Research article

Journal of Sports Analytics
Vol. 12(0): 1–27

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/22150218261468843
journals.sagepub.com/home/san
Sage | IOS Press

WINSCORE revisited: A model-based evaluation of player performance in the NBA and EuroLeague

Gabriele Carta[1] and Carlo A Favero[1]

Abstract

In professional basketball, despite progress in Sport Analytics, widely used methods of player evaluation still rely on box-score statistics aggregated without formal justification for the weights used. The popular Performance Index Rating (PIR) is a prominent example. This paper revisits WINSCORE, a possession-based player evaluation methodology and reinterprets it within a structural, model-based framework that links individual statistics to team outcomes through the efficient management of possessions. In this framework, the weights assigned to player statistics are derived endogenously from the data, providing a coherent foundation for measuring performance. Using NBA and EuroLeague data, we evaluate the model at the team level through goodness-of-fit evidence, residual analysis, encompassing tests against PIR, and comparisons with alternative advanced measures, including Four Factors and adjusted box-score plus-minus indicators. The results show that the model accounts well for variation in team performance, displays no major residual patterns, and adds explanatory content beyond existing metrics. We then extend WINSCORE to construct player performance measures for season-level and single-game analysis. These measures incorporate opponent adjustments and address playing-time endogeneity. The resulting indicators provide transparent tools for player assessment, coaching, and tactical analysis, with a measurement-oriented rather than causal interpretation.

Keywords

Basketball Analytics, box-score statistics, EuroLeague, NBA, player Performance Evaluation, team Efficiency, WINSCORE Model

Received: 19 March 2026; accepted: 25 June 2026

Introduction

The question "How good is this player, really?" is central to basketball and, more generally, to all sports. Despite the progress made in Sport Analytics and its applications to professional basketball, the most popular methods of player evaluation still rely on box-score statistics aggregated without a formal justification for the weights used and without a clear theoretical link to team performance. The Performance Index Rating (PIR), widely used in leagues such as the NBA and the EuroLeague, is a prominent example of this approach.

This paper revisits the WINSCORE methodology, originally proposed by Berri et al. (2006), and reinterprets it within a structural, model-based framework that explicitly links individual statistics to team outcomes through the efficient management of possessions. This reinterpretation allows the weights assigned to player statistics to be derived endogenously from the data, rather than imposed a priori, and provides a coherent foundation for measuring player performance.

We show that a model-consistent measure of player performance, explicitly linked to team performance, is simple, transparent, and easily applicable both to single-game evaluation and to season-level analysis, given the availability of standard box-score statistics.

[1]Department of Economics and Bocconi Sport, Bocconi University, Milan, Italy

Corresponding author:
Carlo A Favero, Department of Economics and Bocconi Sport, Bocconi University, Milan, Italy.
Email: carlo.favero@unibocconi.it

Although our approach is based on a structural model built around individual box-score statistics, it can also be extended to richer settings in which a player's value depends on the system around him, including statistical models based on modern tracking data and Machine Learning methods.[1]

It is important to clarify the scope of this comparison. The purpose of the paper is not to evaluate whether a box-score-based structural model can outperform modern deep-learning, player-tracking, or spatio-temporal analytics frameworks to evaluate single players' performance. These approaches use richer data and are designed to answer partly different questions, especially those involving spatial positioning, lineup interactions, shot quality, defensive assignments, and possession-level dynamics. Our objective is instead to ask whether the same information set used by widely adopted box-score indexes such as PIR can be organized within a transparent model-based framework that links individual statistics to team outcomes. In this sense, modern tracking-data and machine-learning methods are complementary to the present approach rather than direct empirical competitors. As clearly illustrated by Page (2018) in his insightful book on model thinking, organizing and interpreting data through models has become a core competency across many domains, including business strategy, urban planning, economics, medicine, engineering, actuarial science, and environmental science, among others. In this paper, we take a model-based approach to constructing a synthetic indicator of individual basketball performance by optimally weighting several individual statistics while explicitly taking team performance into account.

The increasing availability of detailed performance data has transformed the analysis of professional basketball. However, the most popular and widely used metrics outside the scientific community still rely on the aggregation of box-score statistics (Zhou and Li, 2024), often without a clear theoretical foundation linking individual actions to team success.

PIR provides a convenient summary of player activity, but it assigns equal weights to fundamentally different actions and does not explicitly account for opponent performance or for the relationship between individual statistics and team outcomes. As a result, it may fail to capture the true contribution of players to winning.

This paper adopts a model-based approach grounded in the structure of the game and capable of delivering simple and intuitive ways of aggregating individual statistics. The key insight is that wins are generated by the efficient management of possessions, as originally emphasized by Hollinger (2002) and Oliver (2004), then formalized by Berri and Eschker (2005) and popularized by Berri et al. (2006). This approach, known as WINSCORE, is closely related to the Four Factors framework (Oliver, 2004), which identifies shooting efficiency, turnovers, rebounding, and free throws as the key drivers of team success. WINSCORE can therefore be embedded within a model-based framework that directly links individual player statistics to team performance outcomes, yielding a more grounded and interpretable measure of player value.

A model separates data into two components: structure and noise. The structure captures the essential features of the data-generating process that are independent of the observed sample, while noise reflects accidental features specific to a given sample and should not be fitted when the objective is prediction or out-of-sample analysis. Separating structure from noise is crucial in order to avoid the consequences of in-sample overfitting for out-of-sample performance.

Structure typically relates endogenous variables to exogenous variables through functional specifications and parameters. In our analysis, we impose a common functional specification across different leagues, while allowing the parameters to vary in order to capture the specific features of the game across leagues, namely the NBA and the EuroLeague.

Our model-based approach has the limitation that it ignores spatial and interaction complexity. In fact, Iatropoulos et al. (2025) and Petridis and Pelechrinis (2026) apply regularization techniques within a Machine Learning framework to avoid overfitting in this richer context. However, such methods ultimately lead to analytically constructed indexes of player performance that can be combined with our proposal, since they are built on the same general principles but rely on a different methodology.

Our reinterpretation allows us to organize the analysis along the standard steps of structural modeling: specification, estimation, validation, and simulation. In doing so, we show that WINSCORE provides a coherent framework in which the weights assigned to box-score statistics emerge endogenously from the data, rather than being imposed a priori.

Our empirical analysis, based on data from both the NBA and the EuroLeague, delivers two main findings. First, the model provides an excellent fit to team performance and, through encompassing tests, is shown to dominate PIR, and other well known and widely used advanced box score metrics in explaining variation in wins. Second, we implement a validation strategy that exploits variation across both eras and leagues. Within the NBA, we compare the historical sample studied by Berri et al. (2006) with a recent post-COVID sample. Across competitions, we compare the NBA evidence with EuroLeague data. This design allows us to assess whether the estimated relationship between possession efficiency and wins, and the weights implied by model simulation, are stable across different institutional and competitive environments. We interpret this evidence as a test of empirical robustness within the available data, not as a claim of universal structural invariance.

Building on this framework, we extend WINSCORE from the team level to the individual level. This extension is achieved by exploiting a linear approximation of the model around the mean, which allows us to assign weights to individual statistics in a transparent and tractable way. We further refine this approach by addressing the endogeneity of playing time and by introducing a modified measure that compares players with relevant opponents who have similar court time.

Finally, we show how the model can be applied to single-game data, where players can be matched directly with their opponents and individual contributions are also weighted differently according to the outcome of the game. This extension highlights the importance of context in performance evaluation and provides a tool that is particularly relevant for coaching and tactical analysis.

Overall, the contribution of the paper is twofold. From a methodological perspective, it demonstrates how a widely used performance metric can be embedded within a structural, model-based framework. From an applied perspective, it provides a set of tools that allow coaches and analysts to evaluate player performance in a way that is consistent with the underlying mechanics of the game.

The remainder of the paper is structured as follows. Section "Related Literature" places our contribution in the literature. Section "Wins and Box-Score Statistics: Data-analysis Without a Theory" discusses the limitations of measurement without theory and highlights the presence of a stable structure in basketball data. Section "Reinterpreting WINSCORE as a Model-Based Framework" presents the model-based reinterpretation of WINSCORE. Section "Interpreting the Winscore Model" interprets the model and derives weights for player statistics. Sections "From Team to Players: Season Data" and "From Team to Players: Single Game data" extend the framework to player-level measures using season and single-game data, respectively. Section "Conclusion" concludes.

Related literature

The evaluation of player performance in basketball has generated a large and diverse literature, which can be broadly divided into three main approaches: (i) aggregation-based metrics, (ii) regression-based or plus-minus approaches, and (iii) theory-based structural models.

A first strand of the literature focuses on the construction of performance metrics based on the aggregation of box-score statistics. The most widely used example is the Performance Index Rating (PIR), which combines positive and negative actions into a single index. While such measures are simple and widely adopted in practice, they typically lack a clear theoretical foundation and do not explicitly link individual performance to team outcomes. This limitation has been widely recognized in both the academic literature and among practitioners.

A second strand of the literature attempts to overcome these limitations by directly linking player performance to team success through regression-based methods. Early contributions include (Rosenbaum, 2004), who introduced the adjusted plus-minus framework to measure a player's impact on team performance, and (Ilardi and Barzilai, 2008) , who refined this approach by incorporating box-score information. (Kubatko et al., 2007) provide a systematic framework for analyzing basketball statistics using regression techniques, while (Sill, 2010) proposes improvements based on regularization methods to address estimation challenges.

The most modern applications of this approach are proposed by Iatropoulos et al. (2025) and Petridis and Pelechrinis (2026).

Iatropoulos et al. (2025) use data-mining methods to study how NBA player performance evolves quarter by quarter across regular-season and playoff games. Using Association Rule Mining, they identify recurring links between offensive, defensive, ball-handling, tempo, and overall-impact metrics and game success. Their results suggest that defensive metrics become especially important late in games, while offensive efficiency is particularly relevant in playoff contexts.

Petridis and Pelechrinis (2026) propose Lineup-Regularized Adjusted Plus-Minus (L-RAPM). L-RAPM estimates lineup quality, rather than individual player quality, by regressing points per possession on the offensive and defensive lineups involved in each possession. To address the sparsity and noise of lineup data, it regularizes lineup estimates toward informed priors built from individual player RAPM ratings. The method is therefore predictive and shrinkage-based: it adjusts for opponent lineup strength while improving estimates especially for lineups with few or no observed possessions.

These approaches represent an important step forward, as they move beyond simple aggregation and seek to measure player contributions in relation to team outcomes. However, since they are typically based on reduced-form relationships, it is useful to consider alternative strategies that embed performance measurement within a structural model of the game.

A different empirical exercise would be to benchmark WINSCORE against fully predictive machine-learning models trained on box-score statistics. Such a comparison would require specifying a prediction target, training-validation-test splits, tuning rules, and loss functions. We do not pursue this route because the purpose of the paper is not to optimize predictive accuracy with a flexible learner, but to construct a transparent and interpretable measurement framework based on the same information set used by widely adopted indexes such as PIR.[2]

A third strand of the literature adopts such a theory-based perspective, emphasizing possessions as the fundamental unit of analysis in basketball. This approach originates from the work of (Hollinger, 2002) and (Oliver, 2004), who highlight the importance of efficiency in managing possessions. Building on this insight, (Berri and Eschker, 2005) and (Berri et al., 2006) develop the WINSCORE methodology, which links team wins to efficiency measures derived from box-score statistics. This framework provides a more direct connection between individual actions and team success, but its interpretation has remained largely operational rather than explicitly model-based.

More recently, the increasing availability of high-frequency and tracking data has led to further developments in performance evaluation, including approaches based on spatial and dynamic modeling (e.g., (Cervone et al., 2014)). While these contributions offer richer descriptions of player behavior, they often come at the cost of reduced transparency and greater data requirements.

This paper contributes to the literature by bridging the gap between regression-based and theory-based approaches. We reinterpret WINSCORE within a structural, model-based framework that explicitly links player statistics to team outcomes. In doing so, we show that the weights assigned to box-score statistics can be derived endogenously from the model, rather than imposed exogenously or estimated in a reduced-form setting. Moreover, we extend the framework to the evaluation of individual players, both at the season and single-game level, while explicitly addressing issues such as opponent adjustment and the endogeneity of playing time.

By combining a structural interpretation with empirical validation across different leagues and samples, our approach provides a unified framework for player performance evaluation that is both theoretically grounded and practically applicable.

Wins and Box-Score statistics: Data-analysis without a theory

Leagues like the NBA and the Euroleague make Box-Score statistics available online for the different seasons t, the different teams i, and the different players j, for each team all available statistics are matched by the opponent statistics.[3] The following statistics are available:

- Field Goals (FG): Number of field goals made.
- Field Goal Attempts (FGA): Number of field goal attempts.
- Three-Point Field Goals (3P): Number of three-point shots made.
- Three-Point Attempts (3PA): Number of three-point shots attempted.
- Free Throws (FT): Number of free throws made.
- Free Throw Attempts (FTA): Number of free throw attempts.
- Offensive Rebounds (ORB): Number of offensive rebounds secured.
- Defensive Rebounds (DRB): Number of defensive rebounds secured.
- Assists (AST): Number of assists made.
- Steals (STL): Number of steals.
- Blocks (BLK): Number of shots blocked.
- Turnovers (TOV): Number of turnovers committed.
- Personal Fouls (PF): Number of personal fouls. In addition to PF the Euroleague makes Fouls Drawn (FD) available.

The research question relevant here is how these data can be used to build a single measure of efficiency that can be associated to each player to measure productivity. In this section we consider the construction of measures without theory, by analysing first the PIR approach and by then assessing the difficulties in using simple correlation analysis.

The PIR approach

The Performance Index Rating (PIR) is a widely used basketball metric designed to summarize a player's overall performance through a simple aggregation of box-score statistics. It adds positive actions, such as points, rebounds, assists, steals, blocks, and fouls drawn, and subtracts negative actions, such as missed shots, turnovers, and personal fouls:

$$PIR = PTS + REB + AST + STL + BLK + FD$$
$$- FGMISS - FTMISS - TOV - PF.$$

PIR is available for NBA teams and players. In Europe, the index was originally developed by the Spanish ACB League in 1991 and was subsequently adopted by major European competitions, including the EuroLeague, where it is commonly used to determine weekly MVPs and to assess player contributions. The press and sports commentators frequently rely on PIR to highlight standout performances and compare players across games and seasons. Agents may also refer to PIR in contract negotiations, since it provides a simple and easily interpretable measure of a player's statistical production (Wen et al., 2023).

However, coaches and analysts have expressed reservations about relying solely on PIR for player evaluation. According to the World Association of Basketball Coaches[4], official game statistics such as PIR provide useful information, but they may not fully capture a player's effectiveness or contribution to team dynamics.

These reservations are well founded. PIR aggregates positive and negative statistics without assigning theoretically or empirically justified weights to different actions. For instance, a missed free throw, which costs the team one point, and a missed field goal, which may cost the team either two or three points, are treated equally despite their different opportunity costs. Similarly, rebounds, assists, steals, and blocks all enter the index with the same unitary weight, although

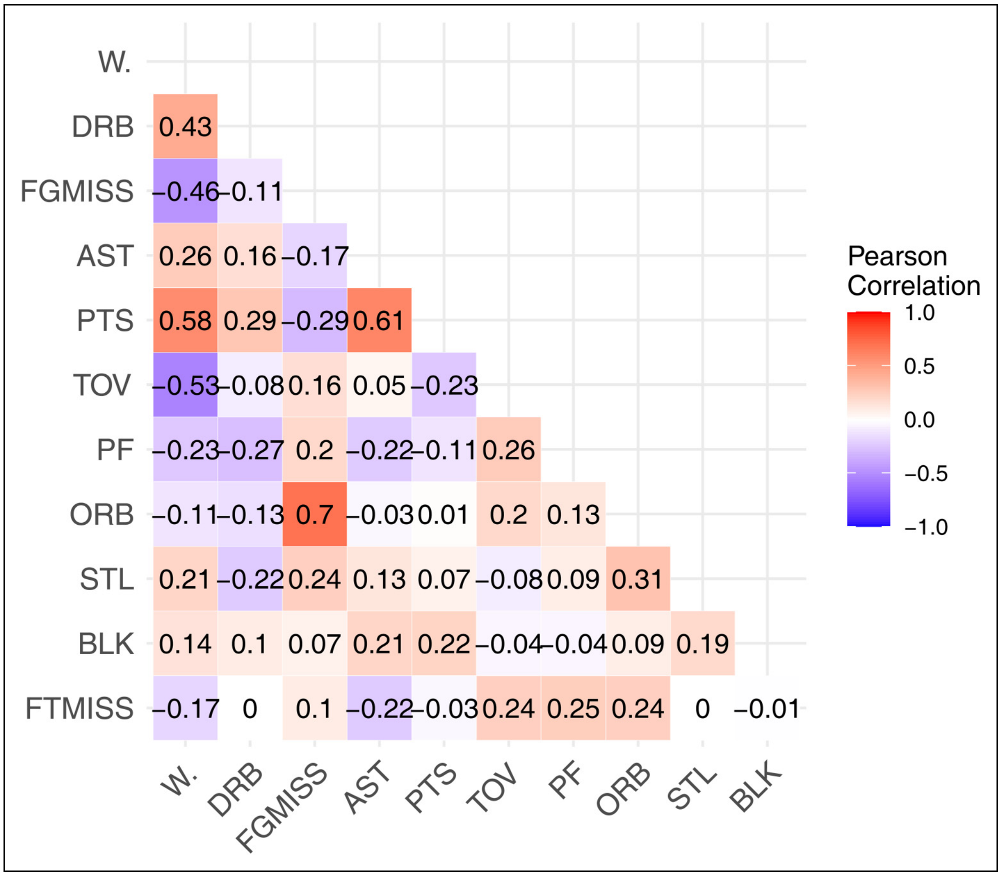

Figure 1. Correlation between Win Percentage and Box-Score Stats for NBA teams, seasons 2021–22 to 2024–25.

their contribution to winning may differ substantially. Moreover, PIR does not explicitly account for opponent performance or for the final outcome of the game. Intuitively, a player's performance should be evaluated relative to the performance of the opponent and weighted according to its impact on team success.

Still, the PIR approach remains very popular, probably because of its simplicity and ease of computation using widely accessible box-score statistics. The relevant question, therefore, is whether the same information set can be used within a more formal method that delivers theoretically grounded performance measures while remaining as easy to implement and interpret as PIR.

Can measurement without theory help?

Given that each game result is important, what are the difficulties of using theory-free measures of the relationship between statistics and game results to weight them? Consider the simplest measure of linear association between statistical variables, namely correlation, and apply correlation analysis to wins and box-score statistics.

We report in Figures 1–3 the correlation between win percentages and box-score statistics for the Euroleague on a post-COVID sample and for NBA data from two different samples: a post-COVID sample and an earlier sample covering seasons from 1992–1993 to 2004–2005.

The evidence highlights the limitations of relying on simple correlations between win percentages and box-score statistics to evaluate team performance. For example, offensive rebounds exhibit a negative correlation with the percentage of wins. Interpreting this result at face value and assigning a negative weight to offensive rebounds in a composite performance metric would be misleading. This is because offensive rebounds are highly correlated with missed field goals, which are themselves negatively associated with winning. Thus, offensive rebounds may partly act as a proxy for poor shooting rather than being intrinsically detrimental to team performance.

Despite these pitfalls, the comparison of correlation patterns across different data samples reveals a striking degree of robustness. Our findings in Figure 1 align closely with those of Figure 2, which is constructed on an earlier sample of the NBA data studied by Berri et al. (2006), and with those from the Euroleague reported in Figure 3. This robustness suggests the presence of a stable underlying structure, or data-generating process, within basketball statistics. Measurement without theory can therefore provide useful statistical background for the implementation of theory-based measurement.

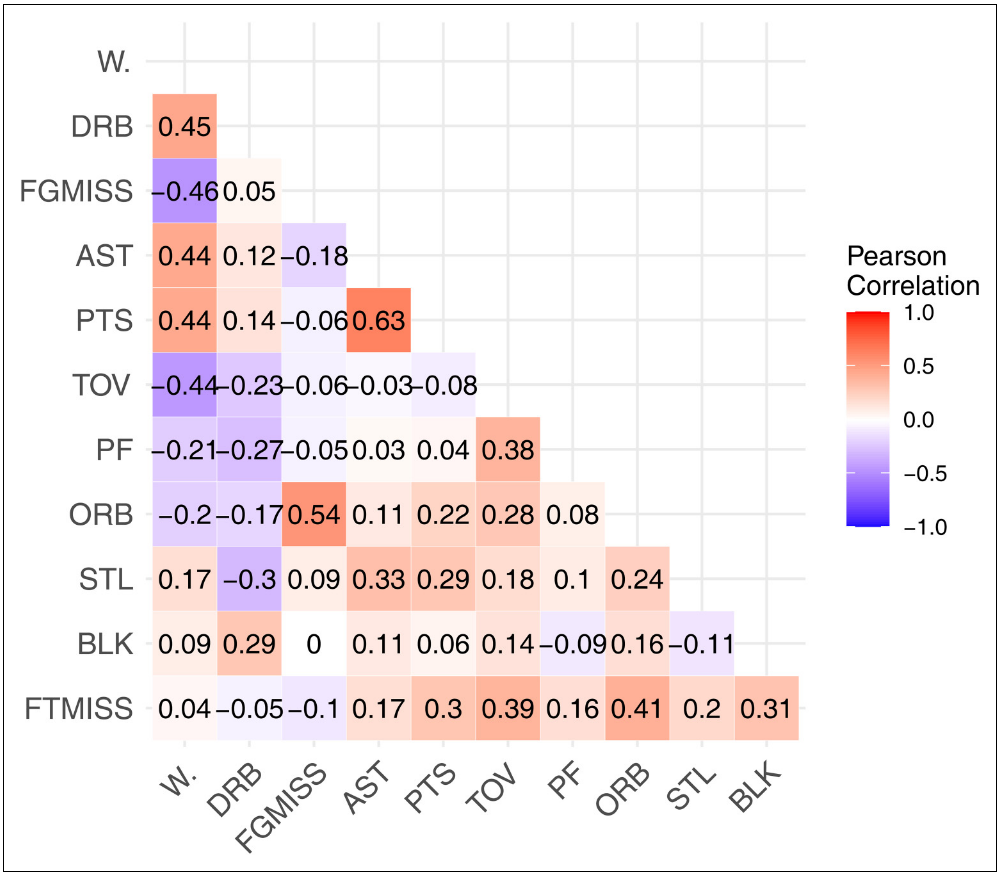

Figure 2. Correlation between Win Percentage and Box-Score Stats for NBA teams, seasons 1992–93 to 2004–05, omitting season 1998–99.

A further difficulty arises from the high correlation among many box-score statistics. When explanatory variables are strongly correlated with one another, their individual contribution to wins cannot be easily separated by means of a simple linear regression. In such cases, regression coefficients may become unstable, sample-dependent, and difficult to interpret as structural weights. Consequently, a purely data-driven linear weighting scheme may produce weights that reflect the particular correlation structure of the sample rather than the true marginal contribution of each statistic to winning. This problem calls for a parsimonious parameterization. Parsimony can be achieved in at least two ways. One possibility is to specify a theoretical model of basketball production that imposes restrictions on the way different statistics enter the performance measure, thereby reducing the number of free parameters to be estimated. An alternative possibility is to adopt statistical machine-learning methods based on regularization, such as ridge regression, lasso, or elastic net, which shrink or select coefficients in order to deal with multicollinearity and improve out-of-sample stability (Lakhloufi, 2025). In both cases, the central point is that high-dimensional and highly correlated box-score data cannot be weighted reliably by unrestricted regression alone: either theory or regularization is needed to discipline the parameterization.

Reinterpreting WINSCORE as a Model-Based framework

Theory-based measurement relies on a structural model to account for the interactions among the relevant variables in a comprehensive way. In our case the model is structural in the sense that it is grounded in a theoretical representation of the game based on possessions and explicitly specifies the mechanism linking player actions to team outcomes, rather than estimating reduced-form correlations between aggregated statistics and wins.

The model-based approach unfolds through a sequence of well-defined steps.

The first step is model specification, where variables are classified into exogenous or endogenous and their interdependence is captured in equations. In this stage, the data for endogenous variables are conceptually decomposed into two components: structure and noise (James et al., 2013). The structural component captures systematic relationships that are invariant to the specific sample and can thus be used for prediction and policy analysis, particularly to evaluate how changes in exogenous variables affect endogenous outcomes. In contrast, the noise component is sample-specific and does not inform point predictions or simulations

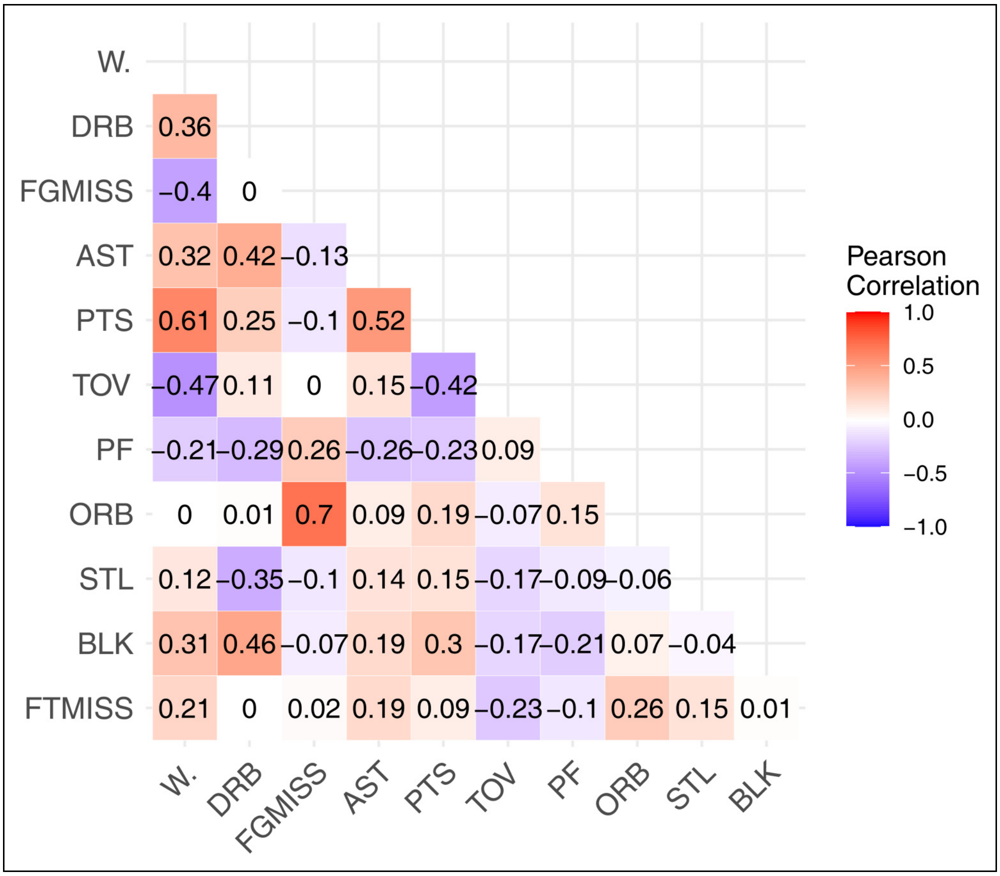

Figure 3. Correlation between Win Percentage and Box-Score Stats for Euroleague teams, seasons 2022–23 to 2024–25.

directly. However, it plays a crucial role in quantifying uncertainty around predictions, for instance, through the construction of confidence intervals. The model may also incorporate relationships among exogenous variables.

Model equations are typically formalized through functional equations involving a set of unknown parameters. The second step involves assigning numerical values to these parameters. This can be achieved through estimation, using the available data on the model's variables, or through calibration, which leverages information external to the dataset (Cooley, 1997).

Once the parameters have been estimated or calibrated, it becomes possible to recover the empirical counterpart of noise in the form of fitted residuals. These residuals serve a critical role in model validation. Validation involves examining whether the statistical properties of the residuals align with the assumptions underpinning the estimation process. Deviations may indicate model misspecification (Favero, 2001). Further validation can be pursued by comparing the specification of the model to alternative specifications or by assesing the model properties across different subsamples.

The final step involves putting the model to work via simulation. By simulating the model under a baseline scenario—where all exogenous variables are held at their average values—and comparing it to a counterfactual scenario in which one or more exogenous variables are altered, it is possible to assess the impact of such changes on the endogenous variables. This step is made possible by the fact that the model makes explicit the causal links between variables (Terner and Franks, 2021).

The remainder of this section demonstrates how the Winscore approach can be reinterpreted within a model-based framework, and how each of the modelling steps outlined above can be operationalized in the context of Winscore.

Winscore specification and estimation

The Winscore approach uses the number of WINS in a regular season as the measurable counterpart of performance. The main theoretical hypothesis is that the key concept to determine performance is how efficiently teams use **possession**.

A possession starts when one team gains control of the ball and ends when that team gives it up (in other words, an offensive rebound would start a new play, not a new possession).

Given a database with t seasons and i teams, possessions can be categorized into Employed Possessions (EP) and

Acquired Possessions (AP):

$$EP_{i,t} = FGA_{i,t} + 0.45 * FTA_{i,t} + TOV_{i,t} - ORB_{i,t}AP_{i,t}$$
$$= OTOV_{i,t} + DRB_{i,t} + TEAMR_{i,t} + OFG_{i,t} + 0.45$$
$$* OFT_{i,t}EP_{i,t}$$
$$\approx AP_{i,t}$$

EP and AP are guaranteed to be approximately the same for the two teams in a game, approximately because ends of quarter possessions might not be evenly distributed across a team and its opponent.

However, the relation $EP_{i,t} \approx AP_{i,t}$ is not a behavioral assumption. It is a possession-accounting restriction implied by the rules of basketball and by the definition of a possession. When one team gives up control of the ball, the opponent acquires a possession. At the game level, small discrepancies may arise from end-of-period possessions and from the conventional box-score formula used to reconstruct possessions. At the season level, however, these discrepancies average out. Appendix B.1 shows that the difference between acquired and employed possessions is virtually zero in both the NBA and EuroLeague samples. As noted by (Berri et al., 2006) $TEAMR_{i,t}$ are not available from the Box-Score statistics, but they can be reconstructed by exploiting the fact that $EP_{i,t} \approx AP_{i,t}$. A measure of performance is constructed by taking the difference of points made per employed possession and points made by the opponents per acquired possession as:

$$PTSxEP_{it} - PTSAxAP_{i,t}$$

The Winscore model is then specified as follows:

$$W_{it} = \beta_0 + \beta_1\left(PTSxEP_{it} - PTSAxAP_{i,t}\right) + u_{it}$$

$$u_{it}N.I.D\left(0, \sigma^2\right)PTSxEP_{it} = \frac{PTS_{i,t}}{EP_{i,t}}PTSAxAP_{i,t}$$
$$= \frac{PTSA_{i,t}}{AP_{i,t}}EP_{i,t}$$
$$= FGA_{i,t} + 0.45 * FTA_{i,t}$$
$$+ TOV_{i,t} - ORB_{i,t}AP_{i,t}$$
$$= OTOV_{i,t} + DRB_{i,t} + TEAMR_{i,t}$$
$$+ OFG_{i,t} + 0.45$$
$$* OFT_{i,t}PTS_{i,t}$$
$$= 1 * FT_{i,t} + 2 * 2PFG_{i,t} + 3$$
$$* 3PFG_{i,t}PTSA_{i,t}$$
$$= 1 * OFT_{i,t} + 2 * O2PFG_{i,t} + 3$$
$$* O3PFG_{i,t}$$

The model is made up of seven equations, but only the first one is stochastic and depends on unknown parameters. The interpretation of the unknown parameters in this equation is very intuitive: β_0 captures the performance over a season of the average team, and is expected to be estimated at approximately one-half of the number of games played, n, while β_1 measures the impact of the efficiency differential in determining the deviation of each team's performance from that of the average team.

A key advantage of this specification is that it avoids the problem of assigning separate regression weights to a large number of highly collinear box-score statistics. Rather than entering field goals, free throws, rebounds, turnovers, and opponent statistics as separate regressors in an unrestricted linear model, the Winscore model combines them through theoretically motivated nonlinear transformations. These transformations produce two efficiency measures: points made per employed possession, $PTSxEP_{it}$, and points allowed per acquired possession, $PTSAxAP_{i,t}$. The difference between these two objects summarizes the team's net efficiency and reduces the dimensionality of the empirical problem to a parsimonious specification involving only two parameters, β_0 and β_1.

The model is therefore nonlinear in the way box-score statistics affect the explanatory variable, but linear in the unknown parameters to be estimated. This distinction is important. The effect of a given statistic is not captured by an independent regression coefficient attached to that statistic alone. Instead, its contribution depends on how it changes the nonlinear efficiency ratio. For example, one additional made three-point shot affects the model through the numerator of $PTSxEP_{it}$, because it increases $PTS_{i,t}$ by three points. At the same time, it also affects the denominator, because the shot attempt is part of employed possessions through $FGA_{i,t}$. Therefore, the contribution of a made three-point shot is not simply a fixed marginal coefficient on $3PFG_{i,t}$. It depends on its joint effect on both points scored and possessions employed. This is precisely how the model incorporates the interaction among correlated statistics without estimating a separate coefficient for each of them.

It is therefore important to distinguish the present specification from an unrestricted regression of wins on all box-score statistics. In the latter case, variance-inflation diagnostics would be directly relevant because each statistic would enter as a separate regressor. In the WINSCORE specification, by contrast, the highly correlated statistics are not assigned separate free coefficients. They are combined through the possession-based structure into the net efficiency variable $PTSxEP_{it} - PTSAxAP_{it}$. The resulting stochastic equation is identified by the time-team variation in this net efficiency measure and contains only two unknown parameters.

The relevant question is how well this stochastic equation separates structure from noise. In the context of a full season, the difference between points made per employed possession and points allowed per acquired possession is structurally close to the natural object that mechanically drives wins. It therefore allows the decomposition of wins

Table 1. Wins and efficiency in the NBA and the euroleague.

	Dependent variable: Wins in Regular Season		
	NBA 1994–2005	NBA 2021–2024	Euroleague 2022–25
(a) Regression of wins on efficiency: NBA 1994–2005 (Berri et al., 2006), NBA 2021–2024, and Euroleague 2022–25			
Intercept	41.00***	41.00***	17.00***
	(0.17)	(0.33)	(0.26)
$(PTSxEP_{i,t} - PTSAxAP_{i,t})$	257.16***	237.38***	78.46***
	(3.40)	(7.11)	(4.34)
Observations	316	90	54
Multiple R^2	0.95	0.93	0.86
Adjusted R^2	0.95	0.93	0.86
F Statistic	5718***	1115***	326.2***

	Dependent variable: Percentage of Wins in Regular Season.		
	NBA 1994-2005	NBA 2021–2024	Euroleague 2022–25
(b) Regression of Percentage of Wins on efficiency: NBA 1994-2005 (Berri et al., 2006), NBA 2021–2024, and Euroleague 2022–25.			
Intercept	0.50***	0.50***	0.50*** .
	(0.002)	(0.007)	(0.04) .
$(PTSxEP_{i,t} - PTSAxAP_{i,t})$	3.13***	2.89***	2.30*** .
	(0.041)	(0.086)	(0.127) .
Observations	316	90	54 .
Multiple R^2	0.95	0.93	0.86 .
Adjusted R^2	0.95	0.93	0.86 .
F Statistic	5718***	1115***	326.2*** .

***$p < 0.001$.

into a structural component and a noise component. To assess the validity of this point, we estimate the stochastic equation across different data sets and different periods, focusing on the stability of both the estimated parameters and the noise/structure decomposition.

Table 1 reports the results of estimating the WINSCORE model on a pooled cross-section of box-score data for all teams for a benchmark NBA sample used by Berri et al. (2006), for the post-COVID seasons in the NBA, and the Euroleague.[5]. As the number of games played in the regular season is 82 in the NBA and 34 in the Euroleague to facilitate comparability Table (1(a)) reports the results with total Wins as dependent variable while Table (1(b)) reports the results with the percentage of Wins as dependent variable.

The estimated parameters are highly significant for all samples considered. The intercept in all regressions, as expected, captures the performance of the average team,[6] the slope coefficient captures the importance of the adopted efficiency measures in explaining deviations of each teams' performance from that of the average team efficiency. This coefficient is positive, highly significant and of a size which is robustly estimated across different samples in the NBA. However, there is a significant difference in the impact of efficiency on Wins between the NBA and the Euroleague, which features a smaller coefficient in the projections of percentage of Wins on efficiency. The statistical significance of the different impact is documented by the implementation of a Chow-type test in Appendix B.2 This smaller coefficient is explained by the fact that the variance of the efficiency measure in the Euroleague is higher than that of the equivalent variable for the NBA. This evidence could be interpreted as a signal of higher "competitive balance" (Zimbalist, 2002) in the NBA.[7] Interestingly, the average points scored per possession are higher in the Euroleague than in NBA, debunking the myth that European basketball features better defenses than the NBA.

Winscore validation

Once the model has been estimated, it can be subjected to validation.

The validation strategy adopted here is cross-sample and cross-league rather than algorithmic cross-validation. This choice follows from the objective of the paper. We do not seek to optimize a flexible forecasting algorithm over random folds; instead, we assess whether the possession-based WINSCORE structure, originally developed for NBA data, remains empirically useful when applied to a different NBA era and to EuroLeague data. The relevant robustness evidence is therefore provided by the stability of the estimated relationship across samples, the residual diagnostics, the encompassing tests against PIR, and the comparison with alternative performance measures. We begin with the assessment of the capability of Winscore to outperform alternative models in predicting W_{it}, and here the obvious candidate is a prediction of W_{it} based on the PIR approach. To this end, after the estimation of the Winscore equation (1), an alternative equation, based on PIR, can be specified as:

$$W_{it} = \gamma_0 + \gamma_1 \left(PIR_{it} - OPIR_{i,t} \right) + v_{it} v_{it} \sim \text{i.i.d.N}(0, \sigma^2)$$

where PIR_{it} is the PIR of team i in season t, and $OPIR_{it}$ is the PIR of the opponents. Table 2 reports the results from the estimation of equation (2) on different leagues and samples.

In both the NBA and Euroleague specifications, the estimated coefficient on the PIR factor is positive and highly significant, indicating a robust relationship between player impact rating and team wins. The NBA regression explains 91 percent of the variance in wins compared to 76 percent in the Euroleague.

Table 2. Regression of wins on PIR factor: NBA (berri), NBA 2021–2024, and euroleague 2022–25.

	Dependent variable: wins		
	NBA (Berri et al., 2006)	NBA 2021–2024	Euroleague 2022–25
Intercept	41.00***	41.00***	17.00***
	(0.169)	(0.367)	(0.335)
PIR factor	0.016***	0.014***	0.011***
	(0.000285)	(0.000475)	(0.000842)
Observations	316	90	54
Multiple R^2	0.91	0.91	0.76
Adjusted R^2	0.91	0.91	0.76
F Statistic	3219***	902.7***	168***

***$p < 0.001$.

The first stage of validation is residual analysis, aimed at assessing whether the model has successfully captured the systematic structure of the data, so that the remaining residuals can be regarded as noise. The residual diagnostics reported in the following Table 3 support this interpretation.

Across the three samples, Winscore residuals fluctuate randomly around zero and display lower dispersion than the PIR residuals. In the NBA 1994–2005 sample, the residual standard deviation is 0.0366 for Winscore and 0.0478 for PIR; in the recent NBA sample, it is 0.0366 for Winscore and 0.0402 for PIR; and in the EuroLeague sample, it is 0.0546 for Winscore and 0.0716 for PIR. The same pattern holds for RMSE and MAE. Moreover, the Jarque–Bera tests do not reject normality for the Winscore residuals in any of the three samples. These results indicate that the possession-based model captures the systematic component of team performance more effectively than the PIR benchmark, while leaving residuals that do not display major systematic departures from the maintained specification.

In Appendix B.3 we report an extensive residual analysis both for the Winscore and PIR regressions. For each sample, we compare the distribution of residuals using summary diagnostics, residual dispersion measures, and Q–Q plots. Overall, these diagnostics support the adequacy of the Winscore specification and indicate that the model is well behaved across the NBA and Euroleague samples.

After estimation of equations (1) and (2), a further validation step can be conducted through an encompassing test, aimed at assessing whether one specification contains all the relevant information captured by the alternative model, or whether the latter provides additional explanatory power beyond that already accounted for by the former. The following encompassing models (Mizon and Richard, 1986) is estimated:

$$W_{it} = \delta_1 \hat{W}_{it}^{WS} + \delta_2 \hat{W}_{it}^{PIR} + \epsilon_{it} \epsilon_{it} \sim \text{i.i.d.N}(0, \sigma^2)$$

where $\hat{W}_{it}^{WS}$ and $\hat{W}_{it}^{PIR}$ are the predicted wins from the two alternative models. The estimated parameters in equation (3) indicate the weights of the prediction from the two models in their optimal combination to forecast the common dependent variable. The more WINSCORE dominates, the closer we are to a situation in which $\delta_1 = 1, \ \delta_2 = 0$.

The results of the encompassing test reported in Table 4 show that the WINSCORE model uniformly dominates PIR.

For all data-sets considered encompassing regressions, the coefficient on the WINSCORE prediction is positive statistically different zero and not statistically different from one, whereas the coefficient associated with the PIR prediction is always not statistically different from zero. Moreover, the joint hypothesis $\delta_1 = 1, \ \delta_2 = 0$ is never rejected.

The benchmarking exercise is designed around the information set of the paper. Since WINSCORE is constructed from standard box-score statistics, the primary benchmark is PIR, which is the most widely used index based on the same type of data. To provide broader evidence on the Winscore model that goes beyond its capability to encompass PIR, in Appendix B.4 we extend the analysis by comparing the WINSCORE model with additional advanced performance measures. These include the Four Factors model (Oliver, 2004), which captures shooting efficiency, turnovers, offensive rebounding, and free-throw rate, as well as adjusted box-score plus-minus measures such as OBPM, DBPM, BPM, and VORP. The first two decompose player impact into offensive and defensive contributions, BPM provides an overall box-score-based estimate of contribution per 100 possessions, and VORP converts this contribution into a cumulative value-over-replacement measure that incorporates playing time. These are stronger box-score-based benchmarks, while tracking-data, lineup-based, and spatio-temporal models are complementary approaches based on richer information sets and are therefore not direct empirical competitors in the present design.

Table 3. Residual diagnostics for winscore efficiency and PIR models.

(a) NBA 1994–2005, excluding 1999

Model	Mean	SD	RMSE	MAE	Min	Median	Max	JB stat.	JB p-value
Winscore efficiency	0.0000	0.0366	0.0365	0.0291	-0.1027	-0.0008	0.0967	0.4074	0.8157
PIR	0.0000	0.0478	0.0477	0.0390	-0.0998	-0.0004	0.1502	6.6935	0.0352
(b) NBA 2022–2025									
Winscore efficiency	0.0000	0.0366	0.0364	0.0287	-0.0976	0.0034	0.1087	0.7917	0.6731.
PIR	0.0000	0.0402	0.0400	0.0332	-0.0948	-0.0039	0.0989	1.4411	0.4865.

(c) Euroleague 2022–2025

Model	**Obs.**	**Mean**	**SD**	**RMSE**	**MAE**	**Min**	**Median**	**Max**	**JB stat.**	**JB p-value.**
Winscore efficiency	54	0.0000	0.0546	0.0541	0.0443	-0.1216	0.0011	0.1063	1.3272	0.5150.
PIR	54	0.0000	0.0716	0.0710	0.0576	-0.1567	0.0040	0.1566	0.4533	0.7972.

Notes: Residuals are computed from win-percentage regressions. JB stat. denotes the Jarque–Bera test statistic for residual normality.

The results, reported in Tables .3–.6, confirm the same pattern. Across both the NBA and Euroleague samples, the fitted values from the WINSCORE model remain statistically significant in the encompassing regressions, while the fitted values from the alternative models generally do not add explanatory power.

Overall, the empirical validation of the specification is therefore based on parameter significance, goodness of fit, residual diagnostics, and out-of-sample or cross-sample stability, rather than on variance-inflation diagnostics for an unrestricted model that is not being estimated. The Winscore specification is very different from an unrestricted regression of wins on all box-score statistics.

Winscore simulation

After estimation (or calibration) of all unknown parameters, the model can then be used to attribute weights to statistics by simulating their impact on predicted wins in the following steps:

- Generate via the model a predicted value for wins in the case all statistics are kept at their average. This is called the baseline scenario simulation.
- Generate via the model a predicted value for wins in case all the statistics are kept at their average except the one, whose effect is to be evaluated.
- The difference between wins in the alternative scenario and wins in the baseline scenario gives the impact of the statistic on WINs.
- Point estimate of the impact are obtained via deterministic simulation, while the statistical distribution of the impact can be obtained via stochastic simulation by taking uncertainty around the estimated coefficients into account.

Note that the model-based procedure takes all feedbacks into account: one more three-points shots made gives the team three points more at the cost of employing a possession.

Table 5 reports the weights derived by model simulation after model estimation based on the Euroleague data and NBA data in the Post-Covid periods and compares them with the weights reported by Berri et al. (2006) derived by the same technique after estimation based on the NBA data from seasons 1994 to 2005 excluding season 1999 due to the players' strike.

The weights reported below are deterministic simulation point estimates. Given our point estimates prediction-level uncertainty can be obtained in the standard way from the estimated regression variance and the variance-covariance matrix of the estimated parameters, or by stochastic simulation drawing (β_0, β_1) from their asymptotic distribution. In the present application, the estimated slope coefficients are highly significant across all samples, and the simulated weights are stable across the two NBA samples. We

Table 4. Encompassing regressions combining WINSCORE and PIR predictions: NBA (berri), NBA model, and euroleague model (rounded to two decimals).

	Dependent variable: wins		
	NBA (Berri et al., 2006)	NBA 2021–2024	Euroleague 2022–25
Pred. WINSCORE	0.92***	0.77***	1.08***
	(0.06)	(0.17)	(0.18)
Pred. PIR	0.08	0.23	-0.08
	(0.06)	(0.17)	(0.18)
Observations	316	90	54
Multiple R^2	0.9952	0.9947	0.9892
Adjusted R^2	0.9952	0.9946	0.9888
F Statistic	32500***	8314***	2391***

***$p < 0.001$.

Table 5. The value of statistics in terms of wins.

Statistics	Euroleague	NBA 2021–2024	NBA (Berri et al., 2006)
Scoring Statistics			
Three-point field goals made	+0.059	+0.054	+0.066
Opponent's three-point field goals made	–0.059	–0.054	–0.066
Two-point field goals made	+0.027	+0.025	+0.033
Opponent's two-point field goals made	–0.027	–0.025	–0.033
Free throws made	+0.016	+0.015	+0.018
Opponent's free throws made	–0.016	–0.014	–0.018
Missed field goals	–0.036	–0.032	–0.034
Missed free throws	–0.016	–0.014	–0.015
Possession Statistics			
Offensive rebounds	+0.036	+0.032	+0.034
Turnovers	–0.036	–0.032	–0.034
Defensive rebounds	+0.036	+0.032	+0.034
Team rebounds	+0.036	+0.032	+0.034
Opponent's turnovers	+0.036	+0.032	+0.034
Steals	+0.036	+0.032	+0.034
Personal Fouls, Blocked Shots and Assists			
Personal fouls	–0.016	–0.014	–0.018
Blocked shots	+0.017	+0.015	+0.021
Assists	+0.018	+0.020	+0.022

Table 6. Edge cases in the 2024–2025 euroleague projected *WINSEFF* ($WINSEFF_{proj}$).

(a) Average $WINSEFF_{proj}$ by role

Role	$WINSEFF_{proj}$	$WINSEFF_{proj}$ excl. edge cases
Guard	-0.506	0.039
Forward	5.136	5.131
Center	8.744	8.951

(b) Best full-season $WINSEFF_{proj}$ player by role

Role	Player	Team	Minutes	$WINSEFF_{proj}$.
Guard	Weiler-Babb, Nick	MUN	841.7	7.03.
Forward	Vezenkov, Sasha	OLY	973.0	14.46.
Center	Tavares, Walter	MAD	812.1	15.10.

(c) Low-minute Winscore projection edge cases

Role	Player	Team	Minutes	$WINSEFF_{proj}$.
Center	Koumadje, Khalifa	BER	96.8	18.254.
Forward	Yilmaz, Erkan	IST	20.0	17.665.
Center	Yebo, Kevin	MUN	99.5	4.290.
Guard	Wong, Isaiah	ZAL	85.9	2.708.
Guard	Laprovittola, Nicolas	BAR	88.1	0.957.
Forward	Begarin, Juhann	MCO	42.6	-6.829.
Center	Pleiss, Tibor	PAN	33.3	-8.912.
Guard	Smith Jr, Dennis	MAD	20.7	-12.579.
Guard	Wallace, Tyrone	ZAL	22.4	-16.064.
Guard	Querejeta, Joseba	BAS	5.3	-38.598.

Notes: $WINSEFF_{proj}$ is the player's Winscore contribution projected to full-season minutes. The edge cases are selected among players with at most 100 minutes. The top and bottom 5 edge cases are reported in decreasing order.

therefore report the deterministic weights as the main object of interest and use cross-sample stability as the primary robustness evidence. An important implication of the simulation exercise is that the small differences in the estimated coefficients across the two NBA samples do not generate economically meaningful differences in the weights assigned to the underlying statistics. Hence, while coefficient estimates may vary slightly across eras, the practical evaluation of box-score actions implied by the model remains highly stable across the different samples considered. Moreover, the small differences between the NBA and Euroleague parameter estimates can be reconciled with the logic of the model. For example, in the league with greater competitive balance, namely the NBA, the contribution of a made three-point shot to wins is expected to be higher than in a league with lower competitive balance, such as the Euroleague.

Note that the Winscore model assigns weights to all statistics entering the PIR, with the exception of personal fouls, blocked shots, and assists, which are not included directly in the model. This exclusion should not be interpreted as evidence of model misspecification. It follows instead from the structure of Winscore, which is based on employed and acquired possessions. Assists and personal fouls, in particular, do not directly create, consume, or transfer possessions.

Nevertheless, these statistics may still carry indirect information about player contribution. Their implied values can therefore be recovered through auxiliary regressions that relate them to the variables already included in the Winscore specification. In Appendix B.5, we show that augmenting the Winscore model with the omitted variables provides no evidence of misspecification, since the coefficients on the additional variables are not statistically different from zero.

This finding has a clear interpretation. In a regression, the coefficient on a variable measures the effect of the component of that variable that is orthogonal to the other regressors.[8] Hence, the insignificance of the omitted statistics indicates that, conditional on the Winscore variables, they do not provide an independent contribution to wins. Their value is therefore not omitted from the model, but is already reflected in the included possession-based statistics. The auxiliary regressions provide a natural way to redistribute part of this value to the corresponding PIR statistics, rather than a correction for model misspecification.

Interpreting the winscore model

After the attribution of weights by simulation to each statistic, the outcome of the model simulation has the following intuitive linear interpretation:

$$W_{it}^{WS} = \frac{n}{2} + WINSEFF_{it} - OWINSEFF_{it}$$

$$\begin{aligned}
WINSEFF_{it} &= SCSTAT_{it} + POSSTAT_{it} \\
&\quad + PFBLKA_{it} SCSTAT_{it} \\
&= 3PFG_{i,t} * w^{3P} + 2PFG_{i,t} * w^{2P} + FT_{i,t} \\
&\quad * w^{FT} + +(FGA_{i,t} - 3PFG_{i,t} - 2PFG_{i,t}) \\
&\quad * w^{FGM} + (FTA_{i,t} - FT_{i,t}) \\
&\quad * w^{FTM} POSSTAT_{it} \\
&= w^{POS}(ORB_{i,t} + DRB_{i,t} + STL_{i,t} \\
&\quad - TOV_{i,t}) PFBLKA_{it} \\
&= w^{BLK} * BLK_{i,t} + w^{PF} * (PF_{i,t} - FD_{i,t}) + w^{A} \\
&\quad * (A_{i,t}) OWINSEFF_{it} \\
&= OSCSTAT_{it} + OPOSSTAT_{it} \\
&\quad + OPFBLK_{it} OSCSTAT_{it} \\
&= O3PFG_{i,t} * w^{3P} + O2PFG_{i,t} * w^{2P} + OFT_{i,t} \\
&\quad * w^{FT} + \\
&+ (OFGA_{i,t} - O3PFG_{i,t} - O2PFG_{i,t}) * w^{FGM} \\
&- (OFTA_{i,t} - OFT_{i,t}) * w^{FTM} OPOSSTAT_{it} \\
&= w^{POS} * (OORB_{i,t} + ODRB_{i,t} + OSTL_{i,t} \\
&\quad - OTOV_{i,t}) OPFBLKA_{it} \\
&= w^{BLK} * OBLK_{i,t} + w^{PF} * (OPF_{i,t} - OFD_{i,t}) + w^{A} \\
&\quad * (OA_{i,t})
\end{aligned}$$

where $[w^{3P}, w^{2P}, w^{FT}, w^{FGM}, w^{FTM}, w^{POS}, w^{BLK}, w^{PF}, w^{A}]$ are the weights reported in Table 5.

Deviations of a team's performance from that of the average team are determined by the difference between its weighted performance measure—constructed using Winscore weights—and the corresponding measure of its opponents.

Accordingly, the predicted number of wins for team i over the course of the season is obtained by applying two adjustments to its performance index, $WINSEFF_{it}$: the first accounts for the strength of its opponents, $OWINSEFF_{it}$, while the second captures the benchmark performance of the average team, $\frac{n}{2}$.

This interpretation of the model, based on linearization, allows to apply the measure of performance constructed using data for Teams to Individual Players. Before taking this step in the next section it is probably worth taking a look at Figure 4 which reports a cross plot of the Wins predicted for NBA teams over the seasons 1992-93 to 2004-5 (omitting season 1998-99) by the exact Winscore model (1) and by its linearized version (4). There is clear evidence of a very strong association between the two variables, with some exceptions only on the tails, to be expected as a consequence of linearization. The data for Boston Celtics, highlighted in green in the graph, confirm the very strong association between the predictions of the exact and

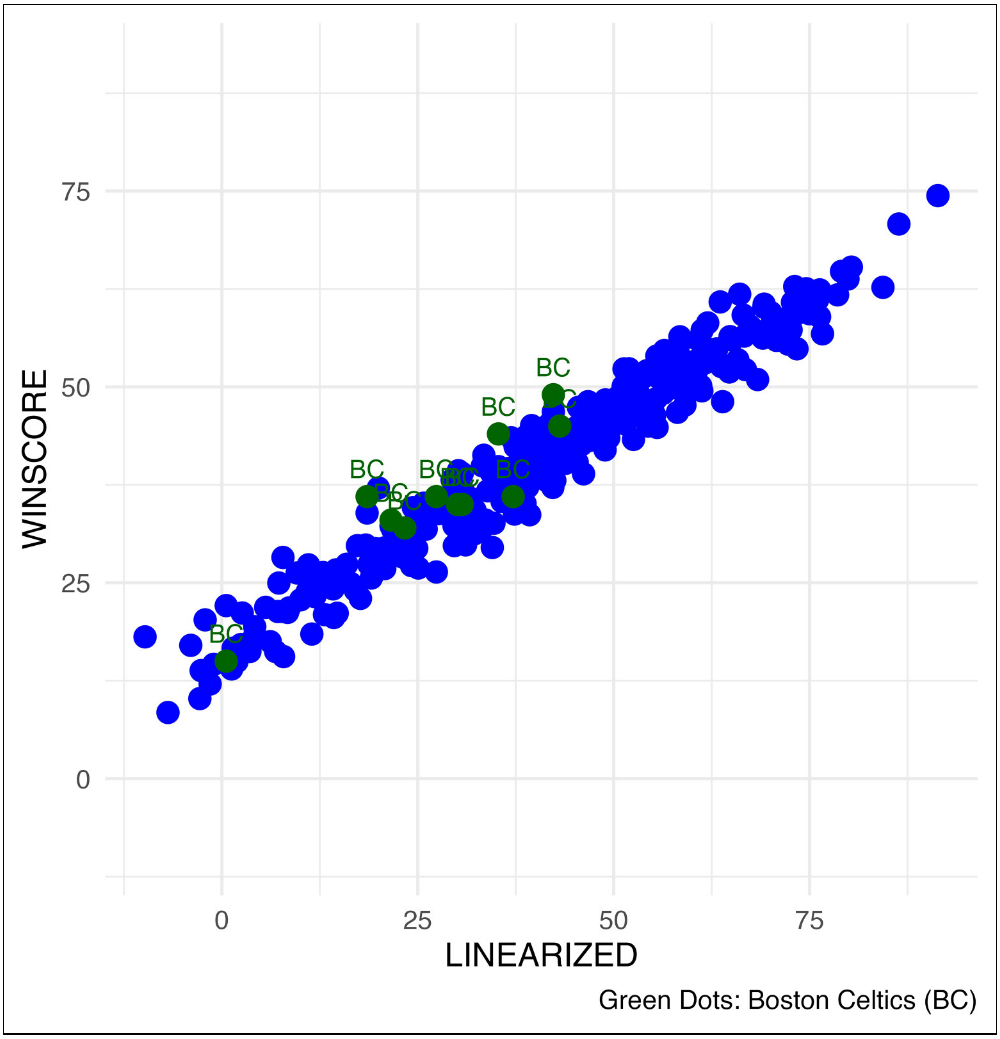

Figure 4. Wins predicted by the Winscore model vs Wins predicted by the linearized Winscore model.

linearized models singling out a specific team over the time-series of seasons. Overall, the strong empirical correspondence between the exact and linearized model predictions validate the linear approximation.

From team to players: Season data

Given the availability of player-level statistics for each player j over a given season, individual performance can be evaluated by applying the Winscore weights derived from the linearized model in equation (4). The original procedure, following Berri et al. (2006), translates team-level efficiency into player-level contributions through three steps.

First, for each player j, one computes $WINSEFF_j$ by aggregating the relevant scoring and possession statistics, together with the statistics required to make the measure comparable to the box-score variables entering the player evaluation. Second, player performance is adjusted relative to the performance of an average player in the same role. Third, the resulting relative contribution is anchored to the wins produced by an average team over the same amount of court time played by player j.

Formally, the Berri-type player-level measure can be written as

$$WINSCORE_j^B = WINSEFF_j - OWINSEFF_j^{adj,B} + \frac{MIN_j}{TOTMIN_j} \cdot \frac{0.5 \cdot G}{5},$$

$$OWINSEFF_j^{adj,B} = \left(\frac{1}{N_{\mathrm{role(j)}}} \sum_{k \in \mathrm{role(j)}} WINSEFF_k \cdot \frac{TOTMIN_k}{MIN_k} \right) \cdot \frac{MIN_j}{TOTMIN_j},$$

where G denotes the number of games in the season, MIN_j is the number of minutes played by player j, $TOTMIN_j$ is the

Table 7. Model predictions vs. actual wins by team.

Team	Wins	WS	Lin. WS	$\sum_{j=1}^{J_i} WINSCORE^{B}_{j,i}$	$\sum_{j=1}^{J_i} WINSCORE^{M}_{j,i}$
ASV	13	12.5	13.5	18.2	13.0
BAR	20	21.2	20.1	27.2	22.8
BAS	14	15.0	20.1	19.7	20.5
BER	5	3.6	0.5	15.5	−1.6
IST	20	22.5	27.0	32.7	24.7
MAD	20	19.6	27.6	20.1	18.7
MCO	21	21.1	15.3	25.5	18.4
MIL	17	16.2	18.3	30.5	19.8
MUN	19	15.8	16.0	25.6	18.6
OLY	24	22.0	22.4	30.3	24.6
PAN	22	22.7	25.4	34.3	23.3
PAR	16	18.5	12.5	22.0	20.2
PRS	19	17.5	9.6	24.4	19.7
RED	18	18.2	15.3	22.0	13.3
TEL	11	13.8	15.4	28.0	15.7
ULK	23	19.4	23.6	23.6	13.4
VIR	9	11.3	8.6	20.9	13.2
ZAL	15	14.7	14.8	17.0	9.1

total number of team minutes available to player j's team over the season, and $N_{\text{role(j)}}$ is the number of players in the same role as player j.

The term $OWINSEFF_j^{adj,B}$ is the role-specific benchmark against which player j's contribution is evaluated. It is obtained by first projecting each player in the same role to full-season minutes and then rescaling the resulting benchmark to the actual playing time of player j. The final term,

$$\frac{MIN_j}{TOTMIN_j} \cdot \frac{0.5 \cdot G}{5},$$

assigns to player j the wins that would be produced by an average player over the same amount of court time. Thus, $WINSCORE_j^B$ measures player j's contribution relative to a role-specific average, while preserving the normalization that an average team wins approximately half of its games.

If the player-level decomposition is internally consistent with the team-level Winscore model, the sum of individual contributions for team i should approximate the number of team wins predicted by the model:

$$\sum_{j=1}^{J_i} WINSCORE^{B}_{j,i} \approx \text{TeamWins}_{\text{i}},$$

where J_i is the number of players on team i.

The key feature of this procedure is that it corrects for differences in playing time by transforming player statistics into full-season equivalents before constructing the role-specific benchmark. This makes players comparable in principle, but it also introduces an important empirical difficulty. Playing time is not randomly assigned. It is the outcome of coaching decisions, roster constraints, injuries, tactical choices, and assessments of player ability. As a result, players with very low minutes are typically a selected group, and projecting their observed performance to full-season minutes can generate extreme values.

The problem is particularly severe because the projection factor is mechanically large when MIN_j is small. For example, a player who plays only one minute and makes a three-point shot would be projected to an implausibly high full-game or full-season scoring rate.[9] Conversely, a player who plays only one minute, misses a field-goal attempt, and records no positive contribution would be projected to an implausibly poor full-season performance. In both cases, the player receives excessive influence in the construction of the role-specific benchmark, despite having provided very little information about sustained performance.

This is not simply a problem of noisy observations. It is a selection problem induced by the interaction between limited playing time and full-season extrapolation. Low-minute players are precisely those for whom observed statistics are least informative about season-long productivity, yet the standard projection assigns them the largest scaling factors. Consequently, a small number of edge cases can distort the role-level benchmark and, through it, the relative contribution assigned to all other players in the same role. Table 6 illustrates this issue directly.

For each player, we compute the projected full-season Winscore contribution, denoted by $WINSEFF_{proj}$, by scaling the player's observed contribution according to his share of team minutes. We then focus on low-minute players, defined as players with at most 100 minutes played, and rank them according to their projected values. The table reports both the most extreme positive and negative cases. For comparison, it also reports the best projected full-season player in each role and the role-level averages computed with and without these edge cases. The evidence shows that a small number of low-minute observations can substantially affect the benchmark used in the standard player-level adjustment.

A simple response would be to impose a minimum playing-time threshold and discard all players below it. However, this solution is unsatisfactory because the choice of the threshold is arbitrary and may discard useful information. We therefore propose an alternative correction that preserves the logic of the role-specific benchmark while reducing the influence of extreme extrapolations.

Instead of projecting all players to full-season minutes, we construct the benchmark for player j using only players in the same role with a similar amount of court time. Specifically, we average $WINSEFF$ across players in the same role whose minutes lie within a 10% window around MIN_j. The modified player-level measure is defined as

$$WINSCORE_j^M = WINSEFF_j - OWINSEFF_j^{adj,M} + \frac{MIN_j}{TOTMIN_j} \cdot \frac{0.5 \cdot G}{5},$$

$$OWINSEFF_j^{adj,M} = \frac{1}{N_j^{\Delta}} \sum_{k \in \mathrm{role}(j) \mathrm{MIN}_k \in [0.9 \cdot \mathrm{MIN}_j,\ 1.1 \cdot \mathrm{MIN}_j]} WINSEFF_k,$$

where N_j^{Δ} is the number of players in the same role as player j whose playing time lies within 10% of MIN_j.

The proposed correction does not claim to fully eliminate the endogeneity of playing time, nor does it identify a causal effect of player quality on team outcomes. Rather, it addresses the most direct source of distortion in the standard procedure: the use of full-season projections for players whose minutes are too limited to provide reliable evidence about full-season performance. The correction can be interpreted as a role-specific matching procedure based on a playing-time caliper. Conditional on role, player j is compared only with players whose realized minutes lie within a 10 percent window around MIN_j. This is analogous to a caliper-matching rule in which players are matched on the observable determinants of their opportunity to contribute, namely role and realized court time. By comparing players only with others who received similar court time, the modified measure reduces the leverage of low-minute outliers while preserving the role-based comparison that underlies the original Winscore construction.

To assess the empirical relevance of the correction, Table 7 reports aggregate data on wins in the Euroleague for the 2024–25 season. For each of the eighteen teams, the table displays actual regular-season victories together with four model-based predictions: the exact efficiency-based Winscore measure computed from team-level data (**WS**); the linearized Winscore measure computed from team-level data (**Lin. WS**); and two player-level aggregations of individual contributions, computed respectively as

$$\sum_{j=1}^{J_i} WINSCORE_{j,i}^B \quad \text{and} \quad \sum_{j=1}^{J_i} \mathrm{WINSCORE}_{j,i}^{M}.$$

Table 7 highlights two main results. First, both the exact and the linearized team-level Winscore models provide reasonable approximations to actual team victories, supporting the use of the Winscore framework as a team-level description of performance. Second, the aggregation of individual contributions based on the modified measure, $\sum_{j=1}^{J_i} WINSCORE_{j,i}^M$, is much closer to team wins than the aggregation based on the original Berri-type adjustment, $\sum_{j=1}^{J_i} WINSCORE_{j,i}^B$.

The difference between the two player-level aggregations is informative. The standard measure tends to overestimate team performance because the role-specific benchmark is affected by players whose limited minutes are extrapolated to full-season equivalents. The modified measure reduces this distortion by comparing each player with others who had similar playing time. The improvement in the aggregate fit therefore supports the view that accounting for the selection of court time is important when translating a team-level efficiency model into player-level season contributions.

From team to players: Single game data

Our proposed $WINSCORE_j^M$ is very naturally extended when single game data are available, and the correction for the opponents performance and for the different amount of time spent on court by the different players can be dealt with in a more precise manner. In this case players can be matched precisely with their opponents to make the correction for the opponent's Winscore.

Consider the case in which statistics from a given game are available for all players of two teams, 1 and 2. The following procedure can be used to compute an individual measure of efficiency

- Group players by role, say back-court and front-court, and compute $WINSEFF_{j,1}^r$ and $WINSEFF_{j,2}^r$, $r = BC, FC$
- rank players by the time spent on court $MIN_{j,1}^r$ and $MIN_{j,2}^r$
- compute the player adjustment by considering the player in the same role who has the same rank with him in terms of minute spent on court and put the two players on court for the same amount of time.

Table 8. Euroleague: April 2025 BARCELLONA - VIRTUS 91-87.

(a) Barcellona: Players (surnames only), WINSCORE and PIR

Player	Seconds	Role	Opponent	Opp. Seconds	WINSCORE	PIR	Points
SATORANSKY	1756	BC	CORDINIER	1522	0.303	19	11
PUNTER	1744	BC	CLYBURN	1409	0.189	22	20
ANDERSON	1489	BC	HOLIDAY	1254	−0.123	12	12
BRIZUELA	1278	BC	MORGAN	1228	0.268	10	10
ABRINES	358	BC	HACKETT	925	0.042	0	0
ABRINES	358	BC	PAJOLA	916	0.013	0	0
PARKER	1366	FC	SHENGELIA	1706	0.045	7	10
PARRA	1251	FC	DIOUF	909	−0.143	13	10
HERNANGOMEZ	1114	FC	ZIZIC	894	0.072	12	9
VESELY	917	FC	POLONARA	755	−0.052	13	9
FALL	369	FC	AKELE	482	−0.072	−2	0
Sum	12000			12000	0.541	106	91

(b) Virtus Bologna: Players (surnames only), WINSCORE and PIR

Player	Seconds	Role	Opponent	Opp. Seconds	WINSCORE	PIR	Points.
CORDINIER	1522	BC	SATORANSKY	1756	−0.136	10	11.
CLYBURN	1409	BC	PUNTER	1744	−0.035	13	15.
HOLIDAY	1254	BC	ANDERSON	1489	0.208	16	16.
MORGAN	1228	BC	BRIZUELA	1278	−0.155	−3	5.
HACKETT	925	BC	ABRINES	358	−0.032	0	0.
PAJOLA	916	BC	ABRINES	358	0.044	3	0.
SHENGELIA	1706	FC	PARKER	1366	0.086	12	17.
DIOUF	909	FC	PARRA	1251	0.180	13	6.
ZIZIC	894	FC	HERNANGOMEZ	1114	0.017	13	12.
POLONARA	755	FC	VESELY	917	0.106	9	5.
AKELE	482	FC	FALL	369	0.134	2	0.
Sum	12000			12000	0.417	88	87.

- compute the contribution of the average player to that of the average team when the average player has been on court the same amount of time with player j

The following measure of efficiency is then computed:

$$WINSCORE^r_{j,1} = WINSEFF^r_{j,1} - WINSEFF^r_{j,2}\left(\frac{MIN^r_{j,1}}{MIN^r_{j,2}}\right) + \frac{MIN^r_{j,1}}{TOTMIN^r_{j,1}} * \frac{0.5}{5}$$

The sum of the Winscore efficiency of each player at the team level gives the expected Wins in the game considered. As expected Wins will not be the same for the two teams, the WINSCORE measure of efficiency takes into account not only performance relative to the opponents and the average players but also the outcome of the game.

To illustrate the measure, Table 8 reports data for the Euroleague game Virtus Bologna-Barcellona, played in April 2025 and won by Barcelona 91–87. Several comments are in order here. First, extracting individual measures of performance from data on a single game differs significantly from doing so over an entire season. In season-long data, errors tend to average out across many games, which is not the case for a single-game analysis. For example, consider the problem of assigning roles to players. With season data, one can rely on the official role classifications—Guard, Forward, and Center in the EuroLeague, or the more detailed classification used in the NBA (Center, Power Forward, Small Forward, Point Guard, and Shooting Guard). While such classifications may include some misassignments, these tend to average out over a season. In contrast, in a single game, due to injuries or tactical adjustments, coaches may assign players to roles different from their official designation. However, because we have detailed data from the game itself, we can reclassify players more appropriately into two functional groups: Front-Court (FC) and Back-Court (BC), and match them accordingly.

Second, the problem of matching each player with an opponent can be addressed by assigning a specific opponent, rather than comparing against an average player in the same role. Likewise, discrepancies in playing time can be handled by applying minor adjustments among players within the same group who played similar minutes during the game.

In Table 8, after classifying players from both teams as either BC or FC, they are sorted by minutes played within their group and then matched with their counterparts. If the two teams have different roster sizes, some players may be matched more than once. In the example at hand, since Barcelona had one fewer player than Virtus, both Pajola and Hackett are matched with Abrines by splitting the total time on court of Abrines of 716 seconds into two Abrines on court for 358 seconds with a Winscore equal to half of that of Abrines.

Once players are classified and matched to opponents with adjusted court time, their WINSCORE can be computed using Equation (10). The final two columns of the table report both WINSCORE and PIR for each player. The total team WINSCORE can be interpreted as the contribution of that game to the team's total wins for the season. For the game considered, Barcelona's contribution is 0.541 of an expected win and Virtus's is 0.417 of an expected win. Each team's total is the sum of individual player contributions. When both teams contribute similarly, it indicates a close game. In unbalanced games, teams' contribution polarizes. Therefore, WINSCORE also reflects the game outcome. In this case, the close values reflect a tightly contested game, with Barcelona winning by four points—a result also driven by the noise component in the data in addition to the structural elements captured by the model.

Notably, PIR and WINSCORE are not correlated. For instance, Parra has the lowest WINSCORE on Barcelona but the fourth-highest PIR. This is because Diouf outperformed Parra in their head-to-head comparison, and WINSCORE is computed relative to a relevant opponent and the average player. PIR, by contrast, is an absolute measure that does not account for opponent context, and thus often gives a very different picture.

The WINSCORE analysis reveals that Barcelona's Back-Court outperformed Virtus's, whereas Virtus's Front-Court outperformed Barcelona's. This contrast is not captured by the PIR measure.

Finally, while scoring has a dominant weight in PIR, the WINSCORE metric offers a more comprehensive assessment of player performance.

Conclusion

This paper has revisited the WINSCORE methodology and reformulated it within a model-based framework for evaluating player performance in professional basketball. The main objective has been to provide a transparent link between team outcomes and individual box-score statistics. Rather than treating performance evaluation as a purely descriptive aggregation problem, the WINSCORE approach starts from a model of team wins based on efficiency in managing possessions. This provides a coherent foundation for assigning weights to individual statistics and for interpreting player contributions in terms of their association with team success.

Empirical estimation using data from both the NBA and the EuroLeague indicates that the model captures an important component of variation in team wins. In particular, the encompassing tests show that WINSCORE contains predictive information that is not fully summarized by the widely used Performance Index Rating (PIR). Simulations based on the estimated model generate interpretable weights for the main box-score statistics. These weights are broadly stable across samples and leagues and are consistent with the logic of the possession-based model. These findings support the usefulness of the WINSCORE framework as an interpretable benchmark for performance evaluation, although they should not be read as evidence that the model captures all dimensions of player value.

The paper also extends the framework to derive player-level efficiency measures at both the season and single-game levels. These extensions incorporate opponent adjustments and playing-time normalization, thereby translating the team-level model into individual performance measures that remain consistent with the underlying structure of the game. At the season level, particular attention is given to the treatment of playing time. Since court time is not randomly assigned, standard full-season extrapolations may give excessive influence to players with very limited minutes. The modified WINSCORE measure proposed in the paper addresses this issue by comparing players with others in the same role and with similar playing time. The resulting player-level aggregation is closer to team-level wins than the standard adjustment, suggesting that accounting for the selection of court time improves the descriptive reliability of individual efficiency estimates.

The single-game application illustrates how the modified WINSCORE measure can be used in applied performance analysis. In the case study, the measure differs from PIR because it evaluates players relative to the context of the specific game, compares them with relevant opponents, accounts for the role of playing time, and incorporates team performance into the assessment of individual contributions. The case study should be interpreted as an illustration rather than as definitive validation. Its purpose is to show how the proposed framework can organize player-level information in a way that is more directly connected to team outcomes than an unweighted box-score index.

The robustness evidence in the paper is organized around the model-based structure. First, the team-level equation is estimated on distinct samples, including the historical NBA sample, a recent NBA sample, and the EuroLeague sample. Second, residual diagnostics are used to assess whether the fitted model leaves systematic patterns unexplained. Third, encompassing tests compare WINSCORE with PIR and with alternative box-score-based measures. Fourth, model simulation is used as a perturbation exercise: starting from a baseline scenario in which statistics are held at their sample averages, each statistic is varied in turn to compute its implied effect on predicted wins. Finally, the exact nonlinear model is compared with its linearized version before the latter is used for player-level decomposition. These exercises provide a systematic assessment of the stability, interpretation, and practical implications of the model within the information set considered in the paper.

Limitations. Despite its advantages, the proposed framework has several limitations. First, the model relies on box-score statistics. These data are widely available, standardized, and easy to compare across leagues, but they do not capture many actions that are central to basketball performance. Off-ball movement, defensive positioning, spacing, screen quality, help defense, shot contests, defensive assignments, and tactical roles are only imperfectly reflected in the box score. As a result, WINSCORE should be interpreted as a measure of contribution based on observable box-score production, not as a complete measure of player quality.

Second, the identification of player contributions rests on a structural relationship between possession efficiency and team wins. Although this relationship is theoretically grounded and empirically informative, it remains a simplification of the complex dynamics of basketball games. Player performance depends on interactions with teammates, opponents, coaching strategies, lineup composition, and game state. These factors are only partially captured by the current specification.

A further limitation concerns the transition from the team-level model to the individual level. The attribution of weights to player statistics is based on a linearization of the model around the mean. This first-order approximation has the advantage of being transparent and tractable, but it may neglect nonlinearities and higher-order interactions among variables. For example, the value of a rebound, assist, or three-point shot may depend on score margin, lineup configuration, opponent quality, or phase of play.

At the same time, the empirical results provide evidence of a meaningful degree of stability in the estimated relationships. The weights are similar across different NBA samples and between the NBA and the EuroLeague, suggesting that the possession-based mechanism captures features that are common across competitive environments. This stability is encouraging, but further work is needed before stronger claims of structural invariance can be made. A more formal assessment over longer time periods, additional competitions, and different institutional settings remains an important direction for future research.

Finally, although the modified measure explicitly addresses the endogeneity of playing time, this issue cannot be fully eliminated with box-score data alone. Coaching decisions reflect information that is not directly observable, including player fitness, tactical instructions, defensive matchups, training performance, and expected complementarities with teammates. The proposed correction reduces the most direct distortion generated by full-season extrapolation, but it does not transform the measure into a causal estimate of player impact.

Beyond the low-minute extrapolation problem, other failure modes should also be noted. First, the measure is less reliable in very small samples, especially at the single-game level, where random shooting variation, foul trouble, injuries, and unusual tactical assignments may have large effects. Second, role classification can become unstable when players are used outside their usual positions or when teams employ non-standard lineups. Third, changes in rules, pace, shot selection, or league style may alter the mapping between possession efficiency and wins, requiring re-estimation of the model. Fourth, because WINSCORE is based on box-score statistics, it may understate contributions that do not appear directly in the box score, such as spacing, screen quality, off-ball movement, defensive communication, and matchup-specific tactical responsibilities. These cases do not invalidate the model, but they define the conditions under which its output should be interpreted as a disciplined box-score-based benchmark rather than as a complete measure of player value.

Directions for future research. Several extensions could further develop the model-based approach proposed in this paper. A first direction is methodological. Future work could move beyond the first-order linear approximation used here by adopting higher-order perturbation methods for the underlying nonlinear model. Such methods, widely used in macroeconomics and dynamic modeling (Schmitt-Grohé and Uribe, 2004), would allow the construction of marginal contribution measures that account more explicitly for nonlinearities and interactions among box-score variables.

A second direction is the integration of richer data sources. Play-by-play data would make it possible to condition player contributions on possession type, score margin, time remaining, lineup configuration, and opponent strength. Tracking and spatial data would allow the model to incorporate shot location, defensive pressure, spacing, off-ball movement, screening actions, help defense, and defensive assignments. In this setting, the value of each statistic could be made context-dependent. For example, the marginal value of an assist, rebound, or three-point shot could differ depending on whether it occurs in transition, in half-court offense, against a set defense, or in high-leverage situations.

A third extension concerns scalability and real-time analytics. The current framework is implemented at the season and single-game levels, but the same logic could be extended to rolling game windows, lineup stints, or possession-level updates. This would allow teams to monitor player contributions dynamically during the season and potentially during games. A scalable implementation could be used to generate real-time or near real-time indicators of which players, lineups, or role groups are creating or losing value relative to opponents.

A fourth direction concerns causal inference. The measures developed in this paper are descriptive: they quantify player contribution as implied by the estimated relationship between box-score statistics and wins. They do not identify the causal effect of a player on team outcomes. A causal extension would require an empirical design that generates plausibly exogenous variation in player availability or lineup composition. One possible route is to combine the WINSCORE framework with lineup-based panel models using player, team, opponent, and game fixed effects. Another is to exploit quasi-experimental variation from injuries, suspensions, staggered rest, or substitution patterns that shift minutes across players for reasons not directly related to contemporaneous performance. A further possibility is to use possession-level or stint-level data to compare team efficiency before and after substitutions, while controlling for score margin, opponent strength, lineup composition, and game state. These extensions would help distinguish observed productivity from causal impact and clarify how much of a player's measured contribution reflects individual ability, teammate complementarities, coaching decisions, or team context. Finally, the external validity of the framework should be assessed using additional competitions and longer time spans. Extending the analysis to other European leagues, domestic cups, international competitions, and women's basketball would provide further evidence on the stability of the estimated weights and on the extent to which the model must be adapted to different playing styles and institutional environments.

Practical implications. From an applied perspective, the results suggest that player evaluation can benefit from metrics that explicitly link individual actions to team outcomes. The WINSCORE framework offers a transparent way to translate box-score statistics into estimated contributions to wins, while making clear which assumptions are required for this translation. This can support decision-making in player comparison, roster construction, rotation analysis, and contract evaluation, especially when richer tracking data are not available.

The opponent-adjusted and time-normalized versions of WINSCORE also highlight the importance of evaluating players relative to their competitive context. Player performance should not be assessed only in absolute box-score terms, since the same statistical line may have different value depending on the opponent, role, minutes played, and team environment. This feature is particularly relevant for scouting and game preparation, where matchup-specific information is often more useful than aggregate season averages.

The single-game implementation provides a practical tool for post-game analysis. By decomposing value across players and role groups, such as back-court and front-court units, the measure can help identify which parts of a team generated or lost value relative to the opponent. The framework should therefore be viewed not as a replacement for expert judgment or richer tactical analysis, but as an interpretable quantitative benchmark that can complement them.

ORCID iDs

Gabriele Carta https://orcid.org/0009-0004-9781-3258
Carlo A Favero https://orcid.org/0000-0002-1668-9426

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. For an extensive and systematic discussion of this literature, see Sarlis (2024).
2. The replication package nevertheless provides the full data-retrieval and estimation pipeline in R, making it straightforward to implement such predictive benchmarks in future work.
3. *https : //www.euroleaguebasketball.net/en/euroleague/* allows to retrieve real-time and historical standard and advanced statistics about competitions, teams, players and for the Euroleague. The same stats for the NBA can be retrieved from *https : //www.basketball − reference.com/*. The Data appendix provides a detailed description of the data we have used in this study.
4. https://about.fiba.basketball/en/organization/recognized-organizations/world-association-of-basketball-coaches
5. The presence of fixed effect for teams and time-effect for season was checked and the relevant coefficients were found to be not statistically different from zero
6. The results in Table 1 are obtained by running all models using a demeaned measure of efficiency, as the sample mean of this variable is very close to zero but not exactly equal to zero.
7. In fact, the institutional framework of the NBA has explicit rules to promote competitive balance, such as the draft mechanism and the salary cap, that the Euroleague does not have.
8. This follows from the Frisch–Waugh–Lovell theorem, according to which a coefficient in a multiple regression can be obtained by residualizing both the dependent variable and the regressor of interest with respect to the other included regressors, and then regressing the former residuals on the latter, see Davidson and MacKinnon (1993).

9. For perspective, the highest number of points ever scored by a player in an NBA regular-season game is 100, achieved by Wilt Chamberlain on March 2, 1962. Chamberlain played the entire game for the Philadelphia Warriors against the New York Knicks, in a 169–147 victory.
10. basketball reference.com
11. https://hackastat.eu/en/players-stats/

References

Berri D, Schmidt M and Brook S (2006) *The Wages of Wins: Taking Measure of the Many Myths in Modern Sport*. Stanford University Press.

Berri DJ and Eschker E (2005) Performance when it counts? the myth of the prime time performer in professional basketball. *Journal of Economic Issues* 39: 798–807.

Cervone D, D'Amour A, Bornn L and Goldsberry K (2014) Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data, in: MIT Sloan Sports Analytics Conference.

Cooley TF (1997) Calibrated models. *Oxford Review of Economic Policy* 13: 55–69.

Davidson R and MacKinnon JG (1993) *Estimation and Inference in Econometrics*. Oxford University Press.

Favero CA (2001) *Applied Macroeconometrics*. OUP Oxford.

Hollinger J (2002) *Pro Basketball Prospectus, 2002*. Brassey's Incorporated.

Iatropoulos D, Sarlis V and Tjortjis C (2025) A data mining approach to identify NBA player quarter-by-quarter performance patterns. *Big Data and Cognitive Computing* 9: 74.

Ilardi S and Barzilai A (2008) Adjusting plus-minus for NBA players using box score data. *Journal of Quantitative Analysis in Sports* 4: 1–14.

James G, Witten D, Hastie T, Tibshirani R et al. (2013) *An Introduction to Statistical Learning*. Vol. 112. Springer.

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

Lakhloufi H (2025) Regularized adjusted plus-minus. https://metricgate.com/docs/regularized-adjusted-plus-minus/. MetricGate (accessed: 2026-05-11).

Mizon GE and Richard JF (1986) The encompassing principle and its application to testing non-nested hypotheses. *Econometrica: Journal of the Econometric Society* 0: 657–678.

Oliver D (2004) *Basketball on Paper: Rules and Tools for Performance Analysis*. Potomac Books, Inc.

Page SE (2018) *The Model Thinker: What You Need to Know to Make Data Work for You*. Hachette UK.

Petridis C and Pelechrinis K (2026) Lineup regularized adjusted plus-minus (L-RAPM): Basketball lineup ratings with informed priors. arXiv preprint arXiv:2601.15000.

Rosenbaum DT (2004) Measuring how nba players help their teams win. Available at http://www.82games.com/comm30.htm.

Sarlis V (2024) Data science for sports analytics.

Schmitt-Grohé S and Uribe M (2004) Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control* 28: 755–775.

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

Terner Z and Franks A (2021) Modeling player and team performance in basketball. *Annual Review of Statistics and Its Application* 8: 1–23.

Wen R, Cao Q and Wang H (2023) Are pay decisions based on pre-contract efficiency necessarily correct? evidence from players contracts of national basketball association. *SAGE open* 13: 21582440231204149.

Zhou Y and Li T (2024) Quantitative analysis of professional basketball: A qualitative discussion. *Journal of Sports Analytics* 9: 273–287.

Zimbalist AS (2002) Competitive balance in sports leagues: An introduction.

The data

Data for the EuroLeague have been retrieved using the R package `euroleaguer`, an unofficial API wrapper for EuroLeague and EuroCup basketball data. Data for the NBA have been web-scraped from Basketball Reference using R code provided in the replication package. The replication package is publicly available at [https://github.com/gabbocarta/winscore_replication.git] and includes the scripts for data retrieval, data cleaning, estimation, robustness checks, and the construction of all tables and figures reported in the paper. The following variables are used in our analysis:

Table A.1. Description of box-score statistics.

Variable	Description
SeasonCode	Identifier for the competition season
GameCode	Unique identifier for each game
TeamCode	Identifier for the team
Seconds	Total seconds played by the player
PTS	Total points scored
2PM	Two-point field goals made
2PA	Two-point field goals attempted
3PM	Three-point field goals made
3PA	Three-point field goals attempted
FTM	Free throws made
FTA	Free throws attempted
OREB	Offensive rebounds
DREB	Defensive rebounds
REB	Total rebounds
AST	Assists
STL	Steals
TO	Turnovers
BLK	Blocks made

(continued)

Table A.1. Continued.

Variable	Description
BLKA	Shots blocked against the player
FC	Personal fouls committed
FD	Personal fouls drawn
PIR	Performance Index Rating
FG%	Overall field-goal percentage
2P%	Two-point field-goal percentage
3P%	Three-point field-goal percentage
FT%	Free-throw percentage
Year	Calendar year of the game

Table B.2. Chow-type tests for equality of efficiency slopes.

Comparison	*F*-statistic	*p*-value
NBA 1994–2005 vs Euroleague 2022–2025	$F(1, 366) = 1322.00$	$p < 0.001$
NBA 2022–2025 vs Euroleague 2022–2025	$F(1, 170) = 438.41$	$p < 0.001$

Diagnostics and robustness

In this section, we report the results of the diagnostic tests and robustness analyses conducted to assess the specification, estimation, and simulation of the Winscore model.

Equality between employed and acquired possessions. The following graphs show that the assumption of possession symmetry over the course of a season is supported by the data. Although possession symmetry may be violated in a single game when one team controls a disproportionate number of end-of-period possessions, these game-level discrepancies are averaged out over the season. As a result, the difference between acquired possessions and employed possessions is virtually zero, and statistically indistinguishable from zero, across teams and leagues.

Testing the Stability of the coefficients in the Winscore model across different leagues. In this section we analyze formally the stability of estimatd coefficients in the Wiinscore model for NBA and euroleague data. We consider the following specification on the pooled sample that combines NBA team-seasons from 2022–2025 and Euroleague team-seasons from 2022–2024,

$$W\%_{i,t} = \alpha + \gamma NBA_{i,t} + \beta(PTSxEP_{i,t} - PTSAxAP_{i,t}) + \delta\big(PTSxEP_{i,t} - PTSAxAP_{i,t}\big) \times NBA_{i,t} + u_i.$$

where i, t denote a team-season observation, the dependent variable is the team win percentage,

$$W\%_{i,t},$$

the main explanatory variable is the Winscore efficiency measure,

$$(PTSxEP_{i,t} - PTSAxAP_{i,t}).$$

and the effect of an NBA dummy variable is allowed on the constant and on the slope of the regression on the pooled sample.

*cases*1ifobservationi, tbelongstotheNBAsample,

0ifobservationi, tbelongstotheEuroleaguesample.cases

The interaction coefficient δ measures the difference between the NBA and Euroleague slopes. A Chow-type test for equality of slopes across the NBA and Euroleague samples is then implemented. This is an F-test comparing a restricted model with a common slope for our efficiency measure to an unrestricted model that allows the slope to differ between NBA and Euroleague.

The null hypothesis is that the slope of efficiency is the same in the NBA and in the Euroleague:

$$H_0 : \delta = 0.$$

Under the null hypothesis, efficiency has the same marginal effect on win percentage in the two competitions.

We run the test for both NBA subsamples with respect to the Euroleague sample. The two models are nested. The restricted model excludes the interaction term, while the unrestricted model includes it. We therefore compare them using an F-test:

$$F = \frac{(SSR_R - SSR_U)/(df_R - df_U)}{SSR_U/df_U},$$

where SSR_R is the residual sum of squares from the restricted model, SSR_U is the residual sum of squares from the unrestricted model, and df_U and df_R are, respectively, the residual degrees of freedom from the unrestricted and restricted model.

In both cases, the null hypothesis of equal slopes is rejected. The evidence indicates that the relationship between efficiency per game and win percentage differs between the NBA samples and the Euroleague sample.

Residual analysis

Alternative efficiency measures. This section defines the alternative efficiency measures used as benchmarks against the Winscore-based model. We consider four box-score plus minus measures: OBPM, DBPM, BPM, and VORP, together with the Four Factors model.

Lineup-based, proprietary, and tracking-based measures such as RAPM, EPM, DARKO, and spatial models are important complementary tools, but they rely on information sets that are not uniformly available for the NBA and EuroLeague samples considered here. The benchmarking exercise is therefore restricted to measures that can be constructed from comparable box-score information across the samples analyzed.

Four factors. The Four Factors summarize team performance through shooting efficiency, turnovers, offensive

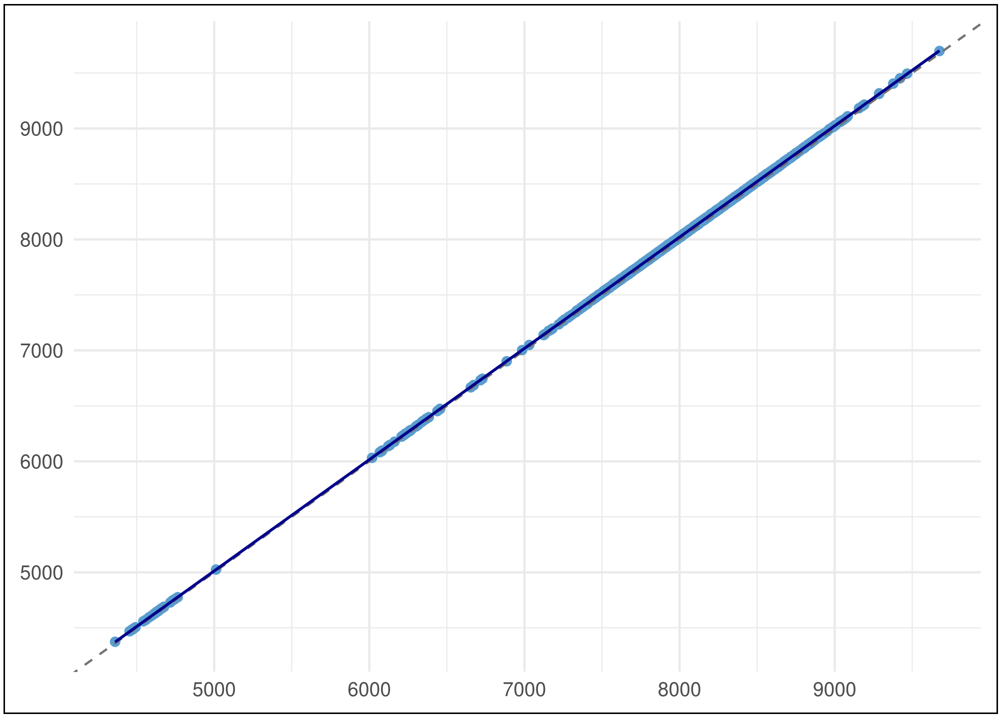

Figure B.1. Employed and acquired possessions by team in NBA seasons, 1980–2025.

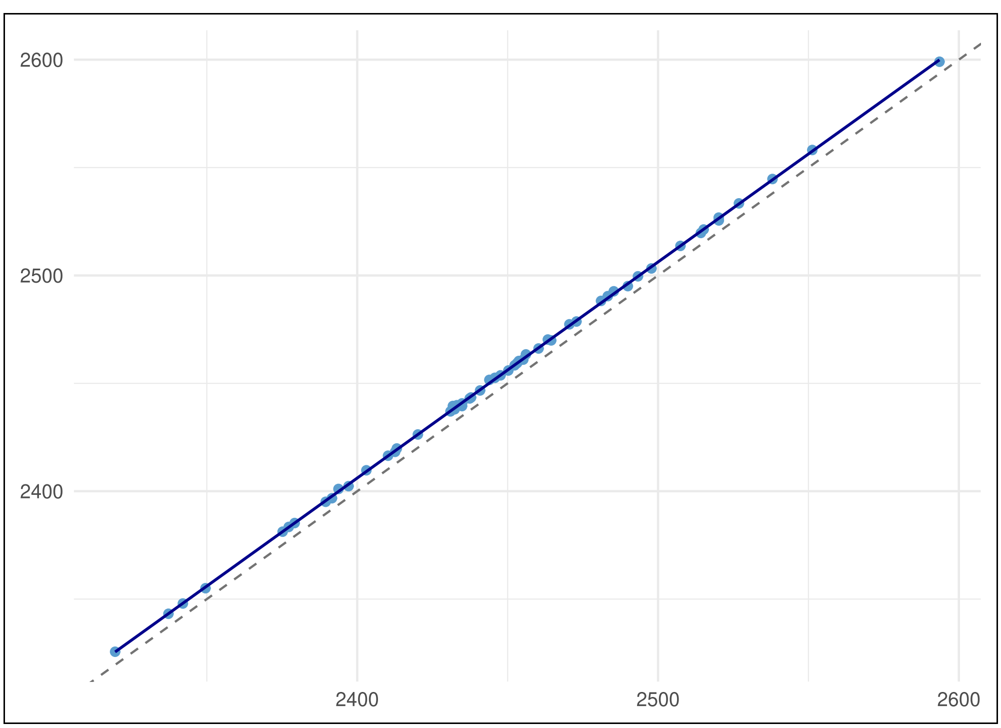

Figure B.2. Employed and acquired possessions in EuroLeague team seasons, E2022–E2024.

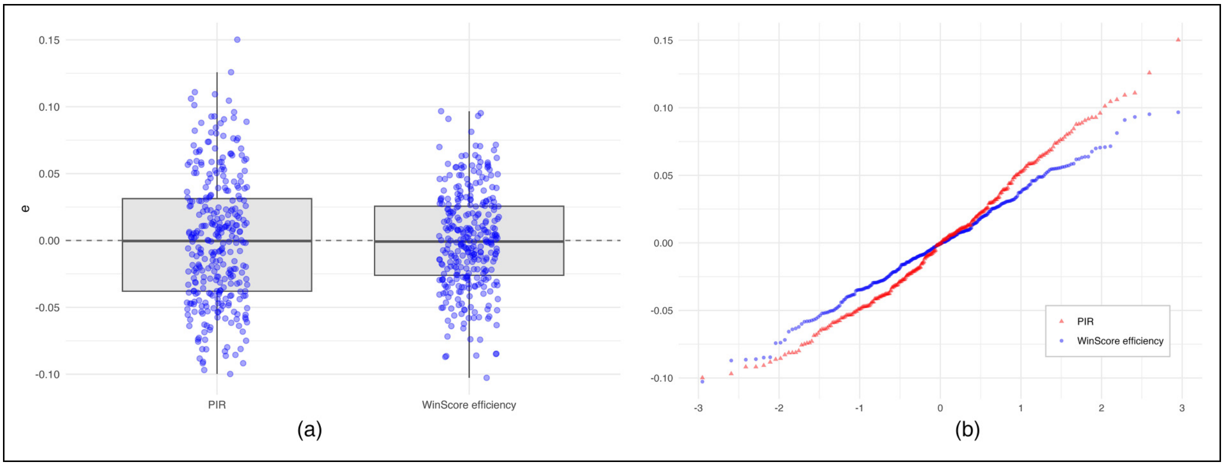

Figure B.3. Residual analysis: NBA 1994–2005, excluding 1999. The left panel compares residuals from the Winscore efficiency and PIR models. The right panel shows Q–Q plots for the same residuals.

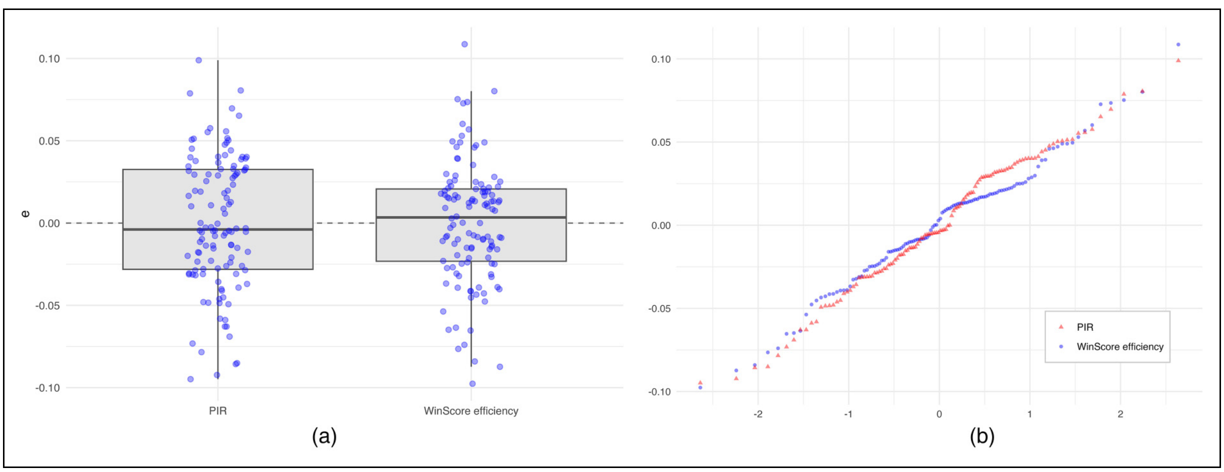

Figure B.4. Residual analysis: NBA 2022–2025. The left panel compares residuals from the Winscore efficiency and PIR models. The right panel shows Q–Q plots for the same residuals.

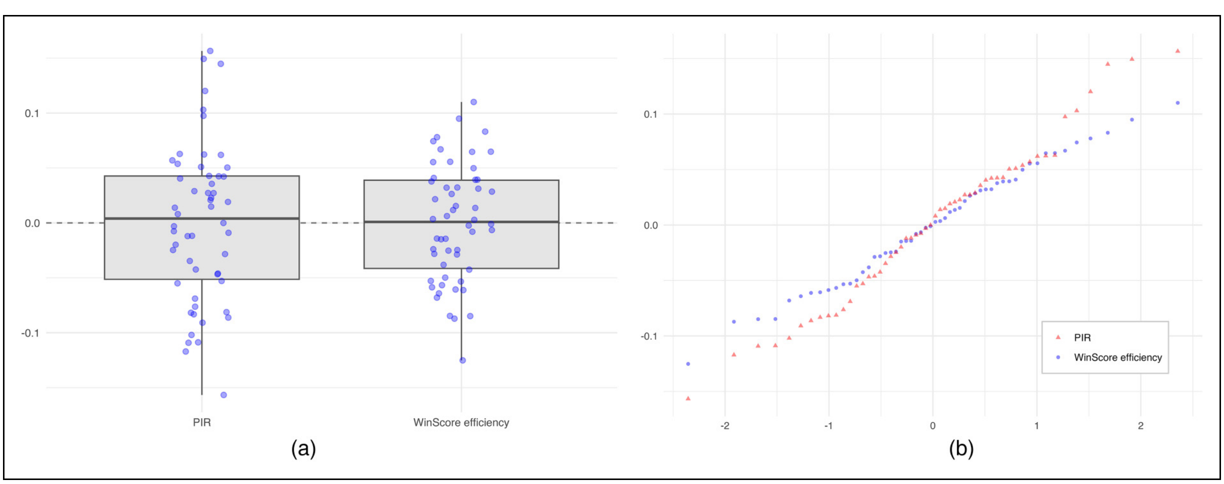

Figure B.5. Residual analysis: Euroleague 2022–2025. The left panel compares residuals from the Winscore efficiency and PIR models. The right panel shows Q–Q plots for the same residuals.

Table B.3. Four factors regressions and encompassing tests: NBA samples.

	NBA 2022–2025	NBA 1994–2005
Intercept	-149.31^{***}	-89.32^{***}
	(26.71)	(15.76)
eFG%	408.19^{***}	343.93^{***}
	(37.49)	(26.96)
TOV%	-522.96^{***}	-468.49^{***}
	(79.08)	(50.85)
ORB%	100.35^{***}	18.23
	(27.63)	(19.51)
FTr	39.22	81.08^{***}
	(33.64)	(17.68)
Observations	120	316
Adjusted R^2	0.626	0.522
F-statistic	50.85^{***}	87.10^{***}
Encompassing test		
$\widehat{Wins}^{WS}$	0.979^{***}	0.988^{***}
	(0.041)	(0.020)
$\widehat{Wins}^{FF}$	0.021	0.013
	(0.042)	(0.020)
Encompassing adj. R^2	0.995	0.995

$^{***}p < 0.001$

rebounding, and free-throw rate. Unlike BPM-based measures, these are team-level box-score measures rather than player-level metrics.

The effective field-goal percentage is:

$$eFG\%_{i,t} = \frac{FGM_{i,t} + 0.5 \cdot 3PM_{i,t}}{FGA_{i,t}}.$$

The turnover percentage is:

$$TOV\%_{i,t} = \frac{TOV_{i,t}}{FGA_{i,t} + 0.44 \cdot FTA_{i,t} + TOV_{i,t}}.$$

The offensive rebounding percentage is:

$$ORB\%_{i,t} = \frac{ORB_{i,t}}{ORB_{i,t} + OppDRB_{i,t}}.$$

The free-throw rate is:

$$FTr_{i,t} = \frac{FTA_{i,t}}{FGA_{i,t}}.$$

We estimate the Four Factors benchmark model as:

$$\begin{aligned} Wins_{i,t} = \alpha + \beta_1 eFG\%_{i,t} + \beta_2 TOV\%_{i,t} + \beta_3 ORB\%_{i,t} \\ + \beta_4 FTr_{i,t} + \varepsilon_{i,t}. \end{aligned}$$

Offensive box Plus/Minus. Offensive Box Plus/Minus, denoted *OBPM*, is a box-score estimate of a player's offensive contribution relative to a league-average player. It measures the number of offensive points per 100 possessions

Table B.4. Single-measure regressions and encompassing tests: NBA samples.

Sample	Measure	Reg. Coefficient	Adj. R^2	$\widehat{Wins}^{WS}$	$\widehat{Wins}^{Alt}$
NBA 2022–2025	OBPM	14.32^{***}	0.723	0.990^{***}	0.010
NBA 2022–2025	DBPM	15.93^{***}	0.628	1.004^{***}	−0.004
NBA 2022–2025	BPM	10.36^{***}	0.936	$0.677^{\cdot}$	0.324
NBA 2022–2025	VORP	2.06^{***}	0.935	0.734^{*}	0.266
NBA 1994–2005	OBPM	14.10^{***}	0.609	0.999^{***}	0.001
NBA 1994–2005	DBPM	14.79^{***}	0.582	1.000^{***}	0.000
NBA 1994–2005	BPM	11.38^{***}	0.941	0.955^{***}	0.045
NBA 1994–2005	VORP	2.27^{***}	0.941	0.958^{***}	0.042

$^{***}p < 0.001$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. .
OBPM, DBPM, BPM, and VORP are demeaned within sample. .
.

Table B.5. Four factors regression and encompassing test: Euroleague 2022–2024.

	Euroleague 2022–2024
Intercept	-71.50^{***}
	(15.46)
eFG%	141.67^{***}
	(19.85)
TOV%	-90.34^{**}
	(32.54)
ORB%	46.94^{*}
	(18.00)
FTr	33.82^{**}
	(12.23)
Observations	54
Adjusted R^2	0.613
F-statistic	21.98^{***}
Encompassing test	
$\widehat{Wins}^{WS}$	0.975^{***}
	(0.107)
$\widehat{Wins}^{FF}$	0.026
	(0.108)
Encompassing adj. R^2	0.989

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

that the player contributes above league average, translated to an average team context.

At the player level, we denote this measure as:

$$OBPM_{j,t},$$

where j indexes the player and t indexes the season.

To compare this measure with team-level Winscore predictions, we aggregate player-level *OBPM* to the team-season level using minutes as weights:

$$OBPM^{team}_{i,t} = \frac{\sum_{j \in i,t} MP_{j,t} \cdot OBPM_{j,t}}{\sum_{j \in i,t} MP_{j,t}},$$

Table B.6. Single-measure regressions and encompassing tests: Euroleague sample.

Sample	Measure	Reg. Coefficient	Adj. R^2	$\widehat{Wins}^{WS}$	$\widehat{Wins}^{Alt}$
Euroleague 2022–2024	OBPM	3.70***	0.601	1.006***	−0.006
Euroleague 2022–2024	DBPM	4.03***	0.328	1.004***	−0.004
Euroleague 2022–2024	BPM	3.38***	0.838	1.216**	−0.216
Euroleague 2022–2024	VORP	0.78***	0.863	0.406	0.594

$^{***}p < 0.001$, $^{**}p < 0.01$.
OBPM, DBPM, BPM, and VORP are aggregated from player-level Hack a Stat data and demeaned within sample.

where $MP_{j,t}$ denotes minutes played by player j in season t, and $j \in i, t$ denotes players on team i in season t.

Defensive box Plus/Minus. Defensive Box Plus/Minus, denoted $DBPM$, is the defensive counterpart of $OBPM$. It estimates the number of defensive points per 100 possessions that a player contributes above a league-average player, translated to an average team context.

At the player level:

$$DBPM_{j,t}.$$

The team-level measure is computed as the minutes-weighted average:

$$DBPM_{i,t}^{team} = \frac{\sum_{j \in i,t} MP_{j,t} \cdot DBPM_{j,t}}{\sum_{j \in i,t} MP_{j,t}}.$$

Box Plus/Minus. Box Plus/Minus, denoted BPM, combines offensive and defensive contributions into a single measure. It estimates the total number of points per 100 possessions that a player contributes above a league-average player.

At the player level:

$$BPM_{j,t} = OBPM_{j,t} + DBPM_{j,t}.$$

At the team-season level, we compute:

$$BPM_{i,t}^{team} = \frac{\sum_{j \in i,t} MP_{j,t} \cdot BPM_{j,t}}{\sum_{j \in i,t} MP_{j,t}}.$$

Value over replacement player. Value Over Replacement Player, denoted $VORP$, transforms BPM into a cumulative value measure. It estimates how many points a player contributes above a replacement-level player, where replacement level is conventionally set at −2.0 BPM.

At the player level, VORP is a playing-time-adjusted transformation of BPM:

$$VORP_{j,t} = \left(BPM_{j,t} - BPM^{replacement}\right) \cdot \frac{MP_{j,t}}{MP_{team,t}} \cdot \frac{G_{team,t}}{G^{season}},$$

where:

$$BPM^{replacement} = -2.0.$$

Here, $MP_{j,t}$ denotes minutes played by player j in season t, $MP_{team,t}$ denotes total team minutes in season t, $G_{team,t}$ denotes the number of games played by the player's team in season t, and G^{season} denotes the reference number of regular-season games. In the NBA case, $G^{season} = 82$, while in the Euroleague regular-season sample, $G^{season} = 34$.

Since $VORP$ is already cumulative, the team-level measure is obtained by summing across players:

$$VORP_{i,t}^{team} = \sum_{j \in i,t} VORP_{j,t}.$$

Results. For each of our samples, we estimate simple regressions of regular-season wins on each alternative measure and then compare each alternative model with the Winscore model using an encompassing regression. The encompassing specification is:

$$Wins_{i,t} = \gamma_1 \widehat{Wins}_{i,t}^{WS} + \gamma_2 \widehat{Wins}_{i,t}^{Alt} + \varepsilon_{i,t}.$$

We report the Four Factors model separately because it contains four regressors. We then report the remaining alternative measures in a more compact table, since OBPM, DBPM, BPM, and VORP each enter as a single demeaned regressor. For the two NBA samples, the player-based OBPM, DBPM, BPM, and VORP measures are downloaded from Basketball Reference[10]. The encompassing regressions show that the fitted values from the Winscore model are always statistically significant in almost all comparisons, which is not the case for the fitted values from the alternative models. This suggests that the alternative measures do not add explanatory power once the Winscore fitted values are included.

Also for the Euroleague sample, we compare the Winscore model with the Four Factors benchmark and with the player-based OBPM, DBPM, BPM, and VORP measures. The latter are downloaded from Hack a Stat[11] at the player-season level, aggregated to the team-season level, and then merged with our Euroleague team-season dataset. As in the NBA samples, the encompassing regressions always favor the Winscore model.

Assists and personal fouls

Table B.7. Regression of wins on efficiency, assists, and fouls: NBA 1994–2005 (Berri et al., 2006), NBA 2022–2025, and euroleague 2022–24.

	Dependent variable: Wins in Regular Season		
	NBA 1994–2005	NBA 2022–2025	Euroleague 2022–24
Intercept	41.00***	41.00***	17.00***
	(0.17)	(0.28)	(0.26)
$(PTSxEP_{i,t} - PTSAxAP_{i,t})$	249.81***	230.36***	77.92***
	(5.87)	(7.59)	(6.19)
$AST_{i,t} - OAST_{i,t}$	0.0018	0.0004	0.0002
	(0.0013)	(0.0020)	(0.0044)
$PF_{i,t} - OPF_{i,t}$	-0.0013	-0.0028	-0.0011
	(0.0014)	(0.0028)	(0.0053)
Observations	316	120	54
Multiple R^2	0.9483	0.9383	0.8626
Adjusted R^2	0.9478	0.9367	0.8544
F Statistic	1908.9***	587.6***	104.7***

$^{***}p < 0.001$.