"He Got Game" Theory? Optimal Decision Making and the NBA*

Matt Goldman[1] Justin M. Rao[2]
[1]Department of Economics, University of California, San Diego
[2]Yahoo! Research Labs, Santa Clara, CA

Feb. 7, 2010

Abstract

The paper examines the optimality of the shooting decisions of National Basketball Association (NBA) players using a rich data set of shot outcomes. The decision to shoot is a complex problem that involves weighing the continuation value of the possession *and* the outside option of a teammate shooting. We model this as a dynamic mixed-strategy equilibrium. At each second of the shot clock, *dynamic efficiency* requires that marginal shot value exceeds the continuation value of the possession. *Allocative efficiency* is the additional requirement that at that "moment", each player in the line-up has equal marginal efficiency. To apply our abstract model to the data we make assumptions about the distribution of potential shots. We first assume nothing about the opportunity distribution and establish a strict necessary condition for optimality, which nearly all players/teams pass. Adding distributional assumptions, we establish sufficient conditions for optimality. In line with dynamic efficiency, we find that the "cut threshold" declines monotonically with time remaining on the shot clock at approximately the correct rate. Most line-ups show strong adherence to allocative efficiency. We link departures in optimality to line-up experience, player salary and overall ability.

*We would like to thank James Andreoni, David Eil, Pedro Rey Biel and Charlie Sprenger for extensive notes. Nageeb Ali, Gordon Dahl, Uri Gneezy, Craig McKenzie, David Miller, William Peterman, Joel Sobel and Joel Watson provided helpful comments. This paper is a merger of Goldman's "Optimizing in the Half-Court" and Rao's "He Got Game Theory:..." (note the lack of "?").

1 Introduction

A fundamental concern of economics is optimality: do agents and firms maximize in accordance with the normative predictions our theoretical models? This paper examines the problem of the allocation of resources across time and different modes of production in order to maximize firm output — a classic line of inquiry, which is taken to a new testing ground. The firms being studied are National Basketball Association (NBA) teams, and the agents making the allocation decisions are the professional players (and coaches) and their objective is the accumulation of points. In NBA basketball, a 5-man line-up has 24 seconds to take the best shot possible. To do so the team must allocate the ball effectively across its factors of production (players) and over the course of the shot clock. We model the game using two important theoretical concepts from game and decision theory: 1) mixed strategy Nash equilibrium (MSNE) 2) weighing current alternative versus the continuation value of a process of random arrivals (optimal stopping). These two concepts underly important economic models that run the gamut from job search to bilateral bargaining.

This paper takes these concepts *seriously*. If we are going to use them as backbones of economic models, experts should adhere to them, especially in high stakes environments. This should apply even for *difficult* decisions and games. Equilibrium play in basketball requires effective randomization across the 5-man line-up, *allocative efficiency*, and accurately weighing the continuation value of a possession versus a current realization, *dynamic efficiency*. We examine NBA teams' adherence to optimality using rich data set comprising all shots taken in the NBA from 2006 to 2010. Our modeling approach is to combine the stopping problem with the allocation problem. At each point on the shot clock, players must randomize over shooters in order to maximize allocative efficiency and also use an appropriate threshold value for continuing the possession versus realizing a shot opportunity. The vectors of shooting frequency should maximize points per possession. Each vector implies a "threshold equilibrium" of play, in the spirit of a purified game (Harsanyi, 1973). It is of course challenging to fit a complex game such as basketball in an abstract game-theoretic model, but we show that we can establish necessary and sufficient conditions for optimality with minimal assumptions. This allows us to fit our model of shot allocation and optimal stopping to observed shooting behavior. We then link adherence to optimality conditions to metrics such as team success, line-up experience and player/coach salaries.

Our key finding is that NBA players are superb optimizers. The average NBA player is shown to adopt reservation shot values *almost exactly equal to the continuation value* of his team's possession throughout the entire range of the shot clock. This "cut threshold" declines monotonically with time remaining on the shot clock at almost the exact rate implied by dynamic efficiency. Most teams core line-ups show impressive allocative efficiency, with the spread of marginal efficiencies within a line-up quite small. Very few players can be shown to individually overshoot from the dynamic (shoot too soon) or allocative (shoot too much overall) perspective. Undershooting is more common and seems to occur primarily in players easily recognized as amongst the NBA elite. Such behavior is suboptimal in our model, but is easily rationalized if we allow for such players to conserve their energy and health for the long-haul of the season at the slight expense of their team's immediate performance.

Past work has tested the predictive power of game theoretic equilibrium concepts in both the laboratory and field. The advantage of the laboratory is that the experiment can be tailored to test the falsifiable implications of the model, while the advantage of the field is that the agents (or firms) under study are performing a familiar task in a natural setting. Indeed the "poor" performance of experimental subjects has led to the use of experts with relevant experience as subjects (Palacios-Huerta and Volij, 2009; Levitt et al., 2010; Palacios-Huerta and Volij, 2008). Studies of MSNE in field settings have typically used 2×2 games such as soccer penalty kicks (Chiappori et al., 2002; Palacios-Huerta, 2003) or tennis serves (Walker and Wooders, 2001). The only work we are aware of that applies an optimal stopping model to professional sports is Romer (2006), which examines NFL coaches' decisions to "go for it" on 4th down on the NFL.[1] These papers generally find that players use the correct mixing proportions and do not serially correlate actions, which are the necessary and sufficient conditions for equilibrium play.

One of the important distinguishing features of this paper from prior work is that the game under study is much more difficult to play optimally. Soccer penalty kicks and tennis serves can be modeled as a game only slightly more complex than matching pennies. In the game of basketball, players face shooting opportunities with variable defensive pressure and shot distance; further complicating matters is the fact that the shot must be taken before the 24 second shot clock expires. Finally, each team's lineup consists of unavoidably heterogeneous players who realize the scoring opportunities via different methods and from different distributions. In shot allocation, the *team* must mix over five pure strategies because there are five players on the court at one time, as compared to an *individual* mixing over two pure strategies. Furthermore, each player must be cognizant of the continuation value of a possession and should shoot only if the value of the shot is in excess of the continuation value.

[1] A key difference between this study and ours is that Romer studies a deliberative decision made about 10 times per season by the coach, not the players, and the low number of occurrences means coaches might intentionally do the wrong thing in order to adhere to the (mistaken) "conventional wisdom". This is precisely what the author finds.

Studying a more complex game does come at a cost — it is more challenging to fit into the confines of an abstract game-theoretic model. The chief hurdle in determining the optimality of player actions is to determine what would have happened had the player shot less or more. Suppose that we observe a player offering higher productivity than his teammates at a given interval of the shot clock. In order to conclude that he is not shooting enough in equilibrium, we have to be sure that his opportunity distribution permits a marginal increase in production at levels similar to those observed. It is entirely possible that such a player's average efficiency is boosted by a number of easy, inframarginal shots. That is, in order to take additional shots, he would have to settle for much less efficient opportunities on the margin. Just as in the theory of the firm, it is the efficiency on these marginal opportunities that should guide optimal allocation.

Unfortunately marginal opportunities are not readily observed (Skinner, 2010; Oliver, 2004; Berri, 1999). Our estimation strategy is to use a telescoping set of assumptions, starting with the case of assuming essentially nothing about the shot opportunity distribution, which provides us with a strict necessary condition of optimality. We then impose more restrictive, but still fairly flexible, assumptions in order to strengthen our optimality tests in the form of a sufficient condition for optimal allocative and dynamic efficiency. The goal is to apply the strictest test of optimality without incorporating a model misspecification. Below we discuss our modeling approach in a bit more detail.

Shot timing is modeled using a finite period optimal stopping model in which players face shot opportunities of varying expected point values and must decide between shooting and continuing the possession. Since a shot must be taken in a 24 second period (the length of the "shot clock") the process has finite opportunities. In this sense, the problem is similar to textbook "secretary" or "car buying" problems, which have been extensively studied in the lab (Seale and Rapoport, 200; Lee, 2006). The wrinkle we add is that not only must a player consider his own future realizations, but also realizations of his teammates on the court. We model this as a repeated, single-shot game of "who will shoot." Intuitively, the offense has to keep the defense guessing about who will shoot the ball so the defense cannot apply extra pressure on the eventual shooter. At a given interval of the shot clock, it should be the case that the offense cannot reallocate the ball and increase team productivity. Put another way, the vector of shooting frequency should maximize points per possession. The vector of shooting frequencies can be thought of as a mixed strategy conditional on defensive response.

Even when we assume *nothing* about the opportunity distributions we can still establish a *necessary* condition for optimality that can detect overshooting: at a given time in shot clock, a player's expected point value per shot should not be lower than the continuation value of a possession, where the continuation value is based on observed play. Exceedingly few players run afoul of this condition. The condition is strong in that is assumption free, but weak in that it will miss many players who take many reasonable shots, but overshoot on the margin. To combat this, we add more structure to the problem by assuming a dependence pattern of shot opportunity arrivals for each player and across different values of the shot clock. This allows us to estimate opportunity distributions (by player and line-up) and provides a stronger test of optimality, because it yields an estimation of the value of the marginal (worst) shot a player is selecting in each interval. Using these estimates, we construct the implied threshold values for continuing a position and find that they closely match those implied by dynamic efficiency. As the shot-clock winds down, players are willing to take monotonically worse shots at nearly the exact rate implied by the model.

We then use these estimates to examine allocative efficiency. The prediction of optimal play is that for a given interval of the shot clock, marginal efficiencies within a line-up should be equal. Using a measure of "spread" that does not depend on sample size, we find that most line-ups share the ball very well — spread in marginal output is low. Adherence to allocative efficiency is correlated with overall performance of the line-up (aggregate output per possession) and player talent. Two basic predictions of our model are strongly supported in the data: 1) good players shoot more but with similar marginal efficiency 2) good players raise the continuation threshold for all players (and thus 'make their teammates better' in observational data).

Our formal model excludes learning and the relevance of factors such as within-line-up talent/salary inequality. In reality, these features might impact the adherence to optimality. To address this question we link the sufficient statistics from our tests of optimality to line-up experience (how often the players share the court), player tenure in the league, player salary and salary inequality using OLS regression. We find that adherence to allocative efficiency significantly positively related to line-up output and salary level. Increased variance in line-up salary leads to more optimal play. On the individual level, star players are more likely to undershoot, perhaps to conserve energy/health of the course of the season. For both measures, player tenure in the league does not appear to have a meaningful impact controlling for these other state measures.

Our paper offers two main contributions. The first, which has already been discussed, is extending past work on MSNE in expert populations to a more difficult, robust game. The second is the analysis of an important stopping problem that has been extensively studied in the lab, due to its association with labor market decisions, but not previously tested in the field, on account of the difficulty in observing relevant decision inputs and outcomes. In prior laboratory study, subjects tend to use fixed stopping thresholds (Lee, 2006) or that are declining but too flat cite (Lee et al., 2004). The experts we study perform far better, even though the decision environment is more complex. Past work using experts has studied firm decisions in dynamic settings such as as harvesting tree stands (Provencher, 1997), renewing patents (Pakes, 1986), replacing bus engines (Rust, 1987) and continuing drives in professional football (Romer, 2006). These problems are similar to the stylized lab version of the problem in one respect — the decision is based on the comparison of expected value of acting versus continuing — but are dissimilar in that they do not involve random arrivals of opportunities (for instance, one can replace a bus engine at any time). NBA shooting presents a situation that is strikingly similar to the classic version of the stopping problem.

The paper proceeds as follows: Section 2 describes the data, Section 3 gives the model estimation and results starting with the minimal assumption set and telescoping out, Section 4 presents a discussion and Section 5 concludes.

2 Data

This analysis is entirely based on play-by-play data for all NBA games from 2006-2010 (four seasons). The game logs detail all the players on the court the outcome of every play. All variables discussed herein are constructed (via some fairly extensive coding) from the raw game logs. Approximately 100 games (out of the 4,920 played during this time period) are missing from this data set. We have no reason to believe their omission is anything other than random.

Table 1: Data Overview and Description

Event/Action	Description
Offensive/defensive line-up	Players on court at given time
Game-time	Minutes and second of each event
Game day	Date of game
Shooter	Player/time of the action
Rebound/assist	Player/time of the action
Foul	Shooting, non-shooting, flagrant, illegal defense
x,y coordinate of shot	Physical location of shot
Turnover	Broken down by bad pass, dribbling error, charge, lost ball

We performed data cleaning to eliminate cases where the assumptions of our model are clearly violated. First, we remove possessions where risk-neutrality is violated, such as end of quarter and end of game situations. Since our model applies only to half-court sets and not to "fast-breaks" (actions taken immediately after steals or offensive rebounds, which involve fairy trivial decision making). We define a fast break as any possession used within the first seven seconds of the offense and have purged all such possessions from the data. Finally, we remove situations when one team is leading by a point margin so wide that the game outcome is no longer in doubt as our model only applies to situations where the outcome of the game can reasonably be effected by chosen offensive strategy.

3 Model and Results

In this section, we first motivate our general model of basketball as one of point maximization per possession. Importantly this requires an assumption of risk-neutrality, which we defend with data and simulations in the Appendix. The basic intuition is that a game involves a large number of possessions, so sacrificing mean efficiency to increase or decrease variance has a first order effect on the mean but only a second order impact on the variance, so it is almost surely a losing strategy. It is trivial to rule out the rare end of quarter situations where this logic collapses. The general model incorporates the intertemporal trade-off of realizing an opportunity versus continuing the possession and the interpersonal trade-off of how frequently each member of the team should shoot at a given time interval. The latter will be interpreted here as cut-points for each player's worst selected shot in a pure strategy equilibrium. But, could also be interpreted as a vector of shooting frequencies in the corresponding Harsanyi purified mixed strategy equilibrium. To fit an abstract model to observed play we must make assumptions that sufficiently simplify the complex game of basketball. The more restrictive the assumptions, the stricter the test of optimality, but the greater chance of misspecification. As such, we adopt a telescoping approach in which we start with a minimal set of assumptions that provide a necessary condition of optimality. We then add assumptions to increase the power our optimality tests and use robustness checks (overidentification tests) to ensure their validity.

3.1 General model

There are 674 distinct players in our four year sample, denote each one by $i \in \{1, ..., 674\}$. $O \equiv \{i\text{: player i is on offense}\}$, $D \equiv \{i\text{: player i is on defense}\}$. Defensive strategy is summarized by the selection of five choice variables $\{d_j\}_{j \in O}$ which represent the average level of attention devoted to each offensive player across the possession. These variables are constrained by the abilities of the defenders according to:

$$g_D(d_{O_1}, d_{O_2}, d_{O_3}, d_{O_4}, d_{O_5}) \leq 0 \quad (1)$$

where g is an arbitrary function with strictly positive derivatives in all arguments.

At every one second long interval of the shot clock, each player has the opportunity to use a possession. With t seconds remaining on the shot clock ($t \in T$), player i draws an unbiased measure of the expected number of points his team would get from his immediate use of the possession, $\eta_{i,t}\tilde{}\Phi_{i,O/i,d_i,t}$. This value represents not just expected points on his immediate shot, but also whatever value his team is likely to get (including points scored after offensive rebounds or from foul shots) before ending their possessions. $\eta_{i,t}$ is also effected by the likelihood that player i should turn the ball over in his attempt to use the possession.

$\forall i \in O$, player i has the opportunity to use the possession, getting on average a value of $\eta_{i,t}$ for his team if he does so. If no player choses to exercise the possession at interval t, than the one second long period of the shot clock is allowed to pass and each player will realize new and independent scoring opportunities in period t-1. Note that we are not assuming a viable scoring opportunity for every player at every interval of the shot clock, it is entirely possible that during many seconds of the shot clock a player will realize scoring opportunities of arbitrarily small expected value.

In period 0, the shot clock is about to expire. If no player uses the possession during this period, then play is stopped and the ball is awarded to the other team with no points for the offense.

In principal, the distribution of scoring opportunities could vary over player, teammates, interval of the shot clock, and the level of defensive attention. Understanding the interrelationships between all of these variables is an important direction for our future work, but for now we assume player i's opportunity distribution is invariant to teammates (O/i) and defense ($d_{i,t}$).

Definition of Variables

First, some definitions. For some possibly singleton set $I \in O$ and a possibly singleton and convex $T \in \{0, ..., 24\}$, define $N_{I,T}$ as the number of shots used by players in I over elements of the shot clock in T and define $P_{I,T}$ as the total number of points realized from these possessions. Further, let:

$$E_{I,T} = \frac{P_{I,T}}{N_{I,T}}; T_* = \{t \in \{0, ..., 24\} : \forall t^{'} \in T, t < t^{'}\}; U_{I,T} = \frac{N_{I,T}}{N_{O,T \cup T_*}} \quad (2)$$

E is a measure of average efficiency in the periods in T and U is the hazard rate with which possessions are used over a particular interval T and by a particular subset of offensive players. Note that all of these variables describe observed data, if we wish to describe the "true" efficiency or usage rate of a particular set of players over some interval, we will use a lower case version of the same definitions (e, u).

General Solution

Each player i chooses a cut-off level, $c_{i,t}$, such that he will use the possession at interval t if and only if $\eta_{i,t} \geq c_{i,t}$. If $\forall i, \eta_{i,t} < c_{i,t}$ than no player choses to use the possession and the team proceeds to the next period of the shot clock. In order for a player's choice of cutoff to be optimal it must meet two basic criteria.

Allocative efficiency: At any given t, the team cannot reallocate the ball to increase productivity on the margin. The frequency at which each player shoots generates equal marginal productivity. *Allocative efficiency generates the best set of shot opportunities, because randomizing effectively is a best response to selective defensive pressure.* Formally, allocative efficiency requires that $\forall i, j \in O, c_{i,t} = c_{j,t}$.

Dynamic efficiency: Conditional on allocating efficiently in future periods, a shot is realized only if its expected value exceeds the continuation value of a possession. Formally, dynamic efficiency requires that $\forall i \in O, c_{i,t} = e_{O,t*}$.

It is straightforward the lineup of Dynamically efficient players would nessecarily satisfy Allocative efficiency. However, we may be interested in lineups which allocate shots poorly across time, but appear to do so in an equitable manner across the players on the team. We now apply these conditions to observed behavior.

3.2 A Necessary Condition for Dynamic Optimality

In this section we develop a minimalistic necessary condition for dynamic optimality that makes no assumptions about our opportunity distributions. We address the question "do players take shots too frequently and/or too soon in the shot clock."

Player i only shoots at t if $\eta_{i,t} > c_{i,t}$. Thus:

$$e_{i,t} = E(\eta_{i,t} | \eta_{i,t} \geq c_{i,t}) > c_{i,t} \tag{3}$$

So long as we observe non-decreasing possession continuation values in t, it is easy to extend the above condition to convex intervals of time. Thus we may take as a null hypothesis of dynamic efficiency that:

$$e_{I,T} \geq e_{O,T_*} \tag{4}$$

where a convenient test is implied by application of a central limit theorem, to give:

$$\begin{aligned}
E_{I,T}\tilde{}N(e_{I,T}, \frac{\sigma_I}{N_{I,T}}) \\
E_{O,T_*}\tilde{}N(e_{O,T_*}, \frac{\sigma_O}{N_{O,T_*}}) \\
T \cap T_* = \emptyset \rightarrow E_{O,T_*} \perp E_{I,T} \\
E_{I,T} - E_{O,T_*}\tilde{}N(e_{I,T} - e_{O,T_*}, \sqrt{(\frac{\sigma_O}{N_{O,T_*}})^2 + (\frac{\sigma_I}{N_{I,T}})^2})
\end{aligned}$$

We ran this test for all players in our sample with $T = \{12, ...17\}$ and were only able to find a single player who violated it at $\alpha = 0.05$. The lone violator was Joel Przybilla (t=-2.463291). These results indicate that NBA players are not taking patently wrong actions. However, we note that the power of this test is quite weak. Typically there is a divergence between a player's "worst shot" (based on shot frequency) and average shot, with the average shot necessarily offering higher efficiency. Put another way, there may be much slack between their average efficiency ($e_{i,t}$) and their marginal efficiency ($c_{i,t}$). If a player's average is consistently below the continuation value, the team would be better off if he did not shoot at all for the interval of the shot clock. Our results show that this optimality condition is rarely violated in NBA play.

3.3 Parametric Uniform Shot Distributions

The conditions of the previous section are quite permissive of potentially suboptimal play for two reasons: 1) they provide no tests for undershooting 2) they give no insight toward allocative efficiency. Furthermore, they offer no convenient way to characterize the overall distribution from which a player realizes scoring opportunities. In this section we address these difficulties. In order to do so w must make a somewhat restrictive assumption. Namely that $\forall t \in T$ player i draws his opportunities from a common distribution. This will allow us to compare player i's performance at the beginning of the shot clock (when he will rarely shoot) to the end of the shot clock (when he will have to shoot a much higher fraction of the time). Intuitively, we are using the shot clock as an instrument to identify player i's opportunity distribution and implied usage curve.

We model player i's scoring opportunities as drawn from a uniform distribution along the interval $[B_i, A_i]$. Additionally, the player selects a shot in period t if his draw lies in the interval of $[c_{i,t}, A_i]$. For the vast majority of players, B_i will take a negative value. This does not mean that players have opportunities to lose points, but merely reflects the fact that the vast majority of players do not realize a good scoring opportunity in most periods of the shot clock. As long as $c_{i,t} > 0$ the part of our theoretical uniform opportunity distribution that lies below zero is irrelevant.

Conditional on the parameters $\theta_i = \{A_i, B_i, \{c_{i,t}\}_{t \in T}\}$, it is straightforward that:

$$e_{i,t} = \frac{A_i + c_{i,t}}{2}, u_{i,t} = \frac{A_i - c_{i,t}}{A_i - B_i}, \frac{de_{i,t}}{du_{i,t}} = A_i - B_i$$

$$Prob(\{E_{i,t}, U_{i,t}\}_{t \in T} | \theta_i, N_{i,t}) = \prod_{t \in T} \phi\left(\frac{U_{i,t} - \frac{A_i - c_{i,t}}{A_i - B_i}}{\sqrt{N_{i,t} \frac{A_i - c_{i,t}}{A_i - B_i}(1 - \frac{A_i - c_{i,t}}{A_i - B_i})}}\right) \phi\left(\frac{E_{i,t} - \frac{A_i + c_{i,t}}{2}}{\frac{\sigma_i}{\sqrt{N_{i,t}}}}\right)$$

We achieve parametric identification because the likelihood matrix is non-singular for all players who choose at least some variation in cutoff levels across periods of the shot clock. This should not be a surprise, if a player did the same thing in every period of the shot clock, our shot clock instrument is irrelevant and we have on hope of learning about his tradeoff between usage and efficiency. It is doubtful that any NBA player's shooting behavior is truly invariant to pressure from the shot clock. However, lack of identification at this point has the potential to create distortions in standard Wald statistics if the truth is in a small enough neighborhood of the unidentified region. We do not believe this difficulty is relevant to the vast majority of players in our data and in this version we will not explicitly address it. But we do note that recent advances in the econometric literature (Dufour et al. 2010; Andrews et al. 2010) may enable us to better understand and be robust to any nonstandard properties of our estimators.

Additionally, our likelihood equation is dependent upon the normality of the distribution of points from an individual shot. In small samples, this is trivially incorrect as the actual distribution of points is discrete and slightly skew. Additionally, we have no reason to believe that our uniform specification is exactly correct. As such we shall not assume that the true data generating process falls within our model and apply consistently estimated covariance matrices from White (1982). Further we shall only perform Maximum Likelihood Estimation (MLE) for each player with at least 15 used possessions in every relevant period of the shot clock. Despite the fact that we are fitting a uniform distribution, the support of observed efficiency and usage rates over any finite sample does not change so we do not have a regularity problem and are able to calculate standard asymptotic variances from the likelihood matrix.

Figure 1 provides the aggregate results of this estimation procedure. We see that indeed NBA players use a monotonically declining cut threshold consistent with the predictions of an optimal stopping problem with finite periods. More impressive is the fact that the cut-thresholds are nearly identical to the continuation values of the possession and the functions have the same shape. We present the results in two panels to enhance the contrast of the slope, while still showing it for all periods of the shot clock. Overall, Figure 1 is strong evidence in favor of near optimal play. NBA players appear to be well-tuned to the continuation value of the possession and adjust their shot choice to reflect it. This is precisely the mechanics of optimal stopping. Not only do they get the mechanics right, but the rate at which the players lower their cut threshold matches the continuation value nearly exactly!

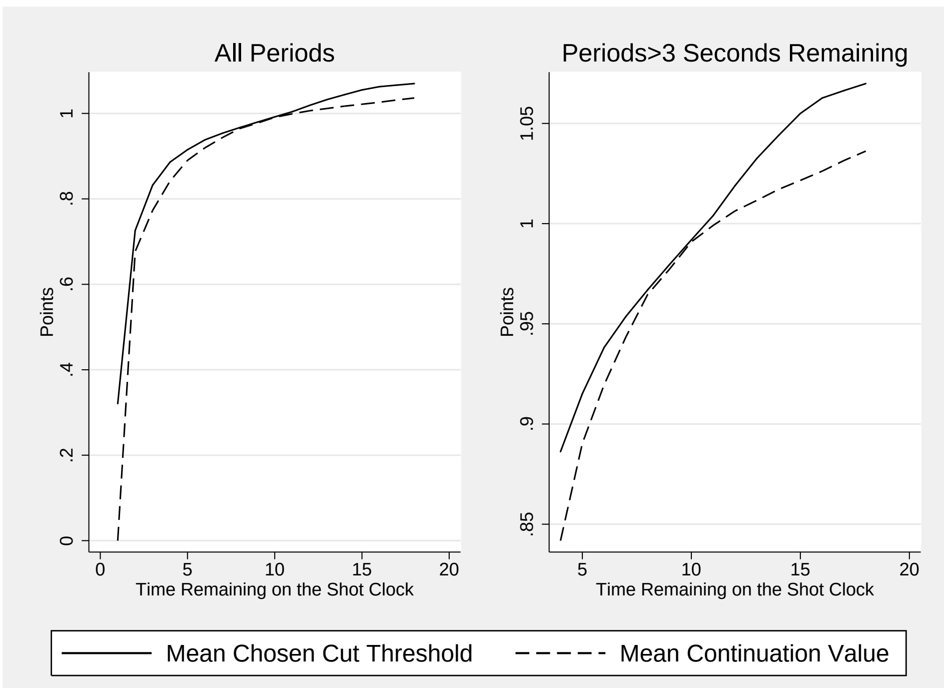


Figure 1: Estimated shooting thresholds (cut-points) and possession continuation values. Left panel: all periods of the shot clock. Right panel: restricted to periods of the shot clock with 4 or more seconds remaining to visually display the slope more clearly.

We do note, however, that in Figure 1 the cut-threshold does lie slightly above the continuation value, which is evidence that undershooting is more common that overshooting. To extend the analysis, we now examine which player's tend to overshoot or undershoot on average. We take as a null hypothesis that each player is dynamically efficient on average. Namely that for each player i, $\sum_{t\in\{2,..,17\}} c_{i,t} = \sum_{t\in\{2,..,17\}} e_{O,t_*}$. To maximize power to detect deviation, we test this hypothesis with a weights inverse to the variance of our estimated cutoffs. Figure 2 provides a histogram of the resulting t-statistics. A negative t-statistic indicates overshooting, a positive t-statistic indicates undershooting.

The results again indicate that undershooting is much more common than overshooting. In fact, only 5 players are found to be significant overshooters — less than we would expect to find by chance alone (although if we take the mean to be 1, not 0, the evidence these players overshoot strengthens considerably.[2] In line with this reasoning, the distribution appears standard normal but shifted over about 1 (the mean 0.98) Most players appear to be optimizing and mistakes tend to come in the form of undershooting early in the shot clock. In this sense, some players wait too long to shoot or do not expend maximum effort on each possession. In contrast, lab subjects tend to pull the trigger too early, typically through the use of a fixed threshold (Lee et al., 2004). We also compute the loss in surplus due to sub-optimal shooting decisions (intuitively integrating between the two lines in Figure 1). The median value of DWL across players is 4%, consistent with nearly optimal play.

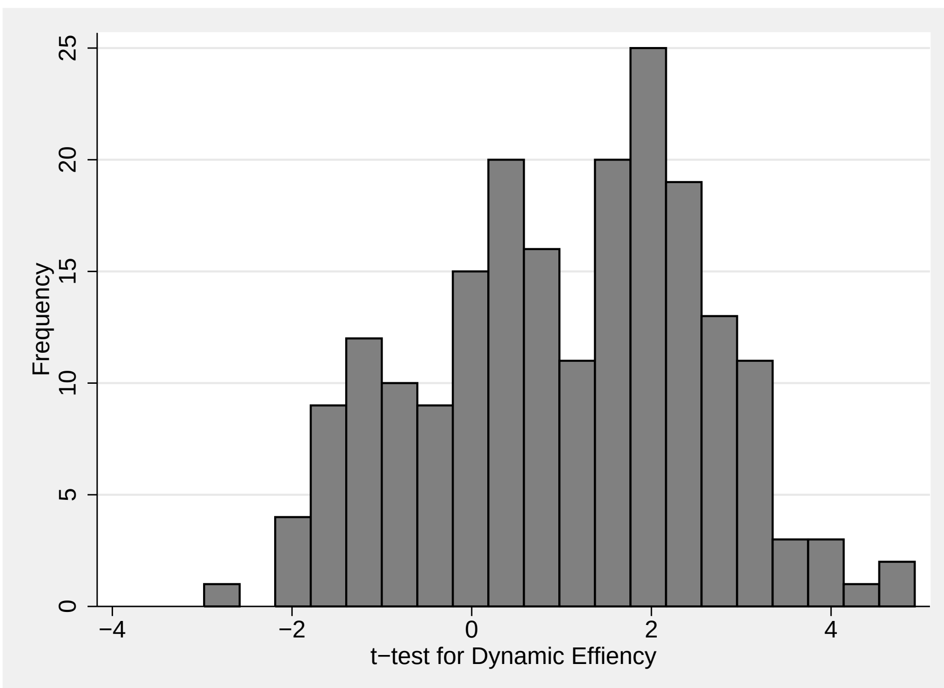


Figure 2: Player-by-player t-statistic for deviations from dynamic optimality. Positive values indicate "under-shooting."

In the discussion section we examine how much of a role effort conservation plays in this finding. While we cannot observe effort directly, we can observe which type of players tend

[2]Lamar Odom, Monta Ellis, Rafer Alston, Russell Westbrook and Tyrus Thomas are the guilty parties.

to undershoot. We find that under-shooters are frequently highly paid “star” players. Our intuition is that some of undershooting is driven by effort conservation by star players who play relatively more minutes-per-game.

3.4 Allocative Efficiency

Allocative Efficiency is the hypothesis that the players on the court, (I) use the same cut-point in all periods of the shot clock. In our estimation we allow the the set I to be an entire five man lineup or a three or four man ”core” that more frequently shares the court. The concept of cores is convenient to improve power in estimation Because our estimation is more precise for these periods, we will focus on cut-points in the first ten seconds of the shot clock ($T = \{9...18\}$). Let c be the $\|T\|\|I\| \times 1$ vector of relevant cut points, sorted first by period of the shot clock. Based on our parametric procedure we have:

$$\hat{c} \sim N(c, \tfrac{V_c}{N})$$

We define the ”true” deviation from Allocative Efficiency as *spread*.

$$S_{I,T} = \sum_{t \in T} \sum_{i \in I} (c_{i,t} - \bar{c}_t)^2 = (c - \bar{c})' I (c - \bar{c}) = z' I z$$

Where $c_{i,t}$ is the cutoff chosen by player i in period t and $\bar{c}_t$ is the average over the five teammates of the cutoff chosen in period t and $\bar{c}$ is the appropriate corresponding $\|T\|\|I\| \times 1$ vector of averages.

Note that, if we define $M = I_T \otimes (I_5 - \frac{J_5}{5})$, we may also simply write $z = Mc$ and $S = c'M'Mc = c'Mc$.

The ”seemingly” natural empirical analog to S is:

$$\hat{S} = \sum_{t \in T} \sum_{i=1}^{5} (\hat{c_{i,t}} - \hat{\bar{c}}_t)^2 = \hat{z}' I \hat{z}$$

However, a little bit of calculation reveals that this measure is biased upwards, especially for small samples.

$$E(\hat{z}) = c - \bar{c}$$

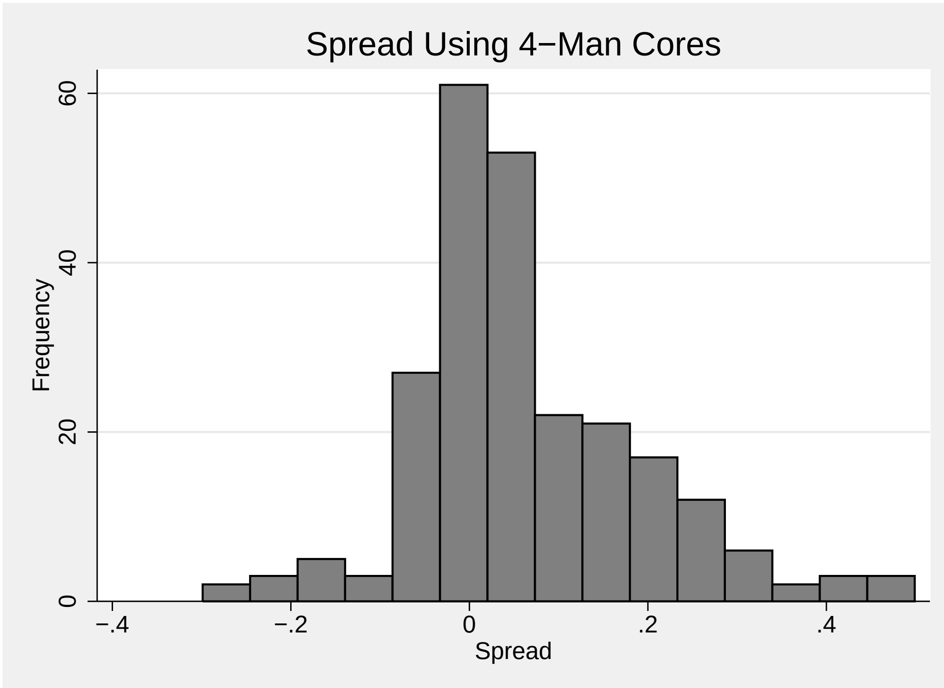


Figure 3: Distribution of *spread* for 4-man cores. Higher values indicate larger deviations from optimality.

$V(\hat{z}) \equiv V_z = \frac{MV_cM}{N}$

Thus: $E(\hat{S}) = \sum E(z_{i,t}^2) = \sum E(\hat{z_{i,t}})^2 + V(\hat{z_{i,t}}) = S + \sum diagV_z = S + \frac{\sum diag(MV_cM)}{N}$

turns out to be upward biased and especially so for small n. In order to correct for this, we essentially subtract off the bias and define:

$\hat{S}^* = \hat{S} - \frac{\sum diag(M\hat{V}_cM)}{N}$

We use this as our measure of dispersion in the marginal output of players sharing the court and refer to simply as *spread*. It turns out that the same 5-man line-ups share the court relatively in frequently. We thus need to define the concept of "4-man core." A 4-man core includes all line-ups that share the same 4 players *and* and the fifth player is occurs less frequently than the core-4. An example is if 4 starters typically share the court with 1 of 3 back-ups (the fifth player). In this case, each back-up plays less than the starters with the line-up, so is not in the 4-core. We then compute *spread* as defined above.

Figure 3 presents the distribution of *spread* for the 246 most common 4-man cores in our data. The median is exceedingly small; it comes in at 0.036 (the mean is slightly larger at 0.06). We also see a large spike at 0 in the distribution. In most line-ups, the players show near equal marginal efficiencies at each period of the shot clock. Again, this is strong evidence of near optimal play, but again play is not perfect. Players seems to be sharing the ball incredibly well, but still about 10% of line-ups are estimate to have spread in excess of 0.25. These line-ups show clear room for improvement in terms of ball allocation.

In the Appendix we present the distribution of spread for 3-man cores. The results confirm the results presented in Figure 3 and in fact the distribution is even more closely centered around 0. This makes sense, in a 3-man core we compute spread over 3 players, instead of 4, that are more familiar with each other and have more similar characteristics. In the discussion section we delve into how *spread* correlates with features of the line-up.

4 Deeper analysis and discussion

The analysis in the preceding section gave sufficient statistics dynamic and allocative efficiency. We now link these measures to features of the line-up such as aggregate efficiency, experience and features of the players.

4.1 Line-up performance, line-up attributes and allocative efficiency

In this section we examine how line-up performance and other attributes of the line-up correlate with adherence to allocative efficiency. We view these regressions and figures as at the very least informative correlations. Clearly we cannot casually improve adherence to optimality and gage the impact on performance. Similarly we cannot exogenously impose an a more experienced line-up and measure the impact on shot allocation. Indeed our estimates themselves provide answers to the question "how much meat is left on the table" (our estimates of DWL are small). One potential endogeneity problem is that a coach might continue to play a high output line-up (driven by talent), despite relatively poor decision making, because the results are still better than experimenting with something new. This would tend to dampen the impact of offensive output on adherence to optimality.[3]

Figure 4 presents a scatter plot of 4-man core *spread* on average output per possession. There is a detectable, but not overwhelming, pattern of more productive line-ups showing

[3] One might expect salary to be a nice control for line-up output, but unfortunately the correlation is not strong. Regressions available from the authors. Essentially there are just too many highly paid lousy line-ups that weaken the signal.

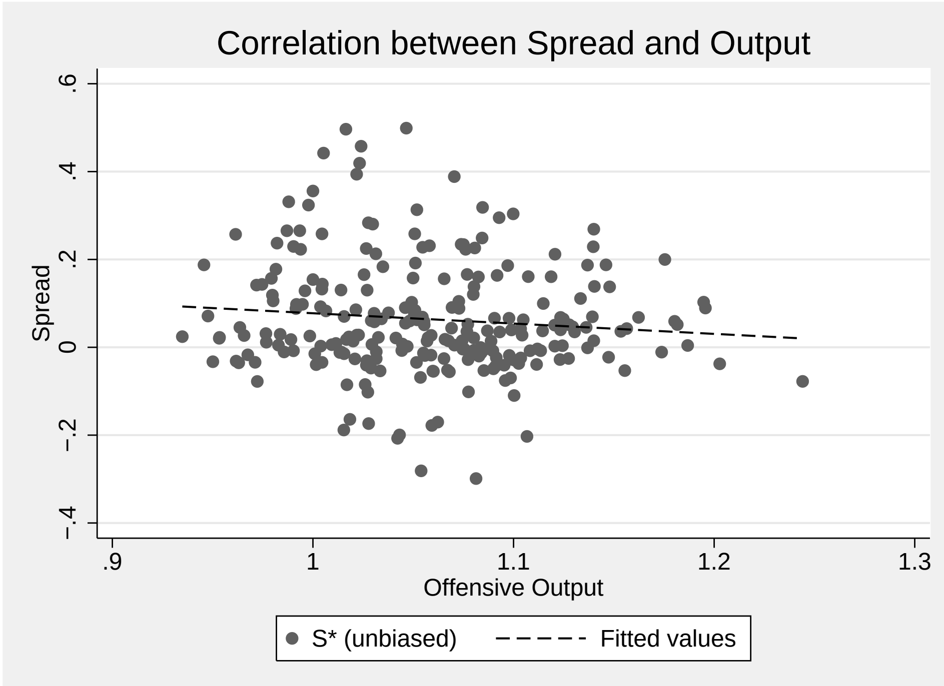


Figure 4: Correlation between spread and output.

lower spread. Table 2 demonstrates that this relationship is significant at the 0.10 level using a two-tailed test. Given that our hypothesis is in fact one-tailed, this significance level can be adjusted to 0.05. Table 2 also gives estimates of other correlates of *spread*. N gives the number of possessions the core shared in the sample. The impact of N is weak and positive. This is not necessarily evidence of a negative impact of repetition because of numerous endogeneity issues based on coaches' decisions. Column (1) includes the average salary of the line-up in logs. Column (2) includes the simple arithmetic mean. For both cases, higher paid line-ups perform significantly better in terms of spread. Higher salary inequality, as measured through the standard deviation, is associated with larger values of *spread*. This result is potentially of interest to labor economists concerned with salary equality and production. "Experience" gives the number of years played in the NBA at the start of 2006. The mean of playing experience across the line-up does not appear to impact adherence to allocative efficiency. The standard deviation of experience seems to have a slight negative impact on *spread*.

Table 2: Impact of line-up characteristics on allocative efficiency (spread).

	(1)	(2)
Output	-0.214*	-0.205
	(0.124)	(0.124)
N	4.41e-05***	4.88e-05***
	(1.61e-05)	(1.69e-05)
ln(Mean Salary)	-0.0862**	
	(0.0425)	
Mean Salary		-1.16e-08*
		(6.23e-09)
s.d. Salary	1.38e-08*	1.30e-08*
	(7.56e-09)	(7.38e-09)
Mean Experience	0.00714	0.00658
	(0.00598)	(0.00625)
s.d. Experience	-0.0171	-0.0182*
	(0.0103)	(0.0104)
Line-ups	238	238

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.2 Aggregate player characteristics and efficiency

In Table 3, we regress the t-stat from adherence to dynamic optimality on individual player characteristics. Consistent with the reasoning offered in Section 3.3, salary is positively related to the t-stat. Recall higher t-stats indicate under-shooting. Indeed the league's star players such as Chris Paul, Lebron James and Kobe Bryant have high t's. Table 3 shows that this the case generally, salary is highly significant in the both the linear specification (1) and linear-log specification (2).

Table 3 also shows that over-shooters tend to offensive rebound specialists; defensive rebounding does not appear to have a reliable impact, although the estimate is noisy. The player position dummies are not significant and are thus suppressed for brevity. The fact that higher paid players are more likely to *under* shoot is perhaps surprising at first. For instance, some readers might have the intuition that NBA players interests diverge from team interests in that they have the incentive to raise their point average through suboptimal play. Under this view, the labor market rewards the wrong attributes (points per game as opposed to efficiency, for example). Our results are inconsistent with the view of labor market. Boosting individual production at a cost to the team is not a strategy employed by NBA players. We think this is interesting in its own right. Teams still have a principle-agent problem in that

Table 3: Impact of player characteristics on dynamic efficiency (t-stat for optimality)

	(1)	(2)
Mean Salary	8.61e-08***	
	(3.13e-08)	
Log Salary		0.502***
		(0.160)
Years Experience	-0.0529	-0.0544
	(0.0382)	(0.0367)
Offensive rebound rate	-30.05***	-29.18***
	(7.432)	(7.648)
Defensive Rebound Rate	4.537	5.369
	(4.186)	(4.130)
Position Dummies	X	X
Players	198	198
R-squared	0.142	0.146

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

long-term contracts create a moral hazard for effort, but it is interesting that very few players exhibit "selfish play." Our belief is that the prevalence of under-shooting among the higher paid players is evidence that the better players conserve energy at times due to their high playing time and long season (over-shooting would be impossible to rationalize this way and we say *far* less over-shooting).

4.3 MSNE, purified games and our contribution

Past work has studied relatively simple games. We study a much more difficult game and in this way really put the theory to the test. Importantly, while this game is difficult, it is far easier than equilibria implied by many models that have complicated MSNE, such as bidding in common value auctions. If we take our theory seriously, we should apply these "hard" tests.

The results are supportive that experts can reach equilibrium in complex games. NBA players appear to be superb optimizers. Furthermore this work incorporate the concept of purified games by modeling MSNE in the context of *dynamic thresholds*. Players adherence to these dynamic thresholds is striking. Exceedingly few are found to over-shoot and within line-ups spread is quite low. We do find that some (especially star) players under-shoot, potentially to conserve energy. Overall all the hallmarks of optimal stopping and MSNE are present.

4.4 Dynamic stopping problems

Past work has also studied stopping problems similar to this one in the lab. It is hard to know how to generalize these studies though, because these are difficult problems that my require experience and training. Perhaps, however, people just do typically have the capacity to solve these complex problems, even with experience. This would be a very important results for models of labor search. There have been papers that have looked at optimal stopping problems using highly trained professionals. These studies use *deliberative* decisions made by *firms* harvesting trees (Provencher, 1997), renewing patents (Pakes, 1986) and replacing bus engines (Rust, 1987). In the field of sports, Romer (2006) is most similar to this paper in the study of coaches decisions to "go for it" on 4th down on the NFL. Again this is a deliberative decision made perhaps 10 times per season by the coach (in this sense it is one that he may not have a lot of experience in making).

This study uses a huge volume of quick decision made players observing a random arrival of shooting opportunities. The findings indicate the players do this quite well, far better than lab subjects. Trained individuals appear quite capable of solving problems with the level of complexity of classical optimal stopping problems.

4.5 Players are not perfect

While NBA players do show striking adherence to optimality overall, they are not perfect optimizers. We find that a minority of players over-shoot and some line-ups show significant deviations from allocative efficiency. Furthermore this can persist even for line-ups that play together quite frequently. Other work has established that minor mistakes are made in NBA player. For instance Rao (2010) finds that a minority of players behaviorally respond to past shot success by taking more difficult shots, despite the fact they perform no better on these shots (hot hand fallacy). Our findings here are consistent with the idea that play is near optimal, but that some players and coaches have room for improvement. The better teams adhere more tightly, which is natural.

5 Conclusion

Past work has studied experts playing relatively simple games. We study a much more difficult game and in this way really put the theory to the test. Importantly, while this game is difficult, it is far easier than equilibria implied by many models that have complicated MSNE, such as bidding in common value auctions. If we take our theory seriously, we should apply these "hard" tests. The unique decision environment we study allows us to extend a stylized stopping problem from the lab to a field setting. Furthermore, we extend the analysis of allocative efficiency across pure strategies to a markedly more difficult game. The trade off is that is more challenging to get hard-and-fast optimality conditions, but our modeling approach telescopes in the strength of assumptions and still provides meaningful conclusions about the adherence to the theoretical standards employed.

The paper uses a huge volume of quick decisions made by players observing a random arrival of shooting opportunities. The decision to shoot is modeled as a dynamic allocative stopping problem. By using a realistic modeling approach we are able to derive strict tests of optimality. We find that players overall adhere quite closely to the theoretical predictions; overall they are suburb optimizers, although mistakes are made. In the context of dynamic efficiency, the shot threshold has precisely the correct slope and nearly overlaps the continuation value of the possession. The mistakes that are made tend to be undershooting, in that the continuation value is lower than the marginal shot; these "mistakes" could be rationalized by the conservation of energy across possessions. In allocative decision making, most teams show a very low variance of marginal efficiencies across players on the court for each interval of the shot clock, consistent with near optimal sharing of the ball. Overall our results extend the realm and difficulty level of games that humans can play according to game theoretic equilbrium.

References

Andrews, D. and Cheng, X. (2010). Estimation and inference with weak identification. http://www.princeton.edu/erp/.

Berri, D. (1999). Who is most valuable? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics*, 20(8):411–427.

Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *The American Economic Review*, 92(4):1138–1151.

Dufour, J.-M. and Valery, P. (2010). Wald-type tests when rank conditions fail: A smooth regularization approach. http://www.fields.utoronto.ca/programs/scientific/10-11/actuarialmath/Dufour_Valery_Regularized_Wald_Test_2010_09_15.pdf.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881.

Harsanyi, J. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23.

Lee, M., Oconnor, T., and Welsh, M. (2004). Decision-Making on the Full Information Secretary Problem. In *Proceedings of the 26th annual conference of the cognitive science society*, pages 819–824. Citeseer.

Lee, M. D. (2006). A hierarchal bayesian model of human decision making on an optimal stopping problem. *Cognitive Science*, 30.

Levitt, S., List, J., and Reiley, D. (2010). What Happens in the Field Stays in the Field: Exploring Whether Professionals Play Minimax in Laboratory Experiments. *Econometrica*, 78(4):1413–1434.

McCabe, K., Mukherji, A., and Runkle, D. (2000). An Experimental Study of Information and Mixed-strategy Play in the Three-person Matching-pennies Game. *Economic Theory*, 15(2):421–462.

Metrick, A. (1995). A Natural Experiment in "Jeopardy!". *The American Economic Review*, 85(1):240–253.

Ochs, J. (1995). Games with Unique, Mixed Strategy Equilibria: An Experimental Study. *Games and Economic Behavior*, 10(1):202–217.

Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis*. Brassey's.

O'Neill, B. (1987). Non-parametric Test of the Minimax Theory of Two-person Zerosum Games. *Proceedings of the National Academy of Sciences*, 84(7):2106–2109.

Pakes, A. (1986). Patents as options: Some estimates of the value of holding european patent stocks. *Econometrica*, 54(4):755–784.

Palacios-Huerta, I. (2003). Professionals play minimax. *Review of Economic Studies*, 70(2):395–415.

Palacios-Huerta, I. and Volij, O. (2008). Experientia Docet: Professionals Play Minimax in Laboratory Experiments. *Econometrica*, 76(1):71.

Palacios-Huerta, I. and Volij, O. (2009). Field centipedes. *The American Economic Review*, 99(4):1619–1635.

Provencher, B. (1997). Structural versus reduced-form estimation of optimal stopping problems. *American Journal of Agricultural Economics*, 79(2):357–368.

Rao, J. M. (2010). Experts' Perceptions of Autocorrelation: The Hot Hand Fallacy Among Professional Basketball Players. Working Paper.

Romer, D. (2006). Do firms maximize? Evidence from professional football. *Journal of Political Economy*, 114(2).

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033.

Seale, D. A. and Rapoport, A. (200). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making*, 13:391–411.

Shachat, J. (2002). Mixed Strategy Play and the Minimax Hypothesis. *Journal of Economic Theory*, 104(1):189–226.

Skinner, B. (2010). The Price of Anarchy in Basketball. *Journal of Quantitative Analysis in Sports*, 6(1):3.

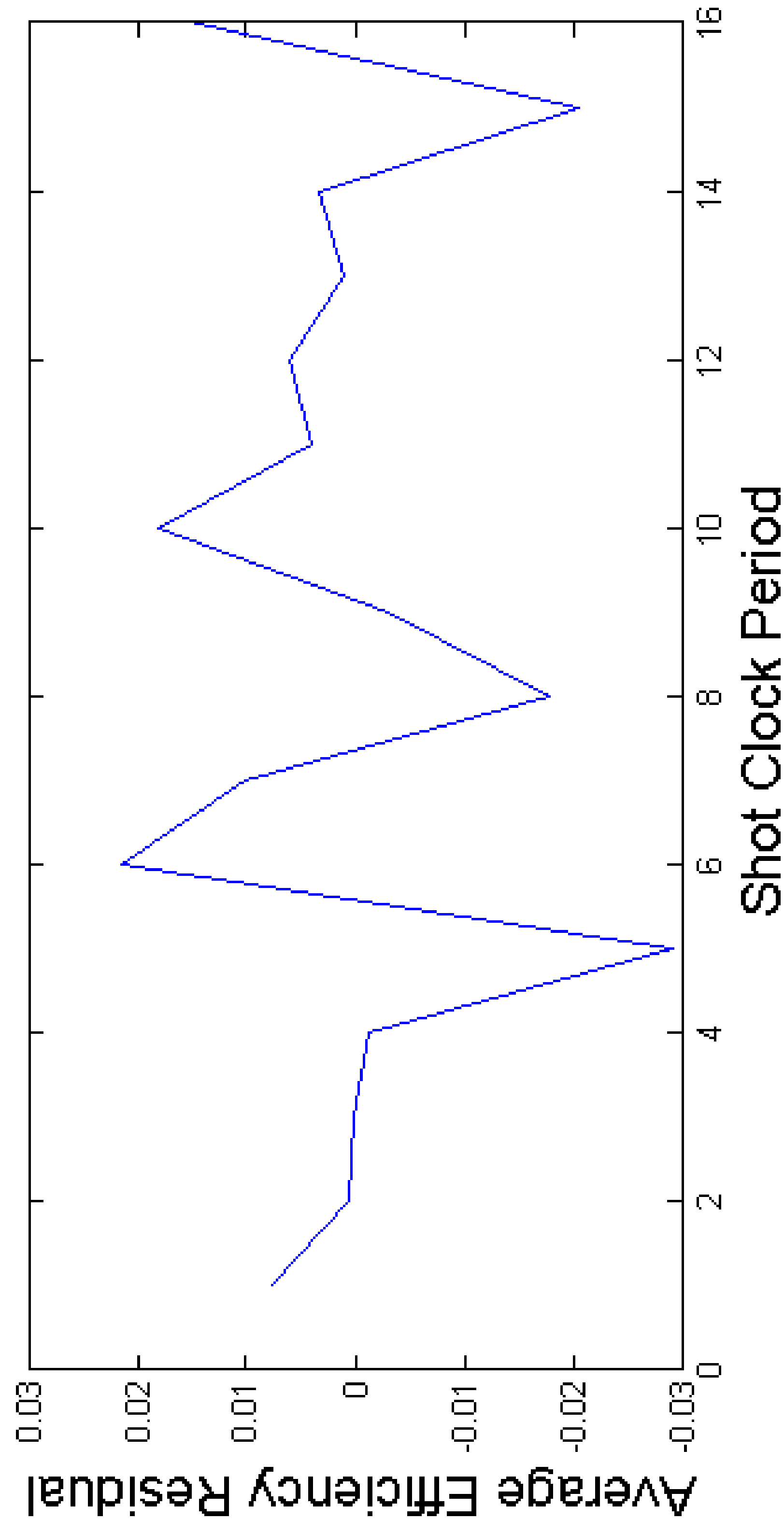
Walker, M. and Wooders, J. (2001). Minimax play at wimbledon. *The American Economic Review*, 91(5):1521–1538.

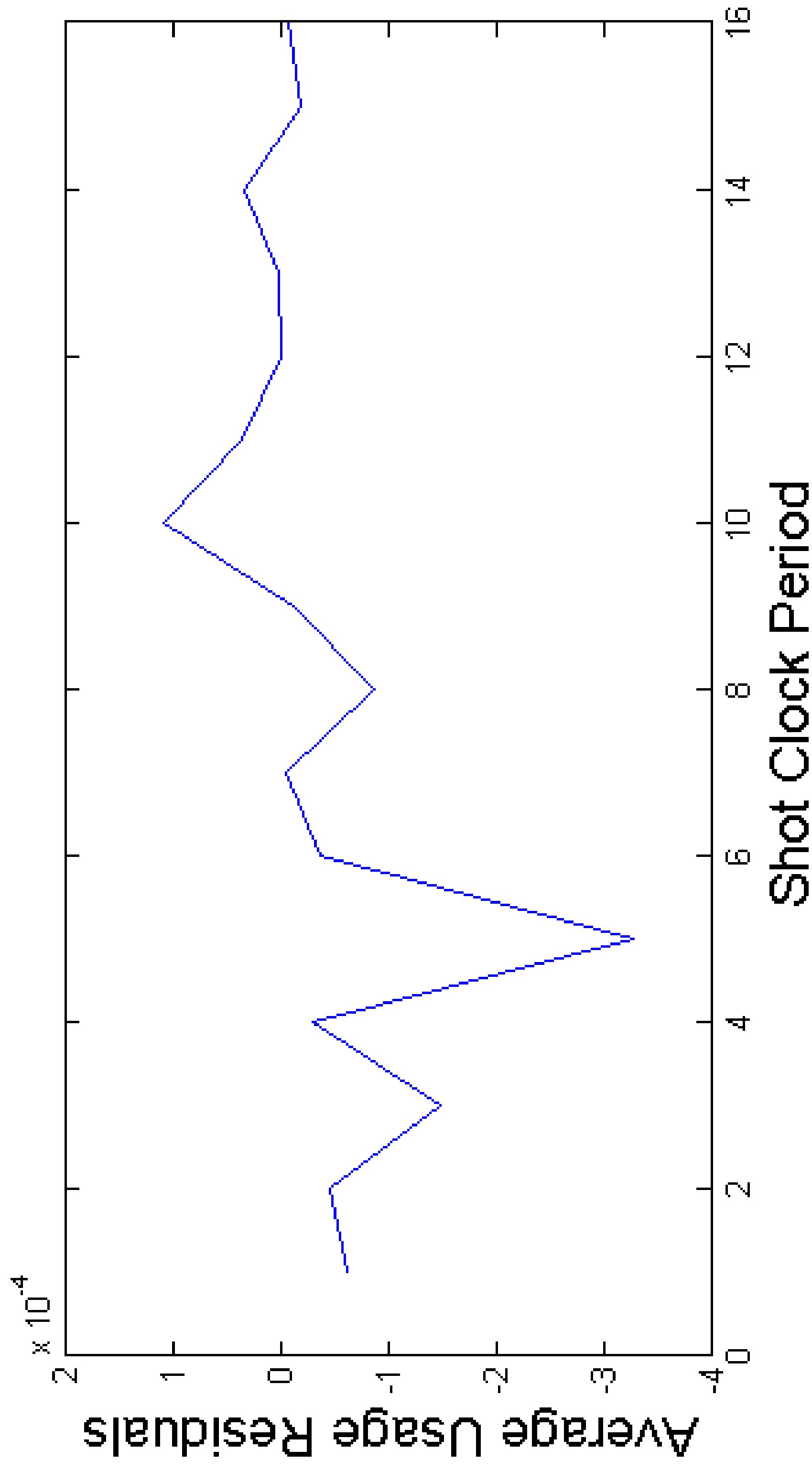
White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, 50(1):1–25.

6 Appendix

6.1 Parametric Model Specification Tests

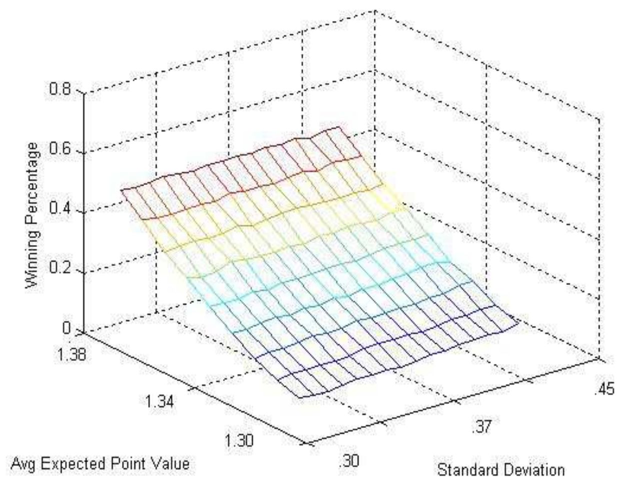
Our model is identified by assuming invariance of a player's ability to realize scoring opportunities across different values of the shot clock. However, because we make this assumption for 16 different periods of the shot clock, our model is overidentified and our assumptions can be tested. Suppose for example that defenses became progressively more tenacious over the course of the shot clock and that player's opportunity distributions generally declined across the shot clock. Then, we would find that a single distribution could not accurately reflect a players ability to score in both the beginning and end of the shot clock. Players would end up shooting less and less efficiently towards the end of the shot clock than our model would predict. In future versions of this paper, we hope to present formal tests of our specification. For now, we present plots showing how our model preforms across the shot clock as an average across all players in our sample. We take solace in the lack of any obvious trends in our residuals.


Average Efficiency Residual
Shot Clock Period
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
0
2
4
6
8
10
12
14
16


x 10^-4
2
1
0
-1
-2
-3
-4
0
2
4
6
8
10
12
14
16
Average Usage Residuals
Shot Clock Period

6.2 Additional support for Risk Neutrality

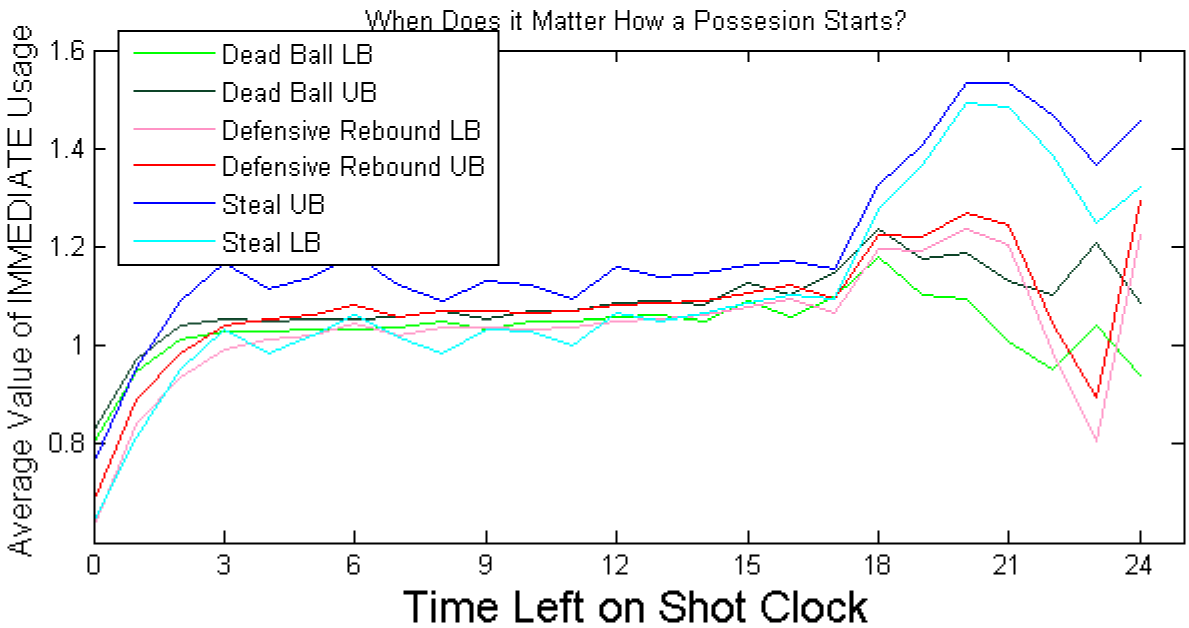
The following plot shows the simulated winning percentage for an underdog with baseline mean expected point value of 1.38 playing a team who averages 1.4 points per possession with standard deviation 0.45. Each "game" was simulated 10,000 times. As evidenced by the figure, although the underdog wants to increase the standard deviation of the expected value of shot attempts it does not want to trade off any meaningful amount of mean to do so.



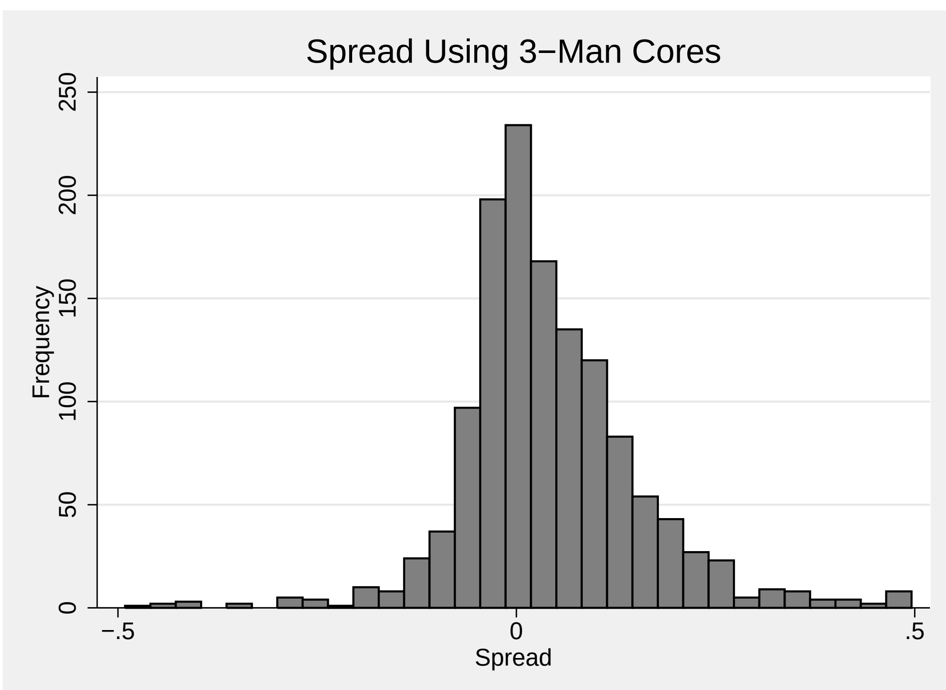
Appendix Figure 1: Underdog winning percentage as a function of standard deviation and mean

6.3 When Does Half-Court Offense Begin?

We decided that half-court offense began with 17 seconds on the shot clock. Our reason for doing so, is that prior to 17 seconds the average value of exercising a possession is found to be strongly correlated with the mechanism by which the possession originated (steal, dead ball, or defensive rebound). However, from 17 seconds on team's are in a half court set and the average value of possession use is now independent of how the possession originated.



6.4 3-man Cores



Distribution of *spread* for 3-man cores. Higher values indicate larger deviations from optimality.