAPM model and compare it with the results of the concurrent NBA season. Table 3 collects the

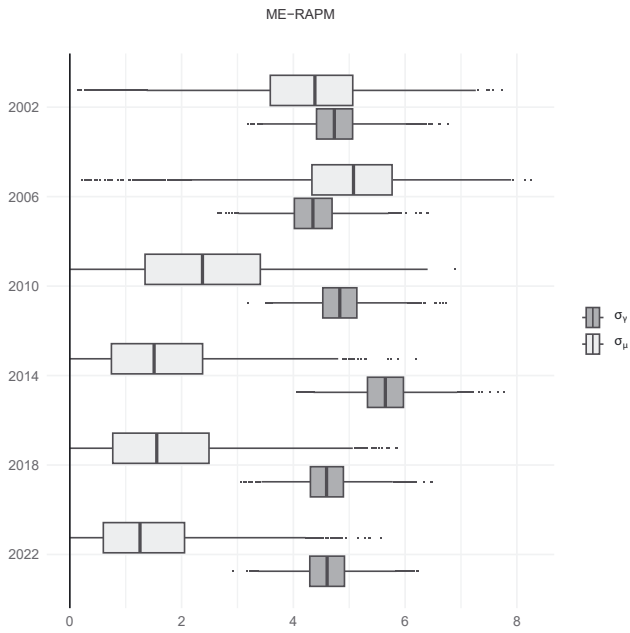


Figure 9: Evolution of scale parameters for lineup and player effects over the NBA seasons.

Table 3: 2018–2019 season. ME-RAPM estimates for the Euroleague and NBA data: the posterior mean and 95 % credible intervals.

Parameters	Euroleague championship			NBA championship		
	Mean	2.5 %	97.5 %	Mean	2.5 %	97.5 %
β_h	5.244	1.230	9.381	2.460	0.635	4.306
β_2	-0.527	-6.384	5.247	1.394	-1.202	4.072
β_3	-0.461	-6.443	5.422	-0.940	-3.500	1.668
β_4	-5.568	-11.485	0.551	0.484	-2.172	3.019
$\log(\sigma_\epsilon)$	5.385	5.368	5.403	5.480	5.473	5.487
$\log(\sigma_\gamma)$	1.500	0.951	1.892	1.523	1.331	1.709
$\log(\sigma_\mu)$	1.822	-0.383	2.474	0.181	-2.632	1.428

model estimation results for the two datasets and highlights the differences between the two cases. European competition’s results present a behavior similar to that observed in the first decade of the 21st century in the NBA. This analysis clarifies that the synergistic lineup effects are more relevant in the Euroleague than in the NBA data. Moreover, the synergistic effect estimated using Euroleague data is even more prominent than in the early 2000s NBA. For the 2002–2003 NBA season, the posterior mean of $\log(\sigma_\mu)$ is 1.401. The graphical representation of the composition of total lineup effects for the Euroleague Champion Team, CSKA Moscow, is given in Figure 10. The reported results show that the synergy among players in this league determines substantial departures from the sum of players’ effects. For instance, the best lineup presents a total players

effect that is lower than the other four best lineups, yet its total expected performance greatly benefits from the interaction among the players.

5.2 The home and away specific results

Considering the estimation of the home and away-specific performance parameters, it is possible to identify some interesting patterns over time. Figure 11 summarizes the results of these estimates for the six seasons analyzed in the previous section. One can notice that the scale of player effects varies slightly over time, but the home parameters are generally larger than the away ones, indicating that the volatility and the mean of the players’ performance is larger when playing at home. This is consistent with the presence of a home advantage in professional basketball previously highlighted in Section 4.1. The opposite result can be observed when considering the lineup-specific scale parameters. In this case, away parameters show values that in most cases are larger than the home ones. Notwithstanding, the magnitude of these effects shows a decreasing trend in the first 10 years of observation, and they are practically negligible in the last seasons. This trend confirms the results highlighted in the previous section, which shows a decreasing importance of teamwork in the modern NBA.

5.3 A focus on COVID-19 lockdown period

The sports analytics literature counts many research papers developed after the COVID period to address the changes observed in sports practice due to the lockdown and the particular environment in which the athletes trained and played. Many papers focus on the so-called bubble effect on players’ and teams’ performances, and propose to treat this period as a pseudo-experimental framework in which the effect of fans’ absence and the concurrent COVID-specific conditions serve as treatment/exposure factors. Focusing on basketball analytics, Alonso et al. (2022) and Lu et al. (2022) use the COVID-games as a natural experiment to test the effect of fan presence on the so-called home advantage. Gong (2022) examines the effect of fans’ presence on referee and home biases. Markwell et al. (2021) studies the effect of bubble-games on the shooting performances of NBA players. Caselli et al. (2024) focuses on the changes in performance of black players related to the lowered racial pressure. McHill and Chinoy (2020) uses this pseudo-experiment to evaluate the changes in players’ performances due to frequent travels and circadian disruptions that characterize the NBA championship. Paulsen and Boucot (2023) studies the psychological effects of the family absence during the NBA bubble games. In this regard, we decided to adopt the

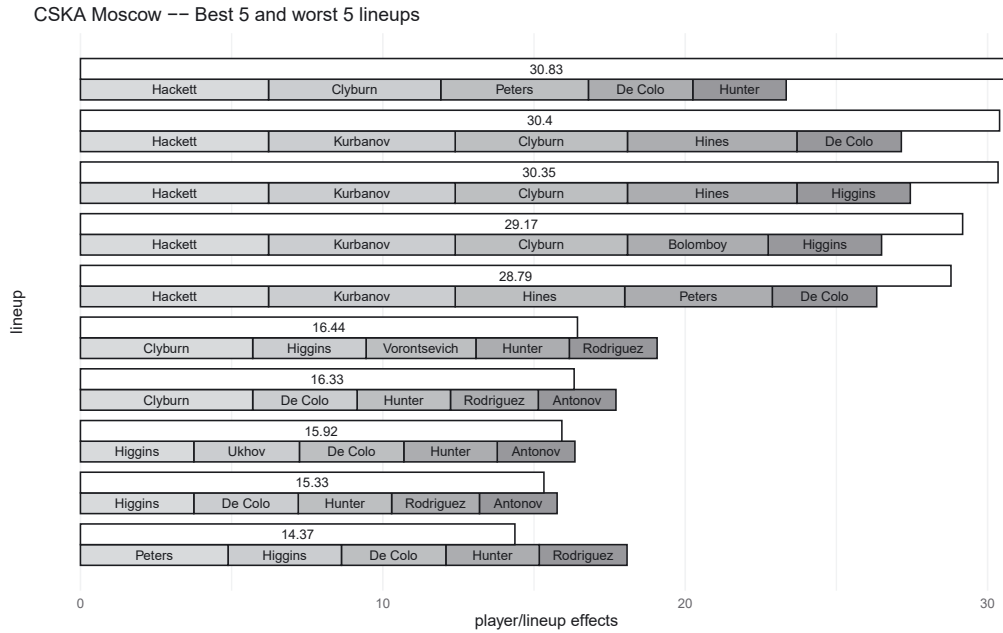


Figure 10: 2018–2019 Euroleague season. The total lineup effects of the five best and worst lineups – Euroleague championship winner CSKA Moscow.

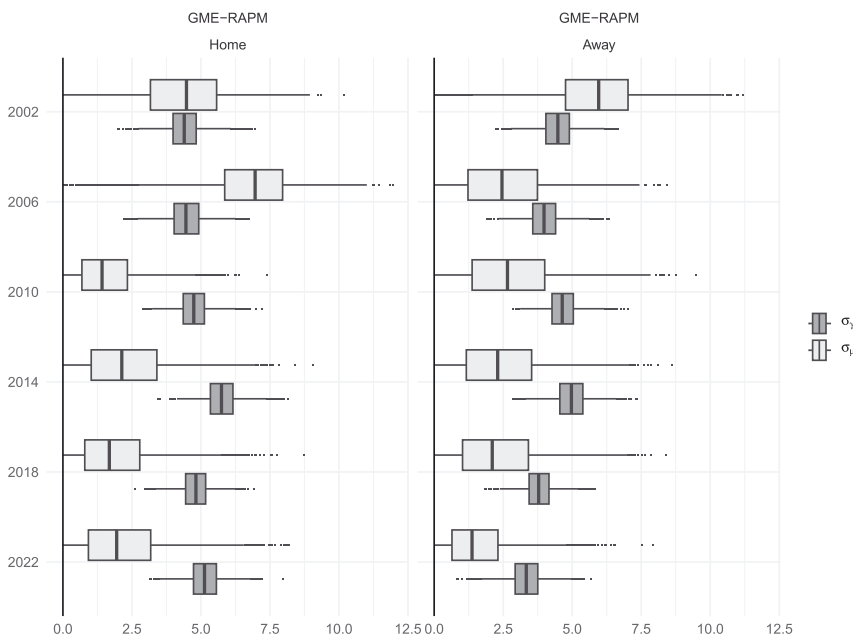


Figure 11: Evolution of scale parameters for home and away-specific lineup and player effects estimated by the GME-RAPM model formulation over the NBA seasons.

proposed models to determine the trend in the sizes of the lineup and players’ effects during the COVID period.

As discussed in the previous section, the estimated values of lineup-specific effects size parameters seem to be constantly low after season 2010/2011, but they show a noticeable trend during the period 2019–2020. These estimates, summarized in Figure 12, reveal a clear increase in

the relevance of lineups’ effects in 2019–2020, and a gradual return to the pre-COVID levels. On the contrary, the size of players’ effects exhibits a mostly constant behavior throughout the entire period. This behavior has also been observed in many other studies focused on the changes in sports performance due to the lockdown and games played without spectators. In our opinion, this peculiar behavior

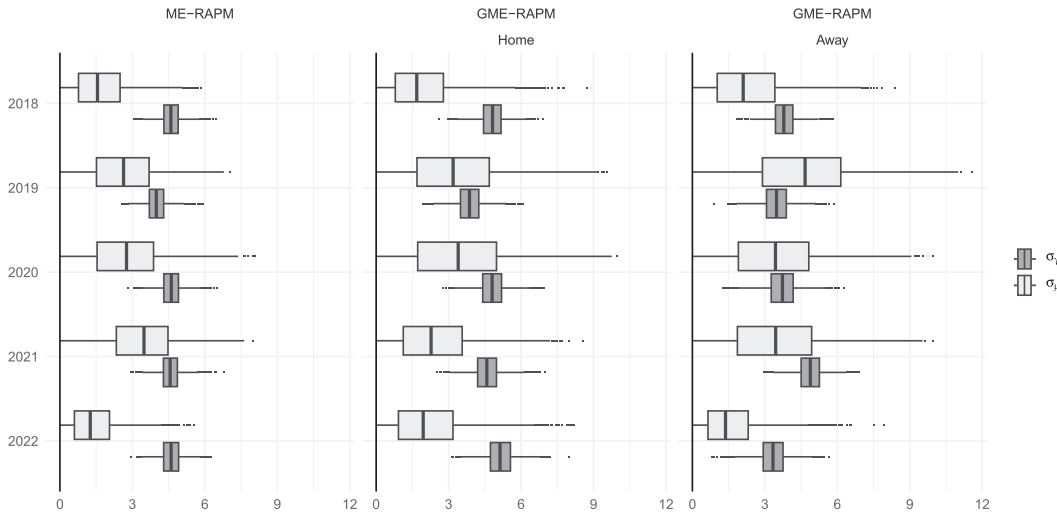


Figure 12: Evolution of scale parameters for generic (ME-RAPM), home and away-specific (GME-RAPM) lineup and player effects from 2018 to 2022, COVID related NBA seasons.

can be connected with the capacity of the top players to assume a dominant role in response to the fans' presence and pressure. Their will to be dominant and involve the fans during the game obviously lowers during the lockdown. From a different perspective, teamwork assumes a more significant role in the absence of fans.

Further insights emerge from comparing home and away-specific estimated parameters. As illustrated in Figure 12 home and away effects allow us to study the difference between the two measures. This analysis suggests that the home advantage effect is captured by the difference observed in the σ_γ parameters. To complement this analysis, the Supplementary Material (Section S4) presents a direct comparison of players' effects at home versus away. The same comparison is also provided for the effects of the lineups. Moreover, in contrast to many other studies, the home effect appears to persist during the COVID period.

6 Conclusions

Applying proposed models to NBA data, we highlighted that players' synergistic effects are substantially irrelevant in today's NBA. The availability of historical datasets, dating back to the early seasons of the 2000s, allowed us to study this phenomenon over the years. When collecting pieces of evidence from separated model estimation results, one can notice that the magnitude of lineup effects, identified with the scale parameter, decreases over time. In other words, in today's NBA, the teams' performance is mainly related to the single players' efforts. By contrast, the interaction

among players was much more relevant in the early years of the new millennium. Moreover, a comparative analysis of the NBA and Euroleague seasons in 2018–2019 shows that players' synergistic behavior was still relevant in a less physical championship, such as the European one. The analysis also considered splitting the performances of players and lineups into home and away-specific parameters. Analyzing these results, one can notice that the general patterns are also observed in this framework, and the difference between home and away parameters is relevant for players' effects only.

These results stimulated further study over the COVID period. The results of this specific analysis are interesting because in accordance with many other studies, the empirical evidence shows that the behaviour of professional players mutates during the lockdown period, showing how much the presence of fans can affect the gameplay of this sport. In particular, during the lockdown, teamwork assumes a more relevant role in the outcome of NBA basketball games.

The proposed model specification, obtained by introducing non-negative constraints on the E-RAPM players' effects, allows us to use and visualize them as parts of a total lineup performance, which is positively or negatively affected by the players' interactions. In fact, under this specification, the total lineup effects can be defined as the sum of non-negative players' contributions and lineup-related adjustments. This solution represents a valuable tool for data summarization. In particular, the estimated performances can be used to evaluate the players in the roster construction phase and choose the best lineups during the games. Moreover, coaches can dynamically choose the best

solution at each shift by comparing opponents' lineups and the available five-man unit efficiencies. Using ME-RAPM results, the identification of single players' contributions is straightforward and can be of practical usefulness in team management and coaching activities.

Future work will investigate the possibility of generalizing the proposed models to a bivariate outcome in a fashion similar to double Poisson models in soccer (Shen 2014). In this context, players and lineups effects can be further extended to include offensive and defensive-specific terms as in Damoulaki et al. (2025). One can find a glimpse of this generalization in the Supplementary Material in Section S5. This generalization is relatively easy for players' effects, but the number of observations for each lineup is limited, and this might lead to mild identifiability issues. In this respect, one may also consider the definition of a more comprehensive outcome that, in addition to the scored points, includes some other relevant offensive and defensive plays, as suggested by the four-factor theory introduced in Oliver (2004), and recently revisited in Poropudas and Halme (2023). Although the current work is built within the basketball framework, it can be easily applied to other team sports, such as ice hockey and soccer (see Macdonald 2011; Schultze and Wellbrock 2018, for some examples of RAPM model application).

Acknowledgments: We thank the anonymous reviewers and the editor for their careful reading of our manuscript and their many insightful comments and suggestions, which helped improve the paper. The authors would like to acknowledge Giuseppe Alfonzetti, Ruggero Bellio and Paolo Vidoni for their insightful conversations on this research.

Research ethics: Not applicable.

Informed consent: Not applicable.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission. L.G. designed the study and performed the main statistical analyses, M.L. di S.M. made the data conditioning, aggregation and data description steps, VM supervised the research and wrote the first version of the paper. All the authors edited and proof-read the manuscript. In particular, L.G. is responsible for Sections 4.1 and 5, M.L. di S.M. for Sections 3, 4.2, 4.3 and 6, and V.M. for Sections 1, 2 and 5.3.

Use of Large Language Models, AI and Machine Learning

Tools: We used Grammarly and the default google search tool to improve basic language aspects.

Conflict of interest: None.

Research funding: This research is funded by PRIN 2022: Project prot. n. 2022R74PLE UGOV code PRIN_2022_MAMELI

_DIES CUP G53D23001870006 funded by the European Union NextGenerationEU PNRR M4C2 inv 1.1.

Data availability: The data that support the findings of this study are accessible from BigDataBall (<https://www.bigdataball.com>), but restrictions apply to these data, which were used under license for the current study, and so are not publicly available.

Software availability: The software used in the estimation of models is R and rstan package. The Stan code is available upon request to the authors.

References

- Aalbers, B. and Van Haaren, J. (2019). Distinguishing between roles of football players in play-by-play match event data. In: *Machine learning and data mining for sports analytics*. Springer, Cham, CH, pp. 31–41.
- Alonso, E., Lorenzo, A., Ribas, C., and Gómez, M.A. (2022). Impact of COVID-19 pandemic on HOME advantage in different European professional basketball leagues. *Percept. Mot. Skills* 129: 328–342.
- Assunção, R. and Pelechrinis, K. (2018). Sports analytics in the era of big data: moving toward the next frontier. *Big Data* 6: 237–238.
- Baumer, B.S., Matthews, G.J., and Nguyen, Q. (2023). Big ideas in sports analytics and statistical tools for their investigation. *WIREs Comput. Stat.* 15: e1612.
- Cabarkapa, D., Deane, M.A., Fry, A.C., Jones, G.T., Cabarkapa, D.V., Philipp, N.M., and Yu, D. (2022). Game statistics that discriminate winning and losing at the NBA level of basketball competition. *PLoS One* 17: e0273427.
- Caselli, M., Falco, P., and Somekh, B. (2024). Inside the NBA bubble: how black players performed better without fans. *J. Popul. Econ.* 37: 39.
- Clay, C.D. and Clay, E.K. (2014). Player rotation, on-court performance and game outcomes in NCAA men's basketball. *Int. J. Perform. Anal. Sport* 14: 606–619.
- Cohan, A., Schuster, J., and Fernandez, J. (2021). A deep learning approach to injury forecasting in NBA basketball. *J. Sports Anal.* 7: 277–289.
- Damoulaki, A., Ntzoufras, I., and Pelechrinis, K. (2025). Lasso multinomial performance indicators for in-play basketball data. *Comput. Stat.* 40: 2157–2181.
- Deshpande, S.K. and Jensen, S.T. (2016). Estimating an NBA player's impact on his team's chances of winning. *J. Quant. Anal. Sports* 12: 51–72.
- Efron, B. and Hastie, T. (2016). *Computer age statistical inference*. Cambridge University Press, Cambridge, UK.
- Engelmann, J. (2017). Possession-based player performance analysis in basketball (adjusted+/- and related concepts). In: *Handbook of statistical methods and analyses in sports*. Chapman and Hall/CRC, New York, NY, pp. 215–228.
- Gabry, J. and Mahr, T. (2024). bayesplot: plotting for Bayesian models. R package version 1.11.1, Available at: <https://mc-stan.org/bayesplot/>.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1: 515–534.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (1995). *Bayesian data analysis*, 1st ed. Chapman and Hall/CRC, New York, NY.

- Gong, H. (2022). The effect of the crowd on home bias: evidence from NBA games during the COVID-19 pandemic. *J. Sports Econ.* 23: 950–975.
- Gong, H. and Chen, S. (2024). Estimating positional plus-minus in the NBA. *J. Quant. Anal. Sports* 20: 193–217.
- Grassetti, L., Bellio, R., Di Gaspero, L., Fonseca, G., and Vidoni, P. (2021). An extended regularized adjusted plus-minus analysis for lineup management in basketball using play-by-play data. *IMA J. Manag. Math.* 32: 385–409.
- Grassetti, L., Bellio, R., Fonseca, G., and Vidoni, P. (2019). Estimation of lineup efficiency effects in basketball using play-by-play data. In: *Book of Short Papers of the 2019 SIS Conference*. Pearson, pp. 363–370.
- Harris, A.R. and Roebber, P.J. (2019). NBA team home advantage: identifying key factors using an artificial neural network. *PLoS One* 14: e0220630.
- Hvattum, L.M. (2019). A comprehensive review of plus-minus ratings for evaluating individual players in team sports. *Int. J. Comput. Sci. Sport* 18: 1–23.
- Jones, M.B. (2007). Home advantage in the NBA as a game-long process. *J. Quant. Anal. Sports* 3, <https://doi.org/10.2202/1559-0410.1081>.
- Kvam, P. and Sokol, J.S. (2006). A logistic regression/Markov chain model for NCAA basketball. *Nav. Res. Logist.* 53: 788–803.
- Lefko, J. (2016). *Spurs nation: major moments in San Antonio basketball*. Trinity University Press, San Antonio, TX.
- Lu, P., Zhang, S., Ding, J., Wang, X., and Gomez, M.A. (2022). Impact of COVID-19 lockdown on match performances in the national basketball association. *Front. Psychol.* 13: 951779.
- Macdonald, B. (2011). A regression-based adjusted plus-minus statistic for NHL players. *J. Quant. Anal. Sports* 7, <https://doi.org/10.2202/1559-0410.1284>.
- Mandić, R., Jakovljević, S., Erčulj, F., and Štrumbelj, E. (2019). Trends in NBA and Euroleague basketball: analysis and comparison of statistical data from 2000 to 2017. *PLoS One* 14: e0223524.
- Manner, H. (2016). Modeling and forecasting the outcomes of NBA basketball games. *J. Quant. Anal. Sports* 12: 31–41.
- Marković, S., Čuk, I., and Živković, A. (2020). The impact of information technologies on the scouting process in sports games. In: *Proceedings of the Sinteza 2020—International Scientific Conference on Information Technology and Data Related Research*. Singidunum University, pp. 240–245.
- Markwell, L.T., Strick, A.J., and Porter, J.M. (2021). No fans, no problem: an investigation of audience effects on shooting performance in professional basketball. *J. Mot. Learn. Dev.* 10: 212–223.
- McHill, A.W. and Chinoy, E.D. (2020). Utilizing the national basketball association’s COVID-19 restart “bubble” to uncover the impact of travel and circadian disruption on athletic performance. *Sci. Rep.* 10: 21827.
- Metulini, R. and Gnecco, G. (2022). Measuring players’ importance in basketball using the generalized Shapley value. *Ann. Oper. Res.* 325: 441–465.
- Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis*. Potomac Books, Inc., Washington, DC.
- Omidiran, D. (2011). A new look at adjusted plus/minus for basketball analysis. In: *Proceedings of the 2011 MIT Sloan sports analytics conference*.
- Özbalta, E., Yavuz, M., and Kaya, T. (2021). National basketball association player salary prediction using supervised machine learning methods. In: *Proceedings of the INFUS 2021 conference*. Springer, Cham, CH, pp. 189–196.
- Papaspiliopoulos, O., Roberts, G.O., and Sköld, M. (2007). A general framework for the parametrization of hierarchical models. *Stat. Sci.* 22: 59–73.
- Paulsen, R.J. and Boucot, O. (2023). Playing in a pandemic: the impact of family on performance in the NBA’s “bubble”. *Manag. Decis. Econ.* 44: 2102–2109.
- Poropudas, J. and Halme, T. (2023). Dean Oliver’s four factors revisited, *arXiv preprints*, <https://arxiv.org/abs/2305.13032>.
- R Core Team (2024). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, Available at: <https://www.R-project.org/>.
- Ribeiro, H.V., Mukherjee, S., and Zeng, X.H.T. (2016). The advantage of playing home in NBA: microscopic, team-specific and evolving features. *PLoS One* 11: e0152440.
- Rosenbaum, D.T. (2004). *Measuring how NBA players help their teams win*, Available at: <https://www.82games.com/comm30.htm>.
- Ruiz, F.J.R. and Perez-Cruz, F. (2015). A generative model for predicting outcomes in college basketball. *J. Quant. Anal. Sports* 11: 39–52.
- Sarlis, V. and Tjortjis, C. (2020). Sports analytics — evaluation of basketball players and team performance. *Inf. Syst.* 93: 101562.
- Sarlis, V. and Tjortjis, C. (2024). Sports analytics: data mining to uncover NBA player position, age, and injury impact on performance and economics. *Information* 15, <https://doi.org/10.3390/info15040242>.
- Schultze, S.R. and Wellbrock, C.-M. (2018). A weighted plus/minus metric for individual soccer player performance. *J. Sports Anal.* 4: 121–131.
- Shen, K. (2014). Data analysis of basketball game performance based on bivariate Poisson regression model. *Comput. Model. N. Technol.* 18: 474–479.
- Sill, J. (2010). Improved NBA adjusted +/- using regularization and out-of-sample testing. In: *Proceedings of the 2010 MIT Sloan sports analytics conference*.
- Stan Development Team (2024). RStan: the R interface to Stan. R package version 2.32.6, Available at: <https://mc-stan.org/>.
- Terner, Z. and Franks, A. (2021). Modeling player and team performance in basketball. *Annu. Rev. Stat. Appl.* 8: 1–23.
- Tiedemann, T., Francksen, T., and Latacz-Lohmann, U. (2010). Assessing the performance of German Bundesliga football players: a non-parametric metafrontier approach. *Cent. Eur. J. Oper. Res.* 19: 571–587.
- Vehtari, A., Gelman, A., and Gabry, J. (2017). loo: efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 1.0.0, Available at: <https://mc-stan.org/loo/>.
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: an improved \hat{R} for assessing convergence of MCMC (with discussion). *Bayesian Anal.* 16: 667–718.
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*, 2nd ed. Springer, Cham, CH.
- Yang, J. and Lu, C.-H. (2012). Predicting NBA championship by learning from history data. In: *Proceedings of artificial intelligence and machine learning for engineering design*.
- Zuccolotto, P. and Manisera, M. (2020). *Basketball data science: with applications in R*. CRC Press, New York, NY.

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/jqas-2024-0138>).