52892
April 18, 2016
MCS 100
Homework 1

Using Expected Point Values per Possession and Plus-Minus to Evaluate Basketball Players

Introduction

In any sport, evaluating the performance impact of a given player towards his or her team's chance of winning typically begins by identifying key performance indicators of winning games. For example, in baseball the most basic object of the game is for a team to score more runs than their opponent and is accomplished by placing and moving men on base while avoiding outs. Time segmented sports, such as baseball, have a distinct advantage over time continuous sports, such as basketball and football, because clear segmentation of the game allows for an easier digestion of the game flow and each player's consequential impact.

In time-continuous sports, the game is typically dynamic and flows around interdependent player-to-player interactions that make it difficult to clearly understand a player's contributions towards his or her team's overall performance. Consider for a moment one of the National Basketball Association's [NBA] all-time greatest three-point shooters Ray Allen. Not known for consistent playmaking abilities, Allen's contribution to his team is being a reliable three-point shooter and is dependent on his teammates passing him the ball. If given space and the ball, Allen will likely make his three-point attempt, but his teammates' pass or potential screen to open up space for Allen to take the shot will not be reflected in the three points he adds to the team's score. Allen's shot will help his team towards the object of the game – score more points than the opposing team – yet could not have happened without his teammates' contributions. So how can all player's actions and subsequent impacts be measured?

In this paper, accessing and valuing a given basketball player's value towards his team is further explored. Current player evaluation statistics, such as John Hollinger's PER or Joe Schaller's TPR, attempt to assign player value by taking traditional box score statistics [PTS, REB, AST, etc.] by using inconsistent mathematical manipulations that place subjective weights on each statistical category. Yet, these methods seem to fail to capture the dynamic and integrated game flow by using statistics that do not capture the entirety of all player interactions and contributions on the court. This paper explores and proposes an alternative method to evaluate a basketball player's on-court value by breaking down a basketball game into its fundamental component of possessions through a subsequent event tree describing every possible outcome. A player's on-court performance can then be valued as an average of their expected point value and their plus-minus rating per possession. Combining individual expected point value per possession along side overall team performance creates an understanding of a player's value.

Methodology

The method proposed in this paper explores player contribution by segmenting a basketball game into its most fundamental aspect: team possessions. Even as a time-continuous event, a basketball game alternates possessions and, as a result, a game's result can be determined by the aggregate of each team's possession efficiency. The object of each possession is to maximize points, but how it happens can come in a variety of ways with varying end possession point values. As described in Figure 1, at its most basic level a possession can result in a turnover, a non-shooting foul or a shot attempt and each action has a subsequent set of sub-actions that could occur. A possession can end in points scored, a continuation of the possession [OREB] or the end of the possession [DREB or TOV]. In basketball, these actions have relative occurrence rates that can be determined through data analysis, which sets ups the ability to create an event tree and subsequent expected point values given a particular action occurs.

Using NBA league data, an expected point value contribution per possession for an average NBA player can be found and then subsequent individual player expected values can then be found and compared to league

averages. By establishing a league average, a given player can be compared to the average and then relatively to his to his peers. Higher EPVs are better, but returning to the Ray Allen scenario above, it does not necessarily capture a total picture of a player's value. To address this shortcoming, a player's plus-minus statistic is aggregated in to the EPV. It is done so by taking the league average plus-minus and then normalizing each individual player's plus-minus using the league average. A player's normalized plus-minus is added to his EPVs for a total value representing his individual impact per possession in conjunction with overall team performance towards winning basketball games.

Comparable Values and Studies

The methodology is partially inspired by Joseph Kuehn's paper, "Accounting for Complementary Skill Sets When Evaluating NBA Players' Values to a Specific Team", that was presented at the 2016 MIT Sloan Sports Analytics Conference. In his paper, Kuehn attempts to estimate a player's value by looking at probable outcomes, corresponding expected value points scored per possession and the corresponding teammates and defenders influence. Combined it presents a substantial method of evaluating a player's impact and corresponding impact given a random set of teammates. Kuehn's work focuses on understanding the impact between a given set of players in their relative skill set and how that impacts a given lineup's expected value of points per possession.

From a statistical perspective, current metrics used to measure a basketball player's value can be separated into two different areas: team performance and calculated value estimates. Player's value in terms of overall team performance is done so with basic/adjusted plus minus, one of the more commonly used metrics. At a basic level, plus-minus evaluates the net game point differential from when a player enters and exits a game. Adjusted plus-minus takes into account relative opposing team strength among other factors. Secondary player evaluation metrics use some type of weighted combination of traditional box score statistics that are added and/or multiplied into a singular number. Examples include John Hollinger's Player Efficiency Rating [PER].

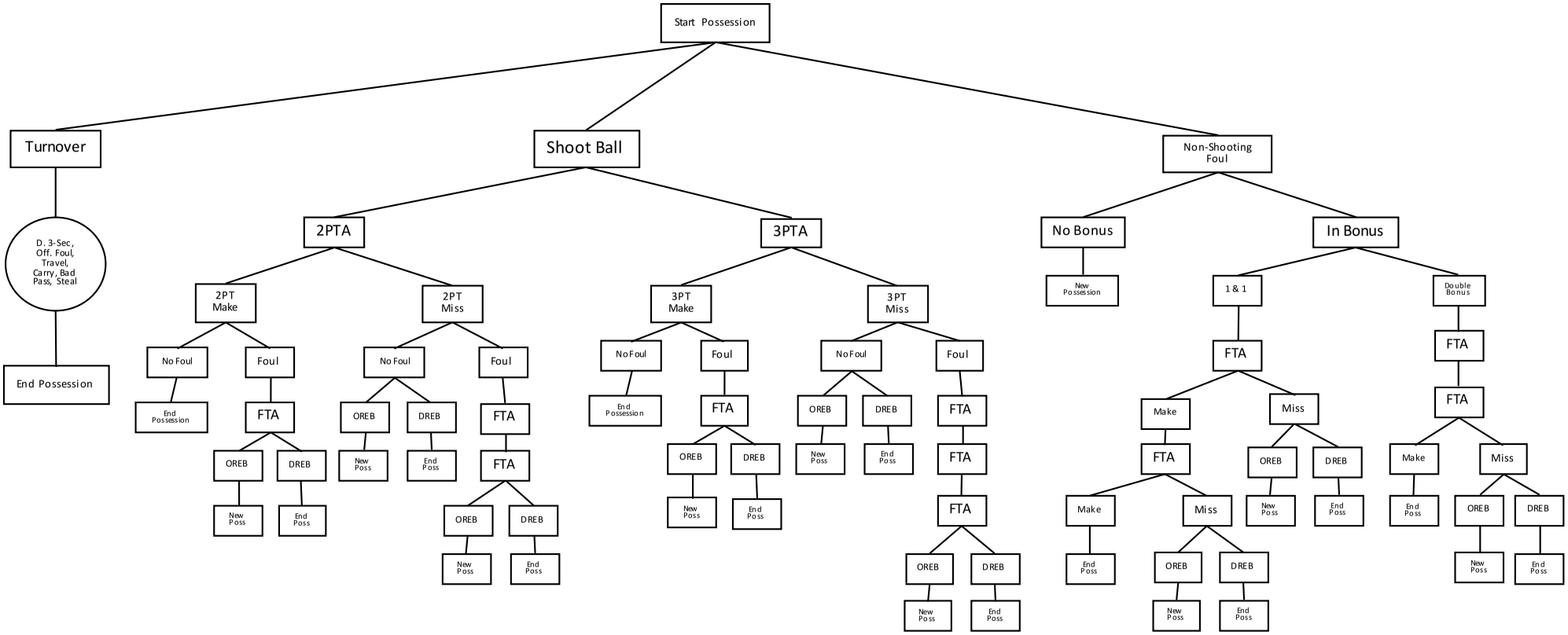
Data Sources

For this method, the most important data required is NBA league and individual player totals for possessions, two-point and three-point attempts, fouls drawn and turnovers. Together these will allow to develop relative percentages for a given possession in the event tree, shown in Figure 1. In addition, subsequent data on league and individual shooting percentages [2PT, 3PT, FT] and rebounding [OREB%, DREB%] for a given possession will finalize the event tree. This data can be tabulated through a thorough analysis of NBA play-by-play data, in R, calculating the occurrence of each action, which can be found on NBAstuffers.com. League plus-minus data can be found at NBA.com/stats. Another potential source could be SportsVU data, this needs more research.

Limitations and Improvements

The most significant limitation of this player evaluation methodology is it still does not take into account in its entirety off-the-ball actions by players. Hypothetically, in this proposed model, a player who only sets ball-screens to open up a three-point shooter on offense will have an EPV of zero, despite being a critical piece to creating space for the three-point shooter to take and make his shot. The current method tries to incorporate a player's normalized plus-minus rating to take into account this limitation. However, this issue with this adjustment is that plus-minus takes into account team defensive performance too. In the hypothetical scenario, if the perpetual ball-screen player plays on a poor defense, his plus minus will not completely reflect his value. In general, this aspect of the plus-minus statistic limits the completeness of this player evaluation metric since the EPV is based strictly on offensive possession performance.

This methodology could significantly improve its utility if the EPV could take into account the defensive of a given player. The resulting reported player evaluation metric could be a net differential, for a total value, or separated out into offense and defensive EPV depending on what type of player a team is looking to evaluate.


Start Possession
Turnover
D. 3-Sec, Off. Foul, Travel, Carry, Bad Pass, Steal
End Possession
Shoot Ball
2PTA
2PT Make
No Foul
End Possession
Foul
FTA
OREB
New Poss
DREB
End Poss
2PT Miss
No Foul
OREB
New Poss
DREB
End Poss
Foul
FTA
OREB
New Poss
DREB
End Poss
3PTA
3PT Make
No Foul
End Possession
Foul
FTA
OREB
New Poss
DREB
End Poss
3PT Miss
No Foul
OREB
New Poss
DREB
End Poss
Foul
FTA
FTA
FTA
OREB
New Poss
DREB
End Poss
Non-Shooting Foul
No Bonus
New Possession
In Bonus
1 & 1
FTA
Make
FTA
Make
End Poss
Miss
OREB
New Poss
DREB
End Poss
Miss
OREB
New Poss
DREB
End Poss
Double Bonus
FTA
FTA
Make
End Poss
Miss
OREB
New Poss
DREB
End Poss