A Bayesian two-stage framework for lineup-independent assessment of individual rebounding ability in the NBA

Nicholas Kiriazis
Department of Mathematics and Statistics
McGill University, Montréal
July, 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Science

© Nicholas Kiriazis, 2023

Acknowledgements

Firstly, I would like to thank my supervisor, Professor Christian Genest of McGill University, and my co-supervisor, Professor Alexandre Leblanc of the University of Manitoba, for their continuous support during my master's journey. Without their guidance when I wasn't sure how to proceed, their experience and knowledge when presented with technical questions, and most importantly, their willingness to take a statistical novice under their wing, I would not have made it very far.

My thanks also go to the Department of Mathematics and Statistics at McGill University as a whole, for nurturing my statistical curiosity and for providing me with the means and tools to work on my research.

Finally, I am particularly grateful to my dad, Tom, for instilling me with enough stubbornness to undertake such a challenge, and my mom, Vivian, for instilling me with enough patience to see it through. Moreover, I would like to thank my brothers Charlie and Lucas for reminding me to laugh a little along the way.

Funding in partial support of this work was provided by the Canada Research Chairs Program and the Natural Sciences and Engineering Research Council of Canada through grants to Professor Genest. Additional funding from the Institut des sciences mathématiques du Québec is also gratefully acknowledged.

Abstract

In basketball, existing methods to assess individual rebounding ability are problematic because they depend on all players present on the court rather than just the player of interest. Although there exist modelling approaches to correct for this dependence, they are generally unsuitable for events with binary outcomes. In this thesis, we propose a Bayesian, two-stage model for predicting conditional rebounding rates in the National Basketball Association (NBA). Although similar in flavour to the popular APM framework, it is different in that it does not assume that individual contributions are linearly additive on the response scale. Furthermore, we improve the regularization approach by using rebounding-specific heuristics. After defining the model, a simulation study is performed to verify its effectiveness, and the parameters are then estimated using data from the 2020–21 NBA season. Predictions are then made for rebounding in the 2021–22 season. It was found that there are some players who are excellent at stealing rebounds away from the opposing team without collecting them themselves, whereas others are simply collecting rebounds that their team has already secured. These subtleties are not captured by relying on traditional rebounding metrics.

Résumé

Au basketball, les évaluations existantes de la capacité de rebond individuelle sont problématiques car elles dépendent de tous les joueurs présents sur le terrain, plutôt qu'uniquement du joueur d'intérêt. Bien qu'il existe des approches de modélisation qui tentent de corriger pour cette dépendance, elles ne sont généralement pas appropriées pour des événements dont les résultats sont binaires. Dans ce mémoire, nous proposons un modèle bayésien hiérarchique à deux degrés pour prédire les taux de rebonds conditionnels dans la National Basketball Association (NBA). Bien que la méthodologie suive dans les grandes lignes l'approche APM, elle est différente car on ne suppose pas que les contributions individuelles sont linéairement additives sur l'échelle de la variable prédite. Nous améliorons en outre l'approche de régularisation en utilisant des heuristiques spécifiques aux rebonds. Une fois le modèle défini, une étude de simulation est réalisée pour juger de son efficacité ; par suite, les paramètres sont estimés à partir des données de la saison 2020–21 de la NBA. Des prévisions sont faites pour les taux de rebonds lors de la saison 2021–22. Il ressort qu'il existe des joueurs qui sont très bons pour empêcher l'adversaire de prendre des rebonds sans nécessairement les prendre eux-mêmes, tandis que d'autres joueurs se contentent de prendre les rebonds qui appartiennent déjà à l'équipe. Ces subtilités échappent aux mesures traditionnelles.

Author contributions

Except as noted below, everything in this dissertation is my original work. None of the content was produced in conjunction with anyone else.

The data used in this thesis were made publicly available by the NBA, and can be accessed at `https://www.nba.com/stats`. The data were accessed using the NBA API, for which the repository can be found at `https://github.com/swar/nba_api`.

I proposed the model formulation, derived the relevant results, and implemented the estimation procedure. I received very helpful guidance throughout the process from Professors Christian Genest and Alexandre Leblanc, who were also kind enough to review this thesis in detail, and suggest improvements, after I wrote the initial draft.

Contents

Acknowledgements . . . i
Abstract . . . ii
Résumé . . . iii
List of Figures . . . ix
List of Tables . . . xi

1 Introduction **1**

2 Contribution **4**

2.1 Introduction . . . 4
2.2 The data set . . . 6
2.2.1 Box score data sets . . . 6
2.2.2 Play-by-play data set . . . 8
2.3 Rebounding models . . . 8
2.3.1 A brief overview of APM and RAPM . . . 9
2.3.2 Limitations of APM based approaches in rebounding . . . 10
2.3.3 High-level ideas . . . 11
2.3.4 Modelling individual contributions to team rebounding . . . 12
2.3.5 Modelling individual rebound collecting ability . . . 26
2.4 Player clustering . . . 30
2.4.1 Heuristics . . . 31
2.4.2 Reformatting the data set . . . 33

2.4.3 Clustering using Gaussian mixture models 36
2.4.4 Clustering approach 37
2.4.5 Results and interpretation 40
2.4.6 Assigning unusable players a position 42
2.5 Estimation Results 46
2.5.1 Team rebounding estimation results 46
2.5.2 Individual rebound collecting estimation results 52
2.6 Results 55
2.6.1 Prediction 55
2.6.2 Player rebounding assessment 59
2.6.3 Example: The Timberwolves acquire Rudy Gobert 62
2.7 Discussion 63
2.7.1 Uncertainty considerations 63
2.7.2 The subtle misleadingness of rebounding numbers 64
2.7.3 Model limitations 66

3 Conclusion **68**

A Description of the variables **71**
A.1 Defense variables 71
A.2 Offense variables 72

B ”Missclassified” centers **73**

C Center classifications **75**

D Non-center classifications **78**

E Replacement player classifications **87**

F Parameter estimates **93**

List of Figures

2.1 Example priors means for the team rebounding ability of some players in the 2019–20 season, as well as 50% central interval. 21
2.2 Comparison between team rebounding rates of observed lineups and generated probabilities. 24
2.3 Comparison between the distribution of observed rebounding rates of players and the simulated individual rebounding rates. 25
2.4 Scatter plot of shifted estimated parameters (y-axis) against their true values (x-axis). 27
2.5 Distribution of possessions played by players in the 2020–21 NBA Season. A total of 540 players were in the data set. 30
2.6 Diagram of the high-level modelling procedure. 32
2.7 Left: Correlation matrix of the offensive feature data set. Right: Correlation matrix of the defensive feature data set. 34
2.8 Left: Scree plot for first 10 principal components of offensive data set. Right: Scree plot for first 10 principal components of defensive data set. . . 35
2.9 Left: Scree plot for average variance proportion retained for a given position in offensive data set. Right: Scree plot for average variance proportion retained for a given position in defensive data set. Note that the labels are those used by the NBA. 35
2.10 Left: Scatter plots of position clusters on offense. Right: Scatter plots of position clusters on defense. 44

2.11 Left: League-wide convergence rate for offensive non-centers. Right: League-wide convergence rate for defensive non-centers. 45
2.12 Distribution of prior means and posterior means. 48
2.13 Standard deviations (y-axis) of each posterior distribution against the posterior mean (x-axis). Given the unequal replacement player partition given in Table 2.3, the difference in replacement player posterior variance is to be expected. 49
2.14 Posterior mean against prior mean. Although correlation is to be expected, the ranking of ability suggested by the model is different than that suggested by the raw rebounding rates. 50
2.15 In-sample prediction of rebounding rate during the 2020–21 NBA season. Lineups were grouped into 10 equally sized bins based on their predicted rebounding rate. 51
2.16 Posterior means of the individual rebound collecting parameters. 53
2.17 Posterior standard deviations (y-axis) plotted against their respective means (x-axis). Color represents the number of minutes of the player in question, and is a proxy for the number of multinomial observations used to estimate the parameter. Replacement players were omitted due their much lower standard deviations, which made the mean-variance relationship less apparent. 54
2.18 Predicted individual rebounding counts conditional on rebound collection (y-axis) plotted against observed individual rebounding counts (x-axis). . . 55
2.19 Predicted probabilities (y-axis) against observed probabilities (x-axis) for seen and unseen lineups during the 2021–22 NBA season. Note that the groups for the unseen lineups contain each about 1940 observations, and the seen lineup groups contain about 560 observations. 57
2.20 Predicted vs observed rebounding counts for each of the 30 teams during the 2021–22 NBA season (in all predictable instances). 58

2.21 Two-stage predicted vs observed rebounding counts for individual players during the 2021–22 season. Note that about ten players were omitted from the offensive plot because they had far more offensive rebounds that those plotted, so their inclusion in the plot ”squished” everyone else together. The fit for those players was comparable to the players retained for plotting. 59

List of Tables

2.1 Five players with greatest classification probabilities for each offensive position. 43

2.2 Five players with greatest classification probabilities for each defensive position. 43

2.3 Replacement players counts for every offensive and defensive position. . . 46

2.4 Top 15 players for offensive team rebounding ability (left) and defensive team rebounding ability (right) 60

2.5 Five most undervalued offensive rebounders (left) and defensive rebounders (right). 61

2.6 Five most overvalued offensive rebounders (left) and defensive rebounders (right). 61

2.7 Timberwolves most common lineup during the 2021–22 season. Note that parameters were estimated using data from the 2020–21 season. 62

B.1 False positive centers on offense. 74

B.2 False negative centers on offense. 74

B.3 False positive centers on defense. 74

B.4 False negative centers on defense. 74

C.1 Centers on offense. 76

C.2 Centers on defense. 77

D.1 Non-centers on offense (1). 79

D.2 Non-centers on offense (2). 80
D.3 Non-centers on offense (3). 81
D.4 Non-centers on offense (4). 82
D.5 Non-centers on defense (1). 83
D.6 Non-centers on defense (2). 84
D.7 Non-centers on defense (3). 85
D.8 Non-centers on defense (4). 86

E.1 Replacement centers on offense. 88
E.2 Replacement centers on defense. 88
E.3 Replacement non-centers on offense (1). 89
E.4 Replacement non-centers on offense (2). 90
E.5 Replacement non-centers on defense (1). 91
E.6 Replacement non-centers on defense (2). 92

Chapter 1

Introduction

All basketball fans agree that rebounding is an important part of the game: whether it be casual fans referring to players as ”walking double-doubles”, analysts commending players for their willingness to do the ”dirty work” and collect rebounds, or coaches criticizing their team’s performance on the boards, there is no debate that effective rebounding helps teams win basketball games. The earliest stars of the game, such as George Mikan and Bob Pettit, were praised for their extraordinary rebounding ability, as were all great centers who followed. Although conventional wisdom always valued rebounding, its importance was further emphasized by Oliver (2004), who formally showed that there is a positive correlation between rebounding efficiency and the probability of winning. This makes plenty of sense: despite rebounds not being directly worth any points, by collecting them effectively, teams can increase their own number of scoring chances, and decrease that of their opponents.

In assessing the viability of a potential lineup for a game, coaches will generally consider its rebounding ability. Therefore, they need to know how well each player rebounds. The importance of understanding these individual contributions to team rebounding is illustrated by the historical evolution of score-keeping and analytics: in the 1950–1951 season of the National Basketball Association (NBA), rebounds made their way onto official scoresheets (*1950–51 NBA Season Summary* n.d.), making them one of the earliest ever box score statistics to be recorded. At the time, "Rebounds per game" (RPG) was the gold standard for assessing individual rebounding ability.

In 1973, scorekeepers began to track offensive and defensive rebounds separately, which led to players being evaluated based on their offensive rebounds per game (ORB) and their defensive rebounds per game (DRB). Although being easy to compute and understand, both ORB and DRB are flawed in that they don't account for how many rebounding opportunities a player had to collect rebounds.

To improve on this, Oliver (2004) proposed the notions of offensive rebounding rate and defensive rebounding rate, which estimate the percentage of rebounds collected by a given player. Despite the obvious improvement, these metrics still don't account for two important factors. Firstly, they don't measure the impact of players who help their team rebound, but who don't collect rebounds themselves. Secondly, given that at most 100% of the rebounds can be collected, rebounding rates convey not only information about the individual's ability, but must also depend on the other players present on the court, which makes individual rebounding rate problematic for player assessment.

A general correction for this first issue was proposed by Rosenbaum (2004), namely the "Adjusted Plus/Minus Framework" (APM): the linear model's only predictor is a list of the players on the court (represented by a sparse matrix containing all lineup combinations used in the data set), and the response is the desired team metric (the original implementation used net rating as the response). Sill (2010) improved on this work by proposing the "Regularized Adjusted Plus/Minus Framework" (RAPM): by introducing ridge regression and cross validation, he obtained more stable parameter estimates, with greater predictive power, making them much easier to implement.

Although Engelmann (2016) suggests applying this framework to rebounding data, the formal literature on the topic is scarce. Despite the improvement of these metrics, they have several limitations, the most notable being that they assume that individual rebounding is additive. Heuristically, this seems unlikely, since only so many shots can be missed during a game.

In this thesis, a modified version of the RAPM framework, which more effectively ”isolates” players, is proposed to better understand both individual and team rebounding. This involves designing a framework which is conditional on the number of missed shots, and therefore allowing to break away from the additive rebounding assumption.

The bulk of the statistical methodology is contained in Chapter 2. Firstly, rebounding numbers are decomposed into two stages, namely team rebounding and individual rebounding, to better understand which players are actually helping their team, versus which players have misleading box score statistics. We model these two probabilities using logistic regression and multinomial regression respectively: a tractable two-stage Bayesian model is proposed, the choice of prior distributions is thoroughly explored, and a simulation study is carried out as a proof of concept.

Secondly, by using Gaussian mixture models, a rebounding-specific version of the ”Replacement Player” methodology is developed so that players appearing infrequently can be be grouped together and represented by fewer parameters, which speeds up estimation times, reduces over-fitting and increases estimation accuracy.

Thirdly, the model parameters are estimated using Markov chain Monte Carlo (MCMC) methods. Fourthly, predictions are made in out-of-sample games to validate the model. Finally, practical applications of the model are briefly explored.

In Chapter 3, which serves as the conclusion, future avenues of development for the proposed framework are discussed.

Chapter 2

Contribution

2.1 Introduction

Long has it been known that traditional NBA box scores can misrepresent the true value of players. One of the earliest alternatives to box-score based metrics in the NBA was developed by Rosenbaum (2004): by only using plus-minus data and a list of players on the court, he attempts to measure "pure" impact, independent of the subjective counting conventions. In this thesis, we adapt this methodology to better evaluate individual rebounding ability in the NBA. This is done by improving Rosenbaum's parameter reduction method, and introducing a prior layer to the parameters of a logistic model.

An already existing improvement to the Rosenbaum original implementation is the adoption of regularization techniques, which is in line with the surge of Bayesian approaches to quantitative sports analysis. The most relevant implementation is that of Sill (2010), who uses cross-validation to determine the prior parameters. Although some applications of Bayesian binomial regression do exist in the quantitative sports literature, such as when Deshpande and Wyner (2017) studied pitch framing for MLB catchers, or when Miskin, Fellingham and Florence (2010) attempted to quantify skill importance in women's volleyball, there appears to be very few applications specifically in basketball.

Furthermore, Rosenbaum's original implementation relies on the notion of a "Reference Player": to reduce the number of parameters to be estimated, he replaces all players having played below some threshold of minutes by a single player. This idea was first suggested by Woolner (2002), when developing "Value Over Replacement Player" (VORP) to evaluate Major League Baseball (MLB) players.

Although parameter reduction is obviously beneficial when data is limited, it seems that reducing a multitude of players to a single one may be too drastic. A middle of the road approach would be to reduce the parameters, but to group players who have similar styles of play.

The idea of player grouping has been popular for NBA players, where the positions tend to be quite artificial. Bornn et al. (2016) suggest using player tracking data and non-negative matrix factorization to group players based on where they like to shoot from (and defend shooters) on the court. Kalman and Bosch (2020) group players using Gaussian mixture models, and Alagappan (2013) has explored the question of positions using topological data analysis. In this thesis, we propose an alternative approach that more heavily relies on styles of play. To this end, these grouped players will be referred to as *replacement players*.

Section 2.2 goes over how the data set was collected, preprocessed, and some subtleties of data collection in the NBA. Section 2.3, briefly reviews the original APM and RAPM frameworks, discusses their limitations and possible ways to mitigate them. Furthermore, in this section, our rebounding model is formally defined, and a simulation study is carried out to assess effectiveness when the true parameter values are known.

In Section 2.4, we explain why modelling every single player is not feasible, outline the heuristics used for the parameter reduction, formally define the classification model, and give the estimation results. In Section 2.5, the model parameters are estimated via MCMC and using data from the 2020–21 NBA season. Predictions are made on the training data, and goodness-of-fit is measured using the Hosmer–Lemeshow test. We also discuss characteristics of the posterior distributions.

In Section 2.6, predictions are made for a subset of the 2021–22 season (out-of-sample) data at both the team and individual levels, and the practical implications for rebounding assessment are briefly overviewed. Finally, in Section 2.7, we discuss methods for incorporating uncertainty in predictions, model limitations, and avenues for future work.

2.2 The data set

All data used in the present project was made publicly available by the NBA. The data were collected using `nba-api` (Patel, 2023), an API Client for `www.nba.com`. Although the data is drawn from a multitude of API endpoints, it can be categorized into two distinct types of data: box score data and play-by-play data.

Data from the 2020–21 NBA season are used to estimate the models, and data from 2021–22 NBA season are used to evaluate model performance.

2.2.1 Box score data sets

Raw data collection and brief description

The NBA has two methods for organizing game-level data: data can either be indexed by date or by game ID. Furthermore, most data sets are exclusively stored using one of the indexing schemes, meaning that one cannot directly combine all measurements into a single observation. This indexing hurdle was overcome by creating a correspondence table between game IDs and dates, allowing for a much richer data set.

The concatenated raw data set contains one row per game per player (hereby referred to as ”performances”). In some very rare instances, game-indexed variables did not have their corresponding data-indexed variables: this occurred when the performance was incredibly short (a few seconds). Such instances were removed from the data set. Practically speaking, there were three types of variables in the data set:

1. Count variables: measures like steals, 3PT shots attempted...

2. Game summary variables: measures like average offensive speed, average defensive rebound distance...

3. Identifying information: measures that were not used for position classification, such as player names or possession counts

Given there are dozens of measured variables for each player, and since many convey equivalent information or even information that says nothing about player similarity, it would be both unnecessary and unwise to include all variables as features. Variable selection for classification was based on practical ideas: wherever possible, measures that are indicative of style of play should be prioritized over measures indicative of efficiency.

The idea is that positions should be used to group players ”who play the same way” rather than to group player ”who are just as good.” For example, rather than retaining 3PT shots made, 3PT shots attempted was retained. Given the practicalities of data collection in the NBA, this wasn’t always possible: for example, by looking at the number of steals, it is impossible to tell whether a player is a smothering defender who steals the ball applying tremendous pressure, or whether the player is a terrible defender who aggressively gambles for steals.

It would be far more informative to look at the ratio of steal to steal attempts, but unfortunately no such data is publicly available. The issue of limited defensive boxscore measures has been more thoroughly explored by Franks et al. (2015). These features were then divided into an offensive data set and a defensive data set.

A full list of the retained variables, as well as links to their descriptions, can be found in Appendix A.

Data pre-possessing

Given that we are interested in classifying players over the course of the season, all performances belonging to a given player were grouped together as follows:

a) count variables were summed together;

b) game summary variables (like average game speed, for example) were combined into a weighted average, where weights depended on the number of possessions in the corresponding game.

Once performances were grouped, all count variables were scaled to be per 100 possessions, in order to adjust for players with varying amounts of playing time.

2.2.2 Play-by-play data set

The play-by-play data set is very straightforward: for every line of the play-by-play log of a given game, we look at the substitution times of players, and determine who was on the court. In some instances, there were errors in the player substitution times, which would leave more or less than five players on the court at a given time. When this occurred, any lines of the play-by-play log that did not contain exactly five players for each team were removed, but if any other rows in the game had the correct number of players on the court, they were preserved. These errors were rare, occurring in fewer than 10 of the 1080 games.

One point that is especially important to note is the following: if the ball goes out of bounds after a missed shot, the team is allocated a *team rebound*. Despite these not being credited to an individual player, they are often generated by players wrestling for position, trying to get their team the out-of-bounds call (for those in the know about the conventions of NBA play-by-play data, also note that the quirky missed first three throw team rebounds were removed from the data set).

2.3 Rebounding models

In a practical context, a rebound is valuable not because it directly adds points to the total, but rather because collecting a missed shot controls how many scoring chances a team, or their opponents, will have on that possession. With this in mind, we believe that modelling rebounding counts is not as useful as modelling the probability of collecting a missed shot, which means that directly applying Rosenbaum's APM approach is not appropriate. Despite this, there are some desirable properties of the APM approach that we are keen on preserving. Firstly, the main appeal of the APM framework (and all derivative models) is that it is designed to allow for prediction in unseen lineups. This is obviously very useful in a practical context when negotiating contract extensions or determining who would be a desirable trade acquisition, for example. Secondly, the APM approach allows to model the interaction between individual players and team response variables, without assuming that the players are making a measureable, direct contribution (like scoring a basket or collecting a rebound, for example).

2.3.1 A brief overview of APM and RAPM

Rosenbaum's APM

The core philosophy of the original APM approach is straightforward: good players, regardless of their box-score stats, will help their team outscore their opponents when they are on the court. Although simple in theory, this becomes complicated in practice since not all players have the same quality of teammate on the court, not all players face the same level of competition, and not all players have the same quality of substitute.

To adjust for this, consider X, a matrix with one column for every player in the league and one intercept column, and where every row is a portion of the game where no substitutions took place. On a given row, the column is set to 1 if the corresponding player is on their home court, -1 if the player is on an away court, and 0 if the player is not on the court. In the original implementation, the response vector, Y, is the net rating of the home team (but the methodology could be applied for arbitrary continuous team-level variables). The weight matrix, W, is a diagonal matrix with entries equal to the number of possessions played during each substitution-less stint. Each player's contribution to the net rating (and the contribution of home court advantage) is estimated using weighted least squares, viz.

$$\hat{\beta} = (X^\top W^{-1} X)^{-1} X^\top Y.$$

There are two main drawbacks with the original APM implementation. Firstly, because of the multicollinearity in X, the standard errors of the estimated regression coefficients are so large that the model is practically unusable on out-of-sample data. Secondly, because of the design of the predictor matrix, there is no distinction between a player‘s separate contributions to offense and defense, only the net difference between the two.

Sill’s RAPM

Sill (2010) built on the APM framework by using ridge regression. The design, response, and weight matrices are identical, but a hyperparameter λ and the identity matrix I are introduced. The estimates for individual contributions to net rating are instead given by

$$\hat{\beta} = (X^\top W^{-1} X + \lambda I)^{-1} X^\top Y,$$

where λ is chosen to minimize the RMSE in out-of-sample games. By instilling a Gaussian prior on the parameters, the standard errors of the estimated parameters are more reasonably sized.

2.3.2 Limitations of APM based approaches in rebounding

Notwithstanding the desirable properties mentioned above, there are still a few drawbacks to APM inspired procedures in the current context, where we are studying rebounding probabilities. Firstly, the most problematic drawback is that the prediction versatility is achieved by assuming player contributions to the response variable are linearly additive. Although this sounds reasonable when modelling variables with infinite support, like point differentials or net ratings, it is not appropriate when modelling probabilities, as we are in this case: we would expect that adding a good rebounder to a good rebounding team will make for a less drastic improvement than adding a good rebounder to a mediocre rebounding team.

There is some anecdotal evidence supporting this claim. For example, Jonas Valanciunas and Steven Adams had a nearly identical OREB% during the 2020–21 NBA season. Despite this, after being traded for each other, Steven Adams saw a relative increase of 25% in his OREB%, and Memphis significantly improved their overall team OREB%. We argue this has to do with the fact that Adams's true contributions to offensive rebounding were obfuscated by playing with other strong individual rebounders on his former team, such as Zion Williamson.

Secondly, APM-like approaches generally achieve numeric stability in the estimates by implementing some form of prior layer or regularization parameter, and hence shrink all players towards some common mean. We suspect that we can improve the shrinkage in the context of team rebounding: we have a rough indicator of rebounding ability by looking at a player's individual rebounding rate, but the difficulty lies in figuring out how these abilities mesh together in a new lineup.

Thirdly, it seems reasonable to assume that offensive rebounding ability and defensive rebounding ability are separate attributes, and can be modelled independently. A direct implementation of APM or RAPM would imply that being a good offensive rebounder is the same as being a good defensive rebounder. This seems unlikely.

2.3.3 High-level ideas

We propose a procedure that preserves the main benefits of Sill's RAPM (i.e., accounting for teammate and opponent quality, predictive versatility, and reasonable standard errors), but use some heuristics about rebounding to overcome some of the drawbacks. The high-level idea is that each player has two latent measures of rebounding ability on both offense and defense (meaning there are a total of four variables of interest). The first latent measure is tied to team rebounding, and the second is tied to individual rebound collecting. Although these latent variables probably depend on each other, we argue that they are not exactly equal. Consider the following three scenarios:

1. Player A boxes out Player B, allowing for a teammate to collect the rebound.
2. Player A boxes out Player B, and collects the rebound themselves.
3. Player A leaves Player B to put themselves in better position to collect the rebound, but also increasing the chance that Player B collects the rebound.

In the proposed high-level rebounding model, Scenario 1 would be indicative exclusively of positive team rebounding ability and negative individual rebound collecting ability; Scenario 2 would be indicative of positive team rebounding ability and positive individual rebound collecting ability; and Scenario 3 would be indicative of negative team rebounding ability and positive individual rebound collecting.

With these notions in mind, we propose factorizing the conditional rebounding rate as follows:

$$\begin{aligned}\Pr(\text{Player A collects rebound} \mid \text{Missed shot}) &= \\ &\Pr(\text{Player A collects rebound} \mid \text{A's team collects rebound})\times \\ &\qquad \Pr(\text{A's team collects rebound} \mid \text{Missed shot}),\end{aligned}$$

and modelling each conditional probability separately. This is why the parameter reduction via clustering is essential: after applying the grouping methodology described in Section 2.4, instead of estimating 2160 parameters, we estimate only 1484 parameters; see Section 2.4 for details.

2.3.4 Modelling individual contributions to team rebounding

Logistic regression

Logistic regression is a natural modelling approach for the team rebounding problem: it allows us to assume that team rebounding ability is additive on the *log odds scale*.

This allows us to preserve the predictive versatility since we can just add up the individual contributions to team rebounding, but also naturally allows for the heuristic of diminishing returns in the team rebounding rate when adding more and more individual rebounding ability to a given lineup. For the sake of tractability, we do not control for things like days of rest or home court advantage, but instead assume that rebounding ability is constant from game to game.

We treat defensive rebounds as Bernoulli random variables, where a defensive rebound is considered a success, and an offensive rebound is considered a failure. Let L denote some arbitrary combination of five offensive players and five defensive players, let $\beta_1^D, \ldots, \beta_5^D$ denote the rebounding ability of the defensive players, and let $\beta_1^O, \ldots, \beta_5^O$ denote the rebounding ability of the offensive players. The most straightforward model for the probability of the defensive team collecting the rebound is given by

$$p^L = \frac{e^{\beta_1^D + \cdots + \beta_5^D - \beta_1^O - \cdots - \beta_5^O}}{1 + e^{\beta_1^D + \cdots + \beta_5^D - \beta_1^O - \cdots - \beta_5^O}} \,. \tag{2.1}$$

Note that the signs are picked so that the greater the value of the parameter, the greater the rebounding ability, regardless of whether we are talking about offensive or defensive rebounding. Although ideal because of its simplicity, a traditional implementation of this model suffers two main drawbacks: multicollinearity and unidentifiability.

Dealing with multicollinearity

When estimating the parameters of a generalized linear model (GLM) by maximizing the likelihood, the asymptotic distribution for the estimator (see Agresti, 2015) is given by

$$\hat{\boldsymbol{\beta}} \sim \mathcal{AN}\left[\boldsymbol{\beta}, (\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X})^{-1}\right],$$

where $\boldsymbol{W}$ is a diagonal matrix with

$$w_i = \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 / \mathrm{var}(y_i),$$

where μ_i is the ith mean response, y_i is the ith observation, and where

$$\eta_i = \ln\left(\frac{\mu_i}{1-\mu_i}\right).$$

As is the case with linear regression, if there is severe multicollinearity in the design matrix, the variance of the components of $\hat{\boldsymbol{\beta}}$ can explode. We could replicate Sill's approach by using ridge regression, but in the context of rebounding, there is reason to believe we can do better.

Although individual rebounding rates may be misleading when determining whether, e.g., Jonas Valanciunas is a better offensive rebounder than Steven Adams, it is probably telling when the difference in individual rates is large, i.e., Jonas Valanciunas (individual OREB% of 13.4%) is almost certainly helping his team's offensive rebounding more than Duncan Robinson (individual OREB% of 0.3%). Although the modelling approach allows to distinguish between individual and team rebounding ability, we argue that instances where there is a drastic difference in these two latent variables are probably rare, and we should only assume that they exist when the evidence is overwhelming.

With this in mind, instead of instilling an identical prior across all players, we propose that individual rebounding ability be reflected in the choice of prior. Although this dictates the stochastic ordering, since we are modelling rebounding ability on the log odds scale, it is not obvious what exactly the priors should be. To better understand measuring rebounding on the log-odds scale, we first discuss model identifiability.

Dealing with unidentifiability

Assume that there exists some underlying quantification of team rebounding ability, denoted by $\boldsymbol{\beta}_*^D$ and $\boldsymbol{\beta}_*^O$. In its proposed form, the model is not identifiable: we could find an equally valid maximum likelihood solution by adding some constant c to every component of $\boldsymbol{\beta}_*^D$ and to every component of $\boldsymbol{\beta}_*^O$. That being said, this alternative model parameterization isn't particularly problematic, since the main goal of the model is to rank the rebounders and predict their performance in unseen lineups. If we add the same constant to every component, these rankings are unchanged, and so is model prediction.

To illustrate a more serious issue, consider the simpler example where we have only two defensive players, denoted by β_1^D and β_2^D, and two offensive players, denoted by β_1^O and β_1^O. Also assume that in this simple example, there is only one player per team.

Assume that we have limited lineup mixture, and therefore, that we have only observed β_1^D against β_1^O, and β_2^D against β_2^O, viz.

$$\ln\left(\frac{p_1}{1-p_1}\right) = \beta_1^D - \beta_1^O, \quad \ln\left(\frac{p_2}{1-p_2}\right) = \beta_2^D - \beta_2^O.$$

In this simple case, we could add some constant c_1 to $\beta_1^D$and to β_1^O, and add some constant c_2 to β_2^D and to β_2^O. We could cleverly pick the constants such that either defensive player can be chosen to be the best rebounder, or either offensive player can be chosen to be the best rebounder. Any such model solution is much more problematic. However, under certain lineup conditions, all model solutions will preserve ordering of the parameter magnitudes.

Theorem 1. *If all possible lineup combinations have been used, all model parameterizations preserve the ranking of the parameters (in terms of their magnitude).*

Proof. Let $\boldsymbol{\beta}_*^D$ and $\boldsymbol{\beta}_*^O$ denote some arbitrary solutions to the model. Consider some alternative solutions $\boldsymbol{\beta}^D$ and $\boldsymbol{\beta}^O$. These can be re-written as

$$\boldsymbol{\beta}^D = \begin{bmatrix} \beta_{*1}^D \\ \vdots \\ \beta_{*n}^D \end{bmatrix} + \begin{bmatrix} \lambda_1^D \\ \vdots \\ \lambda_n^D \end{bmatrix}, \quad \boldsymbol{\beta}^O = \begin{bmatrix} \beta_{*1}^O \\ \vdots \\ \beta_{*n}^O \end{bmatrix} + \begin{bmatrix} \lambda_1^O \\ \vdots \\ \lambda_n^O \end{bmatrix}.$$

Since the probabilities must be equal across all lineup combinations, and using Equation (2.1), we must have that, for any $i_1, \ldots, i_5$ and $j_1, \ldots, j_5$,

$$\beta_{i_1}^D + \cdots + \beta_{i_5}^D - \beta_{j_1}^O - \cdots - \beta_{j_5}^O = \beta_{*i_1}^D + \cdots + \beta_{*i_5}^D - \beta_{*j_1}^O - \cdots - \beta_{*j_5}^O$$

as well as

$$\beta_{*i_1}^D + \lambda_{i_1}^D + \cdots + \beta_{*i_5}^D + \lambda_{i_5}^D - \beta_{*j_1}^O - \lambda_{j_1}^O - \cdots - \beta_{*j_5}^O - \lambda_{j_5}^O$$
$$= \beta_{*i_1}^D + \cdots + \beta_{*i_5}^D - \beta_{*j_1}^O - \cdots - \beta_{*j_5}^O,$$

and

$$\lambda^D_{i_1} + \cdots + \lambda^D_{i_5} = \lambda^O_{j_1} + \cdots + \lambda^O_{j_5}.$$

Assume we were to fix $i_1, \ldots, i_5$. Since the equality must hold for arbitrary $j_1, \ldots, j_5$, it must be that the sum of any subset of five λ^O's is the same. This, in turn, obviously means that all the λ^O's must be equal. Similarly, all the λ^D's must also be equal. Consequently, the relative ordering of the components of $\boldsymbol{\beta}^D$ and $\boldsymbol{\beta}^O$ is identical to the ordering of the components in $\boldsymbol{\beta}^D_*$ and $\boldsymbol{\beta}^O_*$. □

Of course in practice, not all possible lineups are observed, which means that while there may in fact be some true solutions $\boldsymbol{\beta}^D_*$ and $\boldsymbol{\beta}^O_*$, the estimated solution is not guaranteed to preserve parameter ranking.

Despite Theorem 1 not being particularly useful in practice, it does give some insight into the problem at hand. By looking at the simple two-player example, it is obvious that any amount of shift in the defensive parameters must induce an equal (in the aggregate) shift in the offensive parameters, and hence, given an incomplete system, we can find a valid parameterization that makes any player appear as the best rebounder.

It seems heuristically reasonable to believe that the more lineup mixture we have, the more likely it is for the λ^D_i's to be close to each other. This is why introducing replacement players (see Section 2.4) is so important, since it makes it incredibly unlikely for there to be subgroups (containing both offensive and defensive players) who have played exclusively amongst themselves, and allow for these "subsystems" of independent equations to arise. Furthermore, notice that if we were to restrict the parameter space of the model, we limit the values of λ^D_{Min}, λ^D_{Max}, λ^O_{Min}, λ^O_{Max}, and hence limit how extreme any potential re-ordering of the parameters can be. Although there is no mathematically rigorous way to restrict the parameter space, we consider the following thought experiment:

Suppose we could clone the best defensive rebounder, with ability β^D_{Max}, and play him against an average offensive rebounding lineup, who's aggregated offensive rebounding ability is given by c. It seems reasonable to assume that although we don't know the exact defensive rebounding rate for this lineup, it is certainly not greater than 90%. Suppose we could also clone the worst defensive rebounder, with ability β^D_{Min}, and play him against an average offensive rebounding lineup. Again, we cannot say for sure what the true defensive rebounding rate is in this case, but it seems heuristically reasonable that the percentage must be greater that 50%.

Indeed, these bounds seems reasonable if not abundantly cautious: of lineups having played at least 200 possessions, the lowest empirical defensive rebounding rate was 64%, and the greatest empirical defensive rebounding rate was 84%, according to Falk (2021). These restrictions can be written as follows:

$$\frac{e^{5\beta^D_{Max}-c}}{1+e^{5\beta^D_{Max}-c}} \leq 0.9, \quad \frac{e^{5\beta^D_{Min}-c}}{1+e^{5\beta^D_{Min}-c}} \geq 0.5.$$

We are interested in finding a support for the defensive parameters that would allow for predictions as extreme as those laid out above, but that is as "narrow" as possible, to limit the potential for re-ranking of the parameters. This can be formulated as follows in terms of an optimization problem:

$$\begin{aligned}
&\min \quad \beta^D_{\text{Max}} - \beta^D_{\text{Min}} \\
&\text{s.t.} \\
&\frac{e^{5\beta^D_{\text{Max}}}}{1+e^{5\beta^D_{\text{Min}}}} - \frac{e^{5\beta^D_{\text{Min}}}}{1+e^{5\beta^D_{\text{Min}}}} = 0.4, \\
&\beta^D_{Max} \geq \beta^D_{\text{Min}}.
\end{aligned}$$

Note that we can just add $c/5$ to each parameter, re-parameterize, and solve this slightly simpler but equivalent form of the problem, since $(\beta^D_{\text{Max}} - c/5) - (\beta^D_{\text{Min}} - c/5)$ has the same solution, and we are only interested in the *difference* between the two parameters.

Given that the constraint is simply the difference of two independent sigmoids, the gradient can easily be computed, and the optimal solution is readily found using the Lagrangian multiplier. We can, therefore, compute that $\beta^D_{\text{Max}} - \beta^D_{\text{Min}} \leq 0.340$. If we repeat the same thought experiment but for offensive rebounding parameters (and instead allowing for there to be a difference of 35% instead of 40%), we find $\beta^O_{\text{Max}} - \beta^O_{\text{Min}} \leq 0.292$.

Although the assumptions imply that there is some meaningful parameterization of the model in which the offensive parameters are ”close” to each other, and defensive parameters are “close” to each other, we have said nothing about the distance between the collection of offensive parameters and the collection of defensive parameters. However, note that the model can be re-parameterized as follows:

$$p^L = \frac{e^{\beta_1^D+\cdots+\beta_5^D-\beta_1^O-\cdots-\beta_5^O}}{1+e^{\beta_1^D+\cdots+\beta_5^D-\beta_1^O-\cdots-\beta_5^O}}$$

$$= \frac{\exp\{\beta_1^D+\alpha_D+\cdots+\beta_5^D+\alpha_D-(\beta_1^O+\alpha_O)-\cdots-(\beta_5^O+\alpha_O)+-5\alpha_D+5\alpha_O\}}{1+\exp\{\beta_1^D+\alpha_D+\cdots+\beta_5^D+\alpha_D-(\beta_1^O+\alpha_O)-\cdots-(\beta_5^O+\alpha_O)+-5\alpha_D+5\alpha_O\}}$$

$$= \frac{\exp\{\beta_1^{*D}+\cdots+\beta_5^{*D}-\beta_1^{*O}-\cdots-\beta_5^{*O}+\alpha\}}{1+\exp\{\beta_1^{*D}+\cdots+\beta_5^{*D}-\beta_1^{*O}-\cdots-\beta_5^{*O}+\alpha\}},$$

where the new parameterization simply shifts all the defensive parameters by some fixed amount, and all the offensive parameters by some other fixed amount.

As explained above, a solution to this model is equivalent in practice, since it preserves the ordering of the parameters and yields identical predictions. Note that as long as our heuristics about the maximal difference between parameters are correct, and as long as we leave α unrestricted, we can restrict the rebounding parameters to their respective ranges, and still be able to solve the model.

A Bayesian solution

To summarize, to make finding a solution tractable, we have made the following assumptions about the underlying model parameters:

1. Extreme differences in individual rebounding rates likely suggest a difference in team rebounding ability.

2. The defensive model parameters are probably close to each other and the offensive model parameters are probably close to each other, and can be modelled on a similar scale as individual rebounding rates.

3. If we restrict both the offensive and defensive parameters to some subspace, we need an unrestricted parameter to allow for each group to be adequately far apart.

Thus far, we have not formally defined what we mean by "restricted," and we have relied almost exclusively on intuition and heuristics, but a Bayesian model offers a natural, mathematically rigorous framework to implement the idea of parameter restrictions.

Recall that in a Bayesian model, prior beliefs are updated based on observed data. We can effectively "restrict" the parameter space by instilling some informative prior distribution on the parameters (and conversely, leave parameters unrestricted by instilling an uninformative prior).

We propose using the following hierarchical Bayesian framework

$$\beta_i^D \mid DREB\%_i, \sigma \sim \mathcal{N}(DREB\%_i, \sigma^2),$$
$$\beta_j^O \mid OREB\%_j, \sigma \sim \mathcal{N}(OREB\%_j, \sigma^2),$$
$$\alpha \sim \text{improper uniform prior over } (-\infty, \infty),$$
$$Y_{L,k} \mid \beta_{i_1}^D, \ldots, \beta_{i_5}^D, \beta_{j_1}^O, \ldots, \beta_{j_5}^O, \alpha \sim \text{Bernoulli}(p^L),$$

where

$$p^L = \frac{e^{\beta_{i_1}^D + \cdots + \beta_{i_5}^D - \beta_{j_1}^O - \cdots - \beta_{j_5}^O + \alpha}}{1 + e^{\beta_{i_1}^D + \cdots + \beta_{i_5}^D - \beta_{j_1}^O - \cdots - \beta_{j_5}^O + \alpha}},$$

and where $Y_{L,k}$ denotes whether the kth missed shot in the lineup L is a defensive rebound (a success) or an offensive rebound (a failure), and where the subscript i relates to the ith defensive player, and where the subscript j relates to the jth offensive player. Note that in the case of the replacement players, all the rebounds and rebounding opportunities of the grouped players were aggregated to create the rebounding rate.

To justify the choice of mean for each of the parameters, note that the greatest individual offensive rebounding rate was 15.5% (Clint Capela) and the smallest individual rebounding rate was 0.3% (Duncan Robinson). The greatest defensive rebounding rate was 33.6% (Andre Drummond) and the smallest individual defensive rebounding rate was 4.7% (Trey Burke). These offensive and defensive spreads are slightly smaller than the ones implied by the above thought experiments.

Outside of the rough parameter restrictions suggested by the thought experiment, there is no obvious choice for the variance of the prior distributions. Although often times Bayesian applications rely on high variance, minimally informative priors, given the very severe data limitations of the current context, using a broad prior would be akin to just maximizing the likelihood. Ideally, we would like to pick the maximal prior variance that allows for meaningful and useful parameter estimates, to maximize the weight of the data on estimates.

With this in mind, we set the prior variance equal to 4 times the variance of the observed individual rebounding rates, which yields a prior standard deviation of 0.1. There are a few reasons for this. Firstly, since we hypothesized that the effect on the log odds can be modelled using a similar scale to that of the rebounding rates, it follows naturally that scaling up the variance of the rebounding rates is a sufficiently cautious approach.

Secondly, the upper bound on a 95% confidence interval for the best defensive rebounder prior is equal to 0.532, and the lower bound on a 95% confidence interval for the worst defensive rebounder prior is equal to -0.149. If the thought experiment is broadly reasonable, then there is ample ”room” to capture the variability of different players. Similarly, the likely range of the offensive parameters is between -0.193 and 0.351.

Thirdly, from a practical standpoint, the chosen variance seems to be sufficiently cautious when passing judgement about relative player quality. For example, the priors suggest that a priori, there is a 93% chance that Clint Capela is a better defensive rebounder than Trey Burke, which seems a bit too optimistic about Burke’s ability, but only a 59% chance that Andre Drummond is better than Jonas Valanciunas (who had a defensive rebounding rate of 28.9%). Figure 2.1 contains 50% confidence regions for the priors of some notable players, to provide practical justification for the prior construction methodology. Although this doesn’t truly restrict the possible parameter values, it does achieve a similar effect, by making extreme values unlikely.

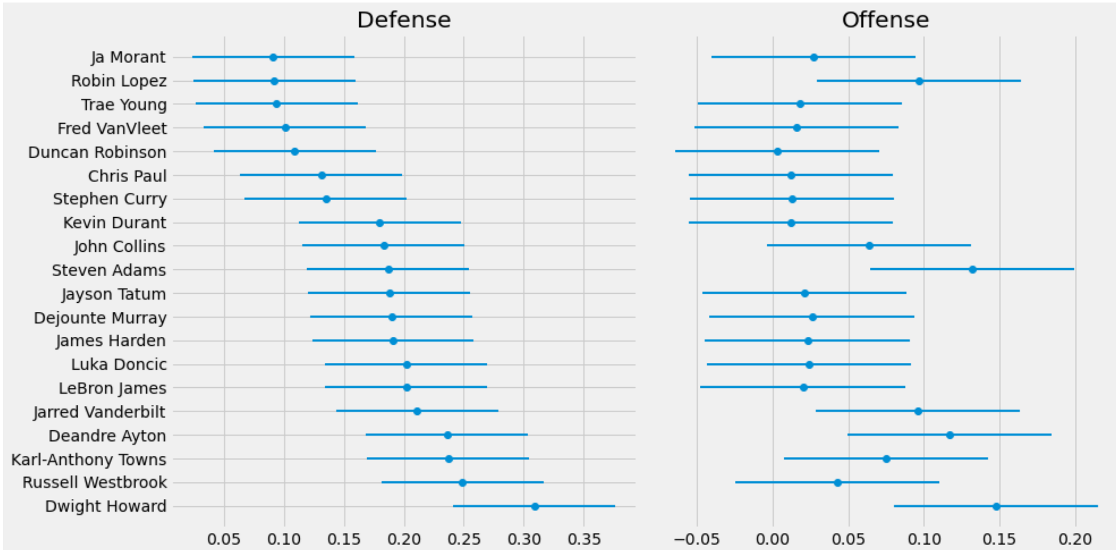

Figure 2.1: Example priors means for the team rebounding ability of some players in the 2019–20 season, as well as 50% central interval.

A simulation study

Given that the proposed model relies heavily on heuristics about how to reduce the parameter space, and that in the case of non-identifiability, the posterior distributions are heavily influenced by the choice of priors, we deem it necessary to see how well the proposed model can recover the parameters under realistic (albeit simplified) conditions. Before running a simulation study, we define the following data structures:

1. **Players**: Each player has a known defensive rebounding attribute, β^D, which is generated from Normal distribution, and a known offensive rebounding attribute, β^O, which is generated from a separate Normal distribution. The choice of parameters for these Normal distributions is discussed below.

2. **Teams**: Each team has eight players. We create a single lineup by sampling five players without replacement. We do this 30 times to create a list of 30 lineups that will be used when playing games. We then assign each lineup a weight by sampling from a symmetric Dirichlet over the 30-dimensional simplex. We denote the weight of the ith lineup of team A by w_i^A.

3. **Games**: For each game, we draw eight lineups from each of the two teams, with the probability of being drawn equal to each lineup weight. We scale the weight of the drawn lineups to determine the proportion of playing time of each lineup, i.e., L_i^A will play $w_i^A/(w_{i_1}^A + \cdots + w_{i_8}^A)$ of the game. Each game consists of 100 missed shots (50 per team) that need to be allocated to a player.

4. **Allocating rebounds**: Within a given lineup, we first allocate the team a rebound using the following probability:

$$p = \frac{e^{\beta_{i_1}^D + \cdots + \beta_{i_5}^D - \beta_{j_1}^O - \cdots - \beta_{j_5}^O}}{1 + e^{\beta_{i_1}^D + \cdots + \beta_{i_5}^D - \beta_{j_1}^O - \cdots - \beta_{j_5}^O}}.$$

Then, based on which team collected the rebound, we randomly assign that rebound to an individual player. The probability that Player i is assigned an individual defensive rebound is given by

$$p_i^D = \frac{e^{4\beta_i^D}}{e^{4\beta_1^D} + \cdots + e^{4\beta_5^D}},$$

where the numerator is the sum across all other players appearing in the same lineup as Player i. Similarly, team offensive rebounds are conditionally allocated to an individual using the following probabilities:

$$p_i^O = \frac{e^{7\beta_i^O}}{e^{7\beta_1^O} + \cdots + e^{7\beta_5^O}}.$$

Note: the inclusion of a scaling factor of 4 and 7 is so that the simulated individual probabilities more closely resemble observed probabilities. Directly using the parameters does not allow for sufficient variability in the individual rebounding rates to match observed rates.

With these structures defined, the actual simulation algorithm is quite straightforward: every team plays each opponent three times (meaning each team plays a total of 87 games instead of the usual 82). During each game, each team will miss 50 shots, and their opponents will miss 50 shots (teams on average miss 47 field goals per game, according to Basketball-Reference), each of which is then allocated to an individual player. Obviously there is nowhere near as much complexity in this simple simulation as there is in actual NBA games, and we therefore briefly discuss what will be referred to as *facilitating simplifications* and *impeding simplifications*.

Facilitating simplifications are simplifications which make parameter estimation easier than real-world conditions. The most notable of the simplifications is that not only are team rebounding ability and individual rebound collecting correlated, they are perfectly dependent. This was done mainly because it was unclear how to link the two latent variables without snooping through the data. Another significant simplification was that lineups are generated randomly, meaning that there is probably less multi-collinearity in the simulation than there is in the true league. Lastly, there are the practical simplifications, like having fewer players per team, players not getting injured, or there not being any team rebounds. These simplifications are probably not that significant.

Impeding simplifications are simplifications which make parameter estimation more difficult than real-world conditions. One key simplification is that there are no replacement-like players and no trades, which greatly decreases the amount of lineup mixture, and makes it more difficult to construct priors that are consistent with each other (since, as shown in Theorem 1, lineup mixture is key to preserving the ordering of rebounding ability). Also, since the lineups and players are truly random, there are probably instances of very unrealistic lineup combinations, which means the variability between lineups is much greater than in the true league.

The hope is that both the impeding and facilitating simplifications roughly cancel each other out, and make for broadly reasonable test conditions.

Parameters for the two Normal distributions used to generate the players were chosen based on the observed rebounding rates of 5-man lineups having appeared in at least 100 offensive possessions and 100 defensive possessions during the 2020–21 NBA season (note that instances with the same 10 players are far too short for the probabilities to be meaningful).

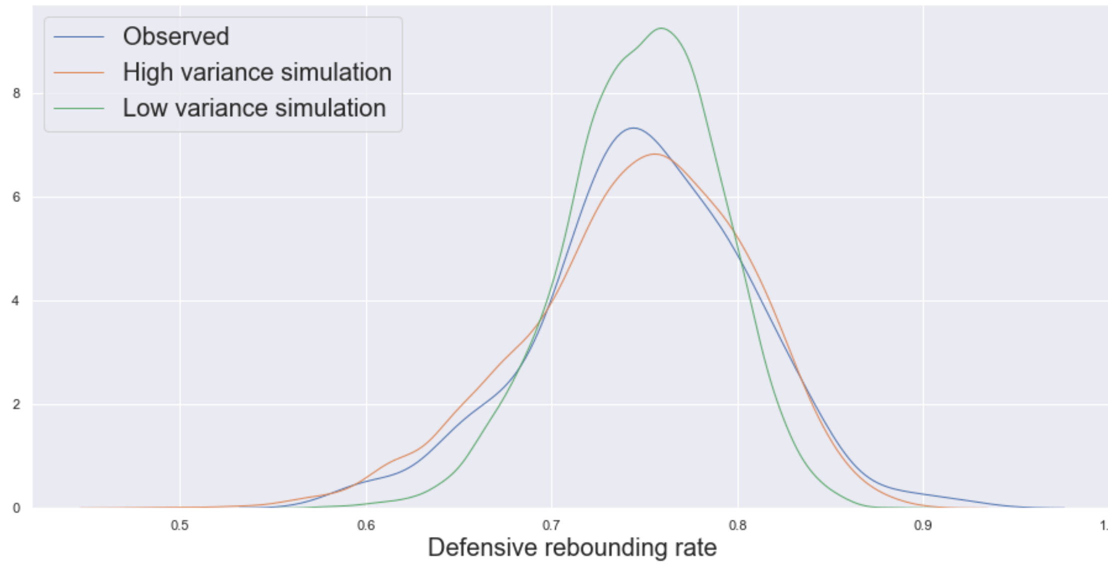

Figure 2.2: Comparison between team rebounding rates of observed lineups and generated probabilities.

To judge whether the Normal parameters were appropriate, 5000 ten-man combinations were created by sampling five players from the defensive distribution and five players from the offensive distribution. The probabilities for these lineups were computed using the logistic function, and the simulated distribution was compared to the observed data. Since the likelihood function is not identifiable, as long as the center of the distributions are adequately spaced, their exact values are insignificant.

The mean of the defensive distribution was chosen to be 0.22, and the mean of the offensive distribution was chosen to be 0. Furthermore, as shown in Figure 2.2, we can achieve nearly identical lineup rebounding rates by setting the standard deviation to 0.1 for each of the two Normal distributions. However, the observed rates are obviously empirical, can contain as few as 50 trials, and are marginalized across all opponents, which means that the variance in the observed lineup rebounding is probably greater than the variance of the true underlying probabilities that will be used in the simulation. With this in mind, we also opted for a second simulation, but with a standard deviation of 0.07.

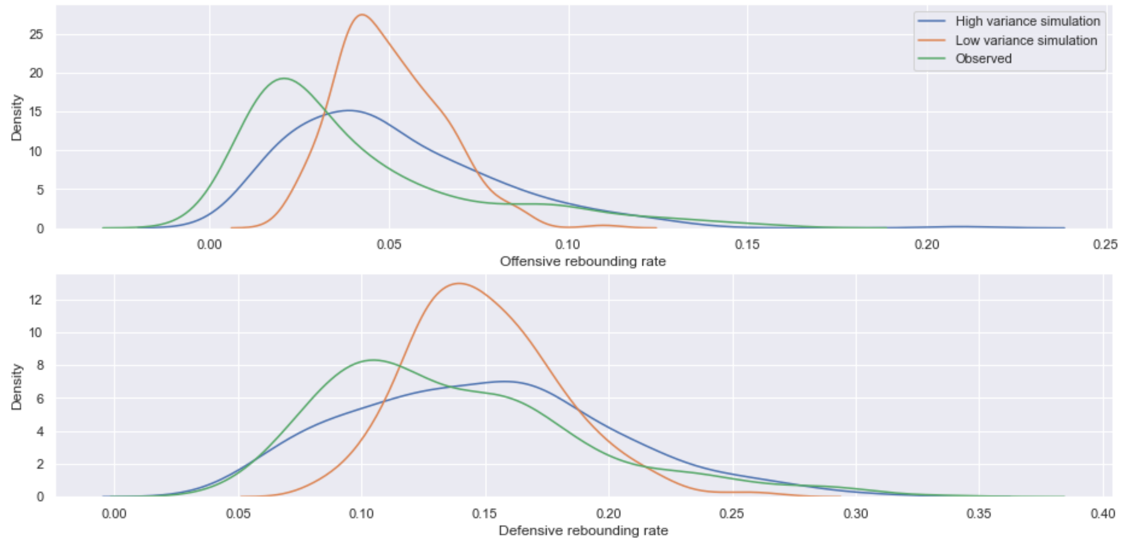

Figure 2.3: Comparison between the distribution of observed rebounding rates of players and the simulated individual rebounding rates.

At this stage, a scale factor was introduced such that the simulated individual rebounding rates roughly resemble the observed individual rebounding rates. The distributions of the observed and simulated data are given in Figure 2.3.

The simulation was carried out for each choice of variance, as described above. The parameters were estimated using `Stan` (Carpenter et al., 2017), which is an implementation of Hamiltonian Monte Carlo. The Markov chains consisted of 1000 warm-up iterations and 1000 sampling iterations. To evaluate the convergence of the process, four separate chains were used, the values of $\hat{R}$, as described by Vehtari et al. (2021), were calculated for each marginal distribution. In all instances, these were close to 1. This was especially important for the estimate $\hat{\alpha}$ of the intercept, since the prior was improper.

To evaluate the accuracy of the model, however, we cannot simply compare the estimated parameters to the known ones: as mentioned earlier, we can shift any parameterization by a constant, and end up with an equivalent model. This is why the inclusion of an intercept term when estimating the parameters was crucial. To make comparisons possible, we assume that the shift across all parameters is constant, and the total shift for each lineup is aggregated into α. We can therefore re-write the linear predictor as

$$\begin{aligned}
\beta_1^D + \beta_2^D + \cdots + \beta_5^D - \beta_1^O - \beta_2^O - \cdots - \beta_5^O + \alpha \\
&= \left(\beta_1^D + \frac{\alpha}{10}\right) + \left(\beta_2^D + \frac{\alpha}{10}\right) + \cdots + \left(\beta_5^D + \frac{\alpha}{10}\right) \\
&\quad - \left(\beta_1^O - \frac{\alpha}{10}\right) - \left(\beta_2^O - \frac{\alpha}{10}\right) - \cdots - \left(\beta_5^O - \frac{\alpha}{10}\right).
\end{aligned}$$

Therefore, we compare the *shifted* parameter estimates with the parameter values used to create the simulated data. Figure 2.4 contains scatter plots of the shifted estimated parameters against their true values.

Although the method seems to work reasonably well in general, there are few outliers. Outliers on the extremes of the cloud seem reasonable enough: because of the shape of the logistic function, for extreme players, a large difference in the parameter can lead to a negligible difference in predicted probability.

The more concerning/interesting outlier is the one player found far above the cloud in the offensive high-variance graph. We suspect that this could be an artefact of lineup construction: since the lineups were completely random, if a weak rebounder just happened to have teammates and/or opposition that were even weaker, it would be nearly impossible to detect that they were weak. In practice, lineup strategy probably makes for more homogeneous lineups that the randomly generated ones. Removing this point alone increases the value of R^2 from 0.42 to 0.48.

Given the satisfactory behaviour of our suggested estimation strategy, we continued by applying the proposed methodology with our NBA data set.

2.3.5 Modelling individual rebound collecting ability

Recall that in traditional multinomial regression, we choose some baseline category, say k, and use a linear relationship to model the log-odds ratio between each non-baseline category and the baseline. That is, if there are k possible categories for the response, we

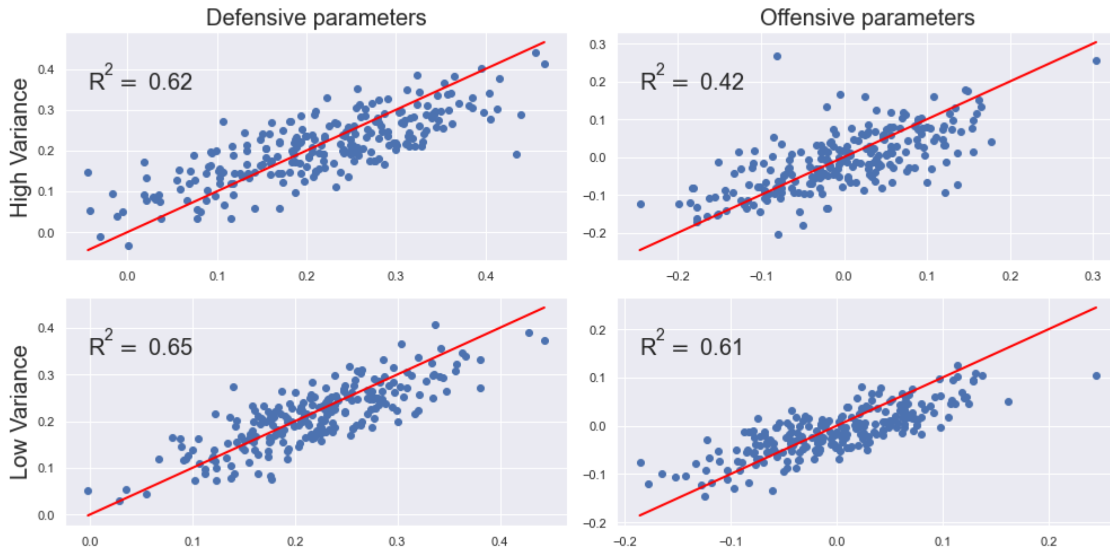

Figure 2.4: Scatter plot of shifted estimated parameters (y-axis) against their true values (x-axis).

have

$$\ln(\pi_{ij}/\pi_{ik}) = \sum_{h=1}^{p} \beta_{jh} x_{ih},$$

where π_{ij} denotes the probability of observation i belonging to category j, and where x_{ih} denotes the hth predictor of observation i, and where p denotes the number of predictors.

Although multinomial regression seems like a natural way to model individual rebounding probabilities, the implementation is not trivial: for every missed shot, the rebound must be either allocated to the team (when the ball goes out of bounds after the missed shot) or to one of the five players *on the court*. This means that although we can treat each rebounding event as a multinomial random variable, the response categories differ for each trial, meaning that traditional multinomial regression, where each outcome can belong to any of the same possible categories, is not suitable.

To remedy this, we assume that each player has some latent rebound collecting ability both on offense and defense, respectively denoted by γ_i^O and γ_i^D for Player i. Similarly to team rebounding, we assume that these variables are constant across lineups, and that rebounding is linearly additive on the log-odds scale.

Furthermore, although players differ between lineups, every lineup can potentially produce a team rebound, meaning that there is at least one common category across all lineups. Therefore, if we model the log-odds ratio between an individual collecting the rebound and there being a team rebound, the resulting framework allows for mixing and matching of never-before seen lineups.

In other words, assume some lineup L, consisting of players 1 to 5, collects a defensive rebound (the same procedure hold for offensive rebounds). We assume that the conditional probability of Player i collecting that rebound is given by

$$p_i^L = \frac{e^{\gamma_i^D}}{1 + e^{\gamma_1^D} + \cdots + e^{\gamma_5^D}},$$

and the probability of there being a team rebound is given by

$$p_T^L = \frac{1}{1 + e^{\gamma_1^D} + \cdots + e^{\gamma_5^D}} = 1 - p_1^L - p_2^L - p_3^L - p_4^L - p_5^L.$$

Because the response categories differ between responses, the likelihood is different from that of traditional multinomial regression.

Let $\boldsymbol{x_1}, \ldots, \boldsymbol{x_N}$ denote the 5-dimensional vectors of multinomial responses where the number of trials is known. Let n_i denote the number of trials for the ith multinomial variable, let $x_i^{(j)}$ denote the number of rebounds collected by the jth player in the ith lineup, and let N denote the total number of observed lineups.

The likelihood of the above model, with the multinomial coefficient omitted for the sake of readability, is given by

$$\begin{aligned} L(\boldsymbol{\gamma^D}; \boldsymbol{x}) &\propto \prod_{L=1}^{N} (p_1^L)^{x_L^{(1)}} \times (p_2^L)^{x_L^{(2)}} \times (p_3^L)^{x_L^{(3)}} \times (p_4^L)^{x_L^{(4)}} \times (p_5^L)^{x_L^{(5)}} \times (p_T^L)^{n_L - x_L^{(1)} - x_L^{(2)} - x_L^{(3)} - x_L^{(4)} - x_L^{(5)}} \\ &= \prod_{L=1}^{N} \frac{(e^{\gamma_1^D})^{x_L^{(1)}} \times (e^{\gamma_2^D})^{x_L^{(2)}} \times (e^{\gamma_3^D})^{x_L^{(3)}} \times (e^{\gamma_4^D})^{x_L^{(4)}} \times (e^{\gamma_5^D})^{x_L^{(5)}}}{1 + e^{\gamma_1^D} + \cdots + e^{\gamma_5^D}}. \end{aligned}$$

Note that a single player can appear across multiple lineups with a different index, meaning that technically, the γ's should also depend on the lineup, and that γ's with different indices can actually represent the same underlying parameter.

It is straightforward to show that the log-likelihood is proportional to

$$\sum_{L=1}^{N} x_L^{(1)}\gamma_1^D + x_L^{(2)}\gamma_2^D + x_L^{(3)}\gamma_3^D + x_L^{(4)}\gamma_4^D + x_L^{(5)}\gamma_5^D \\ - n_L \ln\left(1 + e^{\gamma_1^D} + e^{\gamma_2^D} + e^{\gamma_3^D} + e^{\gamma_4^D} + e^{\gamma_5^D}\right). \quad (2.2)$$

Directly maximizing the log-likelihood in this instance is less problematic than in the team rebounding model: instead of players confounding each other's contributions to team rebounding, we actually have more "complete" information, since we know exactly which player collected the rebound. This means that although player appearances are correlated, they don't lead to the same explosion in the variance of the parameter estimates, although lineup mixture still obviously eases parameter estimation.

However, there a few reasons we avoid this direct maximization. Firstly, since we are not using a traditional GLM, we would have to manually implement a maximization procedure. Secondly, not all variables have the same predictors (the features for each observation are the five players in the lineup), which greatly complicates estimating the parameter variances. Although we could in theory estimate the asymptotic variance using the Fisher information, in practice, we have no feature matrix and therefore cannot use matrix calculus to compute the derivatives, which means the implementation would not be very efficient.

We instead estimate the parameters again using MCMC, but unlike the team rebounding procedure, by instilling (improper) uniform priors over the real line. We note that the posterior distribution is proportional to the product of the prior and the likelihood, i.e.,

$$\pi(\boldsymbol{\gamma} \mid \boldsymbol{x}) \propto \pi(\boldsymbol{\gamma}) \times L(\boldsymbol{\gamma} \mid \boldsymbol{x})$$

for all $\boldsymbol{\gamma}$, $\boldsymbol{x}$. Therefore, by using the posterior mean to estimate the parameters, we are effectively using a weighted average over the parameter space, where the weight is proportional to the likelihood at the given point. Furthermore, when the likelihood function is symmetric and unimodal, the posterior mean is in fact exactly equal to the MLE.

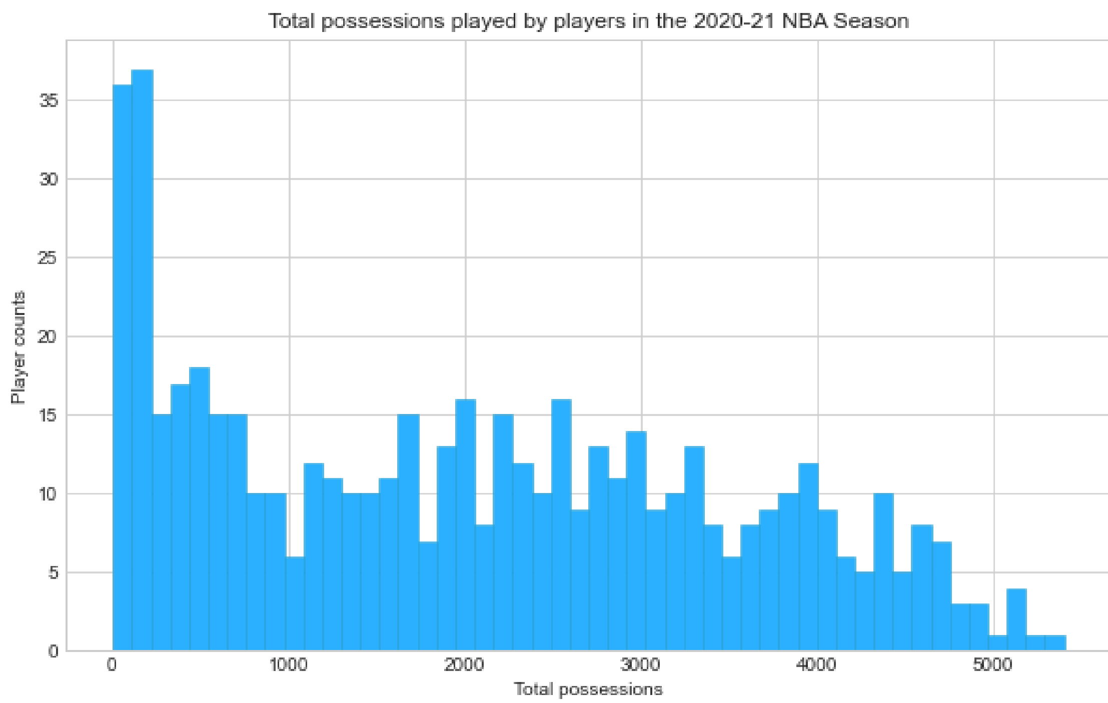

Figure 2.5: Distribution of possessions played by players in the 2020–21 NBA Season. A total of 540 players were in the data set.

2.4 Player clustering

Over the course of the 2020–21 NBA season, 540 different players appeared in at least one game. However, many of the players were merely an injury replacement, were played exclusively to rest starters once a game has been decided, or were simply given a trial run before being cut from the team. A histogram of the distribution of possessions played during the season is given in Figure 2.5. Because of these players' very limited opportunities, it would be impossible to obtain any statistically significant results for them in the limited data set of games. Furthermore, a sizable portion of the players will never make another NBA appearance, meaning that in a practical setting, there is no point in estimating their rebounding ability. These players will henceforth be referred to as *unusable players*, from the perspective of our analysis.

To reduce the number of parameters to estimate and reduce overfitting, all these unusable players could be grouped together, and treated as a single "reference" player, as proposed by Rosenbaum (2004). One issue with this approach is that we would be implicitly assuming that all these players should be considered as equivalent rebounders. This is almost certainly not the case: a 7-foot rim-running big man is almost certainly a different rebounder than a 5'10 scoring guard, despite there not being sufficient data to significantly test for this difference.

Therefore, one similar but less rigid approach would be to group players based on their listed position. However, this would represent further challenges: depending on the data set, not only will player positions not be consistent, but also, the set of available positions may be different.[1] This also completely glosses over positions that many pundits would argue exist, e.g., the *pass-first point guard*, *the point forward*, or *the stretch big*.

Because of the subjectivity and discrepancies in these labels, a more objective approach is explored to group players who are more likely to have similar rebounding ability.

The big picture is as follows: firstly, we retain all players with a lot of playing time, since we have a good idea of their tendencies. We then use these players to learn the underlying positions of the NBA. Then, having learned these positions, we assign each unusable player one of these positions; note that the usable players are not assigned a position, they are just used to *learn* them.

Finally, we assume that all players of the same position are identical, and therefore treat them as a single player, hence greatly reducing the number of parameters to estimate. Figure 2.6 shows how the clustering fits into whole modelling procedure.

2.4.1 Heuristics

Henceforth we distinguish two terms:

[1] For example, the NBA official box scores contains seven unique labels (G, G-F, F-G, F, F-C, C-F, C), but ESPN uses five positions (PG, SG, SF, PF, C) to categorize players.

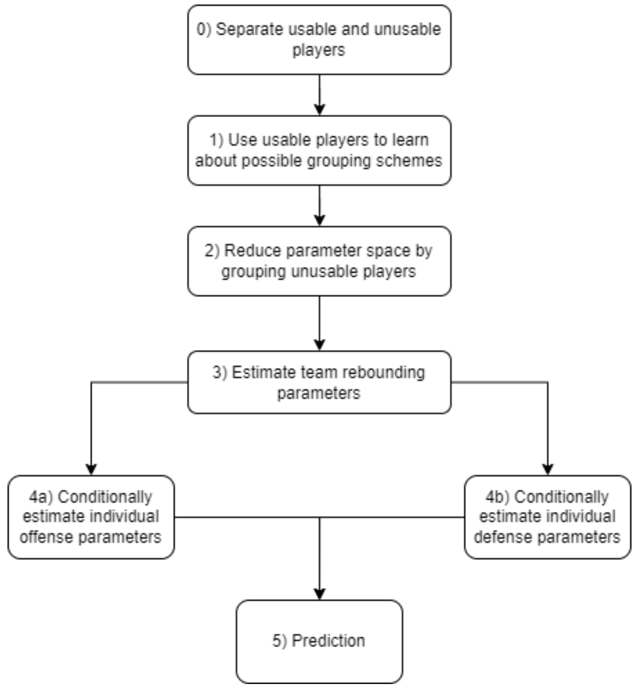

Figure 2.6: Diagram of the high-level modelling procedure.

Position label: the position that a player is assigned in the NBA data set.

Underlying position: the latent position of a player that we are interested in learning.

Moreover, we propose the following postulates:

1. The first grouping heuristic is that there exist separate offensive and defensive positions.

2. The second grouping heuristic is that the position labels do convey meaningful position information in the aggregate: although there may be some ”mislabelled” players (especially amongst players who have played very little), players who are of the same underlying position are more likely to end up with the same position label. Moreover, we assume that the ordering in compound position labels also conveys information about the underlying position, and hence permutations are treated as a unique label. Hence we rely on seven position labels: G, G-F, F-G, F, F-C, C-F, C.

3. The third grouping heuristic is that the set of possible underlying positions is the same for all players, regardless of how much playing time they get: this allows us to learn the underlying positions from players who have a large sample of games played, and then assign these labels to unusable players.

   This third heuristic is important because the small sample sizes of unusable players mean that their features are often extreme, and make them unsuitable for clustering. Because of this, only players having played in at least 1000 possessions are retained for ”positional learning.” This cutoff divided the data set into 366 usable players and 174 unusable players.

4. The final modelling heuristic is that there exists a traditional center underlying position on both offense and defense, and that in general, this underlying position is very easy to identify compared to other positions, meaning that the C position label is more reliable.

   This heuristic is based on the fact that the traits of the traditional center are well captured by the data set, since measurements like average shot distance, shots blocked, or lack of three pointers shot are recorded.

2.4.2 Reformatting the data set

The raw player tendency data sets (we handle offensive variables and defensive variables separately) cannot be used directly to ”learn” the underlying positions: as shown by Figure 2.7, there is significant correlation between the columns of the feature matrices. This can be problematic when dealing with GMMs, since the re-assignment of points to a cluster can be very unstable and vary wildly between iterations. The most obvious solution to this problem is to reduce the dimension of the feature matrices by performing PCA, and retaining only some subset of the principal components. This obviously begs the question of how many principal components should be retained for classification, which will be discussed in Section 2.4.4.

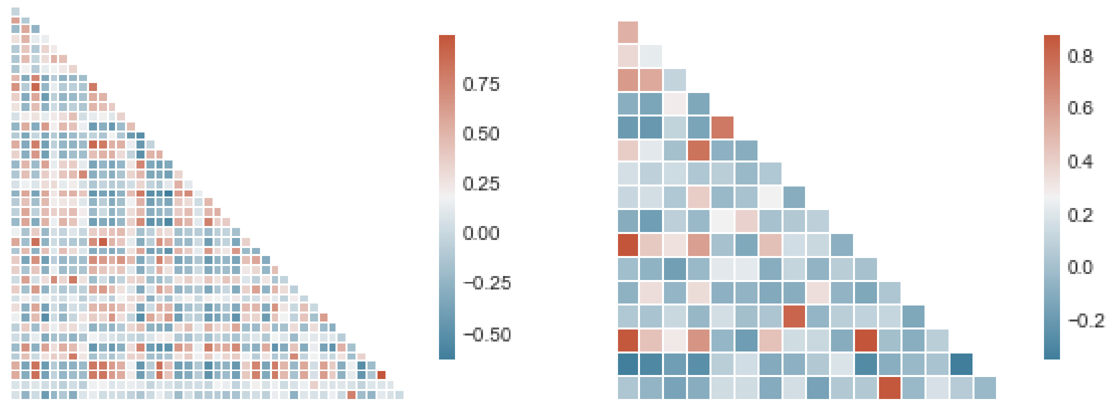

Figure 2.7: Left: Correlation matrix of the offensive feature data set. Right: Correlation matrix of the defensive feature data set.

At a first glance, performing PCA on the normalized data set appears to work well: Figure 2.8 suggests that we can greatly reduce the number the dimensions of the data sets. This plot is however slightly misleading in that some position characteristics are over-represented in the data set because of the ease with which they can be collected. For example, there are multiple features relating to posting up, but very few related to shooting three pointers. This issue becomes obvious when looking at the proportion of variance retained for each position label.

Since the data have been standardized, the proportion of variance retained by the ith principal component for a given player is given by

$$\frac{PC_i^2}{PC_1^2 + \cdots + PC_1^2},$$

where PC_j^2 denotes the eigenvalue associated with the jth principal component. These values can then be averaged across players sharing a position label to get an idea of how well the variability of the said position label is preserved after reducing the dimension. As shown by Figure 2.9, and assuming our heuristics are correct, if we were to reduce the dimension of the data set, we would require a lot of principal components to keep all the non-centers from being ”squished” together.

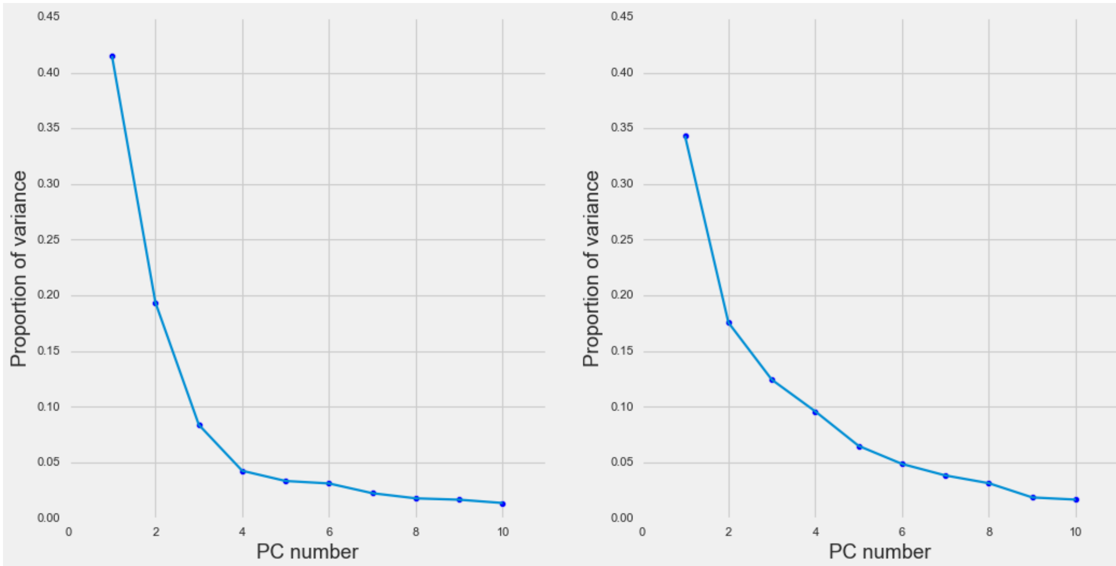

Figure 2.8: Left: Scree plot for first 10 principal components of offensive data set. Right: Scree plot for first 10 principal components of defensive data set.

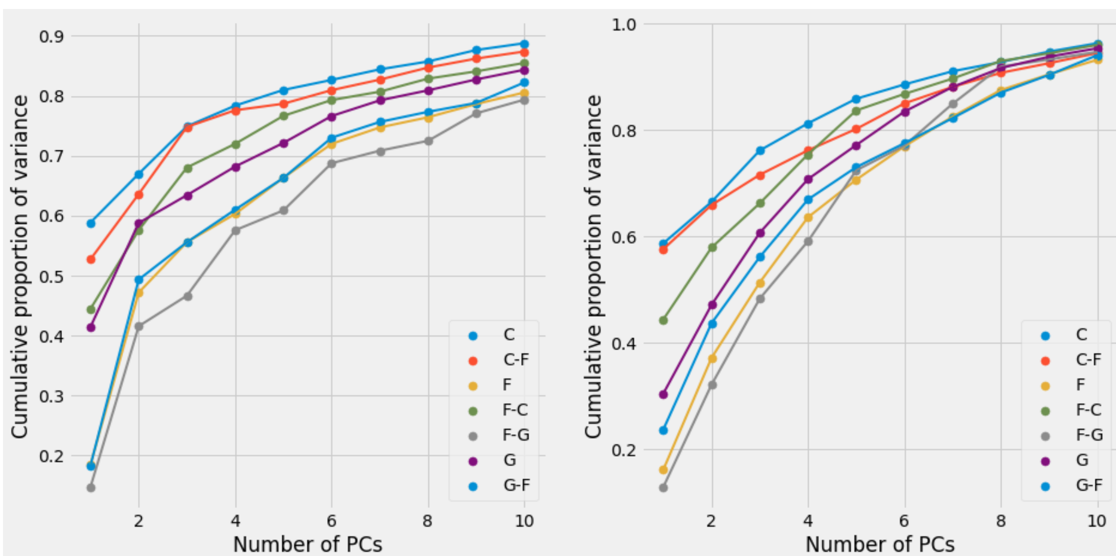

Figure 2.9: Left: Scree plot for average variance proportion retained for a given position in offensive data set. Right: Scree plot for average variance proportion retained for a given position in defensive data set. Note that the labels are those used by the NBA.

2.4.3 Clustering using Gaussian mixture models

The underlying assumption of a Gaussian mixture models (GMM) is that for the n data points, there are K underlying subpopulations, each of which is characterized by some (potentially multivariate) Gaussian distribution. That means that the density for a given observation is given by

$$f(\boldsymbol{x} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{j=1}^{K} \pi_j \, g(\boldsymbol{x} \mid \mu_j, \Sigma_j),$$

where π_j denote the mixture weights, and where μ_j and Σ_j denote the parameters of the jth subpopulation.

In practice, both subpopulation membership and the parameters of each distribution are unknown and must be estimated; in this case, the number of subpopulations is also unknown. For a fixed number of subpopulations, the parameters can be estimated using the expectation maximization algorithm, given by Dempster, Laird and Rubin (1977).

Let $\boldsymbol{x^{(i)}}$ denote the feature vector of the ith observation. Let also $z^{(i)}$ denote the latent cluster membership categorical variable of the ith observation. Furthermore, let $\mu_1, \ldots, \mu_k, \Sigma_1, \ldots, \Sigma_k, \pi_1, \ldots, \pi_k$ be the *current* values of the subpopulation parameters.

First, we have the E-step. For each $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, K\}$, set

$$w_j^{(i)} = P(z^{(i)} = j \mid \boldsymbol{x^{(i)}}; \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\Pr\big(\boldsymbol{x^{(i)}} \mid z^{(i)}\big); \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Pr\big(z^{(i)} = j\big)}{\Pr\big(\boldsymbol{x^{(i)}}; \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\big)}.$$

Then we have the M-step. For each $j \in \{1, \ldots, K\}$, let

$$\pi_{j+1} = \frac{1}{n} \sum_{i=1}^{n} w_j^{(i)}, \quad \mu_{j+1} = \sum_{i=1}^{n} w_j^{(i)} \boldsymbol{x^{(i)}} \Big/ \sum_{i=1}^{n} w_j^{(i)}$$

and

$$\Sigma_{j+1} = \sum_{i=1}^{n} w_j^{(i)} (\boldsymbol{x^{(i)}} - \mu_j)(\boldsymbol{x^{(i)}} - \mu_j)^{\top} \Big/ \sum_{i=1}^{n} w_j^{(i)}.$$

These steps are repeated until the change in the estimated parameters is negligible. Using the final values of $w_j^{(i)}$, we can determine, upon convergence, which subpopulation likely generated each observation by assigning $\boldsymbol{x^{(i)}}$ to cluster k if

$$\arg\max_j w_j^{(i)} = w_k^{(i)}.$$

2.4.4 Clustering approach

Because of this ”squishing” phenomenon, an iterative approach was implemented:

1. Perform PCA on the whole data set.
2. Learn to separate centers and non-centers (based on their underlying position) using some subset of the principal components.
3. Remove underlying centers from the data set.
4. Perform PCA on the original features of the non-centers.
5. Learn the different positions in the reduced data set using some subset of the principal components.

This approach requires the choice of two hyper-parameters: the number of mixture components in the GMM and the number of principal components retained. Also note that the procedures below were repeated for both the offensive and defensive data sets.

Learning traditional centers

For the first iteration of clustering, we need only determine the number of principal components to retain, since we must obviously have exactly two mixture components: centers and non-centers. The number of principal components to retain was determined fitting a GMM to 1, ..., 10 principal components, and then for each model, the F-score was computed by matching with the NBA labels, and then simply picking the number of components which lead to the greatest score. Recall that the F-score (see Dice, 1945 for more details) is given by

$$F_1 = \frac{2 \times T_p}{2 \times T_p + F_p + F_n},$$

where T_p denotes ”true positives”, F_p denotes ”false positives”, and F_n denotes ”false negatives”. The positions labels of C, F-C and C-F were all considered centers when computing the F-score, and all other positions labels were considered non-centers. Also note that since GMMs do not explicitly assign observations a label, to determine which cluster corresponded to the center cluster, the cluster containing the most players with a position label of center was considered to be the center cluster. Therefore, when it came to computing the F-score, all players in the center cluster were predicted to be centers, and all other players were predicted to be non-centers.

For the offensive data set, one principal component was unequivocally found to be most appropriate, with an F-score of 0.837. In practical terms, this clustering also seemed appropriate: some noteworthy false positives were Derrick Favours, Richaun Holmes, and Kevon Looney, and some notable false negatives were Lauri Markkanen and Kelly Olynyk. Note that we use the terms false negatives and false positives to denote players who were clustered incorrectly based on their NBA position label.

As one would expect, the defensive data set was a lot fuzzier: F-scores were nearly identical for 1, 2, 3 and 4 principal components (all hovered around 0.84), before suffering a major drop off. For purely practical reasons, one principal component was retained. In this case, some notable false positives Serge Ibaka, Kevon Looney, and Blake Griffin, whereas some notable false negatives were Aleksej Pokusevski and Larry Nance Jr.

Tables containing all the false positive and false negative centers can be found in Appendix B. Tables containing all centers classifications (based on cluster probabilities) can be found in Appendix C.

Learning the other positions

The current clustering context is slightly different from traditional problems:

1. We need to determine the clustering parameters *as well* as the number of features (i.e., principal components) to retain for clustering.

2. We have some useful, but not entirely accurate, label information that we would like to make use of.

In traditional applications, a common method for picking the number of mixture components (which in this case, corresponds to the number of latent true underlying positions) in a GMM is to use the Bayesian information criterion (BIC). Given that the BIC of a model is a function of the likelihood, it can be used to determine the appropriate number of clusters for some fixed amount of principal components, but it is unsuitable for determining the number of principal components to retain, since the the likelihood will decrease when additional principal components are added. Directly using BIC would also completely ignore the partial label information.

To determine how many principal components to retain, we require an assessment of fit that is independent of the likelihood. This is why the second modelling heuristic, i.e., that sharing an underlying position increases the likelihood of sharing a label, is key.

We first define the following shorthand:

$$\Pr(L_k \mid P_j) = \Pr(\text{Having label } L_k \mid \text{Underlying position } P_j).$$

Assume we have some arbitrary clustering. Let $\pi_j^{(i)}$ denote player i's probability of being assigned position P_j. Based on the heuristic that labels convey meaningful information in the aggregate, we would expect that for an arbitrary underlying position P_j, there exists some label L_k that is much more prevalent amongst players with underlying position P_j, or more formally, that there is some label L_k such that $P(L_k \mid P_j) \gg P(L_h \mid P_j)$ for all $h \neq k$.

For some proposed cluster (i.e., proposed position), we define the following score S_j:

$$L_{k_1} = \arg\max_{\text{Labels}} \Pr(L_h \mid P_j), \quad L_{k_2} = \arg\max_{\text{Labels}\setminus L_{k_1}} \Pr(L_h \mid P_j),$$

and

$$S_j = \Pr(L_{k_1} \mid P_j) - \Pr(L_{k_2} \mid P_j).$$

This score basically favors clusters that have one label that is predominant, since by maximizing it, we maximize the difference in proportions of the most popular label and the second most popular label. In other words, maximizing the score maximizes the homogeneity within the cluster.

Although calculating these conditional probabilities exactly would require knowing cluster membership, we can estimate them using the *predicted* cluster membership probabilities, along with Bayes' Rule, viz.

$$\Pr(L_h \mid P_j) = \frac{\Pr(P_j \mid L_h)\Pr(L_h)}{\Pr(P_j)} = \frac{\sum_i^{\text{Label } L_h} (\pi_j^{(i)}) \times (\frac{N_{L_h}}{N})}{\sum_i^{\text{All players}} \pi_j^{(i)}},$$

where N_{L_h} denotes the number of players with label L_h, and N denotes the total number of players in the data set. To obtain an overall score for the proposed clustering, we then simply compute a weighted average of cluster scores, weighted by the number of players within the respective cluster.

Since the EM algorithm depends on cluster initialization, we note that the clustering can differ from one iteration to the next, meaning that every iteration can potentially have a different score, a different likelihood, and different model parameters, for a fixed amount of principal components and mixture components.

Ideally, we would like a model with a large score (as defined above) and a large likelihood (by large likelihood, we mean relative to iterations with the same number of principal components and mixture components, since otherwise they are not directly comparable). To pick the optimal classification, Algorithm 1 was performed.

Using this procedure, the optimal offensive clustering had an average rank of 99, and the optimal defensive clustering had an average rank of 99.5. Furthermore, the clusterings proposed by the final models seemed heuristically correct.

2.4.5 Results and interpretation

A scatter plot of the non-center positions is given in Figure 2.10 and the five players most likely to belong to each non-center cluster are listed in Table 2.1 and Table 2.2. A full list of all non-center classifications is given in Appendix B. A short practical interpretation and assessment of the clustering is briefly given here.

Heuristically speaking, the offensive clustering is appropriate. Position 0 seemed to contain "shooters with a bit of a handle," i.e., perimeter players who are capable of hand-

Algorithm 1 An algorithm for picking a suitable classification

for i in $1, \ldots, 8$ **do**
  for j in $2, \ldots, 6$ **do**
    for k in 1 to 100 **do**
      Set seed to k.
      Fit a GMM using i principal components and j mixture components.
      Predict the cluster probabilities.
      Compute the score.
      Compute the likelihood.
    end for
    Rank the likelihood for the 100 iterations.
    Rank the scores of the 100 iterations.
    Retain the clustering with the highest average rank.
  end for
end for
  Of all the retained models, pick the one with the highest average rank

ling the ball, but that aren't generally the main ball handler on their team. Some notable examples were players like Jordan Poole, Eric Bledsoe, and Anfernee Simons.

Position 1 seemed to contain ”multi-level” players, i.e., players who tend to operate all over the floor. Although the model assigned all these players the same position, it is interesting to note that Position 1 shows what appear to be sub-clusters. The left sub-cluster contained ”strech-bigs,” like Al Horford, Christian Wood, or Brook Lopez, whereas the right sub-cluster contained players who operate at all three levels by slashing to the basket, like Jimmy Butler, Kevin Durant, or LeBron James.

Position 2 contained ”offensive initiators”, i.e., scoring threats who handle the ball a lot, like Ja Morant, Trae Young, or Luka Doncic. Position 3 seemed to contain players who operated exclusively on the perimeter, without handling the ball much, such as Mike Muscala, Trevor Ariza, and Maxi Kleber.

Position 4 contained exclusively Zion Williamson and Giannis Antetokoumnpo. This agrees with the sentiment that they are (nearly) one-of-a-kind players. Players also appeared to be appropriately placed on the cluster boundaries: for example, Steph Curry and Damian Lillard were both classified as Position 2, but both had a pretty sizable probability of belonging to Position 0, which coincides with their willingness to play off-ball.

Although the defensive clustering was interesting, it was a lot fuzzier than the offensive clustering. Position 0 contained what can only be described as ”limited defenders” such as Trae Young, Ja Morant, and Eric Gordon. Position 1 contained defenders who rely on their size, length and wingspan, such as Grant Williams, Pascal Siakam, and Robert Covington. Position 2 was by far the most interpretable cluster: it aggregated all the defensive pests with a knack for stealing the ball, such as Matisse Thybulle, Alex Caruso, and T.J. McConnell.

Position 3 contained defenders who rely on foot speed to defend on the perimeter, like Marcus Smart, De'Anthony Melton, and Jevon Carter. Given that a big part of defense is about limiting the offensive player's ability, it is obviously quite difficult to describe a defensive performance with count data. One clear avenue for improvement would be to use tracking data, as suggested by Bornn et al. (2016), but unfortunately it is no longer publicly available.

2.4.6 Assigning unusable players a position

Recall that the aforementioned procedure is just for *learning* the positions. It does not actually involve these unusable players. We now explore how to assign unusable players a position.

Table 2.1: Five players with greatest classification probabilities for each offensive position.

Position 0	Position 1	Position 2	Position 3	Position 4
Austin Rivers	Kawhi Leonard	Luka Doncic	Mike Muscala	G. Antetokounmpo
Damyean Dotson	Jimmy Butler	Ja Morant	L. Markkanen	Z. Williamson
Trey Burke	Julius Randle	Paul George	Jeff Green	
F. Campazzo	Ben Simmons	Elfrid Payton	Obi Toppin	
Payton Pritchard	Pascal Siakam	Kyrie Irving	Kenyon Martin Jr.	

For unusable players who are close to the 1000 possession cutoff, assigning a position is straightforward: we simply classify their feature vector, and pick the cluster with the greatest probability. For players who are nowhere near the cutoff, this approach is more problematic since their ”per 100 possessions” stats can be far too extreme. For example, some players played fewer than 20 possessions during the season, which means that if they were to record a single block, they would by far be the greatest shot blocker in the data set. For these extreme cases, we have no choice but to rely on their labels.

To handle both these cases simultaneously, we propose the following high-level idea: a weighted average between the mean classification probabilities of their label and the player’s direct classification probabilities. The more a player has played, the more weight

Table 2.2: Five players with greatest classification probabilities for each defensive position.

Position 0	Position 1	Position 2	Position 3
Eric Gordon	Nassir Little	Matisse Thybulle	De’Anthony Melton
DeMar DeRozan	Josh Hart	Facundo Campazzo	David Nwaba
Stephen Curry	Deni Avdija	T.J. McConnell	Caleb Martin
Trae Young	Aleksej Pokusevski	Gabe Vincent	Victor Oladipo
D.J. Augustin	Nemanja Bjelica	Alex Caruso	Gary Harris

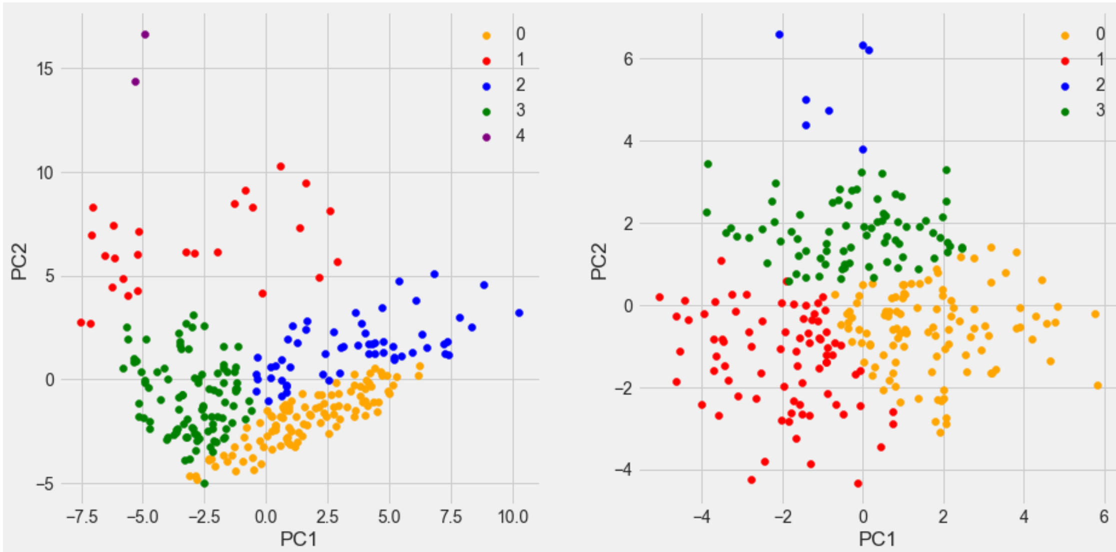

Figure 2.10: Left: Scatter plots of position clusters on offense. Right: Scatter plots of position clusters on defense.

should be placed of their direct classification probability, and vice versa.

Time to position convergence

To determine how long a player must play before their position classification can be considered accurate, we return to the training set of usable players, and study how long it takes for players to be correctly classified under the assumption that their final classification is correct. Recall the Kullback–Leibler divergence, which for discrete distributions P and Q is given by

$$D_{KL}(P, Q) = \sum_{x \in X} P(x) \ln \{P(x)/Q(x)\}.$$

The Kullback–Leibler divergence can be used to measure the "difference" between two distributions. For a given player i, let π_i denote their end-of-season classification probabilities, and let $Q_{i,t}$ denote their classification probabilities using their running feature average at time t. We say that a player has converged by time T if, for some given $\epsilon \in (0, \infty)$,

$$\sup_{t > T} D_{KL}(P, Q_t) < \epsilon.$$

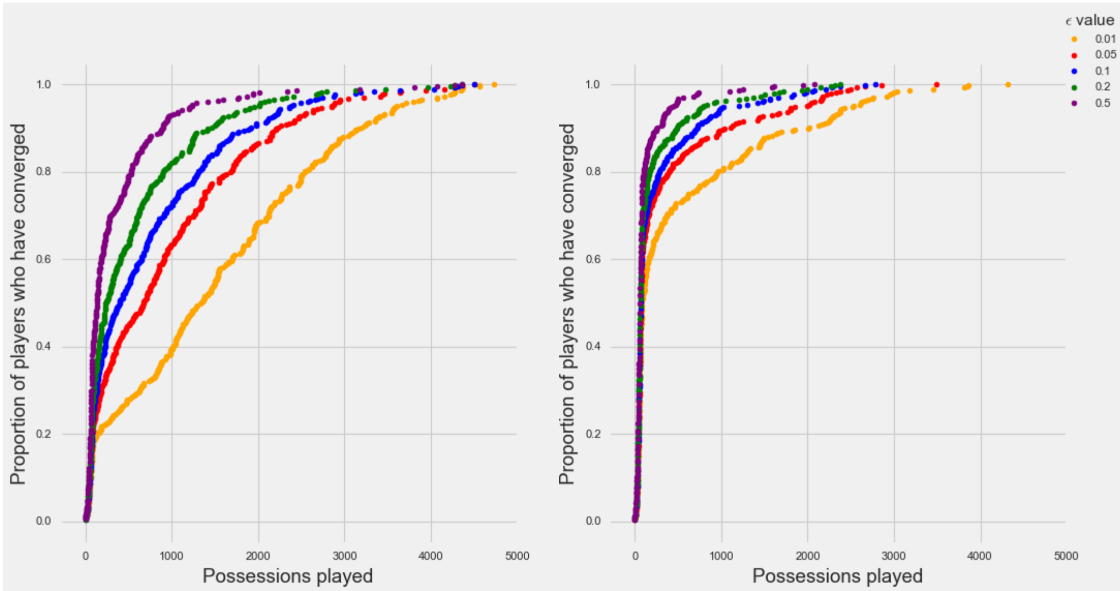

Figure 2.11: Left: League-wide convergence rate for offensive non-centers. Right: League-wide convergence rate for defensive non-centers.

Figure 2.11 shows time to convergence results for the non-center positions. From this plot, it was heuristically decided that 1000 possessions was in fact a sufficient cutoff for both offensive and defensive non-centers. The same procedure was performed for the center position. In line with our heuristics, time to convergence was much shorter when dealing with centers: 500 possessions seemed sufficient to determine whether a player was a center both on offense and defense.

Position assignment

Let $n^{(i)}$ denote the number of possessions played by player i, let $\mathbf{c}^{(L_k)}$ denote the mean cluster membership probabilities for non-center players with label L_k, and let $\mathbf{p}^{(i)}$ denote the predicted probabilities for directly classifying player i's feature vector. We obtained a "smoothed" probability, $\mathbf{p}_S^{(i)}$ by performing the following operation:

$$\mathbf{p}_S^{(i)} = \frac{n^{(i)}}{1000} \times \mathbf{p}^{(i)} + \frac{1000 - n^{(i)}}{1000} \times \mathbf{c}^{(L_k)},$$

and then assigning the position with the greatest value within the $\mathbf{p}_S^{(i)}$ vector. A similar formula was used to to smooth center probabilities, but the denominator was instead set

Table 2.3: Replacement players counts for every offensive and defensive position.

Position	Defense Counts	Offense Counts
0	56	61
1	53	12
2	0	5
3	22	58
C	43	38

to 500. Unusable players were classified as centers if their smoothed center probability was greater than 0.5. Otherwise, they were classified according to their most probable non-center position.

Table 2.3 contains the counts of players who were assigned each underlying position. A full table containing the position assignments of all unusable players can be found in Appendix E. Note that although we estimated six underlying offensive positions and five underlying defensive positions, no unusable players were assigned to offensive Position 4, and no players were assigned to defensive Position 2, which means in the context of rebounding analysis, we effectively had five offensive replacement players and four defensive replacement players.

2.5 Estimation Results

2.5.1 Team rebounding estimation results

MCMC implementation details

As in the simulation study, the parameters for the team rebounding model were estimated using `Stan`, with Markov chains consisting of 1000 warm up iterations and 1000 sampling iterations. Again, the convergence was evaluated by running four chains, and computing the value of $\hat{R}$ for all marginal distributions. Again, all values, including the estimated intercept $\hat{\alpha}$, yielded values of nearly 1.

Practical assessment of parameter posteriors

Before formally evaluating the model fit, we deem in pertinent to explore the posterior distributions from a practical point of view. Further comments on the practical implications of the model are discussed in Section 2.6.

As shown in Figure 2.12, despite the fact that the distribution of prior means was heavily positively skewed, the posterior means appear to be symmetrical and roughly Normal. This is in line with how we would expect traits to be distributed within a population: if we think as team rebounding ability being a function of many latent variables (like strength, size, positioning, effort, age, etc.), the Central Limit Theorem implies that the combination of the factors be approximately Normal. Furthermore, the average of the defensive posterior means is equal to 0.139, the average of the offensive posterior means is equal to 0.0406, and the posterior mean for the intercept is equal to 0.535.

Thus if we were to play five average defensive rebounders against five average offensive rebounders, the predicted defensive rebounding rate of the team would be 73.63%, which is nearly identical to the league-wide average defensive rebounding rate of 73.8%.

We also explore the standard deviations of the posterior distributions, which are given in Figure 2.13. Reassuringly, for all players, the variance of the posterior distribution is lesser than that of the prior, which suggests that the choice of prior distributions was compatible with the likelihood. The average posterior standard deviation was 0.064 for both offense and defense.

Furthermore, note that the posterior variances are much smaller for the replacement players (except for offensive Position 2, which only contained five replacement players). If there were a lot of heterogeneity in the team rebounding ability of all the replacement players who were grouped together, we would expect the variance of the respective replacement player parameter to be larger than if there was homogeneity. Therefore, the fact that the posterior variances are small suggests that the replacement player grouping was appropriate.

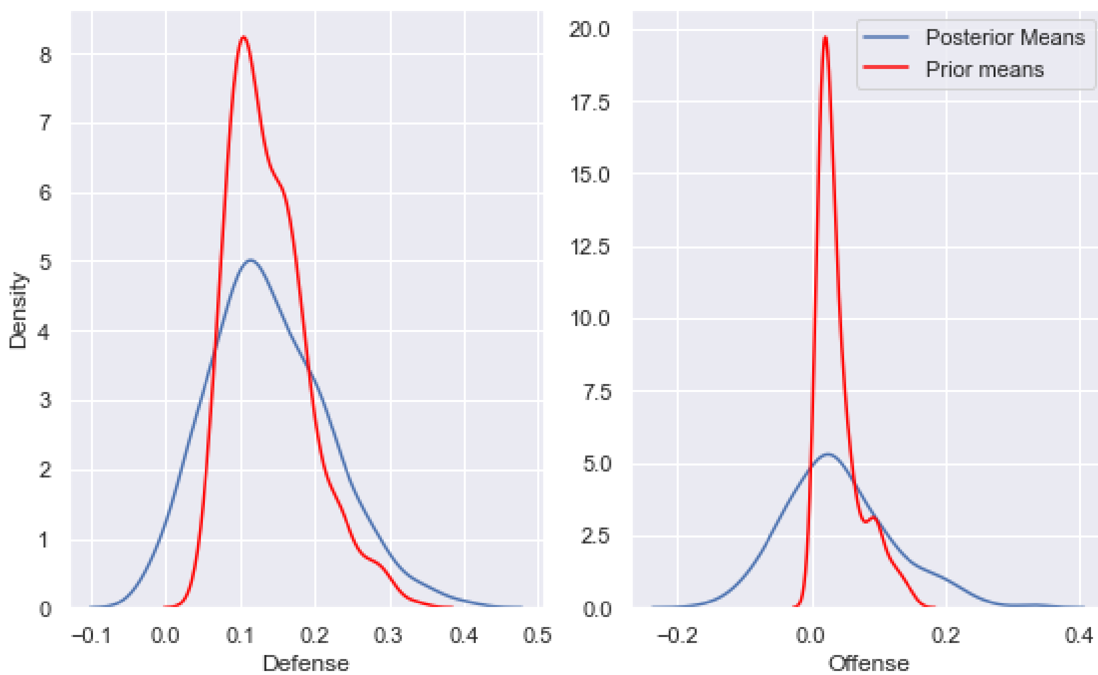

Figure 2.12: Distribution of prior means and posterior means.

One final feature worth exploring is whether the estimated parameters (which are given by the posterior means) tell a different story than the empirical rebounding rates, since otherwise, one could just as well directly use individual rebounding rates to measure contributions to team rebounding. Figure 2.14 plots the posterior mean against the observed individual rebounding rate. There is an obvious correlation between the two, but they are not perfectly concordant. This further supports the idea that individual rebounding rates don't tell the whole story when measuring contributions to team rebounding. This also suggests that the prior variances were not too small, since the likelihood function clearly plays a part in the posterior distribution.

The Hosmer–Lemeshow test

Although we technically have categorical features in the team rebounding model, conventional methods for assessing fit, such as the likelihood ratio test or Pearson's chi-squared test, are inappropriate: since specific 10-man combinations are so rare, the num-

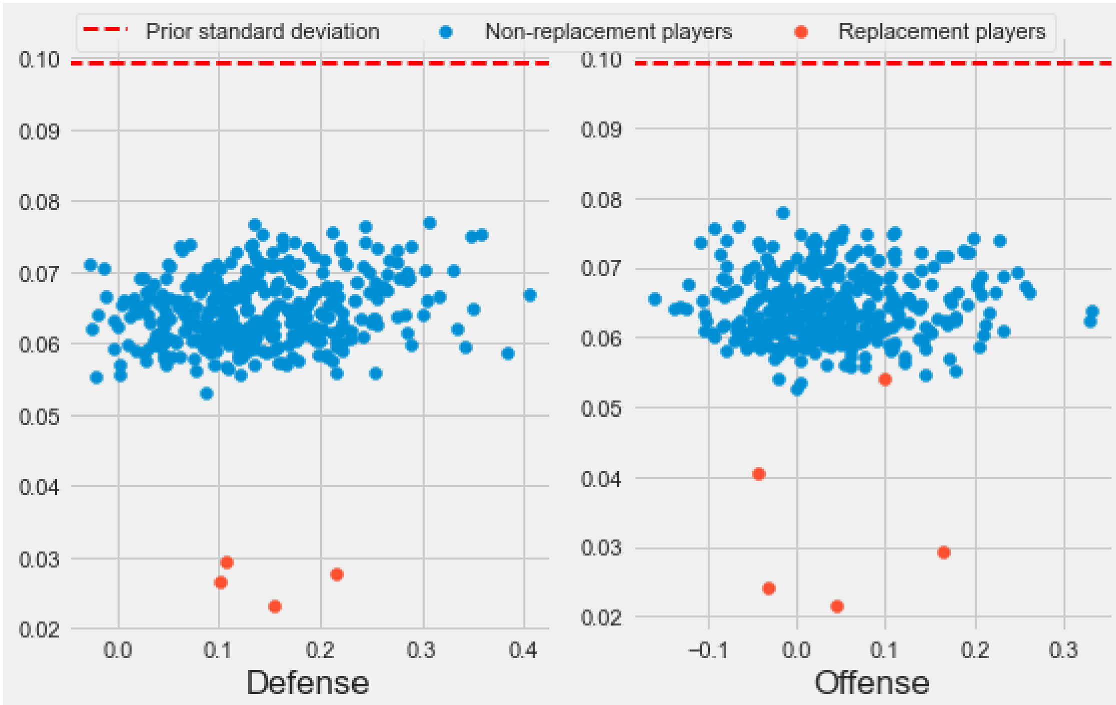

Figure 2.13: Standard deviations (y-axis) of each posterior distribution against the posterior mean (x-axis). Given the unequal replacement player partition given in Table 2.3, the difference in replacement player posterior variance is to be expected.

ber of trials for each observed binomial response is often very low, meaning that the data behave much like in the ungrouped case. We instead opt to use the Hosmer–Lemeshow test to evaluate model accuracy; see Hosmer, Lemeshow and Sturdivant, 2013. Effectively, this involves partitioning observations into equally sized groups based on their predicted probabilities, and then measuring the difference between observed and expected counts within each group. The test-statistic is given by

$$H = \sum_{g=1}^{G} \frac{(O_{1g} - E_{1g})^2}{N_g \pi_g (1 - \pi_g)},$$

where G denotes the number of groups, O_{1g} denotes the observed successes in group g, E_{1g} denote the predicted number of success in group g, N_g denotes the size of group g, and π_g denotes the mean predicted probability of observations in the gth group.

Asymptotically, the test statistic follows a chi-square distribution, with $G-2$ degrees of freedom. This is obviously problematic since the acceptance of the null hypothesis relies

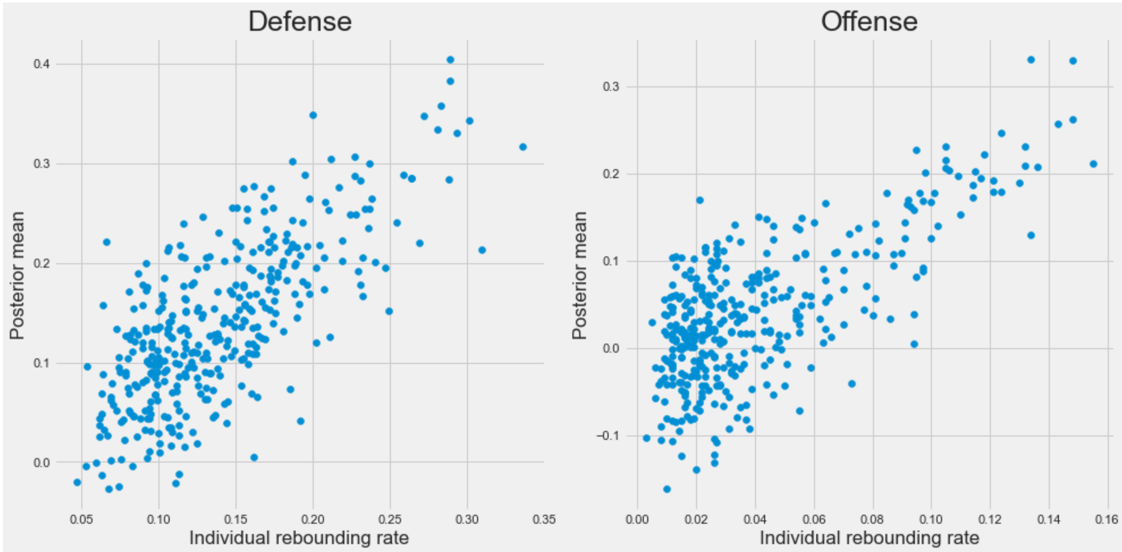

Figure 2.14: Posterior mean against prior mean. Although correlation is to be expected, the ranking of ability suggested by the model is different than that suggested by the raw rebounding rates.

on the arbitrary choice of the number of groups. However, in the case of the training set, this is not a major issue: given that we are forcibly restricting the model parameters, there is a significant difference in the observed and predicted counts of the extreme groups, meaning that for most binning schemes, the null hypothesis is rejected. However, when predicting on the testing set, where we have far fewer observations, the choice of group size can influence the conclusion of the Hosmer–Lemeshow test.

In-sample prediction

To assess whether the model is reasonable, we first assess the fit on the training set. This was done by grouping the observations into ten equally sized groups (each containing approximately 10,660 observations), plotting the observed probabilities against the predicted ones (which is shown in 2.15), and performing the Hosmer–Lemeshow test.

As we would expect, the more extreme the observed probability is, the more the model struggles to predict it accurately. This is also reflected in the Hosmer–Lemeshow test: the test statistic is equal to 68.63 ($p \approx 0$), but the smallest bin and largest bin contribute 15.72

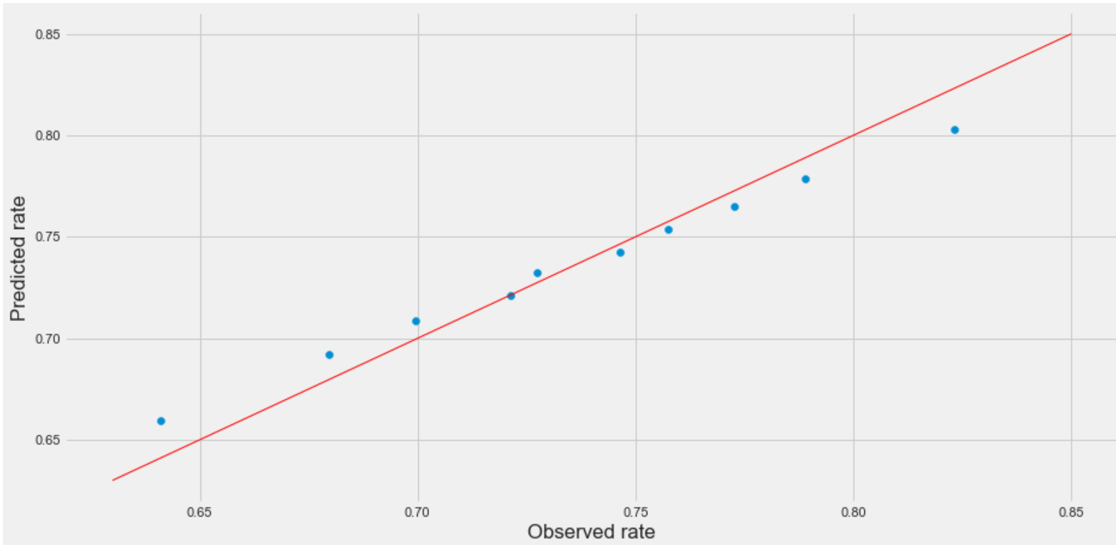

Figure 2.15: In-sample prediction of rebounding rate during the 2020–21 NBA season. Lineups were grouped into 10 equally sized bins based on their predicted rebounding rate.

and 27.30 to the overall statistic, respectively. Although one could obviously improve the fit by increasing the variance of the prior distributions, this would come at the cost of poorer out-of-sample performance.

It is also worth noting that the variance of predictions varies drastically between groups: the smallest group contained predictions between 0.549 and 0.681, and the largest group contained predictions between 0.794 and 0.878, whereas the widest on the non-extreme bins had a width of 0.02. Since the league-wide observed average rebounding rate is 73.8%, the lineups falling into the extreme bins are almost certainly odd lineups that aren't used regularly. With these considerations in mind, it was deemed appropriate to predict on the following season.

2.5.2 Individual rebound collecting estimation results

MCMC implementation details

The MCMC implementation was identical to that of the team rebounding model, except in this instance, since we are modelling individual rebounding probabilities conditional on the team having collected the rebound, we needed to handle the offensive and defensive parameters separately. For both the offensive and defensive models, all marginal distributions yielded a value of $\hat{R}$ near 1.

Practical assessment of parameter posteriors

As with the team rebounding model parameters, we consider practical interpretations of the estimated parameters. As suggested by Figure 2.16, the posterior means seem to be roughly normally distributed, which seems more appropriate for rebounding ability than the heavily skewed empirical rebounding rates (although the offensive parameters still show some positive skew). The average posterior mean is equal to 1.148 for the defensive parameters and -0.785 for the offensive parameters.

The shift between offensive and defensive parameters has a practical explanation: there are a lot more offensive team rebounds than there are team defensive rebounds, because blocked shots are often swatted out of bounds. These values suggest that for the average defensive lineup, only about 6% of defensive rebounds are team rebounds, whereas for the average offensive lineup, that number skyrockets to about 30%. This quirk of the data, and its practical implications, are discussed further in Section 2.7.

The posterior variances have a lot more going on than in the team rebounding section. Figure 2.17, which plots the posterior mean against the posterior standard deviation for each player, has two interesting features. Firstly, the variance of the offensive posterior means is far greater than those of the defensive posterior means. This is to be expected, since we are modelling individual rebounding probabilities conditionally on the team having collected an offensive or defensive rebound, which means that there are far more

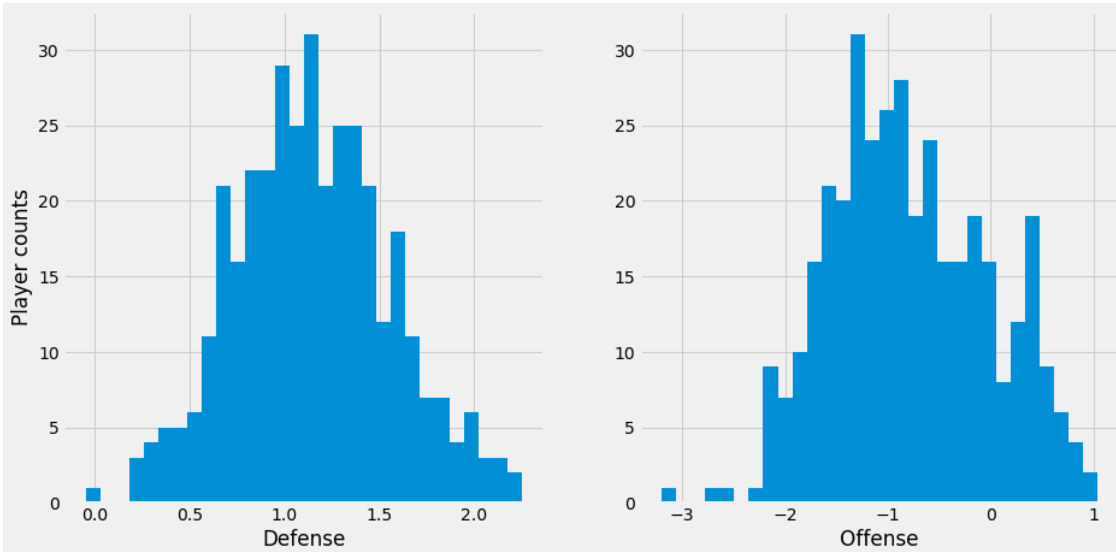

Figure 2.16: Posterior means of the individual rebound collecting parameters.

defensive observations than there are offensive observations. The second interesting feature is that there is obviously a relationship between the posterior mean and the posterior variance.

This relationship can be understood through the Fisher information. Consider a single multinomial observation, whose log-likelihood is given by Equation (2.2), and suppose we are only interested in the variance of γ_1 (the superscript is omitted for readability). The second derivative of the log-likelihood with respect to γ_1 is

$$\frac{\partial^2}{\partial\gamma_1^2}\,\ell = -n_L\left(\frac{e^{\gamma_1}}{1+e^{\gamma_1}+\cdots+e^{\gamma_5}}\right)\left(1-\frac{e^{\gamma_1}}{1+e^{\gamma_1}+\cdots+e^{\gamma_5}}\right).$$

Notice this expression does not depend on the data, which means that the Fisher information is equal to the negative of the second derivative. This further implies that the Fisher information has a global maximum when $e^{\gamma_1}/(1+e^{\gamma_1}+\cdots+e^{\gamma_5})$ is equal to 0.5. It is also strictly increasing in γ_1 on the left of the global maximum, and is strictly decreasing on the right of it.

Conditional individual rebounding rates are nearly always less then 0.5, meaning that in the current context, the Fisher information is essentially an increasing function of the parameter value. Now, recall the Bernstein–von Mises theorem, which states that the pos-

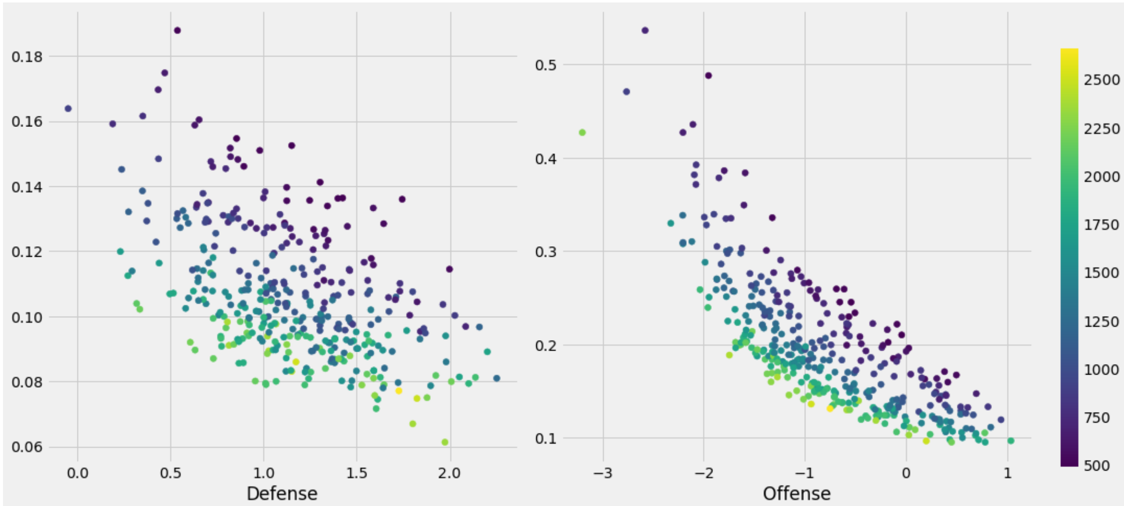

Figure 2.17: Posterior standard deviations (y-axis) plotted against their respective means (x-axis). Color represents the number of minutes of the player in question, and is a proxy for the number of multinomial observations used to estimate the parameter. Replacement players were omitted due their much lower standard deviations, which made the mean-variance relationship less apparent.

terior distribution, under regularity conditions which are satisfied in the present context, converges to a Normal distribution, i.e., as $n \to \infty$,

$$\pi(\boldsymbol{\gamma} \mid \boldsymbol{x}) \rightsquigarrow \mathcal{N}\big[\boldsymbol{\gamma}_0, n^{-1} I(\boldsymbol{\gamma}_0)^{-1}\big],$$

where $\boldsymbol{\gamma}_0$ is the true parameter value and $I(\boldsymbol{\gamma}_0)$ denotes the Fisher Information. In other words, asymptotically, the estimator variance is a decreasing function of the Fisher Information, which is itself an increasing function of the parameter. Therefore, the estimator variance decreases with the parameter value. Also note that this arguments holds when using the full log-likelihood, since we need only add up the second derivatives across all lineups containing the specified player.

In-sample prediction

Figure 2.18 plots the (conditional) predicted individual rebound counts in the training set against their observed values. These is very obvious over-fitting in this case. Although

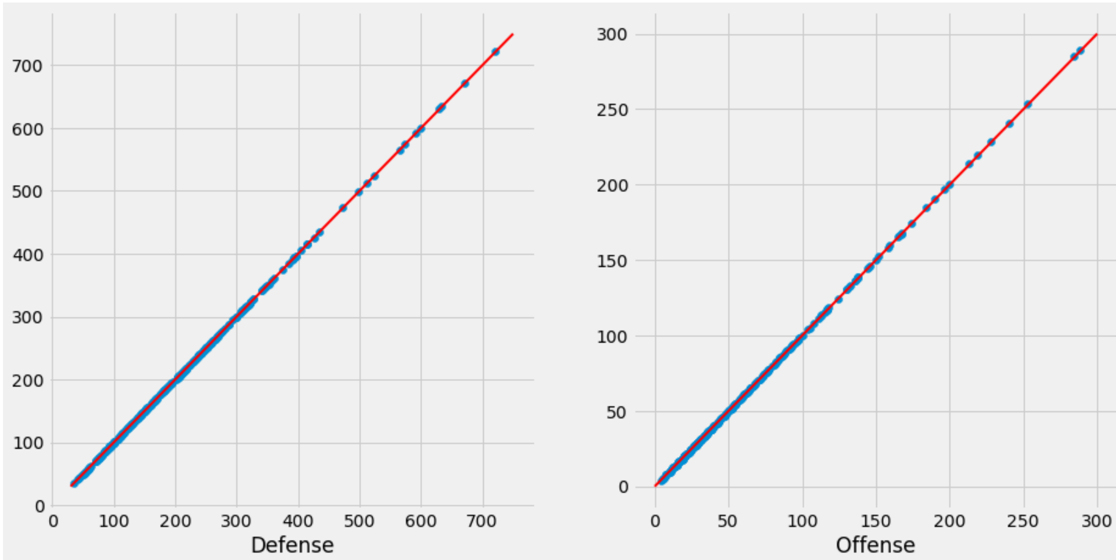

Figure 2.18: Predicted individual rebounding counts conditional on rebound collection (y-axis) plotted against observed individual rebounding counts (x-axis).

the over-fitting could almost certainly be mitigated (by introducing some form of regularization, or by using more seasons for parameter estimation, for example), we opt to leave it untouched because it effectively acts as an empirical, teammate-independent measure of ”perceived” rebounding ability.

Therefore, comparing these parameters with the team-rebounding ones makes it possible to detect ”overvalued” and ”undervalued” rebounders, which is of great practical use. If one were more interested in out-of-sample prediction accuracy rather than in inference, changes to this conditional model would be the most natural place to start.

2.6 Results

2.6.1 Prediction

To asses fit, predictions were made during the 2021–22 NBA season. By virtue of having a new season, we have new players introduced into the data set from two principal sources: rookies who were just signed to their first NBA contract, or players who were formerly replacement players, but who saw a significant increase in playing time relative to 2020–21. This second group was a combination of established players who were returning from long-term injuries (such as Jaren Jackson Jr. or Spencer Dinwiddie), and players who had improved enough to warrant more playing time (such as Gary Payton II or Isaiah Joe).

Although we could have just used the replacement player parameters to predict rebounding rates in such lineups, in general, this did not seem like an adequate assessment of model fit: a superstar coming back from injury or a first overall pick are probably not comparable to a player who is signed to a ten-day contract as an injury replacement.

Given that the main goal is to assess rebounding ability, and that the purpose of predicting is to ensure that we haven't just picked up random noise in the training data, we opted to predict only in instances where all players on the court were non-replacement players in the training set. Thus we predicted on approximately 25,000 missed shots, which represented a bit more than 20% of all shots missed during the 2021–22 season.

Within this subset of ”predictable” data, we further distinguish between two types of samples: *seen* lineup and *unseen* lineups. Seen lineups represent lineups where the exact five-man defensive lineup combination also appeared in the training set. We distinguish between these two types of samples so that we can better detect over-fitting: given the data limitations and the simplicity of the model, it seems likely that formal assessment of fit will deem the model inadequate. However, if the model performs far better in seen lineups than unseen ones, the model was probably over-fit to the training data.

The testing data contains about 19,000 instances with unseen lineups, and about 6000 instances with seen ones. Furthermore, when assessing fit graphically, on top of plotting observed and predicted counts based on the groups used for the Hosmer–Lemeshow test, we also sum predictions over teams and players, to allow for a more practically interpretable assessment of fit.

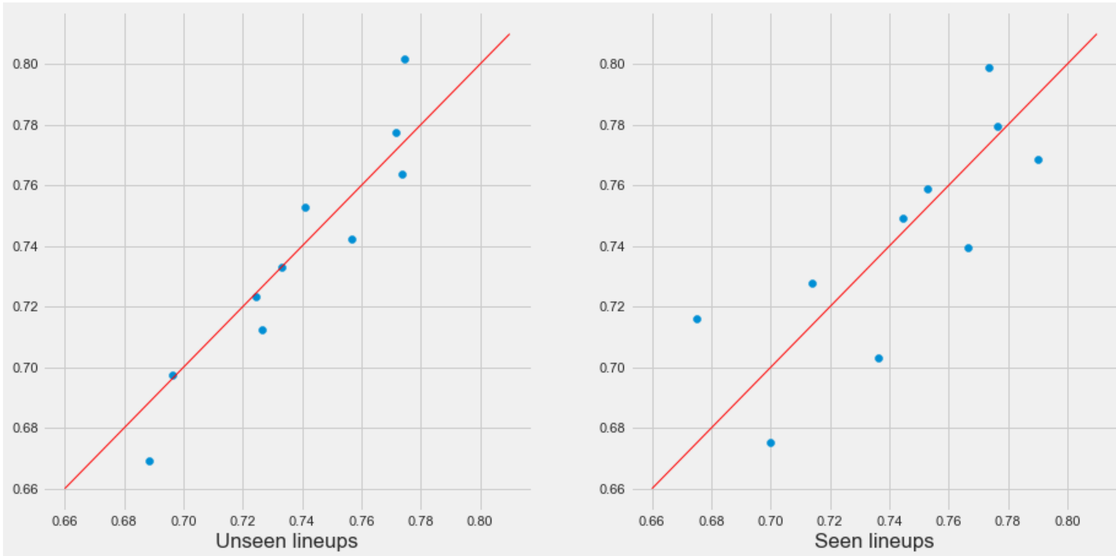

Figure 2.19: Predicted probabilities (y-axis) against observed probabilities (x-axis) for seen and unseen lineups during the 2021–22 NBA season. Note that the groups for the unseen lineups contain each about 1940 observations, and the seen lineup groups contain about 560 observations.

Team rebounding prediction

Figure 2.19 shows the observed team rebounding rates against the predicted ones, for the 10 groups used to conduct the Hosmer–Lemeshow test (note the differences in group sizes). Visually, it does not seem as though we have over-fit to the training data, and the formal test seems to support that impression: the p-value for the unseen lineups is approximately 0.015, and is 0.0483 for the seen lineups, which seems comparable given the difference in sample size. Unlike the in-sample prediction, the subjectivity of the grouping scheme is much more problematic in this case: by increasing the number of groups from 10 to 11, the p-values change to 0.043 and 0.024, respectively. Furthermore, if we increase the number of groups to 20, for both observation types, we fail to reject the null at the 5% level. In short, although the model doesn't seem to explain all the rebounding variability, it does seem to have least captured some meaningful information.

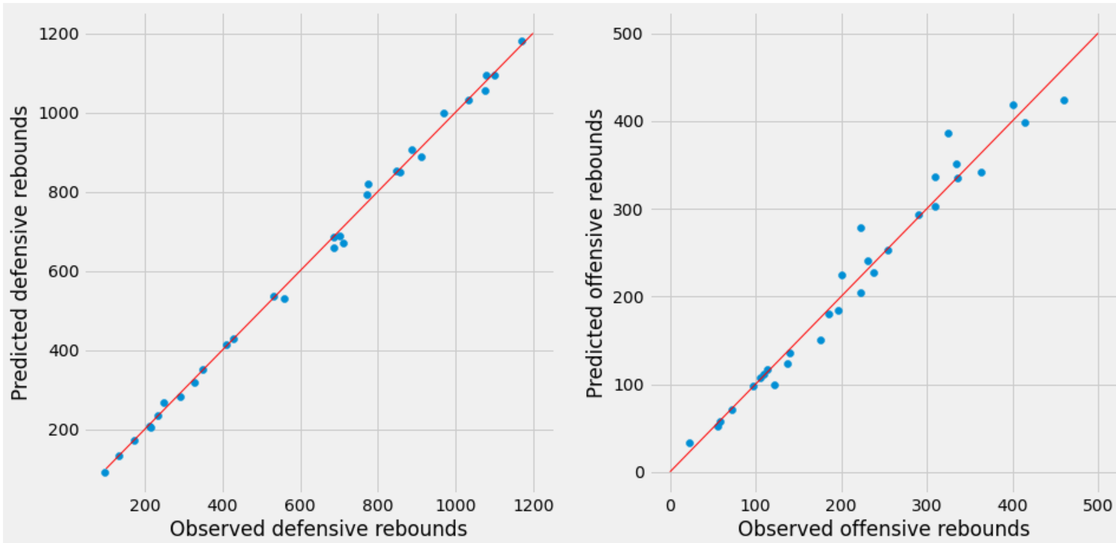

Figure 2.20: Predicted vs observed rebounding counts for each of the 30 teams during the 2021–22 NBA season (in all predictable instances).

We also compare predictions aggregated across teams, which are shown in Figure 2.20, for a more practically interpretable assessment of fit. Also note that the drastic difference in predictions is due to the variable number of replacement players found across all teams: for example, the Houston Rockets decided to rebuild, which meant most of their players were rookies and hence have very few predictable instances, whereas the Los Angeles Lakers made a point of acquiring established veteran players, which means effectively all of their missed shots were predictable. The global performance of the team-rebounding model appears to be reasonably good.

Two-stage individual rebounding prediction

We further attempted to predict individual rebound allocation. For each missed shot, we computed the expected number of rebounds for every player on the court. For each player in the testing data set, we then summed up fractional rebounds across all lineups they appeared in, and compared that to the observed counts across those same lineups.

The resulting scatter plot is in Figure 2.21. Given the greater uncertainty in the offensive individual rebounding parameters and the much smaller number of multinomial

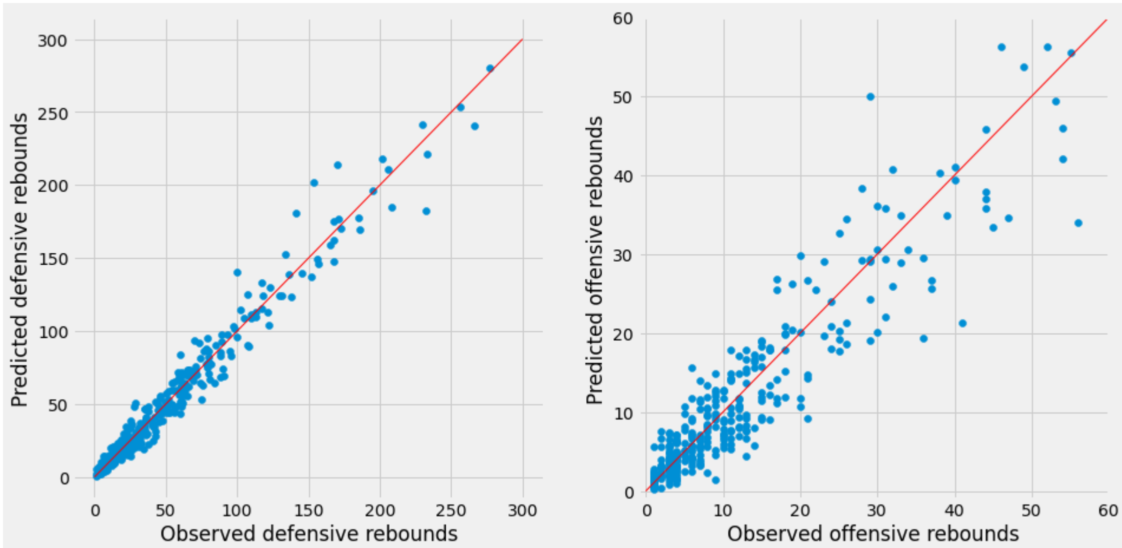

Figure 2.21: Two-stage predicted vs observed rebounding counts for individual players during the 2021–22 season. Note that about ten players were omitted from the offensive plot because they had far more offensive rebounds that those plotted, so their inclusion in the plot ”squished” everyone else together. The fit for those players was comparable to the players retained for plotting.

trials, it is no wonder that the fit is much poorer than in the defensive rebounding case.

2.6.2 Player rebounding assessment

Given the goal of accurately assessing players’ ”true” ability to steal rebounds from opponents rather than from teammates, we include a list of the top 15 offensive and defensive team rebounders in Table 2.4. These rankings are close to those suggested by looking at raw individual rebounding rates, but interesting re-orderings occur.

For example, Andre Drummond has a reputation for being an overvalued rebounder due to the large number of uncontested rebounds he collects. The model suggests that this view is at least somewhat correct. Furthermore, Steven Adams has a reputation for being an undervalued rebounder because of his willingness to let his teammates collect rebounds. Again, this view is supported by the model. The posterior means for team

Table 2.4: Top 15 players for offensive team rebounding ability (left) and defensive team rebounding ability (right)

Player	Posterior mean	Prior mean
Jonas Valanciunas	0.3316	0.134
Enes Freedom	0.3300	0.148
Dwight Howard	0.2624	0.148
Moses Brown	0.2570	0.143
Mitchell Robinson	0.2474	0.124
Steven Adams	0.2314	0.132
Thaddeus Young	0.2313	0.105
Goga Bitadze	0.2282	0.095
Daniel Gafford	0.2226	0.118
Isaiah Stewart	0.2157	0.105
Clint Capela	0.2121	0.155
Andre Drummond	0.2094	0.132
Robert Williams III	0.2081	0.136
JaVale McGee	0.2068	0.105
Jarrett Allen	0.2040	0.106

Player	Posterior mean	Prior mean
Jonas Valanciunas	0.4046	0.289
Nikola Vucevic	0.3831	0.289
Jusuf Nurkic	0.3579	0.283
Ivica Zubac	0.3491	0.200
Kevin Love	0.3479	0.272
Clint Capela	0.3430	0.301
Enes Freedom	0.3336	0.281
DeMarcus Cousins	0.3306	0.293
Andre Drummond	0.3171	0.336
Derrick Favors	0.3071	0.227
Tristan Thompson	0.3039	0.212
Steven Adams	0.3020	0.187
Nikola Jokic	0.3003	0.237
Al Horford	0.2883	0.195
Domantas Sabonis	0.2880	0.259

rebounding parameters and individual rebounding parameters for all players are given in Appendix F.

We also look at players who will appear to be *undervalued* or *overvalued*: these terms refer to players whose individual rebounding parameter is very discordant from their team rebounding parameter. Given that these parameters are on different scales, we measure discordance by subtracting the team rebounding parameter rank from the individual rebounding parameter rank. Table 2.5 contains the five most undervalued offensive and defensive rebounders, and Table 2.6 contains the five most overvalued offensive and defensive rebounders.

Table 2.5: Five most undervalued offensive rebounders (left) and defensive rebounders (right).

Player	Rank difference
Kira Lewis Jr.	267
Furkan Korkmaz	250
Jayson Tatum	245
Stephen Curry	237
Eric Bledsoe	236

Player	Rank difference
Isaac Okoro	303
Robin Lopez	243
Raul Neto	242
Joe Harris	223
Andrew Wiggins	221

One interesting thing to note when looking that the full discordance rankings is that there are clearly some player archetypes which are consistently overvalued or undervalued. In general, it seems that long-range threats (like Desmond Bane, Davis Bertans or Damian Lillard) positively impact their team's offensive rebounding far more than their individual rates would suggest. Perhaps this is due to the fact that their shooting ability forces opposing defenders onto the perimeter and away from the basket, which allows their teammates to more easily collect offensive rebounds for themselves.

Furthermore, there appears to be a subset of centers that could be overvalued on the offensive glass, such as Willie Cauley-Stein, Nikola Vucevic or Al Horford. This suggests that regardless of rebounding ability, the center will collect a significant amount of offensive rebounds by virtue of occupying prime rebounding real estate, which sounds

Table 2.6: Five most overvalued offensive rebounders (left) and defensive rebounders (right).

Player	Rank difference
Derrick Jones Jr.	−268
Willie Cauley-Stein	−256
Eric Paschall	−246
Michael Carter-Williams	−238
Kelly Oubre Jr.	−237

Player	Rank difference
R.J. Hampton	−279
Torrey Craig	−260
Chris Boucher	−258
Tyler Herro	−214
Nerlens Noel	−212

Table 2.7: Timberwolves most common lineup during the 2021–22 season. Note that parameters were estimated using data from the 2020–21 season.

Player	β_D	γ_D
Patrick Beverley	0.1257	1.0883
D'Angelo Russell	0.0422	0.5959
Anthony Edwards	0.1092	1.0665
Karl-Anthony Towns	0.2540	1.7681
Jarred Vanderbilt	0.1254	1.7108
Rudy Gobert	0.2843	2.0078

intuitively reasonable. On the defensive end, there is a trend of ballhandlers generally being overvalued (such as Russell Westbrook or Devin Booker, for example). Perhaps teams are "artificially" funnelling more rebounds to their ballhandlers, so that teams can more efficiently begin their fast break, or maybe these players are matched against opposing perimeter players, meaning that they have fewer boxing out responsibilities.

2.6.3 Example: The Timberwolves acquire Rudy Gobert

To illustrate the relevance of the model, consider the following hypothetical situation: suppose the Minnesota Timberwolves feel like their defensive rebounding needs to be improved. They consider replacing Jarred Vanderbilt, whose defensive rebounding rate is 21.1%, with Rudy Gobert, whose defensive rebounding rate is 28.8%. We explore the impact of this change on their most frequent lineup during the 2021–22 season against average offensive rebounding competition, whose parameter values are in Table 2.7.

Using the fact that the average estimated offensive team parameter across the league is 0.0406 and the estimated intercept term is 0.535 (see Section 2.5), we find that before swapping Vanderbilt out for Gobert, the predicted lineup defensive rebounding rate is 72.8%, and increases to 75.9% after the acquisition, far less than the direct difference between their individual defensive rebounding rates.

The especially noteworthy fact is that in spite of such a move making sense from a team rebounding perspective, looking at predicted individual defensive rebounding rates suggests that the acquisition does not make sense: the predicted individual rebounding rate of Towns goes from 21.3% to 20.2%, and Gobert's predicted rate after the acquisition is 25.7% (his empirical rate was 28.4%).

When this trade ended up actually being made after the 2021–22 season, the observed individual defensive rebounding rates for both Towns and Gobert showed a comparable decline. Note that because of injuries and other players being traded, it is difficult to directly compare predicted rates versus observed ones.

2.7 Discussion

2.7.1 Uncertainty considerations

Despite the posterior mean being a very natural way to assess team rebounding ability, one could exercise more caution by considering the posterior distribution as a whole. Suppose we are interested in quantifying the uncertainty in the predicted team defensive rebounding rate of a specific ten-man lineup combination (i.e. a specific defensive lineup and a specific offensive lineup), denoted by p_L.

Recall that by the Bernstein–von Mises theorem, the posterior distribution asymptotically follows a multivariate Normal distribution. Firstly, this implies that the 11-dimensional marginal distribution (10 players and the intercept) of the lineup parameters follow themselves a multivariate Normal distribution. Let $\boldsymbol{\mu_L}$ and $\boldsymbol{\Sigma_L}$ denote the mean vector and covariance matrix of this marginal distribution.

Secondly, linear transformations preserve normality, which means that the sum of all 11 components of the marginal distribution follows a univariate Normal distribution, where the mean is given by $\mu = \mathbf{1}^\top \boldsymbol{\mu_L}$ and the variance is given by $\sigma^2 = \mathbf{1}^\top \boldsymbol{\Sigma_L} \mathbf{1}$. This in turn implies that the linear component of the model follows a Normal distribution. By using the cumulative distribution function of the Gaussian distribution to approximate the logistic function, along with the Taylor approximation of the logistic function, as suggested by Daunizeau (2017), we can get approximate analytical solutions for the mean and variance of the lineup rebounding rate, which are given by

$$\mathrm{E}(p_L) \approx s\left(\frac{\mu}{\sqrt{1+c\sigma^2}}\right)$$

and

$$\mathrm{var}(p_L) \approx s\left(\frac{\mu}{\sqrt{1+c\sigma^2}}\right)\left\{1-s\left(\frac{\mu}{\sqrt{1+c\sigma^2}}\right)\right\}\left\{1-s\left(\frac{1}{\sqrt{1+c\sigma^2}}\right)\right\},$$

where c is equal to $\pi/8$ and where s denotes the sigmoid function, which is given by

$$s(x) = 1/(1+e^{-x}).$$

Of course, the catch is that these closed-form approximations rely themselves on the covariance of the posterior, which must be estimated via sampling. This means that if one were interested in the variance of a single ten-man lineup, one might as well directly estimate the mean and variance by sampling from the posterior.

The closed-form approximations are only useful if one is interested in the variance of multiple lineups: the covariance matrix is 366×366, whereas there are roughly $\binom{366}{5} \times \binom{366}{5}$ possible lineup combinations. Although estimating the league-wide covariance matrix is still probably not tractable, one could, e.g., realistically estimate all the covariance terms for a specific playoff series (roughly $10 \times 10 = 100$ terms), and then approximate the means and variances of the rebounding rates of the $2 \times \binom{10}{5} \times \binom{10}{5} = 127{,}008$ possible lineups.

2.7.2 The subtle misleadingness of rebounding numbers

Although rebounding is obviously a coveted skill amongst NBA players, simply measuring team rebounding rate may obfuscate ”practical” rebounding ability. For example, consider two players, A and B, who are identical when it comes to corralling down a missed shot. However, assume that player A is an excellent shot blocker, whereas B is a terrible one.

In a practical sense, it is natural to expect any measure of rebounding ability to value these two players equally as they are identical at collecting missed shots. But, since most blocked shots are sent out of bounds (recall that a missed shot sent out of bounds results in a team rebound for the inbounding team), in a technical sense, A is a worse defensive rebounder than B because A is generating offensive rebounds for the opposing team. We also note that some well-known shot blockers, such as Rudy Gobert, Richaun Holmes, and Hassan Whiteside, have a surprisingly low team rebounding parameter, perhaps due to this quirk.

One possible remedy to this problem is to simply remove blocked shots from rebounding opportunities, but we end up with a similar issue: if a shot blocker is able to keep the ball inbound, and tip it to a teammate, should they not be considered a superior rebounder? This appears to be the case for some great shot blockers, like Clint Capela and Jakob Poeltl, who are considered great team rebounders by the model.

On the offensive end, there is potential for the opposite problem to occur: perhaps there are players who are ”extremely good at getting blocked”, and who are therefore technically incredible offensive rebounders. This is perhaps the case for players like Ja Morant, Kira Lewis Jr. or Eric Bledsoe: the model views these players as surprisingly good offensive rebounders, but they all love to attack the basket and are not afraid to challenge players at the rim.

Since this was a data collection issue rather than a modelling one, it was ignored during the estimation procedure. But perhaps future work could explore how to isolate rebounding ability even further. For the time being, we simply suggest being wary of player valuations in the case of exceptional shot blockers and ”blockees.”

2.7.3 Model limitations

Validity of the constant rebounding ability assumption

The idea that team rebounding can be explained solely as an interactionless combination of players on the court is almost certainly false. For example, one popular strategy to mitigate the effectiveness of Rudy Gobert has been to force him to guard capable three point shooters, hence forcing him away from the basket and impacting his defensive rebounding. This suggests that there is an important interaction that is being ignored. Obviously, modelling all such interactions is not tractable given the very limited amount of data, and is omitted in most APM based approaches. One potentially feasible way to incorporate these interactions would be to include an interaction term based on the positions (using the positions assigned in Section 2.4) of the players in question, and assuming that the interaction is identical across all players of the given positions.

Furthermore, although the idea that all players have some constant intrinsic value of rebounding ability is probably approximately true, it is almost certainly not exact, since players may adapt their play-style based on lineup composition. For example, despite the model predicting an increase in individual rebounding rate for Draymond Green when the Golden State Warriors play him in place of Kevon Looney, it is probably the case that Green's contributions to team rebounding are underestimated in that specific lineup, since he is probably more aggressively pursuing rebounds than he would if Looney were still on the court. Accounting for such a difference is obviously impossible given the amount of data, and we therefore suggest that model predictions perhaps be viewed as a "lower bound" on team rebounding ability.

Omitted covariates

For the sake of tractability and ease of implementation, some important intra-season covariates were omitted. Examples include home-court advantage, days of rest between games, and whether or not the game was in "garbage time" (i.e. when the outcome of the game is already decided and game intensity drastically drops). The current model could be expanded by including such features. This would of course require that a greater amount of training data be used, and hence would greatly increase computation time.

Furthermore, given only one season of data was used when estimating the parameters, inter-season covariates were obviously ignored. For example, we would expect young players to practice during the off-season and physically develop, which would improve their rebounding in future seasons. Conversely, for older players, since rebounding is such a physically demanding task, we would expect them to decline as time goes on.

As currently defined, the model would not be able to account for these trends, even if multiple seasons were used, and the parameter would simply be the averaged performance of the player across all training seasons. Future work could involve accounting for expected player development and expected age-related declined, perhaps by adapting the methodology suggested by Vaci et al. (2019), who used age curves to account for change in player performance between seasons.

Chapter 3

Conclusion

As mentioned in Chapter 1, assessing individual contributions to team success is of great importance in professional basketball, since player acquisitions are such an important part of the sport. Despite there being many APM based approaches, these are almost exclusively confined to modelling variables with infinite support, like net rating or point differential. The object of this thesis was to adapt this methodology to model a probability rather than a net rating for individual rebounding specifically, by both reducing the number of parameters, and using rebounding specific knowledge to construct the priors.

To achieve this, data were collected from the 2020–21 NBA season. Firstly, play-by-play data were collected so that we could know exactly who was on the court when a rebound was collected, and assess impact that cannot be directly measured. Secondly, player tracking data were collected to learn player tendencies, so that we could reduce the number of model parameters by grouping underused players with similar play styles.

The first step (but presented second) in the modelling procedure was to use basketball heuristics to sensibly reduce the dimension of the player tracking data set, thus making player clustering using GMMs feasible. This clustering procedure allowed us to replace 348 parameters (two for each of the 174 players having played less than 1000 possessions) with five offensive reference parameters and four defensive reference parameters for each of the two model stages.

To model rebounding in such a way that was both consistent with observations and allowed for impact invisible on the box score, we assumed that the probability of a player collecting a rebound could be written as

$$\begin{aligned}&\Pr(\text{Player A collects rebound} \mid \text{Missed shot}) = \\ &\qquad \Pr(\text{Player A collects rebound} \mid \text{A's team collects rebound}) \times \\ &\qquad\qquad \Pr(\text{A's team collects rebound} \mid \text{Missed shot}),\end{aligned}$$

and we then modelled these probabilities separately.

We assumed that the probability of the team collecting a rebound (the second term of the above expansion) could be modelled using Bayesian logistic regression: each player was assumed to have some intrinsic team rebounding contribution parameter, and a normal prior was constructed on a player-by-player basis based on their individual rebounding rate. The relationship between the lineup defensive rebounding rate and the players present on the court is given by

$$p^L = \frac{e^{\beta_1^D + \cdots + \beta_5^D - \beta_1^O - \cdots - \beta_5^O + \alpha}}{1 + e^{\beta_1^D + \cdots + \beta_5^D - \beta_1^O - \cdots - \beta_5^O + \alpha}}.$$

The adequacy of this approach was verified by performing a simulation study.

We then assumed that, conditionally on the team having collected the rebound, the probability of a specific individual collecting follows a multinomial distribution. The key idea was that again, every player has some intrinsic weight that is constant across lineups, and that it could be measured against the probability of there being a team rebound, since teams rebounds are always a possible outcome, regardless of who is on the court. The probability of player i collecting the rebound is given by

$$p_i^L = \frac{e^{\gamma_i^D}}{1 + e^{\gamma_1^D} + \cdots + e^{\gamma_5^D}}.$$

The posterior distributions were estimated using MCMC, and their properties were studied to verify firstly, that the choice of prior distributions was sensible, and secondly, that the estimation was specific enough to be useful in practice. The model was then used to predict rebounding outcomes in a subset of the 2021–22 NBA season, where it performed reasonably well. Furthermore, the model was used in practical contexts, to both rank players and evaluate potential player acquisitions.

On top of improving the current model, as suggested in Section 2.7.3, we propose a few avenues for future work. Firstly, basketball has plenty of binary events of interest: making a three pointer, getting a stop, scoring out of a timeout, and so on. The proposed methodology could be used to evaluate how individual players contribute to these events beyond the box score.

For example, excellent shooters are often said to have ”gravity,” i.e., they create higher quality scoring chances for teammates just by virtue of being on the court. Perhaps a similar model for scoring probability, and where individual player priors depend on three point percentage, could be useful for late game situations when a basket is needed.

Another potential avenue for future work is to consider alternatives to linear regression in APM-like models. The notion of ”diminishing returns” seems like it could be applicable in contexts besides rebounding, and cannot be modelled by assuming performance is directly linearly additive. For example, on a team with plenty of scoring ability, it seems unlikely that adding another talented scorer would be beneficial, since all players must share the one ball.

Appendix A

Description of the variables

A glossary giving the definition of most statistics can be found at `https://www.nba.com/stats/help/glossary`. Note that some of the tracking data are not actually included in the glossary, but they generally have quite descriptive names. Some examples of such data can be found at `https://www.nba.com/stats/players/drives`.

A.1 Defense variables

- AVG DREB Distance
- Avg Speed Def
- Contested 2PT Shots
- Contested 3PT Shots
- Deflections
- DEF Boxouts
- DEF Loose Balls Recovered
- DFGA
- Dist. Miles Def
- DREB Chances
- Uncontested DREB
- PFD
- STL
- matchup3FGA
- matchupTurnovers
- matchup2FGA

A.2 Offense variables

- AVG DRIB PER TOUCH
- AVG OREB Distance
- AVG SEC PER TOUCH
- Avg Speed Off
- CATCH SHOOT FG2A
- CATCH SHOOT FG3A
- CFGA
- DIST MILES OFF
- DRIVE AST
- DRIVE FGA
- DRIVE FTA
- DRIVE PASS NO AST
- DRIVE TOV
- ELBOW TOUCH AST
- ELBOW TOUCH FGA
- ELBOW TOUCH FTA
- ELBOW TOUCH PASS NO AST
- ELBOW TOUCH TOV
- FT AST
- OFF BOXOUTS
- OFF LOOSE BALLS RECOVERED
- OREB CHANCES
- PAINT TOUCH AST
- PAINT TOUCH FGA
- PAINT TOUCH FTA
- PAINT TOUCH PASS NO AST
- PAINT TOUCH TOV
- PASSES RECEIVED
- PERIMETER AST
- PERIMETER FTA
- PERIMETER TOUCHES
- POST TOUCH AST
- POST TOUCH FGA
- POST TOUCH FTA
- POST TOUCH PASS NO AST
- POST TOUCH TOV
- POTENTIAL AST
- PULL UP FG2A
- PULL UP FG3A
- SCREEN ASSISTS
- SECONDARY AST
- UFGA

Appendix B

”Missclassified” centers

Table B.1: False positive centers on offense.

Name	Label	Center Prob.
Bobby Portis	F	0.7774
Brandon Clarke	F	0.5192
Bruce Brown	G-F	0.5765
Derrick Favors	F	0.9979
Jarred Vanderbilt	F	0.9725
Kevon Looney	F	0.855
Marvin Bagley III	F	0.7416
Precious Achiuwa	F	0.9978
Richaun Holmes	F	0.9981
Serge Ibaka	F	0.9771
Taj Gibson	F	0.9918
Thaddeus Young	F	0.9966
Xavier Tillman Sr.	F	0.9165

Table B.2: False negative centers on offense.

Name	Label	Center Prob.
Al Horford	C-F	0.2412
Alekesej Pokusevski	C	0.0
Brook Lopez	C	0.4571
Dean Wade	F-C	0.0053
Jalen McDaniels	F-C	0.0009
Julius Randle	F-C	0.0003
Kelly Olynyk	F-C	0.0385
Kevin Love	F-C	0.068
Lauri Markkanen	F-C	0.0136
Mike Muscala	F-C	0.0153
Myles Turner	C-F	0.3615

Table B.3: False positive centers on defense.

Name	Label	Center Prob.
Blake Griffin	F	0.6853
Bobby Portis	F	0.9995
Christian Wood	F	0.9999
Darius Bazley	F-G	0.63
Derrick Favors	F	1.0
Draymond Green	F	0.9324
Eric Paschall	F	0.9257
Giannis Antetokounmpo	F	0.7901
Isaiah Roby	F	0.8878
Jarred Vanderbilt	F	0.9351
Juan Toscano-Anderson	F	0.5787
Kevon Looney	F	1.0
Marvin Bagley III	F	0.8526
Maxi Kleber	F	0.8928
Nicolo Melli	F	0.7364
Oshae Brissett	F-G	0.967
P.J. Washington	F	0.9618
Precious Achiuwa	F	1.0
Richaun Holmes	F	1.0
Serge Ibaka	F	1.0
Taj Gibson	F	0.9982
Xavier Tillman Sr.	F	0.9997
Yuta Watanabe	G-F	0.5727

Table B.4: False negative centers on defense.

Name	Label	Center Prob.
Alekesej Pokusevski	C	0.102
Dean Wade	F-C	0.2493
JaMychal Green	F-C	0.0448
Jalen McDaniels	F-C	0.0861
Larry Nance Jr.	F-C	0.0195

Appendix C

Center classifications

Table C.1: Centers on offense.

Name	Label	Center Prob.	Name	Label	Center Prob.	Name	Label	Center Prob.
Alex Len	C	0.995	Goga Bitadze	C-F	0.981	Mo Bamba	C	0.9326
Andre Drummond	C	0.9998	Gorgui Dieng	C	0.9033	Montrezl Harrell	F-C	0.9987
Anthony Davis	F-C	0.9597	Hassan Whiteside	C	0.9997	Moritz Wagner	F-C	0.6372
Aron Baynes	C-F	0.9326	Isaiah Hartenstein	C-F	0.9997	Moses Brown	C	0.9922
Bam Adebayo	C-F	0.9858	Isaiah Stewart	F-C	0.9938	Naz Reid	C-F	0.8236
Bismack Biyombo	C	0.9979	Ivica Zubac	C	0.9998	Nerlens Noel	C-F	0.9481
Bobby Portis	F	0.7774	JaMychal Green	F-C	0.5038	Nic Claxton	F-C	0.9834
Brandon Clarke	F	0.5192	JaVale McGee	C-F	0.9983	Nikola Jokic	C	0.995
Bruce Brown	G-F	0.5765	Jakob Poeltl	C	0.9971	Nikola Vucevic	C	0.9822
Chimezie Metu	F-C	0.7262	James Wiseman	C	0.9815	Onyeka Okongwu	F-C	0.998
Chris Boucher	F-C	0.5028	Jarred Vanderbilt	F	0.9725	Precious Achiuwa	F	0.9978
Clint Capela	C	0.9988	Jarrett Allen	C	0.9995	Richaun Holmes	F	0.9981
Cody Zeller	F-C	0.9898	Jaxson Hayes	C-F	0.9474	Robert Williams III	C-F	0.9929
Damian Jones	C	0.9945	Joel Embiid	C-F	0.9855	Robin Lopez	C	0.9971
Daniel Gafford	F-C	0.9945	John Collins	F-C	0.8139	Rudy Gobert	C	0.9993
Daniel Theis	F-C	0.7629	Jonas Valanciunas	C	0.9997	Serge Ibaka	F	0.9771
Dario Saric	F-C	0.9687	Jusuf Nurkic	C	0.9974	Steven Adams	C	0.9989
DeAndre Jordan	C	0.9963	Karl-Anthony Towns	C-F	0.866	Taj Gibson	F	0.9918
DeMarcus Cousins	C	0.5279	Kevon Looney	F	0.855	Thaddeus Young	F	0.9966
Deandre Ayton	C	0.9996	Khem Birch	C	0.9864	Tony Bradley	C-F	0.9951
Derrick Favors	F	0.9979	Kristaps Porzingis	F-C	0.8957	Tristan Thompson	C-F	0.9952
Domantas Sabonis	F-C	0.9919	LaMarcus Aldridge	C-F	0.8844	Wendell Carter Jr.	C-F	0.9993
Drew Eubanks	F-C	0.9924	Larry Nance Jr.	F-C	0.9448	Willie Cauley-Stein	C	0.9733
Dwight Howard	C-F	0.9986	Marc Gasol	C	0.7239	Willy Hernangomez	C-F	0.9977
Dwight Powell	F-C	0.9823	Marvin Bagley III	F	0.7416	Xavier Tillman Sr.	F	0.9165
Enes Freedom	C	0.9984	Mason Plumlee	F-C	0.991			
Frank Kaminsky	F-C	0.5767	Mitchell Robinson	C-F	0.9541			

Table C.2: Centers on defense.

Name	Label	Center Prob.
Al Horford	C-F	0.9973
Alex Len	C	1.0
Andre Drummond	C	0.9999
Anthony Davis	F-C	0.6205
Aron Baynes	C-F	0.9998
Bam Adebayo	C-F	0.9932
Bismack Biyombo	C	0.9999
Blake Griffin	F	0.6853
Bobby Portis	F	0.9995
Brook Lopez	C	1.0
Chimezie Metu	F-C	0.9981
Chris Boucher	F-C	1.0
Christian Wood	F	0.9999
Clint Capela	C	1.0
Cody Zeller	F-C	1.0
Damian Jones	C	0.9993
Daniel Gafford	F-C	1.0
Daniel Theis	F-C	0.9997
Dario Saric	F-C	0.9998
Darius Bazley	F-G	0.63
DeAndre Jordan	C	1.0
DeMarcus Cousins	C	1.0
Deandre Ayton	C	1.0
Derrick Favors	F	1.0
Domantas Sabonis	F-C	0.9999
Draymond Green	F	0.9324
Drew Eubanks	F-C	1.0
Dwight Howard	C-F	1.0
Dwight Powell	F-C	0.9991
Enes Freedom	C	1.0
Eric Paschall	F	0.9257
Frank Kaminsky	F-C	1.0

Name	Label	Center Prob.
Giannis Antetokounmpo	F	0.7901
Goga Bitadze	C-F	1.0
Gorgui Dieng	C	0.9997
Hassan Whiteside	C	1.0
Isaiah Hartenstein	C-F	1.0
Isaiah Roby	F	0.8878
Isaiah Stewart	F-C	1.0
Ivica Zubac	C	1.0
JaVale McGee	C-F	1.0
Jakob Poeltl	C	1.0
James Wiseman	C	1.0
Jarred Vanderbilt	F	0.9351
Jarrett Allen	C	1.0
Jaxson Hayes	C-F	0.9682
Joel Embiid	C-F	1.0
John Collins	F-C	0.9987
Jonas Valanciunas	C	1.0
Juan Toscano-Anderson	F	0.5787
Julius Randle	F-C	0.9336
Jusuf Nurkic	C	1.0
Karl-Anthony Towns	C-F	0.9997
Kelly Olynyk	F-C	0.9395
Kevin Love	F-C	0.9572
Kevon Looney	F	1.0
Khem Birch	C	0.914
Kristaps Porzingis	F-C	1.0
LaMarcus Aldridge	C-F	0.9993
Lauri Markkanen	F-C	0.8389
Marc Gasol	C	1.0
Marvin Bagley III	F	0.8526
Mason Plumlee	F-C	1.0
Maxi Kleber	F	0.8928

Name	Label	Center Prob.
Mike Muscala	F-C	1.0
Mitchell Robinson	C-F	0.9942
Mo Bamba	C	1.0
Montrezl Harrell	F-C	0.9993
Moritz Wagner	F-C	0.9997
Moses Brown	C	1.0
Myles Turner	C-F	1.0
Naz Reid	C-F	1.0
Nerlens Noel	C-F	1.0
Nic Claxton	F-C	0.9999
Nicolo Melli	F	0.7364
Nikola Jokic	C	1.0
Nikola Vucevic	C	1.0
Onyeka Okongwu	F-C	0.9998
Oshae Brissett	F-G	0.967
P.J. Washington	F	0.9618
Precious Achiuwa	F	1.0
Richaun Holmes	F	1.0
Robert Williams III	C-F	0.9987
Robin Lopez	C	1.0
Rudy Gobert	C	1.0
Serge Ibaka	F	1.0
Steven Adams	C	0.9963
Taj Gibson	F	0.9982
Tony Bradley	C-F	1.0
Tristan Thompson	C-F	0.9989
Wendell Carter Jr.	C-F	1.0
Willie Cauley-Stein	C	1.0
Willy Hernangomez	C-F	1.0
Xavier Tillman Sr.	F	0.9997
Yuta Watanabe	G-F	0.5727

Appendix D

Non-center classifications

Table D.1: Non-centers on offense (1).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Aaron Gordon	F	0.0	0.9956	0.0	0.0044	0.0
Aaron Holiday	G	0.9296	0.0	0.0667	0.0037	0.0
Aaron Nesmith	G-F	0.0344	0.0	0.0472	0.9183	0.0
Al Horford	C-F	0.0	0.9994	0.0	0.0006	0.0
Alec Burks	G	0.8604	0.0	0.1336	0.006	0.0
Alekesej Pokusevski	C	0.8906	0.0	0.0444	0.065	0.0
Alex Caruso	G	0.0251	0.0002	0.9026	0.0722	0.0
Andre Iguodala	G-F	0.0	0.0007	0.0615	0.9378	0.0
Andrew Wiggins	F	0.0	0.0039	0.7374	0.2588	0.0
Anfernee Simons	G	0.9091	0.0	0.0333	0.0575	0.0
Anthony Edwards	G	0.159	0.0	0.84	0.0009	0.0
Armoni Brooks	G	0.6087	0.0	0.0148	0.3765	0.0
Austin Rivers	G	0.9466	0.0	0.0363	0.0171	0.0
Avery Bradley	G	0.645	0.0	0.0395	0.3155	0.0
Ben McLemore	G	0.4168	0.0	0.0203	0.563	0.0
Ben Simmons	G-F	0.0	1.0	0.0	0.0	0.0
Blake Griffin	F	0.0	0.9988	0.0	0.0012	0.0
Bogdan Bogdanovic	G	0.6469	0.0	0.3221	0.031	0.0
Bojan Bogdanovic	F	0.0012	0.0004	0.6728	0.3256	0.0
Brad Wanamaker	G	0.815	0.0	0.1846	0.0003	0.0
Bradley Beal	G	0.0	0.005	0.995	0.0	0.0
Brandon Goodwin	G	0.8899	0.0	0.1101	0.0	0.0
Brandon Ingram	F	0.0225	0.0001	0.9774	0.0	0.0
Brook Lopez	C	0.0	0.8817	0.0	0.1182	0.0
Bryn Forbes	G	0.4711	0.0	0.0082	0.5207	0.0
Buddy Hield	G	0.9227	0.0	0.0288	0.0485	0.0
CJ McCollum	G	0.8042	0.0	0.1958	0.0	0.0
Caleb Martin	F	0.041	0.0	0.3664	0.5926	0.0
Cam Reddish	F-G	0.0151	0.0	0.1986	0.7863	0.0
Cameron Johnson	F	0.2441	0.0	0.0386	0.7172	0.0
Cameron Payne	G	0.8204	0.0	0.1796	0.0	0.0
Caris LeVert	G	0.09	0.0	0.9099	0.0	0.0
Carmelo Anthony	F	0.0	0.1072	0.0105	0.8824	0.0
Cedi Osman	F	0.773	0.0	0.1967	0.0303	0.0
Chasson Randle	G	0.9305	0.0	0.0655	0.004	0.0
Chris Paul	G	0.4006	0.0	0.5994	0.0	0.0

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Christian Wood	F	0.0	0.9999	0.0	0.0001	0.0
Chuma Okeke	F	0.0	0.0009	0.0541	0.9451	0.0
Coby White	G	0.7815	0.0	0.2185	0.0	0.0
Cody Martin	F	0.0	0.0004	0.1865	0.8131	0.0
Cole Anthony	G	0.6183	0.0	0.3817	0.0	0.0
Collin Sexton	G	0.0	0.0024	0.9976	0.0	0.0
Cory Joseph	G	0.3979	0.0	0.6016	0.0005	0.0
D'Angelo Russell	G	0.0154	0.0001	0.9845	0.0	0.0
D.J. Augustin	G	0.88	0.0	0.12	0.0	0.0
DaQuan Jeffries	G-F	0.0278	0.0	0.044	0.9282	0.0
Damian Lillard	G	0.2425	0.0	0.7575	0.0	0.0
Damion Lee	G-F	0.6257	0.0	0.0118	0.3625	0.0
Damyean Dotson	G	0.9467	0.0	0.0364	0.0169	0.0
Danilo Gallinari	F	0.0	0.1577	0.005	0.8373	0.0
Danny Green	G	0.467	0.0	0.0142	0.5188	0.0
Danuel House Jr.	F-G	0.3192	0.0	0.1522	0.5286	0.0
Darius Bazley	F-G	0.0174	0.0001	0.5531	0.4294	0.0
Darius Garland	G	0.5777	0.0	0.4223	0.0	0.0
David Nwaba	G-F	0.0	0.0002	0.0422	0.9575	0.0
Davis Bertans	F	0.472	0.0	0.0077	0.5203	0.0
De'Aaron Fox	G	0.2152	0.0	0.7848	0.0	0.0
De'Andre Hunter	F-G	0.0	0.0198	0.1543	0.8259	0.0
De'Anthony Melton	G	0.2983	0.0	0.6442	0.0575	0.0
DeAndre' Bembry	G-F	0.0	0.002	0.3302	0.6678	0.0
DeMar DeRozan	G-F	0.0	0.0222	0.9778	0.0	0.0
Dean Wade	F-C	0.0	0.0001	0.0477	0.9522	0.0
Dejounte Murray	G	0.7145	0.0	0.2855	0.0	0.0
Delon Wright	G	0.2953	0.0	0.7046	0.0001	0.0
Deni Avdija	F	0.0271	0.0	0.0436	0.9293	0.0
Dennis Schroder	G	0.3095	0.0	0.6905	0.0	0.0
Denzel Valentine	G	0.8773	0.0	0.0428	0.0799	0.0
Derrick Jones Jr.	F	0.0	0.0008	0.0354	0.9638	0.0
Derrick Rose	G	0.0981	0.0	0.9019	0.0	0.0
Derrick White	G	0.7524	0.0	0.2464	0.0011	0.0
Desmond Bane	G	0.7307	0.0	0.0212	0.2481	0.0
Devin Booker	G	0.0	0.9244	0.0756	0.0	0.0

Table D.2: Non-centers on offense (2).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Devin Vassell	G-F	0.3218	0.0	0.0268	0.6514	0.0
Devonte' Graham	G	0.8353	0.0	0.1647	0.0	0.0
Dillon Brooks	G-F	0.0944	0.0001	0.8524	0.0531	0.0
Donovan Mitchell	G	0.2583	0.0	0.7417	0.0	0.0
Donte DiVincenzo	G	0.001	0.0002	0.2756	0.7232	0.0
Dorian Finney-Smith	F	0.002	0.0	0.05	0.948	0.0
Doug McDermott	F	0.0344	0.0	0.0737	0.8919	0.0
Draymond Green	F	0.0	0.9984	0.0	0.0016	0.0
Duncan Robinson	F	0.5655	0.0	0.011	0.4235	0.0
Dwayne Bacon	G-F	0.8709	0.0	0.0878	0.0413	0.0
Dylan Windler	G-F	0.0036	0.0	0.0712	0.9252	0.0
Edmond Sumner	G	0.5344	0.0	0.0827	0.3829	0.0
Elfrid Payton	G	0.0	0.0017	0.9982	0.0	0.0
Eric Bledsoe	G	0.9322	0.0	0.0674	0.0005	0.0
Eric Gordon	G	0.9124	0.0	0.0873	0.0003	0.0
Eric Paschall	F	0.0	0.9951	0.0	0.0049	0.0
Evan Fournier	G-F	0.9234	0.0	0.0725	0.004	0.0
Facundo Campazzo	G	0.9462	0.0	0.0516	0.0022	0.0
Frank Jackson	G	0.2249	0.0	0.2682	0.5069	0.0
Fred VanVleet	G	0.7168	0.0	0.2832	0.0	0.0
Furkan Korkmaz	G-F	0.9136	0.0	0.029	0.0573	0.0
Gabe Vincent	G	0.9384	0.0	0.0378	0.0238	0.0
Garrett Temple	G-F	0.3487	0.0	0.0412	0.6101	0.0
Garrison Mathews	G	0.458	0.0	0.0067	0.5354	0.0
Gary Clark	F	0.0002	0.0001	0.0552	0.9446	0.0
Gary Harris	G	0.0	0.0013	0.2195	0.7792	0.0
Gary Trent Jr.	G-F	0.829	0.0	0.0214	0.1495	0.0
George Hill	G	0.5211	0.0	0.4669	0.012	0.0
Georges Niang	F	0.5367	0.0	0.0155	0.4478	0.0
Giannis Antetokounmpo	F	0.0	0.0	0.0	0.0	1.0
Goran Dragic	G	0.0087	0.0002	0.9912	0.0	0.0
Gordon Hayward	F	0.0	0.0076	0.9792	0.0132	0.0
Grant Williams	F	0.0	0.0223	0.0052	0.9725	0.0
Grayson Allen	G	0.8901	0.0	0.0249	0.085	0.0
Hamidou Diallo	G	0.0003	0.0009	0.8781	0.1207	0.0
Harrison Barnes	F	0.0	0.3087	0.0049	0.6865	0.0

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Immanuel Quickley	G	0.881	0.0	0.119	0.0	0.0
Isaac Okoro	F-G	0.0023	0.0001	0.1586	0.839	0.0
Isaiah Roby	F	0.0	0.992	0.0	0.008	0.0
Ish Smith	G	0.8196	0.0	0.1804	0.0	0.0
JJ Redick	G	0.8206	0.0	0.0278	0.1516	0.0
Ja Morant	G	0.0	0.0011	0.9989	0.0	0.0
Jaden McDaniels	F	0.0032	0.0	0.0472	0.9496	0.0
Jae Crowder	F	0.0424	0.0	0.0469	0.9107	0.0
Jae'Sean Tate	F	0.0	0.8245	0.0	0.1755	0.0
Jake Layman	F	0.0089	0.0	0.0518	0.9394	0.0
Jalen Brunson	G	0.8673	0.0	0.1327	0.0	0.0
Jalen McDaniels	F-C	0.0	0.0003	0.0984	0.9013	0.0
Jamal Murray	G	0.004	0.0003	0.9957	0.0	0.0
James Ennis III	F	0.0089	0.0	0.0832	0.9079	0.0
James Harden	G	0.0202	0.0	0.9798	0.0	0.0
James Johnson	F	0.0	0.0012	0.1379	0.8609	0.0
Jarrett Culver	G-F	0.0	0.0018	0.1891	0.8091	0.0
Jaylen Brown	G-F	0.0	0.0276	0.9662	0.0062	0.0
Jaylen Nowell	G	0.7053	0.0	0.2335	0.0612	0.0
Jayson Tatum	F-G	0.0	0.7718	0.2282	0.0	0.0
Jeff Green	F	0.0	0.0041	0.0149	0.9809	0.0
Jeff Teague	G	0.7611	0.0	0.2389	0.0	0.0
Jerami Grant	F	0.0	0.0036	0.9914	0.005	0.0
Jeremy Lamb	G-F	0.6556	0.0	0.2927	0.0517	0.0
Jevon Carter	G	0.8447	0.0	0.0214	0.1339	0.0
Jimmy Butler	F	0.0	1.0	0.0	0.0	0.0
Joe Harris	G-F	0.3295	0.0	0.0452	0.6253	0.0
Joe Ingles	F-G	0.884	0.0	0.116	0.0	0.0
John Konchar	G	0.0	0.0002	0.045	0.9547	0.0
John Wall	G	0.3955	0.0	0.6045	0.0	0.0
Jordan Clarkson	G	0.5584	0.0	0.4414	0.0002	0.0
Jordan McLaughlin	G	0.8211	0.0	0.1789	0.0	0.0
Jordan Poole	G	0.9283	0.0	0.0707	0.001	0.0
Josh Hart	G	0.2081	0.0	0.073	0.7189	0.0
Josh Jackson	G-F	0.0004	0.0009	0.9218	0.077	0.0
Josh Okogie	G	0.0	0.0015	0.0234	0.9751	0.0

Table D.3: Non-centers on offense (3).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Josh Richardson	G	0.7618	0.0	0.1907	0.0475	0.0
Jrue Holiday	G	0.0	0.0251	0.974	0.0009	0.0
Juan Toscano-Anderson	F	0.0	0.0151	0.0203	0.9646	0.0
Juancho Hernangomez	F	0.0	0.0016	0.0215	0.9769	0.0
Julius Randle	F-C	0.0	1.0	0.0	0.0	0.0
Justin Holiday	F-G	0.4647	0.0	0.0135	0.5217	0.0
Justin Jackson	F	0.8415	0.0	0.1054	0.053	0.0
Justise Winslow	F-G	0.0	0.0179	0.3681	0.614	0.0
Kawhi Leonard	F	0.0	1.0	0.0	0.0	0.0
Keldon Johnson	F-G	0.0007	0.0002	0.2975	0.7016	0.0
Kelly Olynyk	F-C	0.0	0.8427	0.0	0.1573	0.0
Kelly Oubre Jr.	F-G	0.0	0.0012	0.0692	0.9296	0.0
Kemba Walker	G	0.8192	0.0	0.1808	0.0	0.0
Kendrick Nunn	G	0.821	0.0	0.1487	0.0302	0.0
Kenrich Williams	G-F	0.0	0.0012	0.3176	0.6812	0.0
Kent Bazemore	G-F	0.0	0.0002	0.1199	0.8799	0.0
Kentavious Caldwell-Pope	G	0.5957	0.0	0.0225	0.3818	0.0
Kenyon Martin Jr.	F	0.0	0.0216	0.0052	0.9732	0.0
Kevin Durant	F	0.0	1.0	0.0	0.0	0.0
Kevin Huerter	G-F	0.8821	0.0	0.0542	0.0637	0.0
Kevin Love	F-C	0.0	0.991	0.0	0.009	0.0
Kevin Porter Jr.	G-F	0.307	0.0	0.693	0.0	0.0
Khris Middleton	F	0.0	0.01	0.9888	0.0013	0.0
Killian Hayes	G	0.7832	0.0	0.2168	0.0	0.0
Kira Lewis Jr.	G	0.8985	0.0	0.1001	0.0014	0.0
Kyle Anderson	F-G	0.0003	0.0007	0.6641	0.3349	0.0
Kyle Kuzma	F	0.0001	0.0002	0.1424	0.8574	0.0
Kyle Lowry	G	0.4878	0.0	0.5122	0.0	0.0
Kyrie Irving	G	0.001	0.0004	0.9986	0.0	0.0
LaMelo Ball	G	0.0228	0.0001	0.9772	0.0	0.0
Lamar Stevens	F	0.0	0.0163	0.0139	0.9698	0.0
Landry Shamet	G	0.6985	0.0	0.0143	0.2871	0.0
Lauri Markkanen	F-C	0.0	0.004	0.014	0.982	0.0
LeBron James	F	0.0	1.0	0.0	0.0	0.0
Lonnie Walker IV	G-F	0.9301	0.0	0.0308	0.0391	0.0
Lonzo Ball	G	0.8978	0.0	0.1002	0.002	0.0
Lou Williams	G	0.6909	0.0	0.3091	0.0	0.0
Luguentz Dort	G	0.8035	0.0	0.1633	0.0332	0.0
Luka Doncic	F-G	0.0002	0.0003	0.9995	0.0	0.0
Luke Kennard	G	0.6782	0.0	0.1004	0.2214	0.0
Malachi Flynn	G	0.9087	0.0	0.0912	0.0001	0.0
Malcolm Brogdon	G	0.388	0.0	0.612	0.0	0.0
Malik Beasley	G	0.7346	0.0	0.1442	0.1212	0.0
Malik Monk	G	0.9387	0.0	0.0603	0.001	0.0
Marcus Morris Sr.	F	0.0	0.022	0.0119	0.9661	0.0
Marcus Smart	G	0.0002	0.0012	0.9618	0.0369	0.0
Markieff Morris	F	0.0	0.3024	0.0003	0.6974	0.0
Matisse Thybulle	G-F	0.3029	0.0	0.02	0.677	0.0
Maurice Harkless	F-G	0.0027	0.0	0.0477	0.9496	0.0
Max Strus	G-F	0.4046	0.0	0.0056	0.5898	0.0
Maxi Kleber	F	0.0	0.0002	0.0492	0.9506	0.0
Michael Carter-Williams	G	0.0032	0.0003	0.9965	0.0	0.0
Michael Porter Jr.	F	0.0	0.0788	0.0017	0.9195	0.0
Mikal Bridges	F	0.0005	0.0	0.0731	0.9264	0.0
Mike Conley	G	0.6672	0.0	0.3328	0.0	0.0
Mike Muscala	F-C	0.0	0.0074	0.0103	0.9822	0.0
Mike Scott	F	0.2891	0.0	0.0208	0.6901	0.0
Miles Bridges	F	0.0	0.0721	0.0057	0.9221	0.0
Miye Oni	G-F	0.0092	0.0	0.0472	0.9436	0.0
Monte Morris	G	0.7315	0.0	0.2669	0.0016	0.0
Mychal Mulder	G	0.4762	0.0	0.0084	0.5154	0.0
Myles Turner	C-F	0.0	0.7785	0.0	0.2215	0.0
Naji Marshall	F	0.5209	0.0	0.2631	0.2159	0.0
Nassir Little	F-G	0.0126	0.0	0.052	0.9354	0.0
Nemanja Bjelica	F	0.0	0.1337	0.022	0.8443	0.0
Nickeil Alexander-Walker	G	0.9312	0.0	0.0677	0.0011	0.0
Nicolas Batum	G-F	0.0	0.0003	0.0341	0.9655	0.0
Nicolo Melli	F	0.0	0.029	0.0042	0.9668	0.0
Norman Powell	G	0.806	0.0	0.1538	0.0402	0.0
OG Anunoby	F	0.0	0.0309	0.0099	0.9592	0.0
Obi Toppin	F	0.0	0.0033	0.0157	0.981	0.0
Oshae Brissett	F-G	0.0	0.0252	0.0064	0.9684	0.0

Table D.4: Non-centers on offense (4).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Otto Porter Jr.	F	0.0	0.0005	0.0659	0.9336	0.0
P.J. Dozier	G-F	0.0002	0.0003	0.3239	0.6756	0.0
P.J. Tucker	F	0.0	0.0003	0.05	0.9497	0.0
P.J. Washington	F	0.0	0.9992	0.0	0.0008	0.0
Pascal Siakam	F	0.0	1.0	0.0	0.0	0.0
Pat Connaughton	G	0.006	0.0	0.0463	0.9477	0.0
Patrick Beverley	G	0.2786	0.0	0.4841	0.2373	0.0
Patrick Patterson	F	0.0	0.0001	0.0498	0.95	0.0
Patrick Williams	F	0.0005	0.0	0.0769	0.9226	0.0
Patty Mills	G	0.9264	0.0	0.0299	0.0437	0.0
Paul George	F	0.0001	0.0012	0.9988	0.0	0.0
Paul Millsap	F	0.0	0.9945	0.0	0.0055	0.0
Payton Pritchard	G	0.9448	0.0	0.0467	0.0085	0.0
R.J. Hampton	G	0.1128	0.0001	0.8309	0.0562	0.0
RJ Barrett	F-G	0.0033	0.0004	0.975	0.0213	0.0
Rajon Rondo	G	0.1224	0.0	0.8776	0.0	0.0
Raul Neto	G	0.9087	0.0	0.0781	0.0132	0.0
Reggie Bullock	G-F	0.5069	0.0	0.0097	0.4834	0.0
Reggie Jackson	G	0.8838	0.0	0.1154	0.0008	0.0
Ricky Rubio	G	0.7579	0.0	0.2421	0.0	0.0
Robert Covington	F	0.0046	0.0	0.0468	0.9486	0.0
Rodney Hood	G-F	0.1723	0.0	0.1087	0.719	0.0
Royce O'Neale	F	0.0673	0.0	0.0559	0.8768	0.0
Rudy Gay	F-G	0.0	0.0028	0.1816	0.8156	0.0
Rui Hachimura	F	0.0	0.9519	0.0	0.0481	0.0
Russell Westbrook	G	0.0	0.0097	0.9903	0.0	0.0
Saben Lee	G	0.061	0.0	0.9389	0.0	0.0
Saddiq Bey	F	0.0174	0.0	0.0961	0.8864	0.0
Sekou Doumbouya	F	0.0	0.1369	0.0009	0.8622	0.0
Semi Ojeleye	F	0.0064	0.0	0.0473	0.9463	0.0
Seth Curry	G	0.9322	0.0	0.0333	0.0345	0.0
Shai Gilgeous-Alexander	G-F	0.0869	0.0	0.9131	0.0	0.0
Shake Milton	G-F	0.6718	0.0	0.3282	0.0	0.0
Solomon Hill	F	0.2236	0.0	0.027	0.7494	0.0
Stanley Johnson	F-G	0.037	0.0	0.1048	0.8582	0.0
Stephen Curry	G	0.3372	0.0	0.6627	0.0	0.0

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Sterling Brown	G-F	0.0753	0.0	0.0627	0.862	0.0
Svi Mykhailiuk	G-F	0.9304	0.0	0.0439	0.0257	0.0
T.J. McConnell	G	0.1387	0.0	0.8613	0.0	0.0
Talen Horton-Tucker	G	0.4684	0.0	0.5316	0.0	0.0
Taurean Prince	F	0.7667	0.0	0.1289	0.1044	0.0
Terance Mann	G-F	0.0007	0.0002	0.2912	0.708	0.0
Terence Davis	G	0.8724	0.0	0.0607	0.0669	0.0
Terrence Ross	G-F	0.3044	0.0	0.6023	0.0933	0.0
Terry Rozier	G	0.8205	0.0	0.1794	0.0001	0.0
Thanasis Antetokounmpo	F	0.0	0.935	0.0	0.065	0.0
Theo Maledon	G	0.9087	0.0	0.0912	0.0001	0.0
Tim Hardaway Jr.	G-F	0.8614	0.0	0.0339	0.1047	0.0
Timothe Luwawu-Cabarrot	G-F	0.3083	0.0	0.0344	0.6573	0.0
Tobias Harris	F	0.0	0.9112	0.0702	0.0186	0.0
Tomas Satoransky	G	0.0041	0.0003	0.9953	0.0003	0.0
Tony Snell	G	0.4562	0.0	0.0142	0.5296	0.0
Torrey Craig	F	0.0	0.0027	0.0176	0.9797	0.0
Trae Young	G	0.0323	0.0	0.9677	0.0	0.0
Trevor Ariza	F	0.0	0.0002	0.045	0.9547	0.0
Trey Burke	G	0.9474	0.0	0.0461	0.0065	0.0
Troy Brown Jr.	G-F	0.0003	0.0001	0.0786	0.921	0.0
Ty Jerome	G-F	0.9312	0.0	0.0685	0.0004	0.0
Tyler Herro	G	0.9073	0.0	0.0926	0.0002	0.0
Tyler Johnson	G	0.6044	0.0	0.0146	0.3811	0.0
Tyrese Haliburton	G	0.9109	0.0	0.089	0.0001	0.0
Tyrese Maxey	G	0.8577	0.0	0.1423	0.0	0.0
Tyus Jones	G	0.8686	0.0	0.1314	0.0	0.0
Victor Oladipo	G	0.5815	0.0	0.4185	0.0	0.0
Wayne Ellington	G	0.4568	0.0	0.0113	0.5318	0.0
Wes Iwundu	F	0.089	0.0	0.0426	0.8684	0.0
Wesley Matthews	G	0.0997	0.0	0.0393	0.8611	0.0
Will Barton	G	0.0048	0.0003	0.8428	0.1521	0.0
Yuta Watanabe	G-F	0.0011	0.0	0.0533	0.9456	0.0
Zach LaVine	G-F	0.0183	0.0001	0.9816	0.0	0.0
Zion Williamson	F	0.0	0.0	0.0	0.0	1.0

Table D.5: Non-centers on defense (1).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Aaron Gordon	F	0.7057	0.294	0.0	0.0003
Aaron Holiday	G	0.072	0.0156	0.0131	0.8994
Aaron Nesmith	G-F	0.3201	0.6761	0.0	0.0038
Alec Burks	G	0.6134	0.3352	0.0	0.0513
Alekesej Pokusevski	C	0.0151	0.9849	0.0	0.0
Alex Caruso	G	0.0	0.0001	0.9821	0.0178
Andre Iguodala	G-F	0.0429	0.1237	0.0007	0.8328
Andrew Wiggins	F	0.2045	0.7886	0.0	0.0069
Anfernee Simons	G	0.6036	0.3928	0.0	0.0036
Anthony Edwards	G	0.9012	0.0578	0.0	0.041
Armoni Brooks	G	0.9551	0.0392	0.0	0.0057
Austin Rivers	G	0.4991	0.0054	0.0001	0.4955
Avery Bradley	G	0.4088	0.083	0.0011	0.5072
Ben McLemore	G	0.8332	0.1233	0.0	0.0435
Ben Simmons	G-F	0.1564	0.0184	0.005	0.8203
Bogdan Bogdanovic	G	0.3389	0.582	0.0	0.0791
Bojan Bogdanovic	F	0.8047	0.195	0.0	0.0003
Brad Wanamaker	G	0.0611	0.0067	0.0198	0.9125
Bradley Beal	G	0.9187	0.0129	0.0	0.0685
Brandon Clarke	F	0.0159	0.9831	0.0	0.001
Brandon Goodwin	G	0.8002	0.1522	0.0	0.0475
Brandon Ingram	F	0.8218	0.1779	0.0	0.0003
Bruce Brown	G-F	0.0657	0.8048	0.0	0.1295
Bryn Forbes	G	0.9691	0.0295	0.0	0.0014
Buddy Hield	G	0.4341	0.5471	0.0	0.0188
CJ McCollum	G	0.8106	0.0624	0.0001	0.127
Caleb Martin	F	0.0012	0.0526	0.0007	0.9454
Cam Reddish	F-G	0.0145	0.007	0.0981	0.8804
Cameron Johnson	F	0.5496	0.3965	0.0	0.0539
Cameron Payne	G	0.5125	0.1883	0.0004	0.2988
Caris LeVert	G	0.3999	0.0353	0.0011	0.5637
Carmelo Anthony	F	0.7152	0.2847	0.0	0.0
Cedi Osman	F	0.9337	0.0243	0.0	0.042
Chasson Randle	G	0.6232	0.0246	0.0002	0.352

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Chris Paul	G	0.8223	0.0045	0.0	0.1732
Chuma Okeke	F	0.0743	0.3236	0.0001	0.6021
Coby White	G	0.788	0.1671	0.0	0.0449
Cody Martin	F	0.0136	0.174	0.0	0.8124
Cole Anthony	G	0.5374	0.3381	0.0001	0.1244
Collin Sexton	G	0.8974	0.0009	0.0	0.1018
Cory Joseph	G	0.1319	0.118	0.0015	0.7486
D'Angelo Russell	G	0.8031	0.024	0.0	0.1728
D.J. Augustin	G	0.9858	0.0014	0.0	0.0128
DaQuan Jeffries	G-F	0.5052	0.4263	0.0	0.0685
Damian Lillard	G	0.9756	0.0022	0.0	0.0221
Damion Lee	G-F	0.2896	0.6328	0.0	0.0776
Damyean Dotson	G	0.9424	0.0558	0.0	0.0018
Danilo Gallinari	F	0.049	0.951	0.0	0.0
Danny Green	G	0.0056	0.1604	0.0	0.834
Danuel House Jr.	F-G	0.6492	0.3334	0.0	0.0173
Darius Garland	G	0.1508	0.0023	0.0016	0.8454
David Nwaba	G-F	0.0091	0.0374	0.0054	0.948
Davis Bertans	F	0.5387	0.4608	0.0	0.0006
De'Aaron Fox	G	0.2079	0.0047	0.0013	0.7861
De'Andre Hunter	F-G	0.3528	0.6231	0.0	0.0241
De'Anthony Melton	G	0.0053	0.0215	0.0239	0.9493
DeAndre' Bembry	G-F	0.0965	0.0066	0.0092	0.8877
DeMar DeRozan	G-F	0.9945	0.0039	0.0	0.0015
Dean Wade	F-C	0.0275	0.9721	0.0	0.0003
Dejounte Murray	G	0.1915	0.6481	0.0	0.1604
Delon Wright	G	0.0103	0.0025	0.2594	0.7278
Deni Avdija	F	0.0137	0.9863	0.0	0.0
Dennis Schroder	G	0.0732	0.0099	0.0146	0.9023
Denzel Valentine	G	0.5438	0.4346	0.0	0.0216
Derrick Jones Jr.	F	0.0847	0.1153	0.0014	0.7987
Derrick Rose	G	0.4238	0.004	0.0001	0.5721
Derrick White	G	0.1603	0.4227	0.0001	0.417
Desmond Bane	G	0.6675	0.0623	0.0003	0.2699

Table D.6: Non-centers on defense (2).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Devin Booker	G	0.8621	0.1357	0.0	0.0022	Immanuel Quickley	G	0.7093	0.0735	0.0002	0.217
Devin Vassell	G-F	0.1552	0.3035	0.0002	0.5411	Isaac Okoro	F-G	0.1093	0.0789	0.0027	0.809
Devonte’ Graham	G	0.4109	0.0038	0.0001	0.5852	Ish Smith	G	0.17	0.0135	0.0041	0.8124
Dillon Brooks	G-F	0.0179	0.0059	0.094	0.8822	JJ Redick	G	0.98	0.0051	0.0	0.0149
Donovan Mitchell	G	0.9847	0.0018	0.0	0.0135	Ja Morant	G	0.9701	0.0233	0.0	0.0066
Donte DiVincenzo	G	0.3955	0.2873	0.0003	0.317	JaMychal Green	F-C	0.0996	0.9003	0.0	0.0001
Dorian Finney-Smith	F	0.3791	0.6083	0.0	0.0126	Jaden McDaniels	F	0.0231	0.9638	0.0	0.0131
Doug McDermott	F	0.6373	0.3618	0.0	0.0009	Jae Crowder	F	0.2627	0.735	0.0	0.0023
Duncan Robinson	F	0.4643	0.5194	0.0	0.0163	Jae’Sean Tate	F	0.0336	0.071	0.0022	0.8933
Dwayne Bacon	G-F	0.8381	0.1238	0.0	0.0381	Jake Layman	F	0.2132	0.0954	0.0019	0.6895
Dylan Windler	G-F	0.0889	0.7509	0.0	0.1602	Jalen Brunson	G	0.9286	0.062	0.0	0.0095
Edmond Sumner	G	0.1344	0.0037	0.0033	0.8586	Jalen McDaniels	F-C	0.0121	0.9526	0.0	0.0352
Elfrid Payton	G	0.0552	0.0355	0.0079	0.9014	Jamal Murray	G	0.15	0.0051	0.0032	0.8417
Eric Bledsoe	G	0.6678	0.0964	0.0003	0.2355	James Ennis III	F	0.2122	0.0814	0.0022	0.7042
Eric Gordon	G	0.998	0.0018	0.0	0.0002	James Harden	G	0.6247	0.3753	0.0	0.0
Evan Fournier	G-F	0.4695	0.1029	0.0008	0.4269	James Johnson	F	0.1542	0.8388	0.0	0.0069
Facundo Campazzo	G	0.0	0.0	1.0	0.0	Jarrett Culver	G-F	0.1693	0.1352	0.0013	0.6942
Frank Jackson	G	0.848	0.0204	0.0	0.1316	Jaylen Brown	G-F	0.7836	0.1066	0.0001	0.1097
Fred VanVleet	G	0.0184	0.005	0.1018	0.8748	Jaylen Nowell	G	0.5451	0.3122	0.0001	0.1425
Furkan Korkmaz	G-F	0.4116	0.0107	0.0004	0.5773	Jayson Tatum	F-G	0.8539	0.1455	0.0	0.0006
Gabe Vincent	G	0.0	0.0	0.9919	0.008	Jeff Green	F	0.0936	0.9064	0.0	0.0
Garrett Temple	G-F	0.1312	0.0578	0.0037	0.8073	Jeff Teague	G	0.4759	0.0228	0.0005	0.5008
Garrison Mathews	G	0.3097	0.0217	0.0016	0.6671	Jerami Grant	F	0.7937	0.206	0.0	0.0003
Gary Clark	F	0.0468	0.9531	0.0	0.0001	Jeremy Lamb	G-F	0.0721	0.8788	0.0	0.049
Gary Harris	G	0.0514	0.0006	0.0143	0.9336	Jevon Carter	G	0.075	0.0087	0.0143	0.902
Gary Trent Jr.	G-F	0.7829	0.0023	0.0	0.2148	Jimmy Butler	F	0.0204	0.0101	0.0522	0.9173
George Hill	G	0.193	0.0143	0.0034	0.7894	Joe Harris	G-F	0.6351	0.3636	0.0	0.0014
Georges Niang	F	0.5502	0.4479	0.0	0.0019	Joe Ingles	F-G	0.9745	0.0238	0.0	0.0016
Goran Dragic	G	0.9116	0.0288	0.0	0.0596	John Konchar	G	0.0197	0.2062	0.0	0.7741
Gordon Hayward	F	0.8036	0.1695	0.0	0.0268	John Wall	G	0.9624	0.0128	0.0	0.0248
Grant Williams	F	0.0466	0.9486	0.0	0.0048	Jordan Clarkson	G	0.8596	0.037	0.0	0.1034
Grayson Allen	G	0.7627	0.1118	0.0001	0.1254	Jordan McLaughlin	G	0.0031	0.0008	0.6418	0.3542
Hamidou Diallo	G	0.5714	0.3244	0.0001	0.1041	Jordan Poole	G	0.9801	0.0044	0.0	0.0154
Harrison Barnes	F	0.3252	0.6747	0.0	0.0001	Josh Hart	G	0.0109	0.989	0.0	0.0001

Table D.7: Non-centers on defense (3).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Josh Jackson	G-F	0.3144	0.1692	0.0008	0.5156
Josh Okogie	G	0.0002	0.0006	0.9053	0.0939
Josh Richardson	G	0.3234	0.0047	0.0003	0.6715
Jrue Holiday	G	0.0161	0.002	0.1837	0.7981
Juancho Hernangomez	F	0.0353	0.9647	0.0	0.0
Justin Holiday	F-G	0.3562	0.371	0.0002	0.2727
Justin Jackson	F	0.8182	0.1466	0.0	0.0352
Justise Winslow	F-G	0.0537	0.9463	0.0	0.0
Kawhi Leonard	F	0.7467	0.1502	0.0001	0.103
Keldon Johnson	F-G	0.1058	0.8942	0.0	0.0001
Kelly Oubre Jr.	F-G	0.2986	0.531	0.0	0.1704
Kemba Walker	G	0.0994	0.0198	0.0086	0.8721
Kendrick Nunn	G	0.443	0.1303	0.0007	0.426
Kenrich Williams	G-F	0.1943	0.2903	0.0003	0.5151
Kent Bazemore	G-F	0.0149	0.2173	0.0	0.7678
Kentavious Caldwell-Pope	G	0.5857	0.0155	0.0002	0.3987
Kenyon Martin Jr.	F	0.211	0.789	0.0	0.0
Kevin Durant	F	0.0524	0.9476	0.0	0.0
Kevin Huerter	G-F	0.2989	0.0359	0.0019	0.6633
Kevin Porter Jr.	G-F	0.9128	0.0636	0.0	0.0235
Khris Middleton	F	0.6959	0.2952	0.0	0.0088
Killian Hayes	G	0.0551	0.0213	0.0133	0.9103
Kira Lewis Jr.	G	0.0472	0.0028	0.0301	0.9199
Kyle Anderson	F-G	0.1286	0.8295	0.0	0.0418
Kyle Kuzma	F	0.2129	0.787	0.0	0.0001
Kyle Lowry	G	0.7299	0.1595	0.0001	0.1105
Kyrie Irving	G	0.5312	0.0142	0.0002	0.4544
LaMelo Ball	G	0.0687	0.0069	0.0165	0.9079
Lamar Stevens	F	0.0379	0.9573	0.0	0.0047
Landry Shamet	G	0.652	0.3282	0.0	0.0197
Larry Nance Jr.	F-C	0.0326	0.5074	0.0	0.4601
LeBron James	F	0.6759	0.324	0.0	0.0
Lonnie Walker IV	G-F	0.9038	0.0893	0.0	0.0068
Lonzo Ball	G	0.2388	0.1186	0.0014	0.6412

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Lou Williams	G	0.9725	0.0009	0.0	0.0266
Luguentz Dort	G	0.2922	0.0627	0.0019	0.6432
Luka Doncic	F-G	0.2139	0.7861	0.0	0.0
Luke Kennard	G	0.4814	0.5185	0.0	0.0001
Malachi Flynn	G	0.0274	0.0095	0.0432	0.9198
Malcolm Brogdon	G	0.8365	0.1282	0.0	0.0353
Malik Beasley	G	0.9601	0.024	0.0	0.0158
Malik Monk	G	0.9704	0.0282	0.0	0.0014
Marcus Morris Sr.	F	0.4021	0.5978	0.0	0.0
Marcus Smart	G	0.0398	0.0029	0.0416	0.9156
Markieff Morris	F	0.0741	0.9259	0.0	0.0
Matisse Thybulle	G-F	0.0	0.0	1.0	0.0
Maurice Harkless	F-G	0.0931	0.1562	0.0008	0.75
Max Strus	G-F	0.1283	0.0435	0.0047	0.8235
Michael Carter-Williams	G	0.2384	0.1078	0.0016	0.6522
Michael Porter Jr.	F	0.0608	0.9392	0.0	0.0
Mikal Bridges	F	0.2222	0.2236	0.0006	0.5537
Mike Conley	G	0.4115	0.0065	0.0002	0.5818
Mike Scott	F	0.1013	0.8985	0.0	0.0003
Miles Bridges	F	0.034	0.9658	0.0	0.0002
Miye Oni	G-F	0.0878	0.8972	0.0	0.015
Monte Morris	G	0.8879	0.023	0.0	0.0891
Mychal Mulder	G	0.3803	0.6122	0.0	0.0075
Naji Marshall	F	0.3517	0.6267	0.0	0.0217
Nassir Little	F-G	0.0073	0.9927	0.0	0.0
Nemanja Bjelica	F	0.0162	0.9837	0.0	0.0
Nickeil Alexander-Walker	G	0.0836	0.0604	0.0039	0.8522
Nicolas Batum	G-F	0.2417	0.5991	0.0	0.1591
Norman Powell	G	0.5623	0.0039	0.0	0.4338
OG Anunoby	F	0.1657	0.4025	0.0001	0.4317
Obi Toppin	F	0.0335	0.9665	0.0	0.0001
Otto Porter Jr.	F	0.5854	0.4143	0.0	0.0002
P.J. Dozier	G-F	0.5534	0.2335	0.0002	0.2129
P.J. Tucker	F	0.3482	0.5371	0.0	0.1146

Table D.8: Non-centers on defense (4).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Pascal Siakam	F	0.0834	0.9113	0.0	0.0053
Pat Connaughton	G	0.0375	0.9618	0.0	0.0007
Patrick Beverley	G	0.128	0.0068	0.0054	0.8598
Patrick Patterson	F	0.0345	0.9641	0.0	0.0013
Patrick Williams	F	0.3234	0.6725	0.0	0.0041
Patty Mills	G	0.6612	0.0011	0.0	0.3377
Paul George	F	0.3458	0.6472	0.0	0.007
Paul Millsap	F	0.0225	0.9342	0.0	0.0434
Payton Pritchard	G	0.7338	0.1333	0.0001	0.1327
R.J. Hampton	G	0.6023	0.3767	0.0	0.021
RJ Barrett	F-G	0.2865	0.6978	0.0	0.0156
Rajon Rondo	G	0.6297	0.0325	0.0002	0.3375
Raul Neto	G	0.0344	0.0047	0.049	0.9119
Reggie Bullock	G-F	0.5925	0.0515	0.0004	0.3556
Reggie Jackson	G	0.6571	0.3038	0.0	0.0391
Ricky Rubio	G	0.1074	0.0014	0.0025	0.8887
Robert Covington	F	0.0349	0.939	0.0	0.0261
Rodney Hood	G-F	0.9363	0.0509	0.0	0.0127
Royce O'Neale	F	0.0764	0.9229	0.0	0.0007
Rudy Gay	F-G	0.179	0.8188	0.0	0.0022
Rui Hachimura	F	0.2185	0.7813	0.0	0.0003
Russell Westbrook	G	0.7497	0.2502	0.0	0.0001
Saben Lee	G	0.0207	0.0123	0.0418	0.9252
Saddiq Bey	F	0.74	0.2549	0.0	0.0051
Sekou Doumbouya	F	0.1576	0.8419	0.0	0.0005
Semi Ojeleye	F	0.4517	0.5463	0.0	0.002
Seth Curry	G	0.9562	0.003	0.0	0.0408
Shai Gilgeous-Alexander	G-F	0.4935	0.504	0.0	0.0026
Shake Milton	G-F	0.9293	0.0392	0.0	0.0315
Solomon Hill	F	0.3503	0.4456	0.0001	0.204
Stanley Johnson	F-G	0.0221	0.1092	0.0005	0.8682
Stephen Curry	G	0.9892	0.0041	0.0	0.0067
Sterling Brown	G-F	0.0461	0.9482	0.0	0.0057
Svi Mykhailiuk	G-F	0.6023	0.0404	0.0003	0.357

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
T.J. McConnell	G	0.0	0.0	0.9999	0.0001
Talen Horton-Tucker	G	0.048	0.082	0.002	0.8681
Taurean Prince	F	0.5538	0.3699	0.0001	0.0763
Terance Mann	G-F	0.1966	0.7771	0.0	0.0263
Terence Davis	G	0.6721	0.0863	0.0003	0.2414
Terrence Ross	G-F	0.7536	0.021	0.0001	0.2253
Terry Rozier	G	0.0836	0.014	0.0115	0.8909
Thaddeus Young	F	0.018	0.9687	0.0	0.0133
Thanasis Antetokounmpo	F	0.0198	0.0693	0.0016	0.9093
Theo Maledon	G	0.8963	0.0724	0.0	0.0313
Tim Hardaway Jr.	G-F	0.8197	0.1651	0.0	0.0153
Timothe Luwawu-Cabarrot	G-F	0.304	0.6643	0.0	0.0317
Tobias Harris	F	0.1042	0.8958	0.0	0.0
Tomas Satoransky	G	0.5889	0.0539	0.0004	0.3568
Tony Snell	G	0.5552	0.4397	0.0	0.0051
Torrey Craig	F	0.0456	0.8886	0.0	0.0659
Trae Young	G	0.9888	0.0003	0.0	0.0109
Trevor Ariza	F	0.0244	0.178	0.0001	0.7975
Trey Burke	G	0.5238	0.0018	0.0	0.4744
Troy Brown Jr.	G-F	0.2473	0.7326	0.0	0.02
Ty Jerome	G-F	0.9569	0.0181	0.0	0.025
Tyler Herro	G	0.408	0.5548	0.0	0.0372
Tyler Johnson	G	0.7192	0.1443	0.0001	0.1363
Tyrese Haliburton	G	0.3525	0.005	0.0003	0.6423
Tyrese Maxey	G	0.6801	0.0629	0.0002	0.2568
Tyus Jones	G	0.2638	0.0047	0.0007	0.7308
Victor Oladipo	G	0.0203	0.0337	0.0086	0.9374
Wayne Ellington	G	0.9722	0.0139	0.0	0.0139
Wes Iwundu	F	0.407	0.4503	0.0001	0.1425
Wesley Matthews	G	0.1358	0.0139	0.006	0.8443
Will Barton	G	0.7728	0.2094	0.0	0.0177
Zach LaVine	G-F	0.6864	0.3112	0.0	0.0024
Zion Williamson	F	0.367	0.557	0.0	0.0759

Appendix E

Replacement player classifications

Table E.1: Replacement centers on offense.

Name	Label	Center Prob.
Alen Smailagic	F	0.5199
Alize Johnson	F	0.8368
Amida Brimah	C	0.9305
Anzejs Pasecniks	C-F	0.8642
Boban Marjanovic	C	0.9999
Bruno Fernando	F-C	0.9887
Chris Silva	F	0.7219
Cristiano Felicio	F-C	0.851
Daniel Oturu	C	0.9951
Devontae Cacok	F	0.7649
Dewayne Dedmon	C	0.9979
Donta Hall	C	0.9974
Ed Davis	C-F	0.9872
Freddie Gillespie	F	0.9774
Harry Giles III	F-C	0.8783
JaKarr Sampson	F	0.641
Jahlil Okafor	C-F	0.9856
Jalen Smith	F-C	0.8459
Jontay Porter	C-F	0.7384
Justin Patton	C	0.6701
Juwan Morgan	F	0.8614
Mamadi Diakite	F	0.8934
Marques Bolden	C	0.9272
Nathan Knight	F-C	0.5402
Nick Richards	C	0.9371
Norvel Pelle	C	0.9007
Paul Reed	F	0.6969
Reggie Perry	F-C	0.9482
Rondae Hollis-Jefferson	F	0.6854
Tacko Fall	C	0.9901
Thomas Bryant	C-F	0.7566
Thon Maker	F-C	0.7815
Tyler Cook	F	0.9487
Udoka Azubuike	C-F	0.9139
Vernon Carey Jr.	F-C	0.6026
Vincent Poirier	C-F	0.8907
Will Magnay	C	0.9139

Table E.2: Replacement centers on defense.

Name	Label	Center Prob.
Alen Smailagic	F	0.5279
Alize Johnson	F	0.6414
Amida Brimah	C	0.9701
Anzejs Pasecniks	C-F	0.996
Boban Marjanovic	C	1.0
Bruno Fernando	F-C	0.996
Cameron Oliver	F	0.5279
Chris Silva	F	0.5458
Cristiano Felicio	F-C	0.8625
D.J. Wilson	F	0.9998
Devontae Cacok	F	0.5637
Dewayne Dedmon	C	0.994
Donta Hall	C	0.99
Ed Davis	C-F	1.0
Ersan Ilyasova	F	0.6713
Freddie Gillespie	F	0.9811
Harry Giles III	F-C	0.9999
Henry Ellenson	F-C	0.8705
JaKarr Sampson	F	1.0
Jabari Parker	F	0.7629
Jahlil Okafor	C-F	1.0
Jalen Smith	F-C	0.9223
Jaren Jackson Jr.	F-C	0.9996
Jontay Porter	C-F	0.99
Justin Patton	C	0.9951
Luka Samanic	F	0.7003
Luke Kornet	F-C	1.0
Mamadi Diakite	F	0.7112
Marques Bolden	C	0.8845
Meyers Leonard	F-C	0.8645
Mfiondu Kabengele	F-C	0.9993
Nathan Knight	F-C	0.9457
Nick Richards	C	0.9661
Norvel Pelle	C	0.9522
Reggie Perry	F-C	0.9821
Tacko Fall	C	0.9841
Thomas Bryant	C-F	1.0
Thon Maker	F-C	0.8924
Trey Lyles	F	0.9119
Udoka Azubuike	C-F	0.996
Vernon Carey Jr.	F-C	0.9183
Vincent Poirier	C-F	0.996
Will Magnay	C	0.9661

Table E.3: Replacement non-centers on offense (1).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.	Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Abdel Nader	F	0.338	0.0795	0.0706	0.506	0.006	Elijah Bryant	G	0.5522	0.008	0.3244	0.1144	0.001
Adam Mokoka	G	0.5359	0.008	0.3038	0.1514	0.001	Elijah Hughes	G	0.5259	0.007	0.3207	0.1454	0.001
Al-Farouq Aminu	F	0.008	0.6779	0.0129	0.2982	0.003	Ersan Ilyasova	F	0.0428	0.2261	0.0857	0.6295	0.0159
Alfonzo McKinnie	F	0.0269	0.1414	0.0697	0.752	0.01	Frank Mason III	G	0.4657	0.007	0.4318	0.0945	0.001
Amir Coffey	G-F	0.7237	0.0089	0.0636	0.2028	0.001	Frank Ntilikina	G	0.6199	0.004	0.1303	0.2448	0.001
Anderson Varejao	F	0.0567	0.3751	0.0945	0.4527	0.0209	Gabriel Deck	F	0.0338	0.5224	0.0557	0.3751	0.0129
Andre Roberson	G-F	0.2587	0.1642	0.197	0.3791	0.001	Gary Payton II	G	0.4841	0.0089	0.3141	0.1918	0.001
Anthony Gill	F	0.0328	0.2068	0.0567	0.6918	0.0119	Glenn Robinson III	F	0.1046	0.0677	0.0608	0.762	0.005
Anthony Lamb	F	0.007	0.0358	0.0507	0.9035	0.003	Grant Riller	G	0.5194	0.008	0.3652	0.1065	0.001
Anthony Tolliver	F	0.1127	0.2542	0.0847	0.5304	0.0179	Greg Whittington	F	0.0746	0.3144	0.0995	0.4896	0.0219
Ashton Hagans	G-F	0.3019	0.0447	0.2244	0.428	0.001	Henry Ellenson	F-C	0.001	0.7781	0.008	0.2119	0.001
Axel Toupane	G-F	0.2584	0.0398	0.2018	0.499	0.001	Ignas Brazdeikis	F	0.0537	0.1542	0.2318	0.5493	0.0109
Bol Bol	C-F	0.1373	0.6169	0.0189	0.2259	0.001	Iman Shumpert	G	0.5339	0.008	0.3367	0.1205	0.001
Brian Bowen II	F-G	0.1095	0.0438	0.2726	0.5731	0.001	Isaac Bonga	G	0.0338	0.001	0.0656	0.8986	0.001
Brodric Thomas	G	0.1781	0.002	0.4438	0.3751	0.001	Isaiah Joe	G	0.3632	0.002	0.0746	0.5592	0.001
Bruno Caboclo	F	0.0557	0.3771	0.0945	0.4517	0.0209	Isaiah Thomas	G	0.4776	0.1075	0.3124	0.1015	0.001
CJ Elleby	F-G	0.0538	0.0259	0.1954	0.7238	0.001	Jabari Parker	F	0.0398	0.5572	0.0677	0.3204	0.0149
Cameron Oliver	F	0.0498	0.3672	0.0836	0.4816	0.0179	Jahmi'us Ramsey	G	0.5755	0.007	0.2763	0.1402	0.001
Cam Reynolds	F	0.0557	0.2945	0.0975	0.5313	0.0209	Jalen Harris	G	0.6836	0.006	0.2308	0.0786	0.001
Carsen Edwards	G	0.6458	0.004	0.1522	0.197	0.001	Jalen Lecque	G	0.5194	0.008	0.3622	0.1095	0.001
Cassius Stanley	G	0.4269	0.007	0.3095	0.2557	0.001	James Nunnally	F	0.0568	0.2898	0.0966	0.5369	0.0199
Cassius Winston	G	0.5298	0.007	0.3191	0.1431	0.001	Jared Dudley	F	0.0507	0.2684	0.0895	0.5726	0.0189
Chandler Hutchison	F-G	0.0229	0.011	0.1086	0.8566	0.001	Jared Harper	G	0.5483	0.008	0.3343	0.1085	0.001
Charlie Brown Jr.	G	0.503	0.006	0.243	0.247	0.001	Jaren Jackson Jr.	F-C	0.001	0.9204	0.004	0.0736	0.001
Chris Chiozza	G	0.6109	0.005	0.3254	0.0577	0.001	Jarrell Brantley	F	0.0428	0.2271	0.1125	0.6016	0.0159
D.J. Wilson	F	0.006	0.7582	0.009	0.2239	0.003	Javonte Green	G-F	0.001	0.003	0.0189	0.9761	0.001
Dakota Mathias	G	0.606	0.006	0.2667	0.1204	0.001	Jay Scrubb	G	0.4378	0.007	0.2876	0.2667	0.001
Dante Exum	G	0.4245	0.007	0.327	0.2406	0.001	Jaylen Adams	G	0.5502	0.008	0.3323	0.1085	0.001
Darius Miller	F	0.2607	0.1841	0.0726	0.4697	0.0129	Jaylen Hoard	F	0.0189	0.7602	0.0318	0.1811	0.008
Deividas Sirvydis	F-G	0.1968	0.0328	0.2038	0.5656	0.001	Jeremiah Martin	G	0.591	0.007	0.3055	0.0955	0.001
Dennis Smith Jr.	G	0.7781	0.002	0.202	0.0169	0.001	Jerome Robinson	G	0.79	0.004	0.1433	0.0617	0.001
Devin Cannady	G	0.498	0.007	0.2908	0.2032	0.001	Jordan Bell	F	0.0507	0.4328	0.0856	0.4119	0.0189
Devon Dotson	G	0.498	0.008	0.3926	0.1004	0.001	Jordan Bone	G	0.6833	0.005	0.2131	0.0976	0.001
Didi Louzada	G	0.4831	0.008	0.326	0.1819	0.001	Jordan Nwora	F	0.0328	0.1282	0.1491	0.6799	0.0099
E'Twaun Moore	G	0.4821	0.002	0.4542	0.0608	0.001	Josh Green	G	0.0378	0.003	0.0607	0.8975	0.001

Table E.4: Replacement non-centers on offense (2).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Josh Hall	F	0.1413	0.0866	0.1881	0.5771	0.007
Justin James	G-F	0.1104	0.0149	0.2498	0.6239	0.001
Justin Robinson	G	0.5637	0.007	0.2839	0.1444	0.001
KZ Okpala	F-G	0.009	0.005	0.0657	0.9194	0.001
Karim Mane	G	0.5244	0.007	0.2856	0.1821	0.001
Keita Bates-Diop	F	0.0438	0.1534	0.0807	0.7112	0.011
Kelan Martin	F	0.0836	0.1025	0.0816	0.7244	0.008
Keljin Blevins	G	0.4418	0.007	0.2935	0.2567	0.001
Kevin Knox II	F	0.005	0.0129	0.0617	0.9184	0.002
Khyri Thomas	G	0.3552	0.0796	0.3532	0.2109	0.001
Killian Tillie	F-C	0.001	0.5194	0.0239	0.4547	0.001
Kostas Antetokounmpo	F	0.0527	0.3562	0.0896	0.4816	0.0199
Kris Dunn	G	0.5706	0.008	0.3171	0.1034	0.001
Kyle Guy	G	0.6402	0.005	0.2962	0.0577	0.001
Langston Galloway	G	0.5902	0.001	0.0339	0.3739	0.001
Louis King	F	0.0498	0.2647	0.1075	0.5592	0.0189
Luka Samanic	F	0.0219	0.1164	0.0478	0.805	0.009
Luke Kornet	F-C	0.001	0.9184	0.003	0.0766	0.001
Malik Fitts	F	0.0627	0.3154	0.1015	0.4985	0.0219
Markelle Fultz	G	0.503	0.005	0.4303	0.0608	0.001
Markus Howard	G	0.7012	0.005	0.2271	0.0657	0.001
Marquese Chriss	F	0.0577	0.3632	0.0965	0.4617	0.0209
Mason Jones	G	0.7278	0.003	0.2373	0.0309	0.001
Matt Thomas	G	0.79	0.003	0.1244	0.0816	0.001
Matthew Dellavedova	G	0.6328	0.005	0.3005	0.0607	0.001
Meyers Leonard	F-C	0.001	0.798	0.008	0.192	0.001
Mfiondu Kabengele	F-C	0.001	0.9114	0.003	0.0836	0.001
Mike James	G	0.5234	0.005	0.4129	0.0577	0.001
Nate Darling	G	0.5388	0.008	0.327	0.1252	0.001
Nate Hinton	G-F	0.4289	0.0358	0.1871	0.3473	0.001
Nico Mannion	G	0.7602	0.003	0.21	0.0259	0.001
Noah Vonleh	F	0.0597	0.3383	0.1005	0.4796	0.0219
Patrick McCaw	G	0.51	0.008	0.3247	0.1564	0.001
Paul Watson	G	0.4866	0.004	0.1323	0.3761	0.001
Quinn Cook	G	0.668	0.006	0.2495	0.0755	0.001

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	4 Prob.
Quinndary Weatherspoon	G	0.4	0.007	0.4527	0.1393	0.001
Ray Spalding	F	0.0586	0.3519	0.0984	0.4692	0.0219
Rayjon Tucker	G	0.4771	0.007	0.4163	0.0986	0.001
Robert Franks	F	0.0478	0.2537	0.0866	0.594	0.0179
Robert Woodard II	F	0.0548	0.2888	0.0966	0.5398	0.0199
Rodions Kurucs	F	0.0438	0.2321	0.0876	0.6205	0.0159
Rodney McGruder	G	0.3303	0.006	0.2806	0.3821	0.001
Romeo Langford	G-F	0.1224	0.0189	0.1214	0.7363	0.001
Ryan Arcidiacono	G	0.8884	0.001	0.0508	0.0588	0.001
Sam Merrill	G	0.6935	0.005	0.1811	0.1194	0.001
Sean McDermott	F	0.1662	0.2129	0.0736	0.5323	0.0149
Shaquille Harrison	G	0.6905	0.003	0.2438	0.0617	0.001
Sindarius Thornwell	G	0.3085	0.005	0.2129	0.4726	0.001
Skylar Mays	G	0.2886	0.004	0.6488	0.0577	0.001
Spencer Dinwiddie	G	0.4587	0.007	0.4169	0.1164	0.001
T.J. Leaf	F	0.0557	0.2955	0.0985	0.5294	0.0209
T.J. Warren	F	0.0458	0.247	0.0886	0.6016	0.0169
Terrance Ferguson	G	0.5194	0.008	0.3085	0.1632	0.001
Theo Pinson	G-F	0.2816	0.0418	0.208	0.4677	0.001
Tim Frazier	G	0.5727	0.007	0.3217	0.0976	0.001
Tre Jones	G	0.3755	0.004	0.5697	0.0498	0.001
Tremont Waters	G	0.6213	0.005	0.3171	0.0557	0.001
Trent Forrest	G	0.7177	0.004	0.2336	0.0437	0.001
Trey Lyles	F	0.0845	0.0835	0.0666	0.7584	0.007
Ty-Shon Alexander	G	0.5234	0.008	0.3323	0.1353	0.001
Tyler Bey	F	0.0517	0.2704	0.0895	0.5696	0.0189
Tyrell Terry	G	0.5721	0.008	0.3194	0.0995	0.001
Udonis Haslem	F	0.0607	0.3274	0.1015	0.4886	0.0219
Vlatko Cancar	F	0.0239	0.1243	0.0517	0.7913	0.0089
Wenyen Gabriel	F	0.0299	0.194	0.0527	0.7124	0.0109
Yogi Ferrell	G	0.3831	0.007	0.5284	0.0806	0.001
Zeke Nnaji	F-C	0.001	0.1352	0.0318	0.831	0.001

Table E.5: Replacement non-centers on defense (1).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.	Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Abdel Nader	F	0.2566	0.6802	0.0008	0.0623	E'Twaun Moore	G	0.2932	0.088	0.0183	0.6005
Adam Mokoka	G	0.5233	0.1013	0.0457	0.3297	Elijah Bryant	G	0.446	0.1661	0.0473	0.3405
Al-Farouq Aminu	F	0.1827	0.7276	0.0008	0.0889	Elijah Hughes	G	0.4797	0.156	0.0448	0.3195
Alfonzo McKinnie	F	0.2782	0.6528	0.0008	0.0681	Frank Mason III	G	0.5174	0.1246	0.044	0.314
Amir Coffey	G-F	0.7907	0.0997	0.01	0.0997	Frank Ntilikina	G	0.2176	0.0523	0.2558	0.4743
Anderson Varejao	F	0.2135	0.6794	0.0017	0.1055	Gabriel Deck	F	0.3892	0.5378	0.0017	0.0714
Andre Roberson	G-F	0.3807	0.3408	0.0258	0.2527	Gary Payton II	G	0.4377	0.1047	0.1229	0.3347
Anthony Gill	F	0.1784	0.7386	0.0017	0.0813	Glenn Robinson III	F	0.211	0.7467	0.0008	0.0415
Anthony Lamb	F	0.0989	0.8712	0.0008	0.0291	Grant Riller	G	0.4547	0.1081	0.049	0.3882
Anthony Tolliver	F	0.2195	0.5993	0.0017	0.1796	Greg Whittington	F	0.245	0.6437	0.0017	0.1096
Ashton Hagans	G-F	0.4302	0.2608	0.0282	0.2807	Ignas Brazdeikis	F	0.25	0.6852	0.0008	0.064
Axel Toupane	G-F	0.3779	0.2334	0.137	0.2517	Iman Shumpert	G	0.4643	0.1113	0.0673	0.3571
Bol Bol	C-F	0.0183	0.9801	0.0008	0.0008	Isaac Bonga	G	0.2542	0.4277	0.0116	0.3065
Brian Bowen II	F-G	0.2664	0.5062	0.0066	0.2207	Isaiah Joe	G	0.217	0.0441	0.0233	0.7157
Brodric Thomas	G	0.1754	0.4871	0.0158	0.3217	Isaiah Thomas	G	0.5195	0.1037	0.0465	0.3303
Bruno Caboclo	F	0.2135	0.6802	0.0017	0.1047	Jahmi'us Ramsey	G	0.4938	0.1602	0.0415	0.3046
CJ Elleby	F-G	0.4423	0.3826	0.005	0.1701	Jalen Harris	G	0.3596	0.0831	0.0382	0.5191
Cam Reynolds	F	0.2118	0.6827	0.0017	0.1038	Jalen Lecque	G	0.4635	0.1329	0.0498	0.3538
Carsen Edwards	G	0.4385	0.0606	0.0266	0.4743	James Nunnally	F	0.2929	0.6025	0.0017	0.1029
Cassius Stanley	G	0.4448	0.2108	0.0423	0.3021	Jared Dudley	F	0.1978	0.7007	0.0017	0.0998
Cassius Winston	G	0.5507	0.1071	0.0424	0.2998	Jared Harper	G	0.4826	0.1188	0.049	0.3497
Chandler Hutchison	F-G	0.1404	0.7741	0.0033	0.0822	Jarrell Brantley	F	0.2149	0.6979	0.0017	0.0855
Charlie Brown Jr.	G	0.4534	0.1131	0.0366	0.3968	Javonte Green	G-F	0.0888	0.0473	0.1037	0.7602
Chris Chiozza	G	0.3679	0.0764	0.0316	0.5241	Jay Scrubb	G	0.5087	0.0964	0.0432	0.3516
Dakota Mathias	G	0.3776	0.2996	0.0398	0.283	Jaylen Adams	G	0.4676	0.1337	0.049	0.3497
Daniel Oturu	C	0.0889	0.858	0.0008	0.0523	Jaylen Hoard	F	0.3577	0.5416	0.0008	0.0998
Dante Exum	G	0.478	0.1842	0.0407	0.2971	Jeremiah Martin	G	0.412	0.1047	0.0449	0.4385
Darius Miller	F	0.3104	0.4299	0.0017	0.2581	Jerome Robinson	G	0.235	0.0756	0.0266	0.6628
Deividas Sirvydis	F-G	0.2193	0.6013	0.0058	0.1736	Jordan Bell	F	0.1968	0.7051	0.0017	0.0963
Dennis Smith Jr.	G	0.1545	0.3081	0.015	0.5224	Jordan Bone	G	0.6528	0.0756	0.0332	0.2384
Devin Cannady	G	0.4133	0.0996	0.0573	0.4299	Jordan Nwora	F	0.2377	0.5569	0.0008	0.2045
Devon Dotson	G	0.4294	0.103	0.1395	0.3281	Josh Green	G	0.2076	0.0299	0.015	0.7475
Didi Louzada	G	0.451	0.1096	0.0457	0.3937	Josh Hall	F	0.6071	0.348	0.0008	0.044

Table E.6: Replacement non-centers on defense (2).

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Justin James	G-F	0.3544	0.4523	0.0133	0.1801
Justin Robinson	G	0.4958	0.0963	0.0432	0.3646
Juwan Morgan	F	0.1711	0.745	0.0017	0.0822
KZ Okpala	F-G	0.0739	0.7841	0.0025	0.1395
Karim Mane	G	0.4141	0.2357	0.0432	0.3071
Keita Bates-Diop	F	0.2121	0.5732	0.0008	0.2138
Kelan Martin	F	0.1304	0.8198	0.0008	0.049
Keljin Blevins	G	0.4746	0.1189	0.0432	0.3633
Kevin Knox II	F	0.6055	0.3422	0.0008	0.0515
Khyri Thomas	G	0.3787	0.093	0.0382	0.49
Killian Tillie	F-C	0.0947	0.8912	0.0008	0.0133
Kostas Antetokounmpo	F	0.2041	0.6564	0.0149	0.1245
Kris Dunn	G	0.436	0.1088	0.0465	0.4086
Kyle Guy	G	0.6373	0.1253	0.029	0.2083
Langston Galloway	G	0.8022	0.1031	0.0116	0.0831
Louis King	F	0.1934	0.556	0.1494	0.1012
Malik Fitts	F	0.2243	0.6636	0.0017	0.1105
Markelle Fultz	G	0.5112	0.079	0.0316	0.3782
Markus Howard	G	0.6199	0.1228	0.0315	0.2257
Marquese Chriss	F	0.2168	0.6752	0.0017	0.1063
Mason Jones	G	0.6384	0.1995	0.02	0.1421
Matt Thomas	G	0.5723	0.2517	0.0216	0.1545
Matthew Dellavedova	G	0.6409	0.0998	0.0316	0.2278
Mike James	G	0.6376	0.1131	0.0299	0.2195
Nate Darling	G	0.4576	0.1105	0.0482	0.3837
Nate Hinton	G-F	0.4734	0.2218	0.0241	0.2807
Nico Mannion	G	0.2841	0.0482	0.0208	0.647
Noah Vonleh	F	0.2243	0.6636	0.0017	0.1105
Patrick McCaw	G	0.4515	0.1079	0.049	0.3917
Paul Reed	F	0.1578	0.5947	0.0017	0.2458
Paul Watson	G	0.2882	0.2051	0.0241	0.4826
Quinn Cook	G	0.5718	0.0954	0.0365	0.2963
Quinndary Weatherspoon	G	0.3713	0.0889	0.2566	0.2832

Name	Label	0 Prob.	1 Prob.	2 Prob.	3 Prob.
Ray Spalding	F	0.2201	0.6703	0.0017	0.108
Rayjon Tucker	G	0.4136	0.0988	0.1611	0.3264
Robert Franks	F	0.2434	0.6512	0.0017	0.1038
Robert Woodard II	F	0.2085	0.6877	0.0017	0.1022
Rodions Kurucs	F	0.1909	0.7095	0.0017	0.0979
Rodney McGruder	G	0.3367	0.079	0.0416	0.5428
Romeo Langford	G-F	0.3643	0.1859	0.0158	0.434
Rondae Hollis-Jefferson	F	0.1993	0.7076	0.0017	0.0914
Ryan Arcidiacono	G	0.4639	0.0581	0.0124	0.4656
Sam Merrill	G	0.6218	0.1297	0.0291	0.2195
Sean McDermott	F	0.2633	0.6487	0.0017	0.0864
Shaquille Harrison	G	0.2292	0.064	0.0266	0.6802
Sindarius Thornwell	G	0.358	0.0806	0.0332	0.5282
Skylar Mays	G	0.2791	0.0631	0.0382	0.6196
Spencer Dinwiddie	G	0.4373	0.1983	0.0448	0.3195
T.J. Leaf	F	0.2133	0.6805	0.0017	0.1046
T.J. Warren	F	0.3776	0.522	0.0017	0.0988
Terrance Ferguson	G	0.5104	0.1131	0.0466	0.33
Theo Pinson	G-F	0.3987	0.3106	0.0266	0.2641
Tim Frazier	G	0.485	0.1478	0.0449	0.3223
Tre Jones	G	0.5648	0.0648	0.0274	0.343
Tremont Waters	G	0.2988	0.0656	0.0324	0.6033
Trent Forrest	G	0.4767	0.3007	0.0249	0.1977
Ty-Shon Alexander	G	0.4369	0.1836	0.0465	0.3331
Tyler Bey	F	0.1977	0.7035	0.0017	0.0972
Tyler Cook	F	0.3056	0.6653	0.0008	0.0282
Tyrell Terry	G	0.4261	0.1022	0.142	0.3297
Udonis Haslem	F	0.2274	0.6539	0.0058	0.1129
Vlatko Cancar	F	0.3461	0.5923	0.0008	0.0607
Wenyen Gabriel	F	0.1379	0.7957	0.0008	0.0656
Yogi Ferrell	G	0.5373	0.0871	0.0821	0.2935
Zeke Nnaji	F-C	0.0166	0.9767	0.0008	0.0058

Appendix F

Parameter estimates

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Aaron Gordon	0.055	0.1367	0.0597	0.15	0.1457	0.0588	−0.1257	0.1287	1.3624	0.0861
Aaron Holiday	0.012	0.0172	0.0633	0.06	−0.0004	0.0622	−1.7947	0.2735	0.2755	0.1321
Aaron Nesmith	0.041	0.0811	0.0685	0.146	0.1406	0.0681	−0.5829	0.2189	1.2806	0.1247
Al Horford	0.034	−0.079	0.0738	0.195	0.2883	0.0736	−0.567	0.2212	1.5591	0.111
Alec Burks	0.016	0.0131	0.0655	0.156	0.1419	0.0631	−1.5095	0.2414	1.3452	0.0999
Aleksej Pokusevski	0.025	−0.0057	0.0638	0.155	0.1761	0.0646	−0.9283	0.2022	1.4028	0.0976
Alex Caruso	0.025	0.1002	0.0611	0.114	0.1324	0.0612	−1.1046	0.1964	0.9871	0.1117
Alex Len	0.077	0.0442	0.0651	0.174	0.2557	0.0673	0.3475	0.1449	1.3946	0.0964
Andre Drummond	0.132	0.2094	0.0604	0.336	0.3172	0.0666	0.8405	0.1115	2.2516	0.081
Andre Iguodala	0.03	0.092	0.0634	0.134	0.0894	0.0639	−0.9637	0.1763	1.2114	0.0997
Andrew Wiggins	0.034	0.0301	0.0613	0.106	0.2119	0.0575	−0.5836	0.13	0.8145	0.0911
Anfernee Simons	0.012	−0.0177	0.0627	0.109	0.0302	0.0641	−1.869	0.2697	1.0384	0.1101
Anthony Davis	0.054	0.1391	0.0646	0.198	0.2646	0.0674	−0.3258	0.1512	1.6157	0.099
Anthony Edwards	0.023	−0.0251	0.0568	0.121	0.1092	0.06	−1.1564	0.1459	1.0665	0.0976
Armoni Brooks	0.017	−0.0417	0.073	0.112	0.071	0.0738	−1.3212	0.3359	0.9798	0.1509
Aron Baynes	0.082	0.1232	0.0655	0.192	0.2078	0.0656	0.2424	0.1348	1.4379	0.0958
Austin Rivers	0.009	0.0196	0.0663	0.088	0.0437	0.0668	−2.0927	0.3818	0.7836	0.131
Avery Bradley	0.013	−0.085	0.0719	0.078	0.0221	0.0692	−1.5897	0.3838	0.6541	0.1604
Bam Adebayo	0.07	0.027	0.0651	0.204	0.2181	0.0617	0.0359	0.1114	1.6465	0.079
Ben McLemore	0.015	−0.0099	0.0625	0.091	0.0514	0.0634	−1.5423	0.276	0.737	0.1215
Ben Simmons	0.051	−0.0184	0.0641	0.166	0.2152	0.0651	−0.1867	0.1244	1.5543	0.0944
Bismack Biyombo	0.094	0.1583	0.0627	0.15	0.2212	0.0642	0.3347	0.1199	1.1254	0.104
Blake Griffin	0.029	0.0453	0.0612	0.159	0.1955	0.0611	−0.9184	0.1857	1.2646	0.0894
Bobby Portis	0.085	0.1787	0.063	0.229	0.1919	0.0609	0.3079	0.1201	1.9799	0.086
Bogdan Bogdanovic	0.016	−0.067	0.061	0.104	0.1845	0.0613	−1.3691	0.2355	0.9172	0.1082
Bojan Bogdanovic	0.018	0.0008	0.0669	0.099	0.0778	0.0684	−1.3251	0.1695	0.911	0.0949
Brad Wanamaker	0.016	0.0182	0.0646	0.08	0.1045	0.0651	−1.4245	0.2558	0.423	0.1228
Bradley Beal	0.033	0.1419	0.0595	0.094	0.0545	0.0582	−0.6285	0.1374	0.7306	0.087
Brandon Clarke	0.064	0.0528	0.0596	0.16	0.1515	0.061	−0.0075	0.1227	1.5759	0.102

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Brandon Goodwin	0.012	0.0466	0.0699	0.097	0.1338	0.0725	−1.7968	0.3862	0.8231	0.149
Brandon Ingram	0.016	−0.0506	0.0647	0.126	0.1549	0.0675	−1.3927	0.1842	1.1408	0.0939
Brook Lopez	0.054	0.104	0.0645	0.115	0.207	0.0637	−0.1176	0.1242	1.1999	0.0948
Bruce Brown	0.081	0.1433	0.0585	0.154	0.0771	0.0562	0.2611	0.1233	1.2305	0.084
Bryn Forbes	0.008	−0.0374	0.0592	0.069	0.0618	0.0582	−2.1172	0.3103	0.6796	0.1204
Buddy Hield	0.012	0.0422	0.0609	0.126	0.1177	0.0625	−1.6558	0.1989	1.1511	0.0831
CJ McCollum	0.018	0.0443	0.0617	0.099	0.1337	0.0628	−1.4048	0.1943	0.9598	0.1016
Caleb Martin	0.036	0.0352	0.0664	0.125	0.0192	0.0643	−0.7137	0.2115	1.0179	0.1273
Cam Reddish	0.028	0.028	0.0711	0.109	0.1474	0.0712	−0.813	0.2441	1.0086	0.1382
Cameron Johnson	0.023	0.054	0.0601	0.118	0.1508	0.0612	−1.0486	0.196	1.1275	0.1039
Cameron Payne	0.016	0.0118	0.0699	0.119	0.1281	0.0677	−1.4805	0.2651	1.1415	0.1073
Caris LeVert	0.022	−0.0426	0.0607	0.111	−0.0215	0.0553	−1.0313	0.1932	0.9815	0.0932
Carmelo Anthony	0.017	0.0114	0.0591	0.107	0.0351	0.0605	−1.4716	0.1876	1.001	0.0978
Cedi Osman	0.024	0.0506	0.0594	0.113	0.0268	0.0587	−1.0601	0.1767	1.0071	0.0994
Chasson Randle	0.008	−0.0898	0.0665	0.084	0.1328	0.0691	−2.0751	0.3925	0.7036	0.1291
Chimezie Metu	0.064	0.0784	0.0749	0.175	0.1627	0.0735	−0.0411	0.2103	1.4246	0.1363
Chris Boucher	0.078	0.11	0.0592	0.192	0.0412	0.0597	0.1549	0.1159	1.4986	0.0832
Chris Paul	0.012	−0.0832	0.0684	0.131	0.205	0.0673	−1.5202	0.2141	1.3017	0.0902
Christian Wood	0.055	0.0275	0.0644	0.228	0.2484	0.063	−0.1285	0.1403	1.7223	0.0796
Chuma Okeke	0.033	−0.0206	0.0607	0.117	0.0893	0.0646	−0.5108	0.1697	1.0254	0.0999
Clint Capela	0.155	0.2121	0.0617	0.301	0.343	0.0597	1.04	0.097	2.1011	0.0794
Coby White	0.013	0.0098	0.0583	0.116	0.2397	0.0634	−1.6849	0.2024	0.9557	0.0801
Cody Martin	0.059	0.0602	0.0631	0.126	0.0768	0.0651	−0.1154	0.166	1.003	0.1256
Cody Zeller	0.115	0.2028	0.0679	0.21	0.2534	0.0685	0.5454	0.1292	1.5585	0.1069
Cole Anthony	0.027	0.0954	0.0611	0.142	0.1042	0.063	−0.9021	0.1753	1.2827	0.0904
Collin Sexton	0.027	0.1213	0.0562	0.063	0.0487	0.057	−0.9631	0.1448	0.3189	0.1039
Cory Joseph	0.024	0.0208	0.0578	0.085	0.121	0.0581	−1.0202	0.1817	0.6911	0.1039
D'Angelo Russell	0.012	−0.0016	0.0611	0.077	0.0422	0.0614	−1.799	0.2607	0.596	0.1286
D.J. Augustin	0.017	0.0091	0.0605	0.063	0.069	0.063	−1.3455	0.23	0.5342	0.13

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
DaQuan Jeffries	0.036	0.0364	0.0712	0.103	0.1627	0.0749	−0.5415	0.2454	0.8553	0.1546
Damian Jones	0.064	0.0221	0.0737	0.156	0.1587	0.071	0.0289	0.1963	1.2687	0.1267
Damian Lillard	0.012	0.0603	0.061	0.104	0.1023	0.0626	−1.7459	0.1887	1.0286	0.0894
Damion Lee	0.02	−0.0688	0.0668	0.142	0.1394	0.0626	−1.1484	0.2349	1.0406	0.1052
Damyean Dotson	0.005	0.0299	0.0616	0.095	0.0603	0.0659	−2.7631	0.4708	0.8363	0.1202
Daniel Gafford	0.118	0.2226	0.0666	0.16	0.0692	0.0644	0.6129	0.1326	1.3273	0.1042
Daniel Theis	0.057	0.0611	0.0573	0.164	0.1355	0.0571	−0.2746	0.1226	1.4202	0.0802
Danilo Gallinari	0.015	−0.1241	0.064	0.155	0.1786	0.0626	−1.4905	0.2589	1.3124	0.1
Danny Green	0.029	0.0522	0.0657	0.1	0.0288	0.0672	−0.7762	0.1532	1.0399	0.1048
Danuel House Jr.	0.018	−0.0819	0.0659	0.121	0.1009	0.0626	−1.2851	0.2463	1.0486	0.1089
Dario Saric	0.055	0.037	0.0722	0.168	0.2674	0.0715	−0.2097	0.1798	1.4199	0.1104
Darius Bazley	0.027	−0.0431	0.0617	0.193	0.2406	0.0609	−0.8272	0.152	1.5916	0.0833
Darius Garland	0.013	0.0268	0.0598	0.062	0.0364	0.0599	−1.665	0.2237	0.2721	0.1125
David Nwaba	0.055	0.0176	0.0708	0.108	0.0351	0.0683	−0.0742	0.1868	0.9298	0.1287
Davis Bertans	0.011	0.0583	0.0611	0.095	0.0895	0.06	−1.7754	0.2513	0.7589	0.1003
De'Aaron Fox	0.016	−0.0551	0.0638	0.085	0.0776	0.0625	−1.3233	0.184	0.7298	0.0994
De'Andre Hunter	0.024	0.0234	0.0712	0.125	0.1522	0.0712	−1.0065	0.2658	1.1438	0.127
De'Anthony Melton	0.026	0.0772	0.0614	0.118	0.1138	0.0623	−0.9975	0.1946	1.2674	0.1175
DeAndre Jordan	0.097	0.0889	0.0661	0.231	0.2826	0.064	0.543	0.1253	1.6491	0.0841
DeAndre' Bembry	0.033	0.0497	0.0643	0.112	0.0589	0.0652	−0.7269	0.1902	0.9177	0.1124
DeMar DeRozan	0.019	0.0218	0.0597	0.106	0.1286	0.0589	−1.2003	0.1727	0.9205	0.094
DeMarcus Cousins	0.07	0.09	0.0709	0.293	0.3307	0.0701	0.0487	0.1683	2.0849	0.0969
Dean Wade	0.031	−0.092	0.06	0.154	0.1365	0.0609	−0.7564	0.1832	1.1962	0.0985
Deandre Ayton	0.117	0.1946	0.0722	0.236	0.2345	0.0686	0.7815	0.1055	1.9245	0.0818
Dejounte Murray	0.026	0.0105	0.0602	0.19	0.2151	0.0597	−0.8566	0.1475	1.5551	0.0822
Delon Wright	0.036	−0.0057	0.0581	0.122	0.0304	0.0593	−0.5648	0.1376	1.1218	0.0872
Deni Avdija	0.017	−0.0629	0.0611	0.181	0.2021	0.0621	−1.239	0.2336	1.4328	0.0887
Dennis Schroder	0.016	0.0357	0.06	0.093	0.1001	0.0615	−1.5248	0.1967	0.7883	0.0999
Denzel Valentine	0.022	0.0157	0.0633	0.168	0.2522	0.0634	−1.1154	0.2154	1.303	0.0972

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Derrick Favors	0.13	0.1902	0.0722	0.227	0.3072	0.0769	0.6671	0.1247	1.6731	0.1015
Derrick Jones Jr.	0.055	−0.0712	0.0594	0.094	0.0442	0.0632	−0.0755	0.1352	0.9184	0.1101
Derrick Rose	0.015	0.0183	0.0603	0.081	0.0512	0.0607	−1.5501	0.241	0.6432	0.1146
Derrick White	0.012	−0.0341	0.0661	0.086	0.0251	0.0653	−1.6338	0.298	0.7438	0.1255
Desmond Bane	0.019	0.1028	0.0589	0.116	0.1343	0.0593	−1.3572	0.1878	1.2483	0.1045
Devin Booker	0.016	0.0848	0.0672	0.108	0.0163	0.0653	−1.3102	0.1881	1.1249	0.0913
Devin Vassell	0.021	−0.0718	0.0665	0.138	0.1297	0.062	−1.061	0.2194	1.2262	0.1062
Devonte’ Graham	0.013	0.0615	0.0601	0.075	0.0906	0.0596	−1.761	0.2251	0.4396	0.1163
Dillon Brooks	0.026	−0.0108	0.0612	0.069	0.0659	0.0617	−0.9037	0.1525	0.7164	0.1133
Domantas Sabonis	0.066	0.013	0.0634	0.259	0.288	0.0599	0.0923	0.1088	1.8775	0.075
Donovan Mitchell	0.028	0.0496	0.0655	0.095	0.031	0.0673	−0.8975	0.1582	0.8785	0.1032
Donte DiVincenzo	0.043	0.0331	0.0619	0.155	0.1042	0.0608	−0.3227	0.1327	1.5839	0.0918
Dorian Finney-Smith	0.052	0.0675	0.0601	0.116	0.177	0.0605	−0.2921	0.1239	0.888	0.0931
Doug McDermott	0.036	−0.0179	0.0585	0.096	0.1202	0.0555	−0.5579	0.1488	0.7674	0.1012
Draymond Green	0.028	−0.0416	0.0689	0.184	0.2112	0.066	−0.7859	0.1525	1.4319	0.0862
Drew Eubanks	0.09	0.1095	0.0721	0.207	0.2057	0.068	0.2981	0.1491	1.6705	0.1058
Duncan Robinson	0.003	−0.1022	0.0625	0.109	0.1244	0.0594	−3.2011	0.427	1.0203	0.0886
Dwayne Bacon	0.015	0.0031	0.0566	0.102	0.1559	0.0592	−1.4135	0.1926	0.8916	0.0857
Dwight Howard	0.148	0.2625	0.0665	0.309	0.2128	0.0676	0.6514	0.1096	2.1586	0.0967
Dwight Powell	0.086	0.0343	0.0665	0.158	0.1186	0.0675	0.2831	0.1473	1.225	0.1072
Dylan Windler	0.039	0.0167	0.0734	0.171	0.2115	0.0757	−0.653	0.2455	1.399	0.1362
Edmond Sumner	0.02	−0.0264	0.0657	0.083	−0.0044	0.0633	−1.1758	0.2556	0.6393	0.1335
Elfrid Payton	0.041	0.1512	0.0675	0.098	0.1007	0.0676	−0.4943	0.1503	0.8828	0.1079
Enes Freedom	0.148	0.33	0.0624	0.281	0.3336	0.062	0.7854	0.0955	2.051	0.0814
Eric Bledsoe	0.013	0.0942	0.0614	0.1	0.1057	0.0625	−1.6329	0.1998	0.9147	0.0985
Eric Gordon	0.009	0.0389	0.069	0.062	0.0255	0.0691	−2.0772	0.3714	0.352	0.1615
Eric Paschall	0.038	−0.092	0.0757	0.138	0.1271	0.0701	−0.3268	0.2225	0.9615	0.127
Evan Fournier	0.007	−0.0423	0.0597	0.094	0.0988	0.0627	−2.2073	0.3384	0.8241	0.1035
Facundo Campazzo	0.015	−0.0009	0.0633	0.081	0.1715	0.0613	−1.5561	0.2248	0.6758	0.1141

Name	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Frank Jackson	0.021	−0.0163	0.0686	0.092	0.1351	0.0683	−1.1373	0.276	0.7042	0.135
Frank Kaminsky	0.051	0.0597	0.0715	0.208	0.2613	0.0681	−0.2371	0.1968	1.7317	0.1077
Fred VanVleet	0.016	0.0494	0.0639	0.101	0.1358	0.0632	−1.4882	0.191	0.77	0.0939
Furkan Korkmaz	0.015	0.1035	0.0667	0.09	0.1571	0.0681	−1.71	0.2573	0.8138	0.1306
Gabe Vincent	0.016	0.0372	0.0735	0.068	−0.0271	0.071	−1.6018	0.3493	0.47	0.1747
Garrett Temple	0.019	−0.0295	0.0611	0.082	0.1278	0.0623	−1.2348	0.1936	0.582	0.1024
Garrison Mathews	0.015	−0.0837	0.0615	0.065	0.0327	0.0629	−1.3597	0.2673	0.38	0.1347
Gary Clark	0.044	0.0903	0.0681	0.125	0.1937	0.0723	−0.3949	0.1932	1.0512	0.1273
Gary Harris	0.019	0.0148	0.0636	0.054	0.0966	0.0647	−1.3433	0.2364	0.2385	0.1451
Gary Trent Jr.	0.014	−0.0002	0.0525	0.069	0.0022	0.0557	−1.6313	0.2022	0.5162	0.1071
George Hill	0.023	0.0155	0.0711	0.063	−0.0137	0.0706	−1.1277	0.2705	0.4345	0.1696
Georges Niang	0.024	0.0304	0.0691	0.117	0.1638	0.071	−1.1124	0.2011	1.0516	0.1084
Giannis Antetokounmpo	0.048	0.0074	0.0602	0.264	0.2853	0.0615	−0.2117	0.1243	2.1372	0.0813
Goga Bitadze	0.095	0.2282	0.0739	0.148	0.1688	0.0717	0.1875	0.163	1.1131	0.1287
Goran Dragic	0.019	−0.0244	0.0617	0.107	0.1354	0.0603	−1.3014	0.2151	0.9761	0.1014
Gordon Hayward	0.024	−0.0477	0.0618	0.143	0.1451	0.062	−1.0575	0.1844	1.1253	0.1036
Gorgui Dieng	0.068	0.11	0.075	0.181	0.1313	0.0735	−0.1237	0.1852	1.6448	0.1284
Grant Williams	0.04	−0.0002	0.0621	0.113	0.0982	0.0616	−0.6074	0.1631	1.0201	0.1081
Grayson Allen	0.014	−0.0388	0.0605	0.106	0.1128	0.0637	−1.5249	0.2516	1.161	0.1147
Hamidou Diallo	0.046	0.0673	0.063	0.161	0.1088	0.0629	−0.387	0.1482	1.3812	0.0904
Harrison Barnes	0.031	−0.0087	0.0652	0.152	0.1529	0.0644	−0.7075	0.1436	1.332	0.084
Hassan Whiteside	0.109	0.1972	0.0742	0.269	0.2198	0.0731	0.5003	0.1674	1.998	0.1145
Immanuel Quickley	0.02	0.0804	0.067	0.086	0.0757	0.0641	−1.2883	0.2225	0.7054	0.1164
Isaac Okoro	0.031	−0.0562	0.0584	0.066	0.2208	0.0602	−0.7138	0.1357	0.336	0.1022
Isaiah Hartenstein	0.114	0.1876	0.0724	0.207	0.1732	0.0741	0.4157	0.1598	1.5781	0.1178
Isaiah Roby	0.057	0.1086	0.0585	0.169	0.1236	0.0596	−0.1078	0.1277	1.4997	0.0901
Isaiah Stewart	0.105	0.2157	0.0635	0.193	0.182	0.0625	0.3949	0.1064	1.4743	0.0802
Ish Smith	0.022	0.0096	0.065	0.131	0.1053	0.0663	−1.0599	0.2306	1.0596	0.1104
Ivica Zubac	0.121	0.1924	0.0647	0.2	0.3491	0.0649	0.5524	0.1059	1.6399	0.0896

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
JJ Redick	0.006	−0.0566	0.0684	0.082	0.05	0.0711	−2.5795	0.5364	0.7173	0.1475
Ja Morant	0.027	0.1147	0.0648	0.091	0.0321	0.0665	−0.9053	0.1492	1.0116	0.1046
JaMychal Green	0.072	0.1314	0.0647	0.177	0.19	0.066	−0.0749	0.1328	1.4669	0.0981
JaVale McGee	0.105	0.2069	0.0688	0.231	0.178	0.0688	0.3308	0.1413	1.7847	0.1003
Jaden McDaniels	0.032	0.0383	0.0599	0.118	0.0855	0.0594	−0.9014	0.1544	1.0117	0.109
Jae Crowder	0.02	−0.1387	0.064	0.146	0.1723	0.0638	−0.9924	0.1978	1.3827	0.0932
Jae'Sean Tate	0.06	0.1438	0.0547	0.118	0.1628	0.0571	−0.0917	0.1064	1.0105	0.0795
Jake Layman	0.019	−0.0306	0.0699	0.083	0.1105	0.0714	−1.3795	0.3008	0.6312	0.1587
Jakob Poeltl	0.114	0.1728	0.0651	0.173	0.2747	0.0638	0.6896	0.1026	1.4169	0.0869
Jalen Brunson	0.016	0.0129	0.0599	0.117	0.2055	0.0583	−1.5359	0.2099	0.8612	0.093
Jalen McDaniels	0.047	0.0007	0.0631	0.138	0.0941	0.0659	−0.3155	0.1686	1.0983	0.1212
Jamal Murray	0.023	0.1154	0.0633	0.095	0.0487	0.0643	−1.218	0.1813	0.9015	0.1027
James Ennis III	0.042	−0.0011	0.0646	0.123	0.1113	0.0644	−0.2483	0.1763	1.1168	0.1039
James Harden	0.023	0.0321	0.0589	0.191	0.1592	0.0573	−1.075	0.1833	1.4679	0.0783
James Johnson	0.029	−0.0667	0.0616	0.131	0.0743	0.0604	−0.8127	0.1977	1.1223	0.1029
James Wiseman	0.063	0.0061	0.0699	0.19	0.1454	0.0672	0.1477	0.1724	1.4699	0.1073
Jarred Vanderbilt	0.096	0.1777	0.0623	0.211	0.1254	0.061	0.2853	0.1207	1.7108	0.1059
Jarrett Allen	0.106	0.204	0.0587	0.232	0.2053	0.0581	0.4351	0.0982	1.6128	0.0745
Jarrett Culver	0.059	0.0494	0.0743	0.132	0.0633	0.0729	−0.1751	0.2024	1.1511	0.1524
Jaxson Hayes	0.087	0.1081	0.0654	0.161	0.1932	0.0672	0.1292	0.128	1.3761	0.1061
Jaylen Brown	0.035	0.0501	0.0589	0.136	0.1283	0.0614	−0.728	0.1342	1.2534	0.0829
Jaylen Nowell	0.022	−0.0202	0.0651	0.098	0.0674	0.0681	−1.2559	0.2457	0.7973	0.1454
Jayson Tatum	0.021	0.1707	0.0571	0.188	0.1971	0.0584	−1.3335	0.1601	1.5968	0.0757
Jeff Green	0.019	−0.0802	0.058	0.116	0.1074	0.0567	−1.2736	0.1847	0.8679	0.0856
Jeff Teague	0.016	0.0016	0.0617	0.074	−0.0249	0.062	−1.5291	0.263	0.6926	0.1299
Jerami Grant	0.018	0.0807	0.0629	0.118	0.1523	0.0648	−1.3346	0.1768	1.0511	0.0897
Jeremy Lamb	0.032	0.0559	0.0674	0.137	0.0475	0.069	−0.8439	0.2225	1.1924	0.1228
Jevon Carter	0.022	−0.0553	0.0697	0.107	0.1407	0.0715	−1.0364	0.2728	0.9998	0.1363
Jimmy Butler	0.056	0.0489	0.0609	0.153	0.2173	0.0617	−0.1318	0.1316	1.3624	0.0894

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Joe Harris	0.021	−0.0136	0.0594	0.09	0.1781	0.0577	−1.1697	0.1649	0.6326	0.0896
Joe Ingles	0.014	0.0213	0.0598	0.108	0.1419	0.0591	−1.6071	0.2076	0.9912	0.0966
Joel Embiid	0.078	0.0721	0.0681	0.254	0.2405	0.0709	0.1678	0.1244	1.9774	0.094
John Collins	0.064	0.1659	0.0623	0.183	0.1822	0.0617	0.0052	0.1182	1.5575	0.0877
John Konchar	0.057	0.1093	0.071	0.161	0.1348	0.0701	−0.2261	0.1992	1.5889	0.1333
John Wall	0.012	−0.0229	0.062	0.084	0.0462	0.0612	−1.6986	0.2569	0.6861	0.1082
Jonas Valanciunas	0.134	0.3316	0.0639	0.289	0.4046	0.0668	0.7204	0.098	2.2012	0.0891
Jordan Clarkson	0.025	0.05	0.0611	0.117	0.1205	0.0596	−1.0336	0.1596	1.0599	0.0979
Jordan McLaughlin	0.02	0.0686	0.0679	0.09	0.1146	0.0681	−1.3978	0.2312	0.6995	0.1348
Jordan Poole	0.015	−0.0099	0.0665	0.074	0.1047	0.0642	−1.4319	0.2728	0.3733	0.1293
Josh Hart	0.039	0.046	0.0624	0.234	0.2543	0.0664	−0.5943	0.1551	1.8396	0.093
Josh Jackson	0.035	0.1226	0.0571	0.123	0.0949	0.0569	−0.6936	0.148	1.0647	0.0897
Josh Okogie	0.048	0.0152	0.0642	0.079	0.0909	0.0624	−0.363	0.144	0.5675	0.1324
Josh Richardson	0.028	0.0687	0.0601	0.079	0.127	0.0606	−0.9261	0.1588	0.4971	0.1069
Jrue Holiday	0.037	0.0735	0.0596	0.093	0.1761	0.0606	−0.4993	0.1366	1.0011	0.0997
Juan Toscano-Anderson	0.026	−0.1313	0.0646	0.172	0.2208	0.0629	−0.7956	0.2068	1.2829	0.0991
Juancho Hernangomez	0.047	0.0862	0.0628	0.177	0.1859	0.0652	−0.4938	0.1725	1.4592	0.1144
Julius Randle	0.031	−0.0166	0.0693	0.227	0.2868	0.069	−0.7495	0.1315	1.7264	0.0771
Justin Holiday	0.014	−0.0948	0.0607	0.098	0.0902	0.0588	−1.4691	0.1946	0.8403	0.0914
Justin Jackson	0.026	0.019	0.0741	0.1	0.1518	0.0728	−0.9591	0.2663	0.8621	0.1482
Justise Winslow	0.032	−0.0283	0.073	0.187	0.2434	0.0764	−0.6866	0.2595	1.7442	0.1359
Jusuf Nurkic	0.094	0.0048	0.0748	0.283	0.3579	0.0753	0.3944	0.1465	2.0269	0.1003
Karl-Anthony Towns	0.075	0.1382	0.0642	0.237	0.254	0.0634	−0.0182	0.1128	1.7681	0.0956
Kawhi Leonard	0.034	0.0394	0.0615	0.161	0.2114	0.0616	−0.7102	0.1504	1.4833	0.0925
Keldon Johnson	0.045	0.0211	0.0617	0.163	0.1156	0.0607	−0.2615	0.1254	1.3815	0.0869
Kelly Olynyk	0.043	0.0614	0.0558	0.196	0.2162	0.0557	−0.4424	0.1197	1.6045	0.0716
Kelly Oubre Jr.	0.046	−0.0527	0.061	0.135	0.0451	0.0584	−0.1873	0.1339	1.1087	0.0936
Kemba Walker	0.012	0.011	0.0603	0.113	0.0374	0.0613	−1.8773	0.2587	1.0662	0.1019
Kendrick Nunn	0.013	0.0221	0.0639	0.096	0.0906	0.0612	−1.6902	0.2377	0.8598	0.1009

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Kenrich Williams	0.056	0.15	0.0607	0.129	0.1037	0.0622	−0.1738	0.1316	1.1833	0.0943
Kent Bazemore	0.019	−0.0146	0.0645	0.14	0.1233	0.0603	−1.229	0.2112	1.1117	0.0988
Kentavious Caldwell-Pope	0.015	−0.0475	0.062	0.081	0.0683	0.0617	−1.5731	0.2035	0.6766	0.1052
Kenyon Martin Jr.	0.063	0.0909	0.0622	0.159	0.1681	0.065	−0.0272	0.1432	1.3026	0.0956
Kevin Durant	0.012	−0.0089	0.0653	0.18	0.1978	0.0616	−1.7753	0.3001	1.368	0.0869
Kevin Huerter	0.017	0.0051	0.0591	0.09	0.1196	0.0577	−1.2677	0.1705	0.7984	0.0965
Kevin Love	0.026	−0.0094	0.0703	0.272	0.348	0.075	−0.9451	0.2579	1.8241	0.1044
Kevin Porter Jr.	0.022	0.0636	0.0691	0.098	0.0843	0.0691	−1.1403	0.2409	0.8638	0.13
Kevon Looney	0.098	0.202	0.0673	0.165	0.1273	0.0645	0.4361	0.129	1.2928	0.0991
Khem Birch	0.101	0.1781	0.0551	0.138	0.1716	0.0582	0.5602	0.1002	1.1203	0.0831
Khris Middleton	0.023	−0.0327	0.0621	0.143	0.2018	0.0617	−0.938	0.1556	1.4639	0.0879
Killian Hayes	0.009	−0.0344	0.0714	0.092	0.1719	0.0691	−2.1066	0.4357	0.7278	0.1459
Kira Lewis Jr.	0.012	0.1045	0.0684	0.064	0.1579	0.0687	−1.8688	0.3051	0.4362	0.1484
Kristaps Porzingis	0.062	0.0339	0.0624	0.219	0.2225	0.062	−0.1549	0.1351	1.5756	0.0887
Kyle Anderson	0.026	−0.0641	0.0625	0.173	0.1811	0.0617	−0.9017	0.1502	1.6686	0.0942
Kyle Kuzma	0.054	0.034	0.0562	0.157	0.254	0.056	−0.2069	0.1154	1.316	0.0891
Kyle Lowry	0.021	−0.0621	0.0601	0.133	0.0898	0.0621	−1.1287	0.1796	1.0739	0.091
Kyrie Irving	0.029	0.0321	0.0615	0.101	0.0093	0.0599	−0.8687	0.1577	0.7748	0.0908
LaMarcus Aldridge	0.027	−0.108	0.0735	0.148	0.2559	0.0734	−0.8684	0.2588	1.1542	0.1245
LaMelo Ball	0.041	0.0861	0.0604	0.161	0.119	0.0585	−0.5527	0.1507	1.3063	0.1007
Lamar Stevens	0.044	0.0806	0.0724	0.145	0.0606	0.0737	−0.5633	0.2299	1.3036	0.1412
Landry Shamet	0.008	−0.0228	0.0606	0.067	0.027	0.0576	−2.2047	0.3079	0.2944	0.1139
Larry Nance Jr.	0.049	0.0588	0.0643	0.169	0.1555	0.0654	−0.2456	0.1573	1.4045	0.0961
Lauri Markkanen	0.027	−0.0337	0.0613	0.173	0.1852	0.0637	−0.9421	0.1898	1.3816	0.0857
LeBron James	0.02	0.0142	0.0606	0.202	0.1949	0.0606	−1.3219	0.1992	1.616	0.091
Lonnie Walker IV	0.009	−0.061	0.0596	0.093	0.0441	0.0597	−1.9863	0.2881	0.7945	0.1074
Lonzo Ball	0.018	0.0898	0.0612	0.129	0.1483	0.0643	−1.3198	0.1921	1.1748	0.0986
Lou Williams	0.015	−0.0288	0.0586	0.08	0.0751	0.0578	−1.5844	0.2312	0.648	0.1106
Luguentz Dort	0.023	−0.0079	0.0643	0.092	0.0824	0.0678	−1.0176	0.1847	0.8128	0.1061

Name	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Luka Doncic	0.024	0.0551	0.06	0.202	0.1195	0.0593	−1.079	0.1505	1.5277	0.0789
Luke Kennard	0.013	0.0072	0.0633	0.118	0.0464	0.0602	−1.7486	0.2681	1.1224	0.1082
Malachi Flynn	0.01	−0.0142	0.0672	0.114	0.2175	0.0665	−1.8989	0.3396	0.8717	0.1139
Malcolm Brogdon	0.029	0.1118	0.0636	0.118	0.134	0.0632	−0.814	0.1491	1.0622	0.0922
Malik Beasley	0.021	0.0317	0.0593	0.112	0.0765	0.0607	−1.2934	0.2053	0.9612	0.1163
Malik Monk	0.015	0.0221	0.0665	0.096	0.1154	0.0662	−1.6217	0.3033	0.675	0.1377
Marc Gasol	0.039	−0.0473	0.0702	0.174	0.2265	0.071	−0.6322	0.1834	1.3992	0.1065
Marcus Morris Sr.	0.025	0.0819	0.0645	0.133	0.1271	0.0648	−1.1327	0.189	1.2512	0.1034
Marcus Smart	0.022	−0.0049	0.059	0.084	0.1538	0.0612	−1.2206	0.1819	0.7262	0.1078
Markieff Morris	0.04	0.0817	0.0646	0.176	0.171	0.063	−0.5746	0.1703	1.4575	0.0991
Marvin Bagley III	0.092	0.1704	0.0655	0.188	0.1688	0.0647	0.4121	0.1327	1.576	0.0974
Mason Plumlee	0.095	0.0818	0.0688	0.264	0.2845	0.0687	0.4263	0.1109	1.8969	0.0787
Matisse Thybulle	0.023	0.1156	0.0625	0.07	0.0788	0.0623	−1.296	0.1983	0.5825	0.1304
Maurice Harkless	0.017	−0.0785	0.0681	0.097	0.0974	0.0672	−1.2388	0.2905	0.855	0.133
Max Strus	0.011	−0.0171	0.078	0.073	0.1337	0.0768	−1.9523	0.488	0.538	0.1878
Maxi Kleber	0.035	0.0173	0.0612	0.151	0.2551	0.0627	−0.7013	0.1705	1.1571	0.0961
Michael Carter-Williams	0.044	−0.0423	0.0656	0.123	0.1799	0.0685	−0.0913	0.1819	1.0563	0.1194
Michael Porter Jr.	0.051	0.014	0.0595	0.189	0.2002	0.0604	−0.2711	0.1288	1.627	0.0813
Mikal Bridges	0.035	0.0337	0.0683	0.099	0.0618	0.0656	−0.4679	0.1398	1.0212	0.0939
Mike Conley	0.024	0.0393	0.0614	0.082	0.123	0.0617	−1.1051	0.1808	0.7282	0.1127
Mike Muscala	0.028	−0.0435	0.0737	0.157	0.2426	0.0742	−0.8816	0.2493	1.3348	0.1216
Mike Scott	0.012	−0.1064	0.0654	0.129	0.2465	0.0671	−1.78	0.3348	1.1139	0.1279
Miles Bridges	0.041	0.076	0.0559	0.158	0.1094	0.0564	−0.6292	0.1382	1.2287	0.0939
Mitchell Robinson	0.124	0.2474	0.0695	0.155	0.2751	0.0713	0.5422	0.1333	1.2889	0.1137
Miye Oni	0.043	0.0516	0.0754	0.106	0.1336	0.0739	−0.5035	0.234	0.8945	0.146
Mo Bamba	0.091	0.1267	0.0684	0.24	0.2011	0.0699	0.3989	0.1454	1.8416	0.0953
Monte Morris	0.009	0.0561	0.0611	0.071	0.0571	0.0602	−2.2048	0.3093	0.5527	0.1272
Montrezl Harrell	0.1	0.1262	0.0656	0.167	0.1248	0.0662	0.3843	0.109	1.3561	0.0919
Moritz Wagner	0.035	−0.0238	0.069	0.149	0.1116	0.0678	−0.5898	0.2122	1.2883	0.1096

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Moses Brown	0.143	0.257	0.0672	0.238	0.2647	0.0697	0.942	0.1195	1.86	0.0962
Mychal Mulder	0.01	−0.0803	0.0703	0.062	0.0436	0.0656	−1.8496	0.3785	0.1905	0.1591
Myles Turner	0.043	0.0198	0.0648	0.159	0.1543	0.0626	−0.3671	0.1503	1.3487	0.0919
Naji Marshall	0.02	0.0725	0.0698	0.182	0.223	0.0715	−1.307	0.277	1.54	0.1166
Nassir Little	0.046	0.0563	0.067	0.144	0.1162	0.0727	−0.4408	0.2009	1.3249	0.1253
Naz Reid	0.054	0.0376	0.0617	0.173	0.194	0.0625	−0.3042	0.1306	1.37	0.105
Nemanja Bjelica	0.049	−0.0011	0.0713	0.172	0.1872	0.0734	−0.1829	0.2159	1.45	0.1276
Nerlens Noel	0.094	0.0388	0.0634	0.164	0.0659	0.064	0.4019	0.1179	1.4447	0.0928
Nic Claxton	0.074	0.1078	0.0747	0.187	0.2187	0.0737	0.0031	0.1928	1.324	0.111
Nickeil Alexander-Walker	0.012	0.073	0.0631	0.127	0.0691	0.064	−1.8277	0.2862	1.2062	0.1148
Nicolas Batum	0.029	0.0077	0.0631	0.143	0.0585	0.0604	−0.908	0.1597	1.3831	0.093
Nicolo Melli	0.046	0.1401	0.0696	0.166	0.1876	0.0734	−0.585	0.1993	1.3463	0.1233
Nikola Jokic	0.087	0.0955	0.0619	0.237	0.3003	0.064	0.2022	0.0968	1.8234	0.0748
Nikola Vucevic	0.059	−0.0218	0.054	0.289	0.3831	0.0586	0.0304	0.1032	1.9736	0.0613
Norman Powell	0.018	0.005	0.0535	0.079	0.0866	0.0532	−1.3732	0.1687	0.6068	0.092
O.G. Anunoby	0.037	−0.0825	0.0618	0.125	0.0846	0.0624	−0.5116	0.1582	0.9934	0.0923
Obi Toppin	0.037	−0.065	0.076	0.155	0.1423	0.0754	−0.6332	0.2293	1.2664	0.1205
Onyeka Okongwu	0.1	0.168	0.0717	0.172	0.1526	0.0709	0.3062	0.1637	1.3307	0.1264
Oshae Brissett	0.054	0.0433	0.0745	0.149	0.1359	0.0716	−0.0873	0.2218	1.2475	0.1356
Otto Porter Jr.	0.044	0.109	0.072	0.198	0.1686	0.0703	−0.4538	0.2084	1.5892	0.1158
P.J. Tucker	0.035	−0.0674	0.0618	0.113	0.0862	0.0603	−0.5158	0.1567	1.0289	0.0951
P.J. Washington	0.045	−0.0135	0.0585	0.163	0.1691	0.0598	−0.3964	0.1237	1.2519	0.0935
PJ Dozier	0.031	0.1257	0.0622	0.133	0.2064	0.0639	−0.9634	0.1832	1.1673	0.1013
Pascal Siakam	0.045	0.0815	0.0607	0.156	0.158	0.0597	−0.3759	0.1212	1.2215	0.0789
Pat Connaughton	0.039	0.0758	0.0572	0.152	0.1655	0.0571	−0.5072	0.1418	1.5063	0.0901
Patrick Beverley	0.037	0.0165	0.0737	0.108	0.1257	0.0689	−0.5583	0.209	1.0883	0.1313
Patrick Patterson	0.032	0.0197	0.0721	0.093	0.1019	0.0706	−0.891	0.2522	0.8215	0.1516
Patrick Williams	0.033	0.0408	0.0593	0.131	0.1958	0.0605	−0.6691	0.1388	1.0674	0.08
Patty Mills	0.01	0.0151	0.059	0.053	−0.0042	0.0591	−1.9469	0.2507	0.2316	0.1199

Name	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Paul George	0.026	0.0221	0.0599	0.171	0.2209	0.0588	−0.9804	0.1651	1.5431	0.0913
Paul Millsap	0.067	0.109	0.0614	0.167	0.1605	0.0653	−0.0886	0.1412	1.4586	0.0996
Payton Pritchard	0.025	0.0904	0.0594	0.099	0.1122	0.0624	−1.1593	0.1891	0.8998	0.1081
Precious Achiuwa	0.097	0.0921	0.071	0.174	0.1389	0.0701	0.3274	0.1562	1.4628	0.114
R.J. Hampton	0.03	−0.0496	0.0651	0.162	0.0056	0.0659	−0.7406	0.1997	1.4788	0.0975
RJ Barrett	0.027	0.1075	0.0659	0.13	0.0631	0.065	−0.9362	0.1363	1.174	0.086
Rajon Rondo	0.023	0.1111	0.0648	0.113	−0.0119	0.0666	−1.2723	0.2527	1.1243	0.1225
Raul Neto	0.019	0.0088	0.0589	0.087	0.1899	0.0617	−1.1788	0.1995	0.5976	0.1056
Reggie Bullock	0.008	−0.105	0.0634	0.102	0.1682	0.0647	−2.0362	0.2588	0.9235	0.0973
Reggie Jackson	0.016	−0.0505	0.0625	0.107	0.2152	0.0603	−1.4206	0.2204	0.9838	0.1052
Richaun Holmes	0.081	0.0578	0.0647	0.196	0.1783	0.0637	0.3378	0.1141	1.5918	0.0834
Ricky Rubio	0.014	0.0571	0.0634	0.11	0.1091	0.0632	−1.6838	0.1965	0.9844	0.1075
Robert Covington	0.026	0.0612	0.0581	0.183	0.2287	0.0608	−1.0126	0.1415	1.5907	0.081
Robert Williams III	0.136	0.2081	0.0662	0.232	0.1665	0.0672	0.5943	0.1263	1.8684	0.0948
Robin Lopez	0.097	0.1686	0.0618	0.092	0.2001	0.0644	0.4458	0.1171	0.6509	0.1048
Rodney Hood	0.023	−0.0427	0.063	0.084	0.0456	0.0648	−1.1545	0.2304	0.6866	0.1288
Royce O'Neale	0.039	0.0823	0.0689	0.162	0.2774	0.069	−0.5545	0.1302	1.4028	0.087
Rudy Gay	0.028	0.0052	0.0657	0.176	0.1491	0.0629	−0.8539	0.172	1.4675	0.0921
Rudy Gobert	0.11	0.1538	0.0702	0.288	0.2843	0.07	0.4523	0.0955	2.0079	0.0799
Rui Hachimura	0.026	−0.1041	0.061	0.139	0.2302	0.0617	−0.7642	0.165	1.1468	0.0834
Russell Westbrook	0.043	−0.02	0.0575	0.249	0.1518	0.0588	−0.2961	0.1187	1.801	0.067
Saben Lee	0.026	−0.1221	0.0675	0.099	0.09	0.0682	−0.7777	0.2334	0.8101	0.1289
Saddiq Bey	0.022	−0.0503	0.0588	0.145	0.1225	0.058	−1.1139	0.1694	1.2419	0.0808
Sekou Doumbouya	0.045	0.0186	0.0691	0.117	0.0152	0.0663	−0.3456	0.1861	1.0146	0.114
Semi Ojeleye	0.032	0.0186	0.0619	0.121	0.1783	0.0657	−0.8484	0.2041	1.1084	0.1106
Serge Ibaka	0.08	0.0373	0.0732	0.217	0.2757	0.073	0.2287	0.1572	1.8003	0.1075
Seth Curry	0.006	−0.0227	0.0664	0.074	0.0946	0.0692	−2.3256	0.3299	0.729	0.118
Shai Gilgeous-Alexander	0.015	−0.0392	0.0687	0.119	0.0949	0.0664	−1.4508	0.2427	1.0753	0.1067
Shake Milton	0.022	0.0607	0.0618	0.076	0.0405	0.0658	−1.2998	0.1912	0.6767	0.1212

	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
Name	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Solomon Hill	0.028	0.0106	0.0607	0.111	0.0999	0.0608	−0.8409	0.1758	0.9652	0.1016
Stanley Johnson	0.027	−0.022	0.0629	0.128	0.1448	0.0626	−0.9056	0.196	1.014	0.107
Stephen Curry	0.013	0.1057	0.0686	0.135	0.0719	0.0683	−1.5844	0.202	1.1341	0.089
Sterling Brown	0.029	−0.0218	0.062	0.147	0.1061	0.0608	−0.8156	0.1836	1.2565	0.0893
Steven Adams	0.132	0.2314	0.0687	0.187	0.302	0.0702	0.6941	0.1049	1.4976	0.0952
Svi Mykhailiuk	0.017	0.0531	0.0561	0.104	0.0909	0.0593	−1.4053	0.2176	0.9233	0.0994
T.J. McConnell	0.03	0.0749	0.0594	0.106	0.0452	0.0579	−0.8625	0.1528	0.8674	0.0922
Taj Gibson	0.102	0.1403	0.0721	0.151	0.2048	0.0715	0.3896	0.1353	1.217	0.1128
Talen Horton-Tucker	0.019	0.0012	0.0614	0.104	0.0465	0.0624	−1.281	0.2176	0.8904	0.104
Taurean Prince	0.021	−0.0189	0.0627	0.132	0.1041	0.0625	−1.1601	0.2415	1.1111	0.11
Terance Mann	0.044	0.1487	0.061	0.14	0.0869	0.0613	−0.5517	0.1532	1.3128	0.0987
Terence Davis	0.018	0.0151	0.0614	0.125	0.1101	0.0607	−1.3948	0.235	1.0772	0.1041
Terrence Ross	0.01	0.0372	0.0586	0.104	0.0815	0.0601	−1.8711	0.2704	0.9215	0.0996
Terry Rozier	0.019	0.0293	0.0618	0.105	0.0922	0.0609	−1.2708	0.1652	0.8062	0.0983
Thaddeus Young	0.105	0.2313	0.061	0.154	0.1031	0.0627	0.4625	0.107	1.2227	0.0829
Thanasis Antetokounmpo	0.093	0.163	0.0717	0.113	0.0516	0.0708	0.3831	0.171	1.1242	0.1396
Theo Maledon	0.012	−0.041	0.0586	0.101	0.1004	0.0608	−1.5908	0.2111	0.9394	0.0949
Tim Hardaway Jr.	0.01	−0.0417	0.0585	0.101	0.0185	0.0589	−1.9584	0.2396	0.7713	0.0932
Timothe Luwawu-Cabarrot	0.026	0.0972	0.0609	0.092	0.131	0.0594	−1.0584	0.2057	0.6182	0.1142
Tobias Harris	0.032	−0.0258	0.0646	0.17	0.2338	0.066	−0.7062	0.1445	1.5648	0.0924
Tomas Satoransky	0.024	0.0184	0.0631	0.087	0.1281	0.0648	−1.0722	0.1985	0.6493	0.1118
Tony Bradley	0.121	0.1801	0.0665	0.224	0.2481	0.0698	0.6972	0.1475	1.7074	0.1051
Tony Snell	0.019	0.0554	0.0656	0.09	0.0807	0.0671	−1.3107	0.2583	0.8302	0.13
Torrey Craig	0.07	0.068	0.0668	0.185	0.0733	0.0691	0.1574	0.1645	1.7322	0.1101
Trae Young	0.018	0.0411	0.0648	0.094	0.0113	0.0639	−1.2077	0.1766	0.875	0.0984
Trevor Ariza	0.033	−0.031	0.0708	0.15	0.1785	0.071	−0.5859	0.2226	1.3166	0.115
Trey Burke	0.011	0.0483	0.066	0.047	−0.0204	0.064	−1.9939	0.3364	−0.0487	0.1638
Tristan Thompson	0.124	0.1791	0.0626	0.212	0.304	0.0661	0.5972	0.1102	1.6903	0.0892
Troy Brown Jr.	0.039	0.0675	0.0725	0.158	0.098	0.0712	−0.6105	0.2593	1.343	0.1339

Name	Off. Team			Def. Team			Off. Ind.		Def. Ind.	
	Prior Mean	Post. Mean	Post. SD	Prior Mean	Post. Mean	Post. SD	Post. Mean	Post. SD	Post. Mean	Post. SD
Ty Jerome	0.012	0.0468	0.0688	0.1	0.0967	0.0687	−1.7301	0.3353	0.8386	0.1317
Tyler Herro	0.015	0.06	0.0585	0.144	0.039	0.061	−1.5802	0.2156	1.3329	0.0925
Tyler Johnson	0.009	−0.0115	0.0703	0.096	0.0879	0.0693	−2.2049	0.4271	0.6433	0.1344
Tyrese Haliburton	0.022	0.0697	0.0592	0.076	0.0026	0.0569	−1.1107	0.1754	0.6235	0.1077
Tyrese Maxey	0.011	0.0264	0.0647	0.098	0.0675	0.066	−1.9732	0.328	0.8982	0.1274
Tyus Jones	0.016	−0.0312	0.0639	0.093	0.0034	0.0655	−1.4191	0.2213	1.0482	0.1187
Victor Oladipo	0.01	−0.1616	0.0657	0.137	0.1214	0.061	−1.7417	0.3015	1.215	0.0962
Wayne Ellington	0.012	−0.0207	0.0658	0.073	0.0518	0.0662	−1.7645	0.2975	0.5384	0.1315
Wendell Carter Jr.	0.091	0.1437	0.0616	0.219	0.2016	0.0641	0.341	0.117	1.6417	0.078
Wes Iwundu	0.025	0.0225	0.0702	0.13	0.09	0.0704	−1.0729	0.2797	1.1248	0.1355
Wesley Matthews	0.018	0.0238	0.0628	0.064	0.0878	0.0634	−1.4165	0.2394	0.3501	0.1385
Will Barton	0.024	0.0605	0.0641	0.109	0.0701	0.0642	−1.1886	0.1793	1.0396	0.0971
Willie Cauley-Stein	0.073	−0.0405	0.0675	0.175	0.2161	0.068	0.1685	0.1558	1.3212	0.1023
Willy Hernangomez	0.134	0.1295	0.0679	0.247	0.195	0.0673	0.8133	0.1334	1.9628	0.1036
Xavier Tillman	0.065	0.0585	0.0637	0.163	0.1091	0.0659	−0.0338	0.1386	1.5696	0.1079
Yuta Watanabe	0.046	0.1252	0.0656	0.171	0.1732	0.0694	−0.4448	0.1828	1.3626	0.1105
Zach LaVine	0.018	−0.0331	0.0604	0.121	0.0972	0.059	−1.2693	0.1811	1.0133	0.079
Zion Williamson	0.081	0.1062	0.0581	0.131	0.1096	0.0615	0.2404	0.1039	1.1836	0.0937

Bibliography

1950–51 NBA Season Summary (n.d.). Basketball-Reference.com. URL: https://www.basketball-reference.com/leagues/NBA_1951.html.

Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Hoboken, NJ: Wiley. ISBN: 978-1118730058.

Alagappan, M. (2013). 'From 5 to 13: Redefining the positions of basketball'. In: *MIT Sloan Sports Analytics Conference*. Massachusetts Institute of Technology, Cambridge, MA.

Bornn, L. et al. (2016). 'Studying basketball through the lens of player tracking data'. In: *Handbook of Statistical Methods for Design and Analysis in Sports*. Chapman & Hall/CRC. ISBN: 978-1315166070.

Carpenter, B. et al. (2017). 'Stan: A probabilistic programming language'. In: *Journal of Statistical Software* 76.1, pp. 1–32. DOI: 10.18637/jss.v076.i01. URL: https://www.jstatsoft.org/index.php/jss/article/view/v076i01.

Daunizeau, J. (2017). *Semi-analytical approximations to statistical moments of sigmoid and softmax mappings of normal variables*. arXiv: 1703.00091 [stat.ML].

Dempster, A. P., N. M. Laird and D. B. Rubin (1977). 'Maximum likelihood from incomplete data via the EM algorithm'. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 39.1, pp. 1–38. ISSN: 00359246. URL: http://www.jstor.org/stable/2984875.

Deshpande, S. K. and A. Wyner (2017). 'A hierarchical Bayesian model of pitch framing'. In: *Journal of Quantitative Analysis in Sports* 13.3, pp. 95–112. DOI: doi:10.1515/jqas-2017-0027. URL: https://doi.org/10.1515/jqas-2017-0027.

Dice, L. R. (1945). 'Measures of the amount of ecologic association between species'. In: *Ecology* 26.3, pp. 297–302. URL: http://www.jstor.org/stable/1932409.

Engelmann, J. (2016). 'Possession-based player performance analysis in basketball (adjusted +/– and related concepts)'. In: *Handbook of Statistical Methods and Analyses in Sports*. Chapman and Hall/CRC. ISBN: 978-1315166070.

Falk, B. (2021). 'LINEUPS 2020-21 SEASON'. In.

Franks, A. et al. (2015). 'Characterizing the spatial structure of defensive skill in professional basketball'. In: *The Annals of Applied Statistics* 9.1, pp. 94–121. DOI: 10.1214/14-AOAS799. URL: https://doi.org/10.1214/14-AOAS799.

Hosmer, D. W., S. Lemeshow and R. X. Sturdivant (2013). *Applied Logistic Regression*. Hoboken, NJ: Wiley. ISBN: 978-0470582473.

Kalman, S. and J. Bosch (2020). 'NBA lineup analysis on clustered player tendencies: A new approach to the positions of basketball and modeling lineup efficiency of soft lineup aggregates'. In: *MIT Sloan Sports Analytics Conference*. Massachusetts Institute of Technology, Cambridge, MA.

Miskin, M. A., G. W. Fellingham and L. W. Florence (2010). 'Skill importance in women's volleyball'. In: *Journal of Quantitative Analysis in Sports* 6.2. DOI: doi:10.2202/1559-0410.1234. URL: https://doi.org/10.2202/1559-0410.1234.

Oliver, D. (2004). *Basketball on paper*. Potomac Books, Sterling, VA. ISBN: 978-1574886887.

Patel, S. (2023). *nba api*. https://github.com/swar/nba_api.

Rosenbaum, D. T. (2004). *Measuring how NBA players help their teams win*. URL: http://www.82games.com/comm30.htm#_ftn1 (visited on 18/01/2023).

Sill, J. (2010). 'Improved NBA adjusted +/- using regularization and out-of-sample testing'. In: *Proceedings of the 2010 MIT Sloan Sports Analytics Conference*. Massachusetts Institute of Technology, Cambridge, MA.

Vaci, N. et al. (2019). 'Large data and Bayesian modeling – aging curves of NBA players'. In: *Behavior Research Methods* 51.4, pp. 1544–1564. DOI: 10.3758/s13428-018-1183-8.

Vehtari, A. et al. (2021). 'Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion)'. In: *Bayesian Analysis* 16.2, pp. 667–718. DOI: 10.1214/20-ba1221. URL: https://doi.org/10.1214%5C%2F20-ba1221.

Woolner, K. (2002). 'Understanding and measuring replacement level'. In: *Baseball prospectus 2002*. Brassey's, pp. 455–466.