International Journal of Sport Finance, 2012, **7**, 324-339, © 2012 West Virginia University

How Much Trouble is Early Foul Trouble? Strategically Idling Resources in the NBA

Allan Maymin[1], Philip Maymin[2], and Eugene Shen[1]

[1] AllianceBernstein
[2] NYU-Polytechnic Institute

Allan Maymin is a quantitative developer at AllianceBernstein. He is a passionate fan of the NBA and is enthusiastic about the growing role of quantitative basketball analytics in the professional leagues.
Philip Maymin is an assistant professor of finance and risk engineering at NYU-Polytechnic Institute. His research interests include algorithmic finance, behavioral finance, quantitative finance, and sports analytics.
Eugene Shen is a senior vice president and portfolio manager at AllianceBernstein. His research interests include sports analytics and quantitative finance.

Abstract

We analyze a large and comprehensive play-by-play dataset of professional games in the National Basketball Association using tools from financial economics to explore the optimality of strategically idling resources in the face of uncertain future demand. We find that starters ought to be idled by the coach on a "Q+1" basis, meaning that a starter has one more foul than the current quarter, when the future option value is high or the value of the replacement player is high. We use a novel win-probability approach that can be easily extended to other applications.

Keywords: strategic idling, foul trouble, win probability, NBA

Introduction

In the face of uncertain demand, should productive but fragile resources be strategically idled to preserve their future option value? Directly testing such choices in operations research is empirically difficult, because of the inability to conduct controlled experiments or observe natural ones. To address this empirical challenge, we use a multi-year play-by-play dataset of professional games in the National Basketball Association ("NBA"). In NBA games, a player can commit at most six personal fouls before he is disqualified from the game, where a foul is an infraction charged to the player by a referee. Thus the players, specifically the starting five players on each team in each game, are the productive but fragile resource. Coaches can decide to bench or "yank" their star players early in the game if they are afraid that the player may foul out too early, and coaches may make these substitution decisions within a game at their discretion, without restriction. A player may be substituted into or out of the game, during a stoppage in play, any number of times.

To address the question of idle resources, we ask in our context: should starters in early foul trouble be benched? The decision facing an NBA coach is analogous to the cost-flexibility tradeoffs studied in the operations literature. In the language of Fine and Freund (1990), a firm (coach) may pay a current cost (benching his starter) to acquire flexible capacity (opportunity to bring the starter back in the game) so that he is able to respond to future uncertain product demand (use of the players best suitable for end-of-game matchups).

The advantage of yanking is that the starter will likely be able to play at the crucial end of the game but the disadvantage is that he may not play as many minutes as he otherwise would. On the other hand, if a starter is kept in the game, he may not play at his full potential, as the opposing team tries to induce him to commit another foul. In other words, the player himself may choose to strategically idle himself while playing in order to maximize his minutes played. Thus, the coach must evaluate these tradeoffs between idling starters on the one hand and reaping their productivity on the other.

We use a novel win-probability technique to measure the impact of foul trouble on a team's probability of winning. This win-probability technique is sufficiently general to be useful for other questions by appropriately redefining the state variables. Furthermore, we show that the underlying distribution of the probability of winning is more akin to the jump-diffusion model of Merton (1976) that allows for discontinuous jumps than a pure diffusion model. Thus, our four contributions are to introduce the new win-probability approach for evaluating the impact of earlier decisions on later success, document the jump-diffusion characteristics of NBA games, apply the win-probability approach to the problem of early foul trouble that had remained unaddressed in the academic literature, and shed light on the general question of when it may be optimal to idle productive but fragile resources in the face of uncertain demand.

We find that teams generally perform better if foul-troubled starters are benched. In other words, idling resources can be a good strategy. We propose two possible reasons why yanking can be the correct strategy. First, benching a player preserves option value, since late-game situations may favor using certain players over others. Second, players in foul trouble may play worse than they otherwise would, and in many cases, sufficiently worse that the optimal decision is to bench them until they are no longer in foul trouble. Although both factors may be at work, we argue that the first may be responsible for the second: players who know they are likely to be benched once they are in foul trouble may rationally choose to play more tentatively once they are in foul trouble in an attempt to maximize their own minutes, possibly at the expense of the overall victory.

Background

We frame our discussion with a poignant example. In Game 2 of the 2010 NBA Finals between the LA Lakers and the Boston Celtics, Laker superstar Kobe Bryant picked up his fifth foul with 11:15 remaining in the fourth quarter. Lakers Coach Phil Jackson removed Bryant from the game at the 8:11 mark, and reinserted him in the game with 6:16 remaining and the game tied at 85-85. After he re-entered the game, Bryant appeared to play tentatively.

Aschburner (2010) noted that "[Kobe's] desire to stick around can cause him to back off ever so slightly, avoid a defensive encounter, drop his attack gear. Play c-a-r-e-f-u-l-l-y. And self-consciously, in a way that can distract even an assassin from his task." Adande

(2010) further observed that Kobe "couldn't afford to be aggressive. … Bryant took a little longer to initiate his moves, surveying the defense to determine where the late help would come from. He tried jab-stepping a couple of times to see if he could force their hand. … When he did go on the attack, he had to pull up early, leading to some amazing examples of body control (and yet another head-scratching call, on Kendrick Perkins, for an and-one after Bryant specifically stopped short of him to avoid contact)."

The Celtics won the game 103-94, outscoring the Lakers 18-9 in the final six minutes. After the game, Kobe conceded that when playing with foul trouble, "You've just got to be careful" (Adande, 2010).

It can be argued that Bryant's play runs counter to theoretical notions of how players in foul trouble should optimally behave. For example, Weinstein (2010a) argues that benching a starter is suboptimal if the coach's objective function is to maximize the number of minutes played by a starter. Weinstein (2010b) addresses the counterargument that players play worse when they are in foul trouble. He contends that if playing tentatively is better than playing as if the player has no foul trouble, then it must be the case that the team is even better off by letting the starter play. While Weinstein analyzes the foul trouble problem using a theoretical approach, we approach the problem from an empirical point of view. A coach's objective function is not to maximize the number of minutes played by the team's starters. Instead, the coach's objective function is to win the basketball game. Our win-probability framework directly measures the impact of leaving a starter in the game on the probability of winning a game. We find that in most cases, keeping the starter on the bench preserves option value for the latter stages of a game. Every player has different strengths and weaknesses, and coaches value the flexibility of employing the best matchups late in a game. When a player fouls out, the coach has one less option at his disposal. Thus, preserving option value can be more important than maximizing starter minutes. Finally, even if players theoretically should play as if they have no foul trouble, many times they do not, as our Kobe Bryant example shows.

Coaches may have different definitions of how many fouls constitute "trouble," and whether it is optimal to yank or not. Many coaches seem to bench on a "Q+1" basis, i.e., when players commit one more foul than the current quarter. We find that yanking on a "Q+1" basis is generally a good strategy. There are, however, nuances to be considered. Coaches must also consider the time remaining in the game, and the quality of the player with foul trouble. Early in the game, benching a player preserves "option value" since the coach can reinsert a fresh, non-foul plagued starter back into the game in the fourth quarter. Thus, yanking makes more sense early in the game. Also, the coach must evaluate the quality of the player relative to his potential substitute. If the starter is not sufficiently better than his substitute, then he ought to be benched to preserve his option value. On the other hand, a superstar in foul trouble (e.g., Kobe Bryant) may still be better than a fresh bench player. In this case, the coaches may be better off keeping the starter in the game.

Our results can be placed in context relative to the literature on in-game decision making by coaches, starting with Morrison (1976), which compares the coaching strategy of pulling a hockey goalie (and replacing him with an extra skater) at time versus never pulling him, and Morrison and Wheat (1986), which also allows for pulling the goalie later and investigates the optimal time for pulling the goalie. Here, instead of examining time directly, we consider the conditional effect of fouls, and because there is no special goalie position in basketball, we analyze all positions. Washburn (1991) uses a dynamic programming approach to find the optimal decision to pull a goalie. Here we use tools from finance in application to the problem of yanking foul-troubled starters.

Some research suggests behavioral anomalies in sports decision making. Romer (2006) shows that American football teams fail to maximize their probability of winning by playing too conservatively on fourth downs. Staw and Hoang (1995) show that playing time in the NBA is distributed suboptimally, namely with a heavy impact of sunk costs as measured by initial draft position. In that spirit, our results here help indicate potentially better coaching strategies in basketball, though our focus is not on documenting a similar league-wide failure, but rather on evaluating the differences in performance conditional on different yanking strategies.

Price and Wolfers (2010) use box score summary statistics from NBA games to analyze racial discrimination in the context of fouls called. By contrast, we use the higher frequency play-by-play data and we ignore the racial characteristics of referees and players altogether.

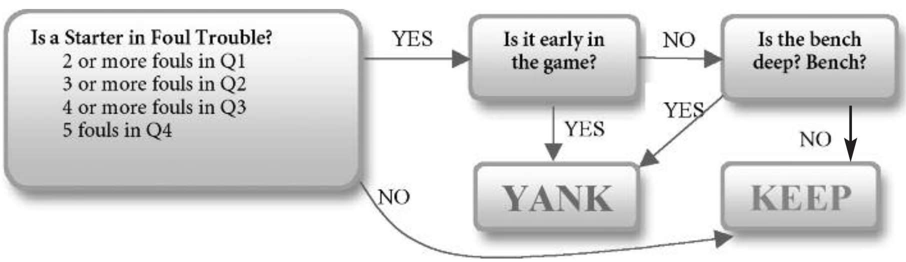

Figure 1: Should a Starter in Foul Trouble be Benched? A Flowchart

Moskowitz and Wertheim (2011) examine the impact of fouls on star players and conclude that they should not be yanked when they are in foul trouble. Their conclusion seemingly runs counter to ours and deserves further discussion. Moskowitz and Wertheim focus on superstars in the fourth quarter and find that they play better with foul trouble. When we use an advanced plus-minus framework (cf. Rosenbaum, 2004), we also find that players seem to do better when they have foul trouble. How do we reconcile this finding with our win probability results?

First, we find the effect of players playing better with foul trouble in the fourth quarter only. We agree with Moskowitz and Wertheim that foul-troubled superstars ought not be benched in the fourth quarter, as shown in the flowchart, because even in foul trouble they play better than their replacements and the remaining option value is negligible. In other circumstances, however, we believe that yanking is optimal.

Second, it is possible that foul-troubled starters cause a team to play better locally but ultimately hurt the team's chances of winning. An OLS method captures only the local increases in effectiveness while our win-probability approach measures the overall impact on winning. A possible reconciliation of the difference in local and global effects may be that when teams fall behind in the fourth quarter, they resort to a "Hack-a-Shaq" strategy to try to get back into the game. This strategy of fouling the opposing team's worst free throw shooter in an attempt to reclaim possession and attempt to score has several effects. The strategy leads to a lot of starters in foul trouble, and helps the team catch up somewhat, but ultimately may fail to help them win the game. This explanation would be consistent with both findings.

Description of Data

We examine processed play-by-play data from the NBA for the 2006-2007, 2007-2008, 2008-2009, and 2009-2010 seasons. This data is from the website http://basketballgeek.com/data, maintained by Ryan J. Parker, and represents a processed version of the play-by-play information from the NBA and ESPN. The data includes such information as the names of all players on the court at each time, the location of the shots taken, the reasons for fouls called, and more. For our purposes, we use the players on the court and the substitutions, as well as the fouls.[1] There are 1,149 games in 2006-2007, 1,183 in 2007-2008, 1,176 in 2008-2009, and 1,215 in 2009-2010 with sufficient data available.[2] We use nearly every tick in the database, eliminating only those that do not lead to a change in a state variable.

We distinguish "threshold" fouls, namely those that move a player into foul trouble, from other fouls with the commonly used "Q+1" measure (i.e., benching players who commit their third foul in the second quarter, or fourth foul in the third quarter, etc.). Relative to other measures such as those that depend on the ratio of the remaining fouls to the remaining game time, "Q+1" is the best match to actual coaching decisions. For a comparable number of threshold fouls, "Q+1" fouls had a much higher percentage of yanks, indicating that it is the more common rule that coaches follow. In addition, preliminary investigations using alternate measures of foul trouble yielded qualitatively similar results, and so the simplicity and ubiquity of the "Q+1" measure was chosen for the remainder of the analysis. In our data set, NBA coaches benched 72% of players who committed threshold fouls, as defined by the "Q+1" measure. In our analysis, we compare the probability of winning the game for the 72% of times these players were benched versus the 28% of times when they remained in the game.

A player is "yanked" by his coach if he is substituted out of the game within some number of seconds of committing a foul. We use 30 seconds to allow for coaches to keep their player in for one more offensive possession, where they may be less likely to commit a foul and more likely to help their team relative to the expected substitution. The results did not differ substantially if the delay was zero seconds or 60 seconds instead. (Notice that in our introductory example, Bryant was benched several minutes after getting into foul trouble in the fourth quarter. His performance, however, was a possible illustration of self-idling.)

Finally, we define the following variables, which are each net (home minus away):

- FTR: net number of starters in the game with foul trouble, where foul trouble is defined by the "Q+1" criteria
- BFTR: net number of starters on the bench with foul trouble, where foul trouble is defined by the "Q+1" criteria
- NWP: difference between the total wins produced per 48 minutes (WP48) by the home players currently on the court and the total WP48 of the away

players currently on the court. Wins Produced are defined in Berri (2008), Berri, Schmidt, and Brook (2006), and Berri and Schmidt (2010)

Model Specification: Win Probability Framework

Our specification can be contrasted with the advanced plus/minus model of Rosenbaum (2004). Rather than using player dummy variables, we use Wins Produced, and rather than estimating the impact on the subsequent point differential, we measure the impact on a team's probability of winning. These differences allow us to incorporate state variables and variance by extending a general statistical model of team performance as described in Stern (1994). Furthermore, by focusing on the probability of winning, we employ a "macro"-level analysis that could complement or be complemented with the "micro"-level analysis of a plus/minus-type model.

Stern models the probability of the home team winning as $N(F_t/\sigma_t)$ where $N(\cdot)$ is the cdf of the standard normal distribution, $0 \leq t \leq 1$ is the fraction of the game elapsed, the variance σ_t^2 is proportional to $1 - t$ such that $\sigma_t^2 = (1 - t)\sigma^2$ for some volatility constant σ, and F_t is the "forward lead" at time t:

$$F_t = l_t + (1-t)\mu$$

where l_t is the current lead and μ is the drift constant.

Our "Win Probability" framework uses Stern's model with a few extensions. Our first and most general innovation is to add other state information to the definition of the forward lead. One advantage of this framework over regressing point differentials on various variables is our methodology automatically dampens (heightens) the effects of early (late) game actions when the variance is greater (lower). For example, a point scored early in the game hardly changes the probability of winning, whereas a point scored at the end of a close game can shift it substantially. Further, our framework highlights the importance of the information ratio in determining net benefit or cost, thus taking into account the risk as well as the return. Our second innovation is to model variance as declining linearly with time at a rate lower than one. These innovations are described in detail below.

First, we add possession, strength of players on the court, a constant, and two foul variables to Stern's definition of forward lead:

$$F_t = \alpha+\beta_l l_t+\beta_P P_t+(1-t)(\mu+\beta_{NWP}NWP_t+\beta_{BFTR}BFTR_t+\beta_{FTR}FTR_t)$$

where α is a constant, β_l is the coefficient for the lead, P_t represents possession at time t and β_P is its coefficient, and β_{NWP}, β_{BFTR}, and β_{FTR} are the coefficients associated with NWP, BFTR, and FTR. We multiply NWP, BFTR, and FTR by the fraction of the game remaining because these variables affect the rate of change of the lead.

WP48 data is provided by David J. Berri. An average player will have a WP48 of 0.100. A star player may be defined as one with a WP48 of at least 0.200, and a superstar as one with a WP48 of at least 0.300. For example, in 2006-2007 Jason Kidd led the league in WP48 with 0.405. A team with Kidd and four average players would have a total WP48 of $4\cdot 0.100+0.405=0.805$. Over a complete 82 game regular season, such a team would be expected to win $0.805\cdot 82=66$ games. (In fact, the Nets only won 41 games that year, because the other players were below average.)

The expected signs of β_l, β_P, and β_{NWP} are all positive because having a lead, possession, or more talent on the floor should help your team.

When the game begins, each team has 5 starters without foul trouble. Focusing on the home team for the moment:

FTR(home) = 0 and BFTR(home) = 0

In the next instant, a player can either enter Q+1 foul trouble or not, and the coach can bench a player or not, so there are four possible states of the world in the next time period. A starter is a player who started the game and therefore is likely one of the best players on the team. A bench player did not start the game and therefore is likely not as good as the starter he would replace. A starter can be in foul trouble or not, and can be benched or not benched (i.e., currently playing on the court).

Table 1: Foul Trouble States

State	FTR(home)	BFTR(home)	WP(player on court)
Player in foul trouble; benched	0	1	WP(bench)
Player in foul trouble; not benched	1	0	WP(starter)
Player not in foul trouble; benched	0	0	WP(bench)
Player not in foul trouble; not benched	0	0	WP(starter)

If the player in foul trouble is benched, then we are in the first state of the world, and the forward lead equation becomes:

$F_t = \alpha+\beta_l l_t+\beta_P P_t+(1\text{-}t)(\mu+\beta_{NWP}(NWP_t+WP(bench)_t\text{-}WP(starter)t+\beta_{BFTR}{}^*1)$

If, on the other hand, the player in foul trouble stays on the court, then we are in the second state of the world, and the forward lead equation is:

$F_t = \alpha+\beta_l l_t+\beta_P P_t+(1\text{-}t)(\mu+\beta_{NWP}(NWP_t)+\beta_{FTR}{}^*1)$

Subtracting these two equations and rearranging we find that it will be optimal to bench if:

$\beta_{FTR}\text{-}\beta_{BFTR}<\beta_{NWP}[WP48(bench)\text{-}WP48(starter)]$

The optimality of yanking will therefore depend on the magnitude of $\beta_{FTR}\text{-}\beta_{BFTR}$, compared to the difference in WP48 of the starter and the bench player. This follows because we need to compare the dropoff in quality from the starter and the bench player (if the starter is yanked) with the dropoff in quality from a starter playing through foul trouble (if the starter is not yanked).

Next, we model variance as a first order function of time: $\sigma_t^2 = 1 - \gamma t$, where γ (gamma) will be estimated below. Volatility at time t=0 is normalized to be 1. This volatility specification is analogous to a "jump-diffusion" model in finance such as Merton (1976)[3], where volatility at time t=1 is greater than zero. This specification seems appropriate for basketball. Variance should be proportional to possessions, which generally increases linearly with time remaining. At the end of a game, however, there is more variance than would be implied by a pure diffusion model, since teams can squeeze many possessions into the final minute of the game with fouls and time-outs.

Why do we use this win probability framework? First, it enables us to measure the strategic value of decisions, not just the tactical value. To measure the tactical value, we could instead run a regression where the dependent variable is change in score and the independent variables are the same. Such a regression would be similar to

Rosenbaum's advanced plus-minus augmented with foul variables. The problem with that approach is that it measures only the local changes in point scored over the next tick. It is possible that keeping a foul-troubled starter in the game will help over the subsequent minute, but also increase the risk that he will pick up another foul and play worse later or even foul out. In that situation, the team performance is improved locally, relative to situations where the starter is yanked, but ultimately harmed globally, especially if the starter eventually fouls out. By contrast, our win probability framework incorporates the future consequences of current actions since the dependent variable is whether the game is ultimately won or not, which is what matters most.

A second advantage of the win probability approach is the ease with which other state information that might affect the probability of winning can be added. For example, one can add lineup dummy variables to see how a three-guard lineup performs against a three-big-man lineup.

A third advantage is that the framework automatically magnifies the importance of late game situations where the game has not been decided and decreases the importance of actions during times when the game has already been decided. The advanced plus-minus framework can adjust for clutch time/garbage time somewhat arbitrarily, but our framework makes those adjustments automatically. At the beginning of a game, changes in the probability of winning are minimal, since there are still many possessions remaining in the game (the variance of possible outcomes is high). The win probability framework will weigh these events less, since each action will have a small impact on the probability of winning. On the other hand, the win probability framework will weigh the actions that occur at the end of a close game more, since the probability of winning can change significantly in those situations.

What are the disadvantages of our framework? We estimate the model with maximum likelihood estimation (MLE), which can become computationally expensive, particularly as the number of variables to be estimated increases, although in practice, we have not found this to be a constraint for our specification above. Also, we need to correct the standard errors of our regressions because the observations overlap. Specifically, even though the plays themselves do not overlap, the probability of winning the game as estimated with (for example) 10 minutes left is virtually identical to the probability of winning the game as estimated with 9:59 left. The MLE estimation procedure sees more than two million observations but in reality we have less than five thousand games in our dataset. The standard errors in this paper have been adjusted for the overlapping observations issue (see Appendix A for details).

Key Research Findings

Summary Statistics and Descriptive Measures

Figure 2 graphs the total number of threshold fouls and yanks for each team across all four NBA regular seasons from 2006-2007 through 2009-2010. The Golden State Warriors are the biggest outlier by two measures: they had the most threshold fouls, so they are further to the right on the graph, and they had the fewest yanks per threshold foul, so they have the lowest slope. Although this graph incorporates all four seasons, the Warriors were also the outlier in each of the four years; Don Nelson was the coach all four years.

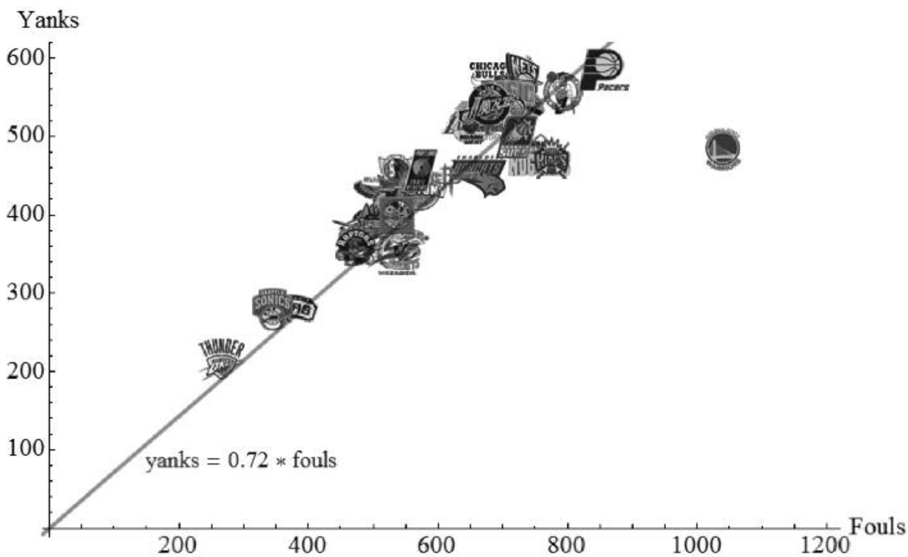

Figure 2: Total Threshold Fouls and Yanks per Team from 2006-2007 to 2009-2010

For illustrative purposes, Table 2 lists the top five in various categories, but only for the 2009-2010 season because the leaders changed year-to-year. The first category is the total number of fouls committed, led by Nene Hilario with 275. The second category is the total number of threshold "Q+1" fouls committed. The third category is percentage of time that a player was benched by his coach immediately after committing a threshold "Q+1" foul. The fourth category is the percentage of time that a player was not benched in that situation.

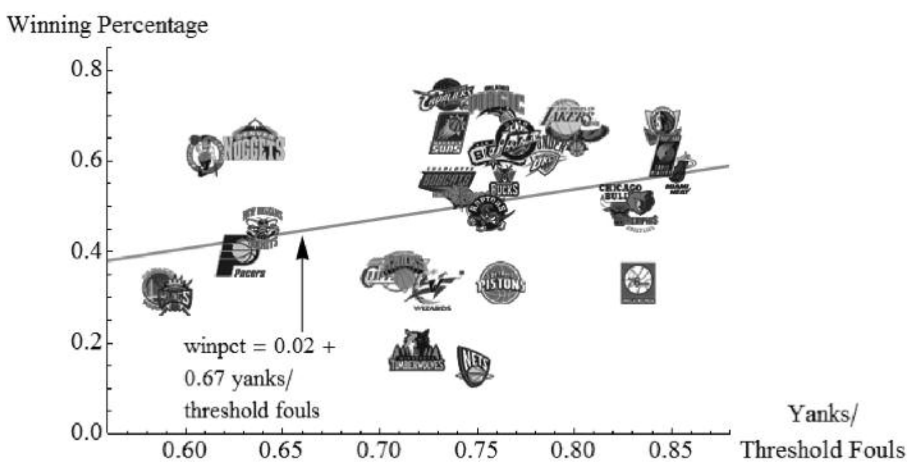

Figure 3: Winning Percentage in 2009-2010 vs. Ratio of Yanks per Threshold Fouls

O.J. Mayo was benched almost every single time he committed a threshold foul; Stephen Jackson was kept in the game 65% of the time he committed a threshold foul. To ensure sufficient data, the ranking methodology for these percentages includes only those players that committed an above-average number of threshold fouls.

Figure 2 showed data across all four seasons because the differences across seasons were slight. Figure 3 shows the winning percentage of each team just for the 2009-2010 season, because winning percentages changed substantially for some teams across years, and plots it as a function of how frequently they benched starters after threshold fouls. On balance, the better teams yanked more frequently.

Table 2: Leaders in Fouls and Threshold Yanks and Non-Yanks for 2009-2010

Rank	Total Fouls	Threshold Fouls	Highest Yanks Per Threshold Foul	Lowest Yanks Per Threshold Foul
1	Nene Hilario (275)	Kendrick Perkins (84)	O.J. Mayo (95%)	Stephen Jackson (35%)
2	Dwight Howard (268)	Nene Hilario (77)	Andre Iguodala (94%)	David West (37%)
3	Samuel Dalembert (248)	Roy Hibbert (77)	Ben Gordon (94%)	Monta Ellis (41%)
4	Shaquille O'Neal (245)	Jason Thompson (71)	Joel Przybilla (94%)	Paul Pierce (42%)
5	Richard Jefferson (241)	Andris Biedrins (64)	Raymond Felton (93%)	Chris Paul (43%)

Win Probability Model

Recall that we define the forward lead as:

$$F_t = \alpha+\beta_l l_t+\beta_P P_t+(1-t)(\mu+\beta_{NWP}NWP_t+\beta_{BFTR} BFTR_t+\beta_{FTR}FTR_t)$$

Because the magnitude of $BFTR$-β_{BFTR} is central to the decision to bench, we can add $\beta_{BFTR}(FTR_t)$-$\beta_{BFTR}(FTR_t)$ and rewrite the equation as:

$$F_t = \alpha+\beta_l l_t+\beta_P P_t+(1-t)(\mu+\beta_{NWP}NWP_t+\beta_{BFTR}(BFTR_t+FTR_t)+(\beta_{FTR}-\beta_{BFTR})FTR_t)$$

where α is a constant, β_l is the coefficient for the lead, P_t represents possession at time t and β_P is its coefficient, and β_{NWP}, β_{BFTR}, and β_{FTR} are the coefficients associated with NWP, BFTR, and FTR, respectively.

Table 3 displays the regression coefficients for all four years individually and combined, along with the adjusted t-statistic in brackets below. The adjusted t-statistics are the raw t-statistics divided by 15.8 to account for the overlapping nature of the observations. See Appendix A for more information. All t-statistics reflect comparisons to a null hypothesis value of zero, except for volatility, which reflects comparison to a null hypothesis value $\gamma = 1$ representing the assumption of Stern (1994) of a diffusion model without jumps.

Note that the variance coefficient is below one, contradicting the assumption of Stern (1994). Professional basketball game scores do not diffuse normally and do appear to have jumps.

We find that β_{FTR}-β_{BFTR} is negative, which indicates that keeping foul-plagued starters in the game can have a detrimental effect on a team's probability of win-

Table 3: Coefficients and Adjusted t-Statistics of Win Probability Regression

	Constant α	Lead βl	Drift μ	Poss. β_P	Volatility γ	NWP48 β_{NWP48}	Fouls 1 $\beta_{FTR} - \beta_{BFTR}$	Fouls 2 β_{BFTR}
2006-2007	-0.0193	0.0559	0.2602	0.0285	0.9928	1.4104	-0.1575	0.0972
	[-0.89]	[18.72]	[4.55]	[1.60]	[1.23]	[9.08]	[-0.72]	[1.33]
2007-2008	-0.0084	0.0558	0.3015	0.0290	0.9973	1.4757	-0.1487	0.0302
	[-0.39]	[19.26]	[5.27]	[1.63]	[0.94]	[11.14]	[-0.64]	[0.37]
2008-2009	0.0029	0.0562	0.3260	0.0299	0.9966	1.8608	-0.2172	0.0424
	[0.14]	[18.56]	[5.59]	[1.68]	[0.83]	[12.52]	[-0.92]	[0.53]
2009-2010	0.0192	0.0539	0.2674	0.0270	0.9967	1.5551	-0.1125	0.0504
	[0.93]	[19.30]	[4.81]	[1.55]	[1.00]	[11.00]	[-0.57]	[0.64]
All 4 Years	-0.0022	0.0553	0.2885	0.0284	0.9958	1.5720	-0.1580	0.0586
	[-0.20]	[37.98]	[10.12]	[3.22]	[2.10]	[21.92]	[-1.44]	[1.50]

ning. Focusing on the four-year results, the estimated value of β_{FTR}-β_{BFTR} is -0.1580, while β_{NWP48} is 1.572. The t-stat for β_{FTR}-β_{BFTR} is -0.1580, which makes the estimate significant at the 15% level. The equivalent of the last row of Table 3 was also run for results excluding Golden State, and the results were nearly the same. The estimate value of β_{FTR}-β_{BFTR} is even higher, but the standard error was higher still, so the t-stat was slightly lower.

Recall that the average NBA player has a WP48 of 0.100, while a star player is commonly defined as one who has value above 0.200. Thus, these results imply that foul trouble makes a star player perform worse than an average player. If a team has a bench with average players, then only players with WP48 over 0.258/1.572 = 0.164 should be allowed to stay in a game with foul trouble.

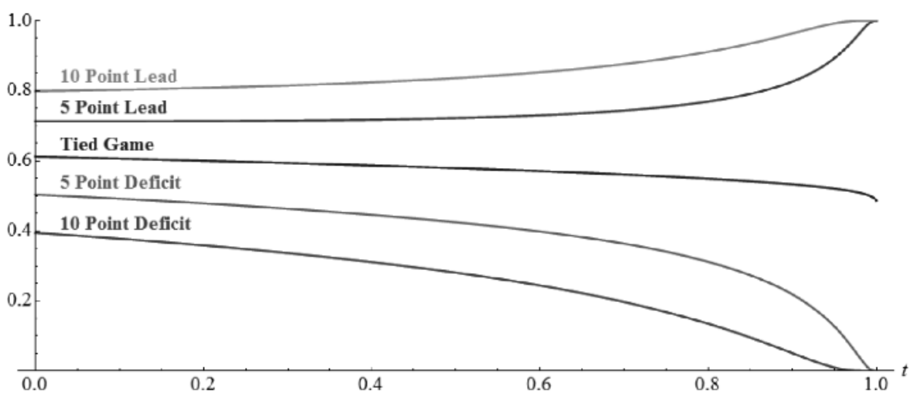

Figure 4: Home Team Win Probability vs. Time Elapsed for Different Leads

The "drift" and "constant" coefficients combine to tell us how a generic team's home-court advantage evolves over time, with "constant" representing the end-of-game home court advantage.

Figure 4 shows the home team's probability of winning a hypothetical game with a variety of constant leads: a 10-point lead, 5-point lead, a tied game, a 5-point deficit and a 10-point deficit. For a game between two evenly matched teams with no possession advantage, the home team has a 62% chance of winning at the start of the game. This home-court advantage has been researched by many (see Moskowitz and Wertheim (2011) for a discussion). As the tied game progresses, this probability of winning decays towards 50%. Indeed, in our data set, the probability of winning actually dips slightly below 50% in the last minute of play. The reason is unclear, but Moskowitz and Wertheim (2011) might explain this phenomenon with referee whistle swallowing.

On the other hand, conditional on a 5- or 10-point lead, the probability of winning approaches one as the game progresses, and symmetrically, conditional on a 5- or 10-point deficit, the probability of winning approaches zero.

In the results for all four years, the lead coefficient β_l is very significantly positive, with an adjusted t-statistic of 38. The possession coefficient is also very significantly positive, and around half the size of the lead coefficient, which indicates that a possession is worth around half a point.

Recall that the decision to bench is largely determined by the size of β_{FTR}-β_{BFTR}. In terms of signs, β_{FTR}-β_{BFTR} is clearly not statistically significantly positive; indeed, it is far more likely negative.

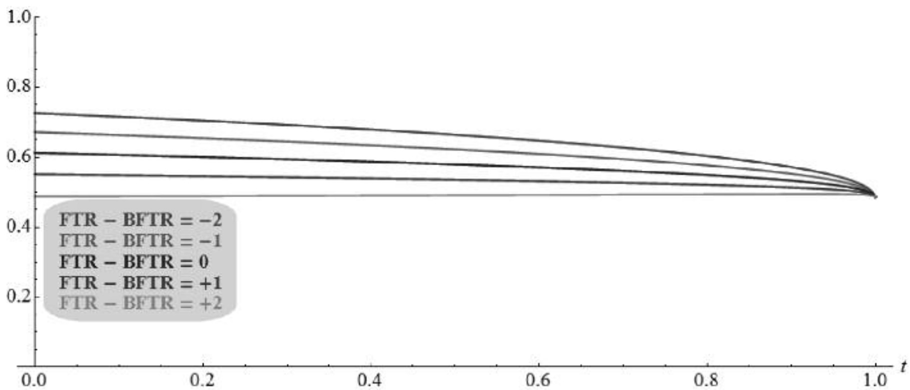

Figure 5: Home Team Win Probability vs. Time Elapsed, Conditional on *FTR-BFTR*

Figure 5 graphs the impact of playing with one or two net starters in foul trouble. The middle line is the same line as the middle line from Figure 4, representing the probability of winning a tied game with no net starters in foul trouble. With one foul-troubled starter playing, i.e., *FTR*=1, *BFTR*=0, the probability of winning at the beginning of the game is reduced from 62% (*FTR-BFTR*=0) to 58% (*FTR-BFTR*= +1); with two foul-troubled starters playing, that probability of winning is further reduced to 53% (*FTR-BFTR*= +2). The impact on the probability of winning by construction decreases as the game progresses, so by the end of the game, it is immaterial whether a foul-troubled starter is playing or not. This observation is consistent with our discussion earlier that yanking a starter has the most option value early in the game.

Finally, we reiterate that our win probability approach is able to capture the global, strategic impact of decisions or state variables, whereas OLS or adjusted plus-minus approaches only capture the local, tactical impact. This may be one other reason why foul-troubled starters appear to play better: indeed, they may play better than their replacement for the next few possessions or minutes, but because the cost of continuing to play now is the erosion of their later option value, the short-term tactical victory often comes at the expense of an ultimate loss in the game.

Is the "Q+1" measure of foul trouble the correct one? Even if it is, shouldn't players who anticipate being yanked for their "Q+1"-th foul also play worse after committing their "Q+0"-th foul? We can explore these questions by rerunning the analysis of Table 3 with alternative definitions of foul trouble. Recall that in Table 3 we defined a player as being in foul trouble if their number of fouls equaled or exceeded "Q+1."

Table 4 reports the relevant regression results from definitions of foul trouble ranging from exactly equaling "Q+0" to equaling "Q+2" for all four years in our sample. All unreported regression coefficients were virtually identical to those of the last row of Table 3.

Table 4: Coefficients and Adjusted t-Statistics for Different Foul Thresholds

	Fouls 1 β_{FTR}-β_{BFTR}	Fouls 2 β_{BFTR}
Q+0	-0.0400	0.0167
	[-1.22]	[0.58]
Q+1	-0.1554	0.0582
	[-1.41]	[1.48]
Q+2	-0.4113	0.0352
	[-0.32]	[0.15]

Players do play a little worse when they reach "Q+0" foul trouble, even though the significance is a little lower. However, since the magnitude of the coefficient is one quarter that of the "Q+1" coefficient, the drop-off in play at "Q+0" is likely not because "Q+0" is a better measure of foul trouble but rather because players anticipate being yanked at their next foul.

Further, the "Q+2" threshold foul is statistically insignificant from zero. In other words, "Q+1" does seem to be the most correct definition of early foul trouble.

Conclusions and Further Areas of Research

We describe a novel and general approach to estimating performance in professional basketball games by using tools from finance and we demonstrate its usefulness by applying it to the problem of early foul trouble, a problem that until now had been absent from the academic literature. Our analysis shows that most of the time, a starter in foul trouble should be benched. If left in the game, the player may become a liability, since he is afraid of picking up another foul. We find that in this situation, NBA coaches do not act irrationally by using the "Q+1" rule. Our conclusions should not be interpreted as a contradiction of Romer (2006) and Staw and Hoang (1995), who find that coaches act irrationally, because our study is limited to the issue of NBA foul trouble only.

A possible area of research would be to examine how a player's statistics (points, rebounds, assists, steals, blocks, charges, plus/minus, etc.) are affected by foul trouble. Such analysis would complement our analysis. Also, we suggest that player fatigue, injury history, team-strategic fouls such as non-penalty end-of-quarter situations, and bench depth be incorporated into the analysis. In addition, the particular reason for why yanking can be locally better but globally worse, such as the Hack-a-Shaq possibility discussed earlier, may be another fruitful area of future research to shed more light on the reconciliation between our results and those of Moskowitz and Wertheim (2011).

Furthermore, the approach outlined here can likely be applied to numerous other questions such as comparing the efficacy of a three-guard lineup, testing for the effects of momentum or streaks, and even evaluating an individual's entire contribution, including defensively, by conditioning on their presence on the court. It could also incorporate other non-basketball-related factors to assess the impact of in-game presentation, attendance, concessions, and other factors on the probability of winning.

Finally, the results of our analysis shows that there exists at least one area with abundant data in which idling resources may be strategically optimal, and it suggests that such situations may be more common than previously assumed.

References

Aschburner, S. (2010). Rare foul trouble for Kobe plays right into Celtics' hands. *NBA.com*. Retrieved from http://www.nba.com/2010/news/features/steve_aschburner/06/07/kobe.fouls/index.html

Adande, J. A. (2010). Unusual night all around in Los Angeles. *ESPN.com*. Retrieved from http://sports.espn.go.com/nba/playoffs/2010/columns/story?columnist=adande_ja&page=Bryant-100607

Berri, D. J. (1999). Who is most valuable? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics*, *20*, 411-427.

Berri, D. J. (2008). A simple measure of worker productivity in the National Basketball Association. In B. Humphreys & D. Howard (Eds.), *The Business of Sport* (pp. 1-40). Westport, CT.: Praeger.

Berri, D. J., & Schmidt, M. B. (2010). *Stumbling on wins: Two economists explore the pitfalls on the road to victory in professional sports*. Princeton, NJ: Financial Times Press.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2006). *The wages of wins: Taking measure of the many myths in modern sport*. Palo Alto, CA: Stanford University Press.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637-654.

Fine, C. H., & Freund, R. M. (1990). Optimal investment in product-flexible manufacturing capacity. *Management Science*, 36, 449-467.

Morrison, D. G. (1976). On the optimal time to pull the goalie: A Poisson model applied to a common strategy used in ice hockey. *TIMS Studies in Management Sciences*, *4*, 137-144.

Morrison, D. G., & Wheat, R. D. (1986). Misapplications reviews: Pulling the goalie revisited. *Interfaces*, *16*, 28-34.

Moskowitz, T., & Wertheim, L. (2011). *Scorecasting: The hidden influences behind how sports are played and games are won*. New York, NY: Crown Archetype.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*, 125-144.

Price, J., & Wolfers, J. (2010). Racial discrimination among NBA referees. *The Quarterly Journal of Economics*, *125*, 1859-1887.

Romer, D. (2006). Do firms maximize? Evidence from professional football. *Journal of Political Economy*, *114*, 340-365.

Rosenbaum, D. (2004). Measuring how NBA players help their teams win. Retrieved from http://www.82games.com/comm30.htm

Staw, B. M., & Hoang, H. (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, *40*, 474-494.

Stern, H. (1994). A Brownian motion model for the progress of sports scores. *Journal of the American Statistical Association*, *89*, 1128-1134.

Washburn, A. (1991). Still more on pulling the goalie, *Interfaces*, *21*, 59-64.

Weinstein, J. (2010a). *Foul trouble*. Retrieved from http://theoryclass.wordpress.com/2010/04/27/foul-trouble

Weinstein, J. (2010b). *Foul follow-up*. Retrieved from http://theoryclass.wordpress.com/2010/05/17/foul%C2%A0follow-up

Endnotes

[1] The running score of the game is not provided in 2006-2007 though it is in later seasons. For 2006-2007, we manually calculate the running score by combining free throws and field goals, with adjustments for three-point field goals. These score calculations occasionally differ from the official final score of the game, for example if a three-point field goal had been incorrectly ruled a two-point field goal during the course of the game. To counteract this discrepancy, we fix the final score of each game to be the official final score, and adjust the initial net score by the difference between our calculated final net score and the official final net score. We do this for consistency of scores within a game, but our results do not differ materially if we use calculated score only. To maintain consistency, we do this for later years as well.

[2] Play-by-play transcripts for games occasionally fail to include sufficient information to determine which players are on the court, and so such games are ignored. Such errors by the NBA and ESPN seem to be random.

[3] Many financial models, such as Black and Scholes (1973), assume that stock prices follow a continuous lognormal diffusion process. Merton (1976) introduced a jump-diffusion model, which assumes that stock prices occasionally jump discontinuously, in addition to their lognormal diffusion process.

Acknowledgments

The authors would like to thank Ashwin Alankar, David Berri, Jeffrey Chuang, Harry Gakidis, Matt Goldman, Eason Hahm, Cheng Ho, Alan Liu, Ted Mann, Benjamin Morris, Kenneth Shirley, Michael Shen, Thomas Shen, Richard Thaler, participants at the 5th Annual MIT Sloan Sports Conference, and the referees and editor for their helpful feedback and comments, and further to David Berri for providing the historical WP48 data.

Appendix A

Stern (1994) points out that approaches that estimate the probability of winning a game at various points throughout the game implicitly use overlapping data points because the probability of ultimately winning the game is highly correlated from one moment to the next, thus the standard errors may be misestimated. He reports simulation results that the t-statistics computed without adjusting for this overlapping phenomenon when the number of observations per game is n=4 (i.e., observations at the end of each quarter) are overstated by 30%.

Figure A1 below displays the results of extending Stern's simulation approach to larger values of n. In the data set of this paper, using tick-by-tick data, the number of observations or ticks per game averages 498, so the appropriate t-statistic multiplier is 15.8.

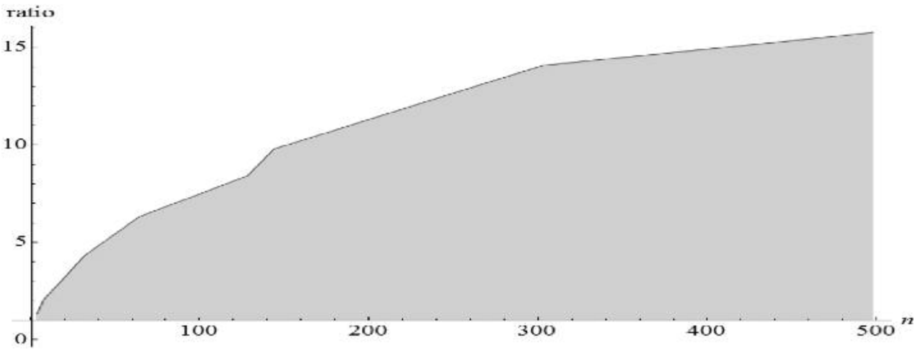

Figure A1: Stern Simulation t-statistic Adjustment Factor