# Replaying the NBA

Nathan Sandholtz
Simon Fraser University
nsandhol@sfu.ca

Luke Bornn
Simon Fraser University, Sacramento Kings
lbornn@sfu.ca

**Abstract**

The Cleveland Cavaliers took 329 contested mid-range jump shots with over 10 seconds remaining on the shot clock during the 2015-2016 regular season. What could've happened if they had taken these shots 20% less frequently over the season? We attempt to answer these types of questions by modeling plays from the 2015-2016 NBA regular season as episodes from team-specific Markov decision processes. Using STATS SportVU optical tracking data, we model the transition probabilities as a tensor indexed in time in order to simulate plays with dynamic probabilities across the shot clock. To culminate, we simulate seasons under altered shot policies of interest within the basketball analytics community and explore the net changes in efficiency and production under these alternative shot policies.

## 1 Introduction

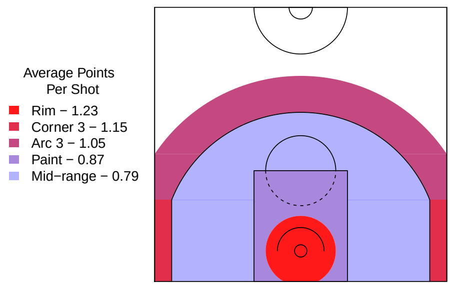
A basketball game is a collection of finite stochastic processes; each play is comprised of a finite number of transitions between players and locations, ultimately terminating in a shot, turnover, or foul. An integral attribute of these processes, however, is that they are non-stationary; the transition probabilities are not constant over the 24 second shot clock. Surprisingly, this non-stationarity often gets overlooked when evaluating measures of efficiency and production.

To illustrate this phenomenon, consider the familiar breakdown of shot efficiency in Figure 1(a). On average across the NBA in the 2015-2016 regular season, three pointers and shots in the restricted area were the most valuable shots, while mid-range shots were the least valuable. Juxtaposed with the corresponding breakdown of shot volumes for these regions shown in Figure 1(b), this appears to signify major inefficiency in the league's overall shot selection.

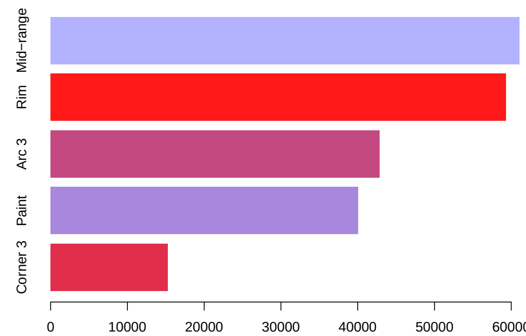
A primary key to understanding this apparent disparity is to recall the rule imposed by the 24 second shot clock. Figure 2 shows the league average empirical shot policies for each court region. We define a shot policy as the probability that any on-ball event (i.e. dribbles, passes, and shots) will be a shot as a function of the shot clock. As the shot clock winds down, the probability of shooting increases — quite dramatically in the final seconds of the shot clock. Taking this into account, the figures above make more sense; naturally teams would rather take a shot, even if it's a poor one, than get no points at all.



2018 Research Papers Competition
Presented by:




**(a)** 2015-16 NBA league average shot efficiency




**(b)** 2015-16 NBA league-wide shot volumes

**Figure 1**

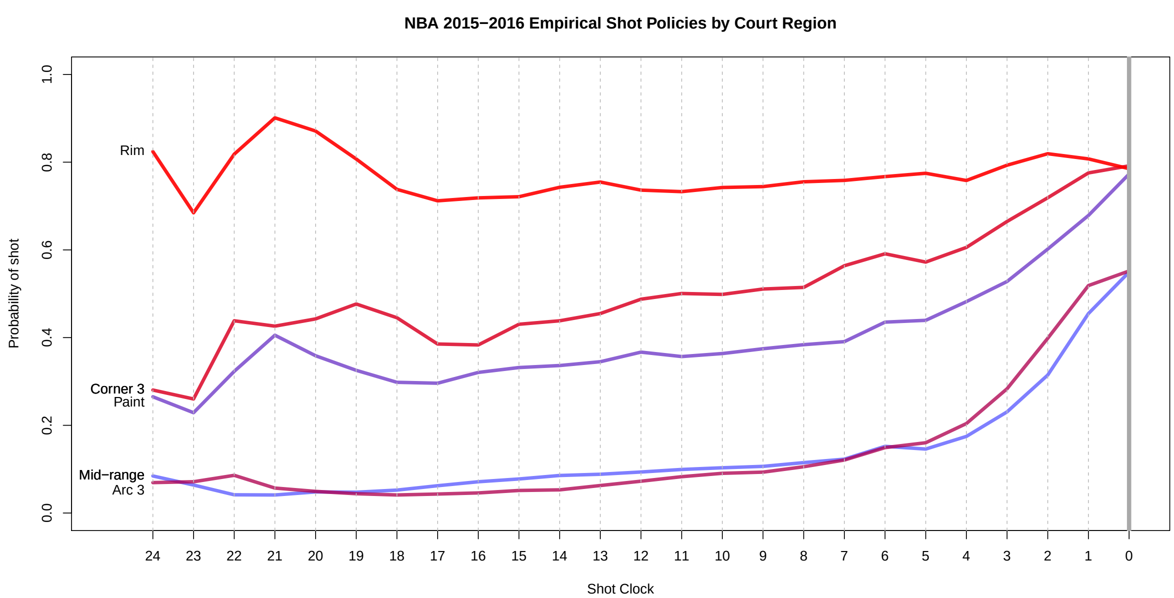



**Figure 2:** 2015-2016 NBA empirical league average shot policies. We see lower probability of shots in the mid-range and arc 3 areas because the on-ball events in these regions are dominated by passes and dribbles. Interestingly, we clearly see the impact of fast breaks in the paint and rim shot policies early in the shot clock.

The existence of the shot clock makes evaluating shot selection a far more difficult task. We cannot simply conclude that teams should take fewer mid-range jumpers — we have to consider *when* and *whom* should take fewer mid-range jumpers and how these changes would effect the team's efficiency and production. This is the key point of interest in this research project. While we focus on shot policies in this paper, narrowing in on a player's choice to shoot or not at any given instant, the framework presented here easily extends across the space of possible decisions players can make on-ball, including movement and passing.

In pursuit of answers, we have developed statistical methods to simulate from finite non-stationary stochastic processes in a basketball context. The main idea is to construct a method to imitate plays with respect to time for any given team. This entails simulating plays not simply by outcome, but rather at the sub-second level, incorporating every intermediary and terminal on-ball event over the course of a play. After the process is calibrated, we explore the consequences of altering shot probabilities over specific intervals in time. To this end, we model plays from the 2015-2016 NBA regular season as episodes realized from team-specific non-stationary Markov decision processes. Since time lapses between events are incompatible with parametric forms, we rely on simulation to ultimately quantify the effects of changes to a team's shot policy.

The rest of the paper is outlined as follows. In Section 2, we give a brief overview of Markov decision processes (MDP), provide a summary of related work, and describe our data and model. In Section 3, we describe our MDP-based play simulator and show the results of our simulations under two altered policies in comparison to the corresponding observed policies for all teams in the league. We provide a brief conclusion in Section 4.

# 2 Methods

## 2.1 Markov Decision Processes

A finite Markov decision process is a framework utilized in many modern reinforcement learning problems which characterizes the interactions between an agent and its environment. The "Markov" qualifier assumes that the agent takes actions based solely on the current state of the environment (which can include a fixed lag). We will represent a MDP as a tuple $\langle \mathbf{S}, \mathbf{A}, T(\cdot), R(\cdot), \pi(\cdot)\rangle$. **S** represents a discrete and finite set of states. **A** represents the set of actions the agent can take. $T(\cdot)$ defines the transition probabilities between states at any given step in the process. $R(\cdot)$, the reward function, defines the reward the agent receives for any given state/action pair. Finally, the policy, $\pi(\cdot)$, governs the probability that the agent takes any given action based on the current state of the environment. $\pi(\cdot)$ is the only aspect of the system which the agent controls. The goal of the agent is to maximize his rewards, which he does by altering his policy. We can define these functions succinctly in mathematical terms:

$$T(s, a, s') = \mathbb{P}[\mathbf{S_{t+1}} = s' \mid \mathbf{S_t} = s, \mathbf{A_t} = a] \tag{1}$$

$$R(s, a) = \mathbb{E}[\mathbf{R_{t+1}} \mid \mathbf{S_t} = s, \mathbf{A_t} = a] \tag{2}$$

$$\pi(s, a) = \mathbb{P}[\mathbf{A_t} = a \mid \mathbf{S_t} = s] \tag{3}$$

In basketball terms, $\pi(s, a)$ simply represents the probability that the ball-carrier takes a shot given his current state. If he takes a shot, $R(s, a)$ dictates the expected point value of that shot. If he decides not to shoot, $T(s, a, s')$ denotes the probabilities of the ball entering any other state given his current state. For example, if LeBron James has the ball at the top of the arc and he decides not to shoot, $T(s, a, s')$ governs the probability of the next state of the ball — perhaps a dribble in the same location, or a pass to a teammate in a different court region. To illustrate the Markov decision process in context of a basketball play, consider the graph in Figure 3.

In most settings the functions $T(\cdot)$, $R(\cdot)$, and $\pi(\cdot)$ are assumed to be stationary, however in this paper we assume that only the reward function $R(\cdot)$ is time-independent. We refer the reader to [10] for a more expansive introduction to reinforcement learning and Markov decision processes.

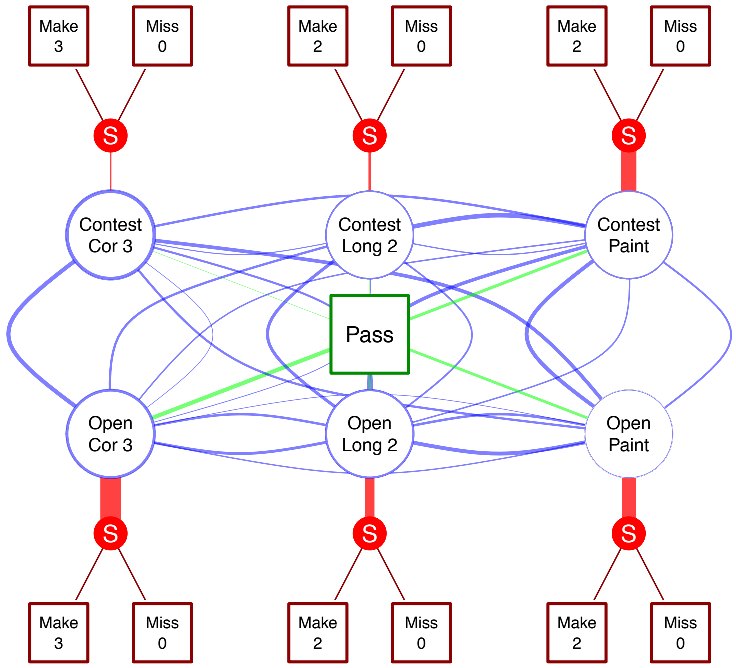



**Figure 3:** A toy graph to illustrate the components of the MDP for a single player in context of a basketball play. The blue circles represent states, the red circles represent actions (shots), the web of curved lines represent transition probabilities between states, and the red squares show the set of possible rewards given a shot. The red lines of varying width connecting the blue state circles to the red action circles represent the policy. Players may pass the ball to another player (not shown) which is also considered a transition to another non-terminal state.

## 2.2 Related Work

Markov models have been utilized in varying contexts in several sporting domains. Through the use of a semi-Markov process, Thomas [11] describes times between goals scored in hockey, and uses this to demonstrate that scoring a goal has an effect of shortening the remainder of an NHL game. Routley and Schulte [9] apply a Markov game formalism to value player actions in hockey, incorporating context and a lookahead window in time. In [6] Goldner uses a Markov model to provide a framework for evaluating plays in American football. In their degree project [3], Damour and Lang model set pieces in soccer using a Markov model to estimate transition probabilities. Peña uses Markov processes to model possessions and their outcomes in the English Premier League in [8].

The landmark work of Cervone et al. in [2, 1] is perhaps most relevant to the methods we explore in this paper. The state space of our model is similar to the coarsened model they use, but the main similarity with their work is the use of a non-stationary Markov model. In contrast to most Markov models in sports, they account for the non-stationarity inherent in a basketball possession by expanding their state-space to include time. Additionally, the hazard model they use explicitly defines transition probabilities as a continuous function of space and time.

The work of Franks et al. in [4, 5] was also influential in this project, particularly regarding our shot efficiency model. Finally, our methods rely on simulating episodes from Markov processes to generate results. In [7] Min-hwan et al. simulate realizations from stationary Markov models under differing line-up combinations to estimate game outcomes.

### 2.3 Description of Data

We use high-resolution spatio-temporal tracking data collected by Stats LLC for the 2015-2016 NBA regular season. These data include the $x, y$ coordinates of all 10 players on the court and the $x, y, z$ coordinates of the ball at 25 observations per second. These data are merged with play-by-play data yielding additional features including play-action events such as shots, passes, dribbles, fouls, etc. Figure 4 shows a two second snapshot of the data during a game in March 2016. For our project we only use observations with tagged ball-events. This includes dribbles, passes, rebounds, turnovers, and shots. This significantly reduces the number of observations while retaining the core structure of a play.

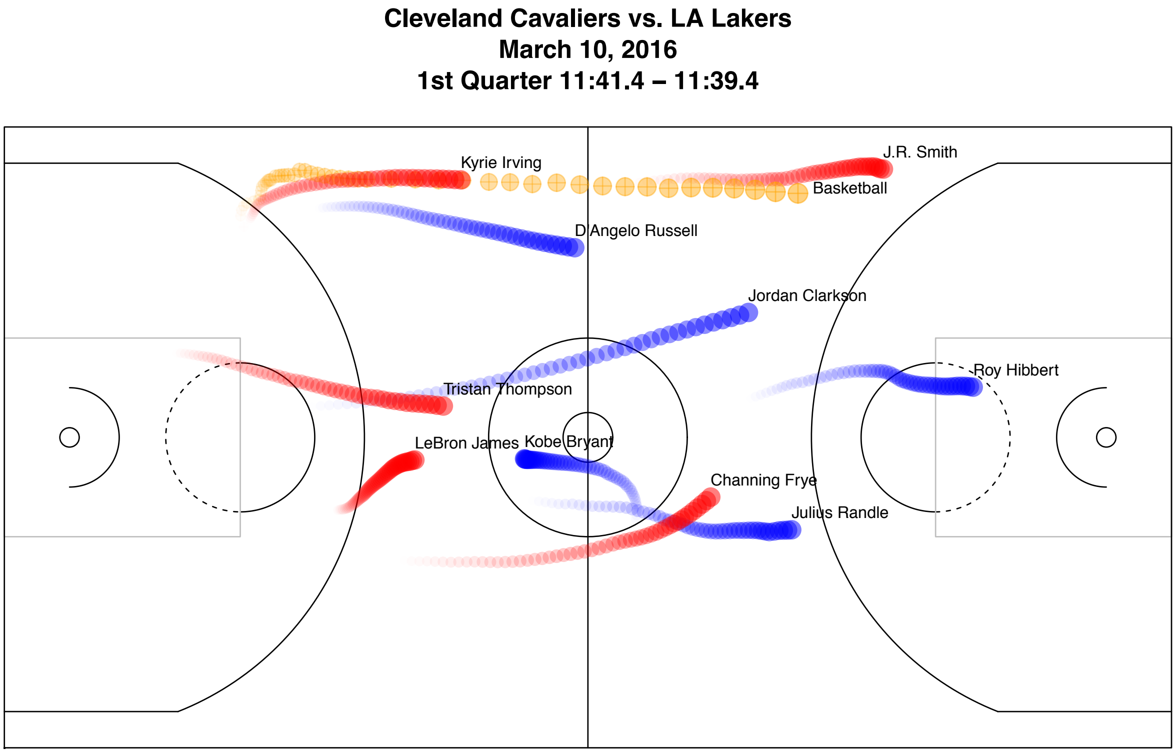



**Figure 4:** Two seconds of data in the first quarter of Cavaliers vs. Lakers on March 10, 2016. Cleveland's Kyrie Irving begins with possession of the ball, then passes to JR Smith who is in position to drive to the basket or take a 3-point shot. Laker Jordan Clarkson scrambles to defend him, while Roy Hibbert is moving to defend the rim.

### 2.4 State and Action Space

The state space of our model, **S**, is defined in context of the ball-carrier. At any time $t$, the state is given by the identity of the ball-carrier, his court region, and an indicator variable of defensive pressure. Court region is a function of the $x, y$ coordinates of the ball-carrier and can take any of the six regions shown in Figure 1. Defensive pressure is determined by the distance of the nearest defender to the ball-carrier and is dependent on the court region of the ball-carrier.

Since we are primarily interested in shot policies, we've chosen a binary action space; at each step in the process, the ball-carrier decides to either shoot or not shoot ($\mathbf{A} = \{\text{shoot, not shoot}\}$) based on their policy $\pi(\cdot)$. If a shot is taken, the play terminates, otherwise the subsequent transition is assumed to come from $T(\cdot)$.

## 2.5 Transition Probability and Shot Policy Tensors

In order to incorporate non-stationarity in $T(s, a, s')$ and $\pi(s, a)$, we propose using a tensor to model the transition probabilities and shot policies. This allows us to consider each transition probability as a non-linear function of time, with the constraint that at any discrete step in time, the slice of the tensor representing time $t$ is a valid transition probability matrix (TPM). Specifically, we model both tensors as a collection of 12 matrix slices, each slice representing a two second interval of the shot clock as illustrated in Figure 5.

Transition Probability Tensor

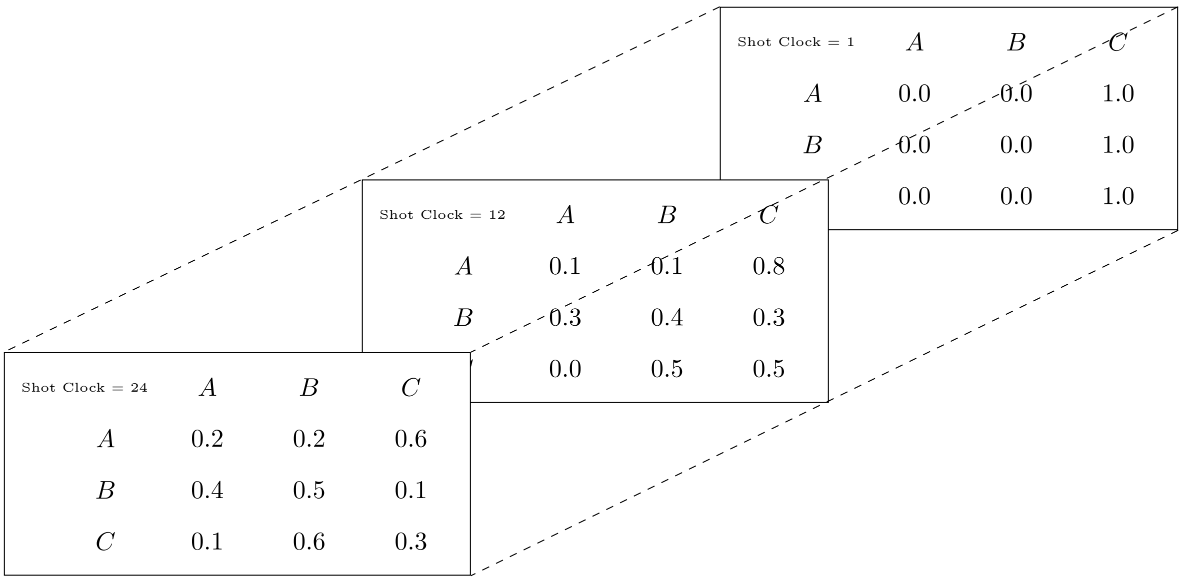



**Figure 5:** The shot policy tensor and transition probability tensor are collections of 12 transition probability matrices each representing a two second interval of the shot clock. The states and corresponding probabilities shown here are purely illustrative.

This tensor framework is the key to accurately exploring the effects of altering inefficient shot policies. As mentioned in the introduction and illustrated in Figure 2, the efficiency of a shot is dependent on the time remaining on the shot clock. The tensor we use allows us to correctly account for the dynamic nature of transition probabilities and tailor our policy alterations accordingly. The shot policy tensor is virtually identical to the transition probability tensor in form. The only difference is the column space; since the agent (player) makes only a binary decision at every step of the process, given any time $t$, the shot policy is a matrix slice with row space equal to the row space of the corresponding transition probability matrix slice and a column space of length two (shot & no shot). We estimate a team's shot policy tensor using the same methods for estimating the transition probability tensor and consider them jointly moving forward.

We estimate a team's transition probability tensor from their observed transition counts indexed at two second intervals of the shot clock. At time $t$, each row of transition counts $\mathbf{y}_{(i,\cdot,k)}$ is modeled using a multinomial likelihood:

$$f(\mathbf{y}_{(i,\cdot,k)}|\mathbf{p}) = \prod_{j=1}^{J} (p_{ijk})^{y_{ijk}} \,. \tag{4}$$

Here, $i$ indexes the starting state, $j$ indexes the new state, and $k$ indexes time. We estimate the tensor probabilities empirically using the team transition count tensor modulated by the league transition probability tensor (also estimated using league wide empirical counts):

$$\hat{p}_{ijk} = \frac{y_{ijk} + \hat{p}_{lmk}^{\text{League Avg}}}{\sum_{k=1}^{K}(y_{ijk} + \hat{p}_{lmk}^{\text{League Avg}})} \tag{5}$$

The purpose of the modulation term $\hat{p}_{lmk}^{\text{League Avg}}$ in (5) is two-fold. First, it ensures that $\hat{p}_{ijk} > 0$ for all cells in a team's transition probability tensor, which in turn makes (4) a valid probability distribution for any potential state of the MDP. We define our league state-space at a lower level, using court region/position combinations, indexed by $l$ and $m$ above (using the classical position assignments of center, power forward, small forward, shooting guard, and point guard).

This leads to the second purpose of the modulation term; modulating transition counts $y_{ijk}$ by $\hat{p}_{lmk}^{\text{League Avg}}$ yields more plausible probabilities for low-minute players. For most players, the transition counts between any two states overwhelm the modulation piece $\hat{p}_{lmk}^{\text{League Avg}}$ which is by definition less than 1. However for low-minute players who may have few (or no) observed transition counts from state $i$ to state $j$, the modulation piece pulls their estimated transition probability toward the league average for their position.

## 2.6 Reward Function

Recall that the reward function is an expected value:

$$R(s,a) = \mathbb{E}[\mathbf{R_{t+1}} \mid \mathbf{S_t} = s, \mathbf{A_t} = a]. \tag{6}$$

In context of a basketball play, (6) can be restated as, "How many immediate points do we expect when a player in state $s$ takes action $a$?" If the action is a shot, then this expected value is his expected points per shot from the given state. If the action is not a shot, this expected value is almost (but not exactly) 0. However, in our analysis we have omitted all plays that incorporated foul shots and offensive rebounds for pedagogical simplicity. By doing so, when $a \neq shot$ this expected value is exactly 0, which in turn allows us to define the reward function of the MDP completely in terms of a shot efficiency model.

### Shot Efficiency Model

Given a shot, we model the probability of a make as a function of the player's latent skill from the location of the shot and whether or not the shot was contested. We estimate these latent skill parameters using a Bayesian logistic regression model with a hierarchy that borrows strength across players with similar shooting characteristics. For shot $n$,

$$p(Y_n = 1|S_n, c_n, \boldsymbol{\theta}, \boldsymbol{\xi}) = \frac{\exp(\theta_{kb} + I(c_n)\xi_b)}{1 + \exp(\theta_{kb} + I(c_n)\xi_b)} \tag{7}$$
$$\theta_{kb} \sim \mathrm{N}(\beta_{gb}, \sigma^2), \;\; \beta_{gb} \sim \mathrm{N}(0, \tau^2), \;\; \xi_b \sim \mathrm{N}(0, \nu^2).$$

$Y_n$ indicates whether the shot was made, $S_n$ denotes the state at the time of the shot (player $k$ in location $b$), and $c_n$ indicates whether the shot is contested. $\theta_{kb}$ represents the latent shooting skill of player $k$ in location $b$ and $\xi_b$ is a global effect for defensive pressure in court region $b$. The skill parameters have a hierarchical prior — each player is member of a group $g$, which groups were created using the k-means clustering algorithm with 8 groups. Shot efficiency is then determined by scaling the estimated make-probabilities for each state by the corresponding point value of the shot (2 or 3 points, depending on court-region). Figure 6 shows an example of the posterior uncontested shot efficiency for Lebron James.

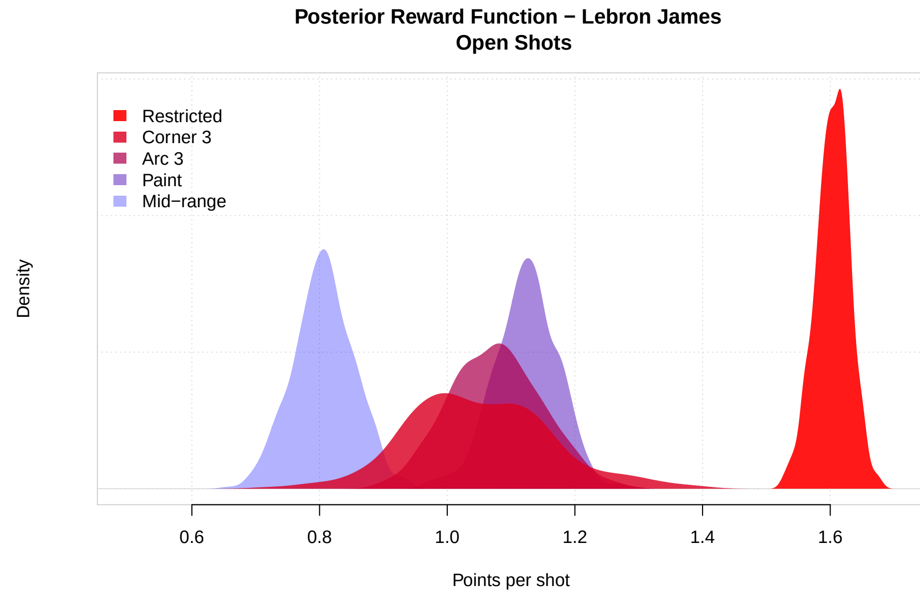



**Figure 6:** Lebron James' posterior uncontested shot efficiency for all court regions. As illustrated here (and for nearly all players) the shot efficiency posterior variance for 3-point shots is considerably greater than it is for 2-point shots.

# 3 Simulating Plays

## 3.1 Building the Simulator

We now come to the crux of the paper — simulating team specific plays under alternate shot policies. The simulation algorithm takes as inputs the components mentioned above: the transition probability tensor, the shot policy, and posterior draws from the shot efficiency model. Additionally, our simulator requires initial states and starting shot clock times for all the plays we want to simulate. For these inputs we use the observed starting states and corresponding times on the shot clock for each team's collection of plays in the 2015-2016 regular season. Lastly, we need a mechanism to take time off the shot clock at each step in the Markov process. This component of the simulator makes an analytic solution to this problem intractable; the distribution of time lapses between events does not lend itself to a parametric distribution. For this reason, we sample the empirical distribution of time-lapses between events (with replacement) as a mechanism to simulate the time between events in our play simulator.

To check the calibration of the simulator, we simulated 50 seasons across the entire league and checked correlation between the simulated transition counts and observed transition counts. The simulations match on multiple metrics; for example two point shots ($\rho$ = 0.997), three point shots ($\rho$ = 0.999), and turnovers ($\rho = 0.998$).

We now have all the tools to answer the questions posed in our introduction. Before providing results, however, we want to stress the potential of the framework we have built at this point. The possibilities we can explore are limitless. For example, we could explore the effects across a season of LeBron taking uncontested arc-threes more frequently. We could test the effects of certain players on a team shooting mid-range jump shots less frequently early in the shot clock while increasing the shot probability for other players at the rim or beyond the arc. With only minor changes to our framework, we could include the decision to pass in our action space and explore changes to both shot and pass policies.

### 3.2 Attenuating mid-range shots: a conservative approach

For now, we restrict our attention to mid-range jump shots taken early in the shot-clock, which are generally regarded as less efficient shots relative to other court regions. Specifically, we modify a team's original shot policy to reduce contested mid-range jump shots by 20% while there are more than 10 seconds remaining on the shot clock. We chose a conservative reduction factor for two reasons. Firstly, we want to give ample respect to players' decisions; players make decisions on the court with far more contextual information than we have in the data. Despite the fact that these shots are the least efficient on average, assuming that they were all poor decisions without looking at the video is presumptuous. Secondly, there are potential game-theoretic consequences at play here that are difficult to anticipate. A major change to a team's offensive strategy would naturally lead to different patterns in how the defense responds, which could in turn render the transition probabilities of our altered transition probability tensor inaccurate. Because of this, we believe that testing minor perturbations to a team's policy will yield more credible results.

After reducing the contested mid-range shot probability in a row of the transition probability tensor, the multinomial model becomes invalid because the row no longer sums to 1. To account for this, we renormalized the perturbed row to make it valid probability distribution again. We will speak more to how the probabilities get redistributed subsequently.

We simulated 500 seasons for each team under their original policy and the altered policy defined above. Figure 7 shows the distribution of transition counts to terminal states aggregated over different indices of interest for both the original and altered policies for the Cleveland Cavaliers.

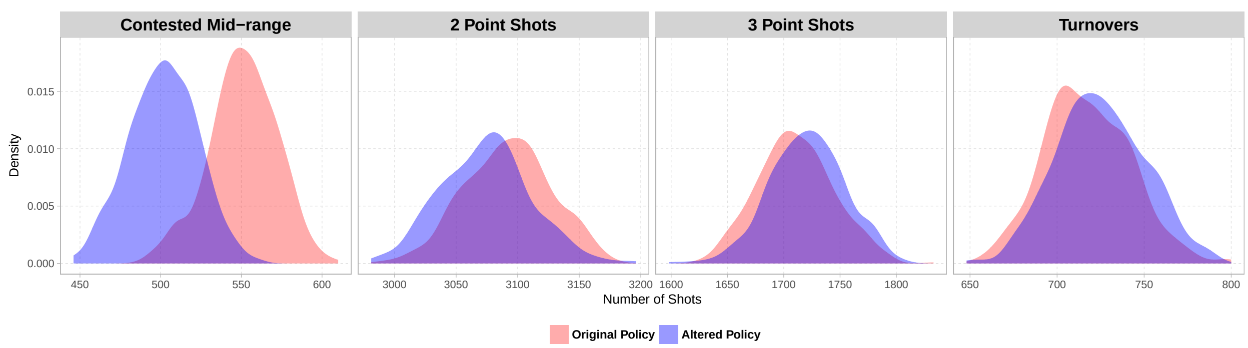



**Figure 7:** Distribution of simulated transition counts to four terminal states (contested mid-range jump shots, total 2-point shots, total 3-point shots, and total turnovers) under Cleveland's orginal and altered policies. The altered policy reduces contested mid-range jump shots by 20% while there are more than 10 seconds remaining on the shot clock.

The most obvious distinction in the four plots in Figure 7 is the distance between the two contested mid-range distributions, which is not surprising since this is the transition probability that we directly tampered with in the altered policy. More interesting are the consequences illustrated in the third and fourth plots. Taking contested mid-range shots less frequently not only leads to an increase in three point attempts, but also an increase in turnovers and shot clock violations. This is because reducing the probability of shooting leads to longer possessions, which in turn create more opportunities for turnovers.

The most important factor to quantify is how the altered policy affects efficiency and production. To measure these effects we restrict our attention to the expected differences in points per shot and points per 100 plays. For each team's 500 simulated seasons under their observed and altered policies we calculated the mean points per shot and mean points per 100 plays and formed confidence intervals for these means using bootstrap samples from the simulations. We then subtracted the sample means of each team's original policy simulations from the alternate policy simulations to clearly illustrate the differences between the policies. The results are shown in Figure 8.

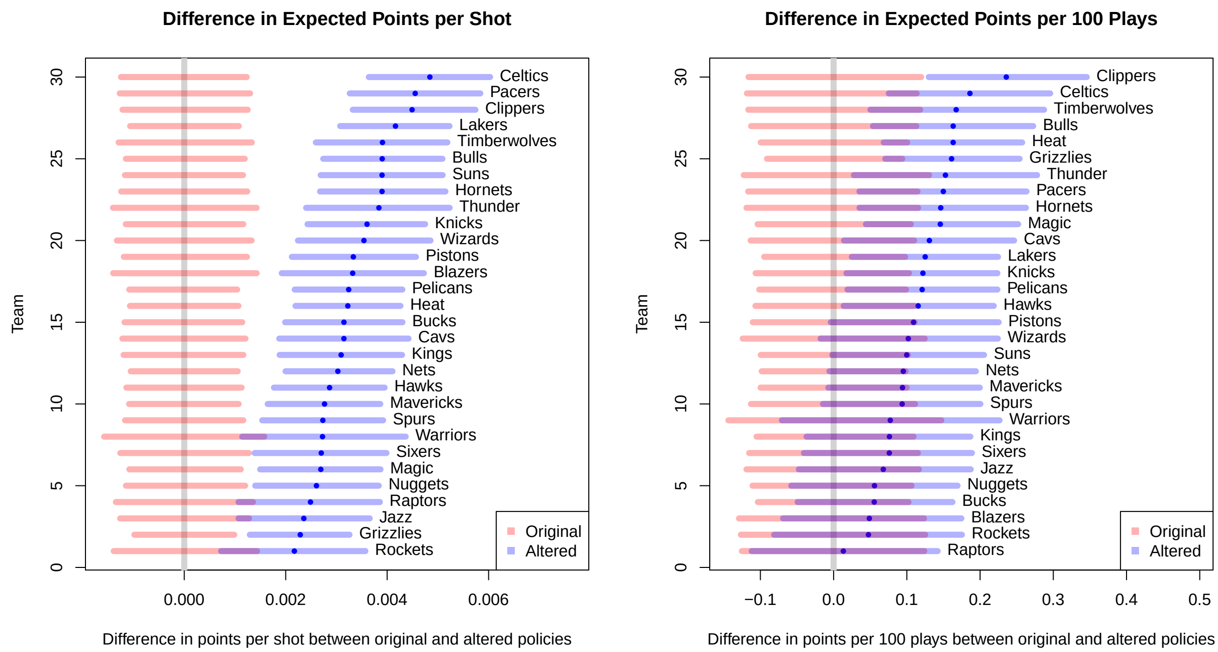



**Figure 8:** Differences in expected points per shot and expected points per 100 plays for all teams from the 2015-2016 NBA season between the original policy (red) and the conservative mid-range alternate policy (blue). Each point estimate has a corresponding 95% confidence interval estimated using bootstrap samples of the simulated season metrics. The altered policy reduces contested mid-range jump shots by 20% while there are more than 10 seconds remaining on the shot clock.

Not surprisingly, there are significant gains in expected points per shot for almost every team. By reducing the frequency of the (on average) least efficient shot, team's expected points per shot should increase. As expected, the Rockets, who are infamous for leading the trend in avoiding mid-range jump shots, are at the bottom of the list in these differences. Since the Rockets take the fewest mid-range shots in the league, it follows that this policy change would have a smaller effect on them relative to other teams. Other teams near the bottom are teams who play at a notoriously slow pace and simply don't take many shots early in the shot clock regardless of court region. Because our altered policy only modified shot probabilities early in the shot clock, these teams would naturally be less affected.

On the other hand, according to our results, the Celtics and Clippers could benefit significantly from this altered shot policy. The Clippers in particular are the only team where the confidence interval for the increase in expected points per 100 plays does not overlap with the confidence interval of the original policy. Despite differences in magnitude, we want to stress that this conservative policy change appears to benefit all teams in expectation.

Notice that the ordering of the teams isn't the same in both plots. This results from the defining feature of our model; non-stationarity. Essentially, each team's pace of play is automatically built into our simulator, which consequently effects turnover rates (particularly shot clock violations) differently across teams despite each team receiving the same policy change.

Finally, we want to make careful note here that it is the right hand plot — the difference in expected points per 100 plays — that contains the most important information on whether the altered policy is beneficial or not. Even if the expected points per shot is much higher under an altered policy, if the corresponding production is less than that of the original policy, then the increase in shot efficiency is not worth the loss in points resulting from increased turnovers. However, in these cases, policy changes could be more nuanced; teams could consider not only *decreasing* an inefficient shot policy, but additionally *increasing* the shot policy in more efficient states, such as threes and shots at the rim. The next section provides an example of how a seemingly beneficial policy change could have adverse effects if the changes weren't so carefully designed.

## 3.3 Eliminating mid-range shots: the dark side of Moreyball

The subtitle for this section is admittedly tongue-in-cheek, but our methods do suggest that when Moreyball (i.e. eliminating mid-range jump shots) is taken to the limit without increasing shot probabilities in more efficient states, the consequences could be detrimental. Consider the following altered (albeit extreme) policy: all mid-range shots get taken 90% less frequently regardless of defensive pressure and time on the shot clock. Figure 9 shows the results of this extreme policy in terms of expected shot efficiency and production.

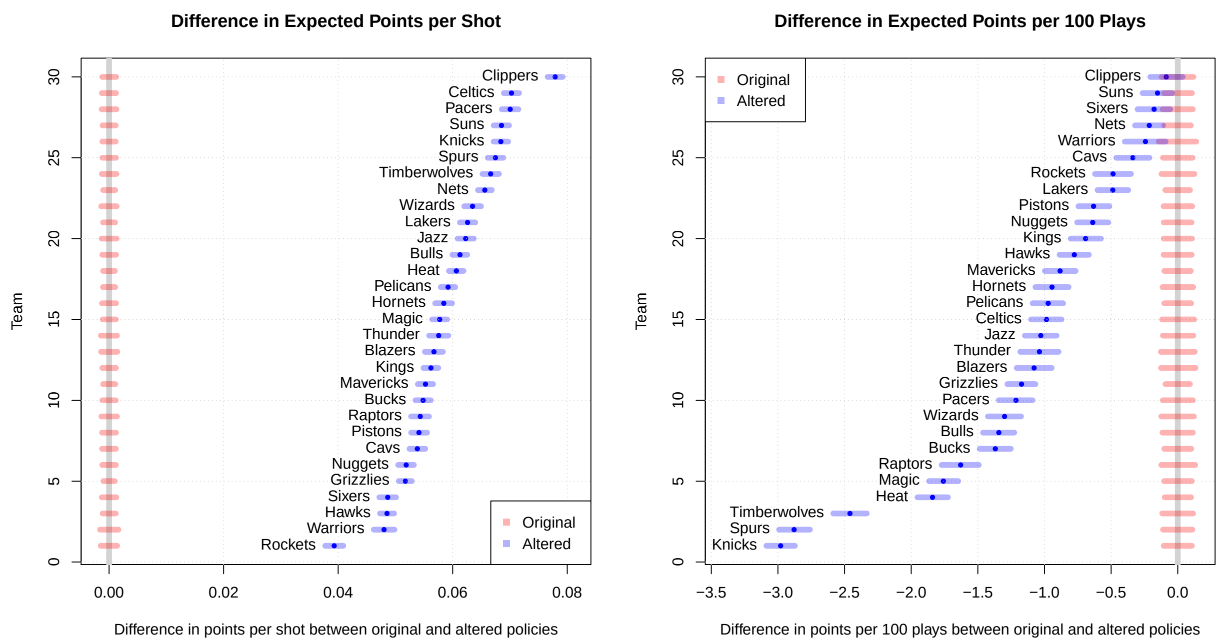



**Figure 9:** Differences in expected points per shot and expected points per 100 plays for all teams from the 2015-2016 NBA season between the original policy (red) and extreme mid-range policy (blue). Each point estimate has a corresponding 95% confidence interval estimated using bootstrap samples of the simulated season metrics. The altered policy in this case reduces all mid-range jump shots by 90% irrespective of defensive pressure and time remaining on the shot clock.

Shot efficiency sky rockets across the board (ironically, yet not surprisingly, the least so for the Rockets), but this time there is an expected overall *decrease* in production for every team! Under this more extreme policy, the increase in turnovers due to waiting for more efficient shots while passing up mid-range jump shots outweigh the benefits of the better shot selection. It's interesting to note which teams are affected the most by this policy change — the offensive production of the Knicks, Spurs, and Timberwolves take the biggest hit, indicating that they rely heavily on the mid-range shot in their offenses. Teams such as the Clippers and Suns would be less negatively affected by this policy change, suggesting that only slightly less drastic changes in their mid-range shot policies could yield more productive offenses.

Again, we are quick to note that these results are by no means deterministic — realistically a policy change of this magnitude would involve significant strategic changes among all players and court regions. In fact, Houston provides an excellent example; their franchise has had success in their offense not simply because they avoid mid-range shots, but because they've specifically designed offensive schemes to *increase* shot probabilities in more efficient states. Our results indicate that coaches cannot simply tell their players to drastically reduce their mid-range shot frequency and expect the team's offense to improve. On the contrary, offensive tactical changes should be crafted with precision and awareness of the downstream consequences in time.

## 4 Conclusion

We have successfully developed and implemented a method to test the impact of team shot policy adjustments over the course of a season at an unprecedented level of detail. With the example of the conservative mid-range jump shot altered policy, we show that even a minor policy change could result in significant improvement in both offensive efficiency and production for every team in the league. In the more extreme example, our results illustrate how policy changes that may seem advantageous on the surface could have unanticipated detrimental consequences.

These examples are only the tip of the iceberg in terms of how these methods could be utilized. Using our methods, teams could assess proposed strategy changes outside of games rather than risking poor results by testing them in games. Without the ability to run controlled experiments coaches have little information to use in estimating the effect of new offensive strategies. Our methods could help bridge the gap in anticipating and predicting the impact of proposed changes.

Another useful application of these methods would be to forecast how injuries (or potential injuries) to players could impact a team's season. Injuries represent perhaps the biggest factor of uncertainty that front offices face when building a roster. Our methods could help them better quantify this uncertainty by allowing them to estimate the potential effects of having to play second and third string players if any starters suffered a long term injury.

Finally, we have considered only shooting decisions in this introductory work, but our methodology could naturally scale to include all different types of basketball decisions, allowing coaches and analysts to explore incredibly nuanced tactical changes. Additionally, with tracking data now available for most major sports including hockey, football, and soccer, our methods could extend to testing decision policies in other sporting environments. In future work, we hope to further illustrate the vast potential our methods could unlock.

## References

[1] Dan Cervone, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data. *SLOAN Sports Analytics Conference*, pages 1–9, 2014.

[2] Daniel Cervone, Alex D'Amour, Luke Bornn, and Kirk Goldsberry. A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes. *Journal of the American Statistical Association*, 111(514):585–599, 2014.

[3] Gabriel Damour and Philip Lang. Modelling Football as a Markov Process: Estimating Transition Probabilities Through Regression Analysis and Investigating it's Application to Live Betting Markets. 2015.

[4] Alexander Franks, Andrew Miller, Luke Bornn, and Kirk Goldsberry. Characterizing the spatial structure of defensive skill in professional basketball. *Annals of Applied Statistics*, 9(1):94–121, 2015.

[5] Alexander Franks, Andrew Miller, Luke Bornn, and Kirk Goldsberry. Counterpoints : Advanced Defensive Metrics for NBA Basketball. *MIT Sloan Sports Analytics Conference*, pages 1–8, 2015.

[6] Keith Goldner. A markov model of football: Using stochastic processes to model a football drive. *Journal of Quantitative Analysis in Sports*, 8(1), 2012.

[7] Min-hwan Oh, Suraj Keshri, and Garud Iyengar. Graphical Model for Baskeball Match Simulation. *MIT Sloan Sports Analytics Conference*, pages 1–8, 2015.

[8] Javier López Peña. A Markovian model for association football possession and its outcomes. 2014. URL http://arxiv.org/abs/1403.7993.

[9] Kurt Routley and Oliver Schulte. A Markov Game Model for Valuing Player Actions in Ice Hockey. *Uncertainty in Artificial Intelligence (UAI)*, pages 782–791, 2015.

[10] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1.

[11] Andrew C Thomas. Inter-arrival Times of Goals in Ice Hockey. *Journal of Quantitative Analysis in Sports*, 3(3), 2007.