Do Basketball Scoring Patterns Reflect Strategic Play or Illegal Point Shaving? *

Jesse Gregory
University of Wisconsin, NBER
jmgregory@ssc.wisc.edu

Current Version: July, 2014

Abstract

This paper develops and estimates a model of college basketball teams’ search for scoring opportunities, to provide a benchmark of the winning margin distributions that should arise if teams’ only goal is to win. I estimate the model’s structural parameters using first-half play-by-play data from college games and simulate the estimated model’s predicted winning margin distributions. Teams’ optimal state-dependent strategies generate patterns that match those previously cited as evidence of point shaving. The results suggest that corruption in NCAA basketball is less prevalent than previously suggested and that indirect forensic economics methodology can be sensitive to seemingly innocuous institutional features.
JEL Codes: C61, L83 K42

*I am grateful to John Bound, Charlie Brown, Morris Davis, Robert Gillezeau, Sam Gregory, Dmitry Lubensky, Brian McCall, Mike McWilliams, Todd Pugatch, Colin Raymond, Lones Smith, Chris Taber, Justin Wolfers, and Eric Zitzewitz for helpful comments. All remaining errors are my own.
Measuring corruption is inherently difficult because law-breakers cover their tracks. For that reason, empirical studies in forensic economics typically develop indirect tests for the presence of corruption. These tests look for behavior that is a rational response to incentives that only those who engage in the particular corrupt behavior face. The validity of these indirect tests depends critically on the assumption that similar patterns do not occur if agents only respond to the incentives generated by the institutions that govern non-corrupt behavior.

Research designs in forensic economics vary to the extent that they are informed by formal economic theory. In a survey of the field, Zitzewitz (2011) proposes a “taxonomy” of forensic economic research designs that ranges from the entirely atheoretical, in which corrupt behavior is measured directly, to the formally theoretical, in which corrupt behavior is inferred from particular violations of price theory or the efficient-market hypothesis. A common intermediate approach is to posit a statistical model of non-corrupt behavior and to measure the extent to which observed behaviors deviate from that model in a manner that is consistent with corrupt incentives. The soundness of a research design of this variety depends on the plausibility of the assumed statistical model and the extent to which the study’s findings are robust to deviations from the assumed statistical model.

A recent application of this “intermediate” forensic economic strategy purports to find evidence of rampant illegal point shaving in college basketball (Wolfers 2006). Point shaving is when a player on a favored team places a point-spread bet that the opposing team will “cover” the point spread – win outright or lose by less than the point spread – and then manages his effort so that his team wins but by less than the point spread. Because this behavior causes some games that would otherwise end with the favored team winning by more than the median winning margin to win by just below the median winning margin, point shaving tends to increase the degree of right skewness in the distribution of favored teams’ winning margins. Under an assumption that the distribution of winning margins would be symmetric in the absence of point shaving, Wolfers (2006) tests for point shaving by measuring skewness in the empirical distribution of winning margins around the point spread. The skewness-based test suggests that point shaving is rampant, occurring in about six percent of games where one team is strongly favored. Wolfers’ study garnered significant attention in the popular media, reflecting public surprise that corruption might be so pervasive in amateur athletics. While point-shaving scandals have been uncovered with some regularity dating as far back as the early 1950s, the public’s perception seems to be that such scandals are fairly isolated incidents.

To obtain a more formal benchmark of the patterns that one should expect under the no-point-shaving null hypothesis, this paper develops and estimates a dynamic model of college basketball teams’ within-game searches for scoring opportunities. Using play-by-play data from the first halves of NCAA games linked to gambling point spread data, I estimate the model’s structural parameters from play during the first halves of games. With the estimated model, I then simulate play during the second halves of games. I find that the sorts of strategic adjustments across game states (current score and time remaining) that are

---

1 A point-spread bet allows a gambler to wager that a favored team’s winning margin will exceed a given number, the point spread, or bet that the winning margin will not exceed the point spread. A typical arrangement is for the bettor to risk $11 to win $10 for a point spread bet on either the favorite or the underdog.

2 Bernhardt and Heston (2008) cite a group of media outlets in which Wolfers’ (2006) study was featured. These include the “New York Times, Chicago Tribune, USAToday, Sports Illustrated and Barrons, as well as National Public Radio and CNBC TV.”
commonly observed in actual games are consistent with an optimal policy. Further, I find that the scoring patterns generated by optimal policies generate skewness patterns that closely match those previously cited as evidence of illegal point shaving. The analysis suggests that the existing skewness-based test greatly exaggerates the prevalence of point shaving.

In the model, teams take turns as the offensive side searching for scoring opportunities. The offensive side faces a sequence of arriving shot opportunities which vary in their probability of success. As in actual NCAA basketball games, the offensive side has 35 seconds to attempt a shot before the opponent is automatically awarded the ball. Like a worker in a job-search model, the searching team compares each arriving opportunity with the value of continued search. Because of the fixed horizon, the optimal strategy within a possession is a declining reservation policy where initially only the most advantageous opportunities are accepted and less advantageous opportunities become acceptable as time goes by. The optimal reservation policy depends on the current relative score and the time remaining in the game. Especially near the end of the game, a trailing team prefers to hurry by taking short possessions, and the leading team prefers to stall by taking long possessions. I show that the direction of skewness that this process introduces to the distribution of the stronger team’s winning margin depends on whether stalling incurs the larger opportunity cost (causing left skewness) or hurrying incurs the larger opportunity cost (causing right skewness). The parameters of the search process determine the opportunity cost of stalling and the opportunity cost of hurrying.

I estimate the model’s parameters using play-by-play data from the first halves of games, and I find that under the estimated parameter values stalling is less costly than hurrying. As a result, the leading team makes larger strategic adjustments than the trailing team, and the score difference tends to shrink (or grow more slowly) on average compared to what would occur if each team chose the strategy that maximized its expected points per possession. These optimal adjustments result in a right-skewed distribution of winning margins in games in which one team is a large favorite. As a false experiment, I apply the skewness-based test for point shaving to simulated data. I find “evidence” of point shaving in the two highest point-spread categories, and the implied prevalence of point shaving is statistically indistinguishable the Wolfers (2006) estimate.\footnote{It should be noted that Wolfers’ (2006) article acknowledges the limitations of the skewness-based test relative to a test more formally grounded in theory, and characterizes the resulting estimates as “prima facie” evidence of widespread point shaving. The structural approach that Wolfers’ article outlines as a possible extension differs from the approach in this study. Wolfers suggests that a structural model of the point shaver’s behavior might allow for a more accurate inference about the prevalence of point shaving based on the observed deviation from symmetry in the empirical distribution. This study adopts a structural approach to more accurately characterize the distribution one might expect under the no-point-shaving null.}

Studies in economics that treat sports as a research laboratory are sometimes criticized as being unlikely to generalize to more typical economic settings. However, for forensic economics, studies involving sports have provided particularly compelling case studies. Participants’ willingness to engage in illegal behavior in the highly monitored environment of sports competitions suggests that corruption likely plays a more prominent role in less well-monitored settings (Duggan and Levitt, 2002; Wolfers, 2006; Price and Wolfers, 2010; Parsons, Sulaeman, Yates, and Hamermesh, 2011). The high-quality data and well-defined rules and institutions that are perhaps unique to sports also provide an excellent opportunity to assess the robustness of forensic economic methodology. This study’s findings suggest that forensic economic studies should take great care to assess the robustness of their methods to unmodelled assumptions about even seemingly
innocuous institutional features. Further, the findings suggest that structural modeling can yield improved predictions of behavior under the no-corruption null hypothesis even in settings in which off-the-shelf price theory and efficient market theory do not yield immediate predictions.

The conclusions of this paper conform with those of Bernhardt and Heston (2008). Bernhardt and Heston find that the patterns attributed to point shaving by Wolfers (2006) are present in subsets of basketball games in which gambling related malfeasance is less likely on prior grounds. The authors conclude that even in the absence of point shaving, asymmetries exist in the distribution of the final-score differentials among games in which one team is a large favorite. Because the approach of this paper and the purely empirical approach of Bernhardt and Heston (2008) are vulnerable to different criticisms, I consider my study a complement to their work.

The remainder of the paper is organized as follows: Section 1 develops a simple illustrative model, Section 2 develops a richer model of the basketball scoring environment, Section 3 describes estimation, Section 4 describes the data, Section 5 presents point estimates of the model’s structural parameters and examines the model’s fit, Section 6 presents the results of simulations calibrated with estimated parameters, Section 7 provides corroborating evidence for the main model-based results, and Section 8 concludes.

1 Illustrative Two-Stage Dynamic Game

Before turning to the full dynamic model, I use a simple two-stage model to illustrate the kinds of model primitives that can rationalize a skewed distribution of winning margins in a dynamic competition.

In this simple model, two competitors A and B accumulate points during a two stage game. Let \( X_1 \) represent the difference between A’s and B’s points during stage 1, and let \( X_2 \) represent the difference between A’s and B’s points during stage 2. Following stage 2, A receives a payoff of one if \( X_1 + X_2 \geq 0 \) and zero otherwise, and B receives a payoff of one if \( X_1 + X_2 < 0 \) and zero otherwise.

A random component influences the scoring process. Before each stage, A and B each select an action that influences the mean and the variance of scoring during that stage. Before stage 1, A selects \( \sigma_{A1} \in [0, 1] \) and B selects \( \sigma_{B1} \in [0, 1] \). Stage 1 then occurs, and both competitors observe the realized value of \( X_1 \). Then before stage 2, A selects \( \sigma_{A2} \in [0, 1] \) and B selects \( \sigma_{B2} \in [0, 1] \). The quantities \( X_1 \) and \( X_2 \) are given by,

\[
X_1 = \mu(\sigma_{A1}) - \mu(\sigma_{B1}) + \left( \sigma_{A1} + \sigma_{B1} \right) Z_1 + \Delta \\
X_2 = \mu(\sigma_{A2}) - \mu(\sigma_{B2}) + \left( \sigma_{A2} + \sigma_{B2} \right) Z_2 + \Delta
\]

where \( \mu() \) is a twice differentiable real-valued function that describes the relationship between a players chosen action and the mean of scoring, \( Z_1 \) and \( Z_2 \) are standard normal random variables that are independent from one another, and \( \Delta \geq 0 \) is a constant that allows for the possibility that A is stronger than B. I assume

Bernhardt and Heston (2008) suggest that the goal of maximizing the probability of winning could induce an asymmetric final-score distribution, but stop short of suggesting a theoretical model.

[4]Bernhardt and Heston (2008) suggest that the goal of maximizing the probability of winning could induce an asymmetric final-score distribution, but stop short of suggesting a theoretical model.
that $\mu()$ is bounded, strictly concave, and reaches an interior maximum on $[0, 1]$. The choice of a very high variance ($\sigma$ near one) or a very low variance ($\sigma$ near zero) involves a lower expected value of scoring. To ensure an interior solution, I assume that the $\mu'(\sigma)$ is unbounded, going to $-\infty$ as $\sigma$ approaches 1 and going to $\infty$ as $\sigma$ approaches 0 (An example of a function satisfying these assumptions is a downward-facing semi-circle).

Because each competitor maximizes a continuous function on a compact set, the minimax theorem ensures that a solution to this zero-sum game exists, and inspection of the players’ best reply-functions finds that the solution is unique. The optimal stage 2 actions $(\sigma_{A2}^*, \sigma_{B2}^*)$ satisfy the first order condition,

$$\mu'(\sigma_{A2}^*) = \frac{X_1 + \mu(\sigma_{A2}^*) - \mu(\sigma_{B2}^*) + \Delta}{\sigma_{A2}^* + \sigma_{B2}^*} = -\mu'(\sigma_{B2}^*) \tag{3}$$

Figure 1 plots $A$’s iso-payoff curves along with the choice set $\mu()$ to illustrate the reasoning behind this result. If $X_1 + \Delta = 0$, $A$ and $B$ each have a equal chance of winning heading into stage 2, and each chooses the strategy that maximizes its expected points. If $X_1 + \Delta > 0$, $A$’s chance of winning is greater than 1/2, and $A$ selects a relatively low-variance action in order to reduce the chances of a come from behind win for $B$. If $X_1 + \Delta < 0$, $A$’s chance of winning is less than 1/2, and $A$ selects a relatively high-variance action in order to increase its chances of a come from behind win.

In stage 1, $A$ and $B$ each have a unique optimal strategy. Define $V(X_1)$ to be $A$’s expected payoff entering stage 2 if the relative score after stage 1 is $X_1$. The optimal stage 1 actions $(\sigma_{A1}^*, \sigma_{B1}^*)$ are the solution to,

$$\max_{\sigma_{A1}} \min_{\sigma_{B1}} \int_{-\infty}^{\infty} V\left(\mu(\sigma_{A1}) - \mu(\sigma_{B1}) + \Delta + \left(\sigma_{A1} + \sigma_{B1}\right)z\right) d\Phi(z) \tag{4}$$

where $\Phi()$ is the standard normal CDF.

Now consider how the players’ optimal strategies influence the shape of the distribution of winning margins. One might expect for the winning margin $X$ to be symmetrically, because $X = X_1 + X_2$ is the sum of two normally distributed random variables. That is not the case in general, because the players’ choices introduce dependence between $X_1$ and $X_2$. In two particular cases, the distribution of $X$ is symmetric. The first case relies on the symmetry of the game when $A$ and $B$ are evenly matched.

**Proposition 1:** The winning margin is symmetric when $A$ and $B$ are evenly matched: Assume that $\Delta = 0,$ which means that the two teams are evenly matched. Then the distribution of winning margins is symmetric.$^5$

Symmetry also holds when $A$ and $B$ are not evenly matched if the menu function $\mu()$ is symmetric. In that case $A$’s and $B$’s strategic adjustments, based on the realized value of $X_1$, have exactly offsetting impacts on the mean and the variance of $X_2$.

---

$^5$This result follows directly from the symmetry of the game. Because the solution is unique, switching the names of $A$ and $B$ cannot change the distribution of winning margins.
Proposition 2: The winning margin is symmetric (normally distributed) when increasing variance and decreasing variance are equally “costly”: Assume that \( \mu() \) is symmetric, that is, that \( \mu(.5) = 0 \) and that \( \mu(.5 + \epsilon) = \mu(.5 - \epsilon) \) for any \( \epsilon \in [0, .5] \). Then, \( E(X_1) = E(X_2|X_1) = \Delta, \sigma_{A1} + \sigma_{B1} = \sigma_{A2} + \sigma_{B2} = 1, \) and \( X \sim N(2\Delta, 2) \). Proof provided in Appendix I.

When \( \mu() \) is symmetric, \( E(X_2|X_1) \) is constant. In general, though, \( E(X_2|X_1) \) is not constant, and dependence between \( X_1 \) and \( X_2 \) can generate skewness in the sum of \( X_1 \) and \( X_2 \). Proposition 3 provides conditions under which \( E(X_2|X_1) \) monotonically increases or decreases in \( X_1 \).

Proposition 3: The monotonicity of \( E(X_2|X_1) \) if strategic adjustment of the variance is more difficult in one direction than the other: Let \( \sigma^* \) be the action that maximizes \( \mu() \). Assume that for any \( \sigma'< \sigma^* \) and \( \sigma'' > \sigma^* \) with \( u'(\sigma') = -u'(\sigma) \) that \( |u''(\sigma')| < |u''(\sigma'')| \) (a player who prefers to increase variance faces a steeper marginal cost than a player who prefers to reduce variance). Then \( E(X_2|X_1) \) monotonically decreases in \( X_1 \). Conversely, assume that for any \( \sigma'< \sigma^* \) and \( \sigma'' < \sigma^* \) with \( u'(\sigma') = -u'(\sigma) \) that \( |u''(\sigma')| > |u''(\sigma'')| \) (a player who prefers to reduce variance faces a steeper marginal cost than a player who prefers to increase variance). Then \( E(X_2|X_1) \) monotonically increases in \( X_1 \). Proof provided in Appendix I.

The results thus far characterize the slope of the expected second-half scoring “drift” \( E(X_2|X_1) \) with respect to \( X_1 \). Proposition 4 suggests that the direction of skewness in the winning margin distribution depends on the curvature of \( E(X_2|X_1) \) in \( X_1 \).

Proposition 4: A sufficient condition for determining the direction of skewness in winning margins: Assume that \( E(X_2|X_1) \) is convex in \( X_1 \). Then the winning margin is right skewed. Conversely, assume that \( E(X_2|X_1) \) is concave in \( X_1 \). Then the winning margin is left skewed. Proof: An application of Zwet (1964). See Appendix I for more details.

Because the function \( \mu() \) is bounded, \( E(X_2|X_1) \) is also bounded and, therefore, may not be convex or concave over all values of \( X_1 \). Nonetheless, propositions 3 and 4 together with the boundedness of \( E(X_2|X_1) \) suggest circumstances under which one might expect to find skewness in the distribution of winning margins. Figure 2 illustrates this idea. When \( \Delta \) is large, \( X_1 \) will typically be large. If \( E(X_2|X_1) \) monotonically decreases in \( X_1 \) but is bounded, one might expect \( E(X_2|X_1) \) to be convex over many large values of \( X_1 \). In that case, one would expect the distribution of winning margins in games with large \( \Delta \) to be right skewed. Similarly, if \( E(X_2|X_1) \) monotonically increases in \( X_1 \) but is bounded, one might expect \( E(X_2|X_1) \) to be concave over many large values of \( X_1 \). In that case, one would expect the distribution of winning margins in games with large \( \Delta \) to be left skewed.

Wolfers’ (2006) statistical test attributes right skewness in the distribution of the favorite’s winning margins to the influence of point shaving. If winning margins are right skewed in the absence of point shaving, then that test overstates the true prevalence of point shaving. On the other hand, if winning margins are left skewed in the absence of point shaving, then that test understates the true prevalence of point shaving.
The remainder of the paper considers a model of basketball teams’ search for scoring opportunities in order to assess whether a skewed distribution of winning margins is consistent with optimal strategies. In basketball games, the two teams alternate turns as the offensive side, and the game ends after a fixed amount of time has passed. The total number of possessions that occurs has a significant influence on the variance of the score. During a given possession, a team can reduce the variance of the scoring that will occur during the remainder of the game by stalling; that is, taking a lot of time before attempting a shot. A team can increase the variance of the scoring that will occur during the remainder of the game by hurrying; that is, attempting a shot very quickly.

The reasoning behind this simple model suggests that the favored team’s winning margin should be right skewed if increasing the variance of points has a lower opportunity cost than reducing the variance of points. In the context of the richer model that I consider, that condition is satisfied when stalling has a lower opportunity cost than hurrying. The estimated model finds that, indeed, the opportunity cost to stalling is lower than the opportunity cost of hurrying in NCAA basketball games. Further, the simulations that I conduct using the estimated model suggest that teams’ optimal end-of-game adjustments generate skewness patterns that closely resemble those previously attributed to point shaving.

2 Full Model

2.1 Model

In this section, I propose a more detailed model of the environment in which basketball teams compete. Teams alternate turns as the offensive side, and the team that is on offense faces a sequence of arriving shot opportunities that vary stochastically in quality. I refer to a single turn as the offensive side as a possession. The team that is searching for a scoring opportunity must compare the potential reward of each opportunity with the option value of continued search. As such, I model the game as a sequence of many alternating periods of finite horizon search.

In this model, the tradeoff between the expected value of points and the variance of points is endogenous. A unique reservation policy exists that maximizes a team’s expected points per possession. If the team chooses to hurry, by selecting a lower reservation policy, its expected points per possession will be lower. Also if it stalls by choosing a very high reservation policy, its expected points will be lower. The duration of a given possession influences the variance in the score in the remainder of the game through its influence on the total number of possessions that will occur. Hence, as in the simple model, an opportunity cost is associated with choosing a strategy that substantially raises or substantially reduces the variance of points.

In this more complex model, the team that is on defense is given the opportunity to intentionally foul the offensive team. I include this feature because the strategy is ubiquitous in the final minutes of actual NCAA basketball games, and the strategy influences the distribution of winning margins in a way that the simulation experiments I conduct find to be important.

In the model, two teams, $A$ and $B$, accumulate points during a single competition with a typical duration of $T$ seconds. In NCAA basketball games, $T = 2400$ seconds (40 minutes). If the score is not tied at time $T$, then the game ends. If the score is tied at time $T$, then play is repeatedly extended by an additional 300
seconds (five minutes) until a winner is determined. Two measures of time are relevant. Let $t \in [0, 1, 2, ..., T]$ measure the time since the game began in discrete periods of one second each. At any $t$, let $s \geq 0$ describe the number of seconds since the current possession began. The variable $s$ corresponds to the “shot clock” in NCAA basketball games.

One team at a time is on offense searching for a scoring opportunity. Let $o \in \{A, B\}$ indicate the team that is on offense. The team that is on offense faces a sequence of scoring opportunities. A scoring opportunity is characterized by a pair of variables, $p$ and $\pi$. The variable $p \in [0, 1]$ describes the probability with which an opportunity will succeed if it is accepted. The variable $\pi \in \{1, 2, 3\}$ describes the number of points that will be awarded if the opportunity is accepted and succeeds. For standard opportunities, $\pi \in \{2, 3\}$, and the particular case when $\pi = 1$ is described below.

Each period, time $t$ increases by one. The team on offense switches if one of three events occurs; the team on offense accepts a scoring opportunity at $t - 1$, the possession duration $s$ reaches a maximum limit or 35 seconds at $t - 1$ (known as a shot clock violation), or an exogenous turnover occurs between period $t - 1$ and $t$. The arrival process for turnovers is described later. If the offensive team does not switch from one period to the next, $s$ increases by one. Otherwise, $s$ is reset to zero.

The variable $X = X_A - X_B$ measures the difference in the two teams’ accumulated point totals. When a team accepts a shot opportunity, the uncertainty is resolved and the offensive team receives either 0 or $\pi$ points.

Scoring opportunities and turnovers arrive stochastically. I impose that no turnovers or shot opportunities arrive during the first five seconds of the possession to reflect the time required in actual games for the offensive team to move the ball from its defensive portion of the court to its offensive portion. Following this initial five-second span, one scoring opportunity arrives each second. The variables that describe a scoring opportunity are drawn from the conditional density function $f(p, \pi|t, s, X, o)$. I impose the simplifying assumption that the distribution from which $p$ and $\pi$ are drawn depends only on the team that is offense but does not otherwise vary. That is $f(p, \pi|t, s, X, o) = f(p, \pi|o)$. I also assume that draws of $(p, \pi)$ are independent across time. Finally, I assume that turnovers arrive with a constant probability for each team, and I let $\upsilon_A$ and $\upsilon_B$ denote the turnover probabilities when the team on offense is $A$ or $B$. I refer to these assumptions jointly as a conditional independence assumption (CIA). These assumptions facilitate estimation of the model and allow me to conduct simulations of end-of-game play using parameters estimated from data on play early in games.

Both the offense and the defense face a choice during each one-second increment. If no turnover occurs, the team on defense has the option to intentionally foul the team on offense. The defense’s choice of whether to foul or not occurs before realizations of the draw of $(p, \pi)$. When an intentional foul is committed, the offense is granted two one-point ($\pi = 1$) scoring opportunities during the same period known as free throws. Free throws succeed with known (to the teams) probabilities $p_{ft}^A$ and $p_{ft}^B$. If no turnover occurs and the defense chooses not to intentionally foul, a shot opportunity $(p, \pi)$ is drawn. The offense must choose whether to accept the current opportunity or to continue searching. I denote the choice spaces of the defense

---

6The assumption that a scoring opportunity arrives every second is not as restrictive as it might seem, because the arrival of a very poor scoring opportunity (one that succeeds with very low probability) is no different from a non-opportunity.
and offense with \( A^D = \{0, 1\} \) and \( A^O = \{0, 1\} \). For the team on defense, 1 represents the choice to intentionally foul. For the team on offense, 1 represents the choice to accept a scoring opportunity.

When the game ends, the teams receive payoffs \( U^A \) and \( U^B \). The team with the higher score receives a payoff of one and the team with the lower score receives a payoff of zero. Because the model is used to benchmark the patterns one should expect in the absence of corruption, teams’ only objective is to win the game. Each team chooses its strategy to maximize its expected payoff. Define the value function \( V^A(\omega) = E[U^A|\omega] \) to be \( A \)'s expected payoff from the state \( \omega \). Because the game is zero-sum, this value function also characterizes \( B \)'s expected payoff. The uniqueness of this expected payoff function follows directly from the uniqueness of the teams’ reservation policies.

By CIA and the assumed process for the state transitions, state transitions are Markovian. That property allows the optimization to be expressed as the following dynamic programming problem.

\[
V^A(\omega) = \max_{a \in A^A(\omega)} \left\{ \min_{b \in A^B(\omega)} \left\{ E[V^A(\omega')|\omega, a, b]\right\} \right\}
\]  

The sets \( A^A(\omega) \) and \( A^B(\omega) \) represent the teams’ choice sets given the current state. A team’s choice set is \( A^O \) if \( \omega \) indicates that the team is on offense and is \( A^D \) otherwise.

### 2.2 Approximate Model Solution

The model developed in the previous section does not have an analytic solution. However, for any parameterization of the density function \( f \), a numerical solution can be computed using backward induction. I compute the full numerical solution to the model when performing dynamic simulations in section 6. In this section, however, I develop an approximate solution method that describes the optimal policy within a single possession. I use this approximate solution as the basis for estimating the model’s parameters.

Define the function \( EV \) to be \( A \)'s expected payoff in a state prior to the realization of the offensive team’s shot opportunity,

\[
EV([t, s, X, o]) = E_{(p', \pi')}\left[ V\left([t, s, X, o, p', \pi']\right) \right]
\]

I now construct a linear approximation of \( EV \) within a single possession that begins in the state \( \omega_0 = [t, s = 0, X, o] \). Let \( s^* \) denote the duration of the possession, and let \( x^* \) denote the change in \( X \) that occurs as a result of the possession. Then the state of the following the possession is \( \omega^* = [t + s^*, s = 0, X + x^*, -o] \), where \( -o \) indicates the opponent of team \( o \). I next note that the expected payoff in the state \( \omega^* \) can be approximated linearly using,
\[
\begin{align*}
EV^A(\omega^*) & \approx EV^A(\omega_0) + EV^A_X(\omega_0) x^* + EV^A_t(\omega_0) s^* \\
EV^A(\omega^*) & \approx EV^A(\omega_0) + EV^A_X(\omega_0) \left( x^* + \phi(\omega_0^*) s^* \right)
\end{align*}
\]  

(7)

where \( EV^A_X = EV([t, s, X + 1, o]) - EV([t, s, X, o]) \), \( EV^A_t = EV([t + 1, s, X, o]) - EV([t, s, X, o]) \), and \( \phi(\omega_0^*) = EV^A_t(\omega_0^*)/EV^A_X(\omega_0^*) \). The term \( \phi(\omega_0^*) \) can be thought of as a marginal rate of substitution between time and points. The partial effect of one second passing on \( A \)'s expected payoff is the same as a \( \phi(\omega_0^*) \) point change in the relative score.

Now note that because \( EV^A(\omega^*) \) is a positive affine transformation of the expression \( \left( x^* + \phi(\omega_0^*) s^* \right) \), the within-possession policy that maximizes \( EV^A \) is the same as the policy that maximizes \( \left( x^* + \phi(\omega_0^*) s^* \right) \). Thus, up to a linear approximation of the value function, all strategically relevant information from the states \( t \) and \( X \) is embedded in \( \phi(\omega_0^*) \).

The optimal policy for the team on offense is a reservation rule that I denote with \( R(s; \phi) \). For each value of \( s \) within the possession, the reservation rule describes the point value \( p\pi \) for which the offense is indifferent between accepting the opportunity and opting for continued search. Using the linear approximation developed above, it is straightforward to show that when \( A \) is on offense the optimal reservation rule is,

\[
R^A(s; \phi) = E(x^* | s^* > s) + \phi E(s^* - s | s^* > s)
\]  

(8)

Given the fixed boundary at \( s = 35 \), the reservation value can be defined recursively. First, define the auxiliary function \( z^A(s; \phi) = E[s^* x^* | s^* \geq s] \). I first construct and expression for \( z^A \), and then use \( z^A \) to construct \( R^A \).

\[
\begin{align*}
z^A(s; \phi) &= 35 u_A \phi + (1 - v) E \max \left( p\pi, 0 \right) & \text{if } s = 35 \\
&= s u_A \phi + (1 - v) E \max \left( p\pi, z^A(s + 1; \phi) \right) & \text{if } s < 35
\end{align*}
\]  

(9)

\[
\begin{align*}
R^A(s; \phi) &= 0 & \text{if } s = 35 \\
&= (1 - u_A) \left[ E \max \left( p\pi, z^A(s + 1; \phi) \right) - s \phi \right] & \text{if } s < 35
\end{align*}
\]  

(10)

Recall that \( v \) is the hazard of a turnover in each one-second period. The optimal strategy is a declining reservation policy. An analogous derivation yields the optimal policy for team \( B \).

To illustrate how the optimal policy varies across game states, figure 3 plots the predicted reservation policies in three different game states holding the model parameters constant; one for which \( \phi = 0 \), one in which the \( \phi < 0 \) (offensive team trails late in game), and one in which \( \phi > 0 \) (offensive team leads late in game). The plots are constructed using the search parameters estimated for the favored team in games with a point spread between 0 and 4. When \( \phi = 0 \), the intermediate reservation policy is implemented. That is the
point-maximizing policy. When \( \phi < 0 \) a lower reservation policy is chosen, which leads to shorter average possessions and fewer points per possession than when \( \phi = 0 \). When \( \phi > 0 \) a higher reservation policy is chosen, which leads to longer average possessions and fewer points per possession than when \( \phi = 0 \).

3 Estimation

I estimate the parameters governing the search process separately for each of six categories defined by the size of the point spread. The first five point-spread categories are each four points wide, and the sixth category includes all games with a point spread above 20. I estimate one set of parameters describing the favorite’s search process and another set of parameters describing the underdog’s search process. Favorites and underdogs in the six point-spread categories make up twelve team types. This approach allows me to conduct separate simulation experiments by point-spread categories.

I estimate the model’s parameters by maximum likelihood using data from the first halves of games. Estimating the model’s parameters from first-half possession data requires two approximating assumptions. I assume that during that first half of the game the team on defense does not intentionally foul.\(^7\) I also impose that \( \phi = 0 \), an approximation that is reasonable because the importance of expected scoring swamps the importance of possession duration early in the game. In addition to the computational savings that result from the \( \phi = 0 \) simplification, restricting attention to first-half possessions is also appealing, because the approach ensures that the model’s parameters are identified from choices that are entirely distinct from the second-half choices that the model is used to investigate.

For estimation, I impose a parametric functional form for the joint density \( f(\pi, p|o) \); the joint density from which shot opportunities come for a given offense. Because \( \pi \) is discrete, it is convenient to express the joint density as the product of a probability mass function and a conditional density; \( f(\pi, p) = f_\pi(p) f(p|\pi) \). I treat \( f_\pi(2) \), the probability that an opportunity is worth two points, as a parameter to be estimated. I next impose that \( f(p|\pi = 2) \) and \( f(p|\pi = 3) \) belong to the family of beta density functions, each described by two parameters to be estimated \((M_2, V_2), (M_3, V_3)\).\(^8\) The final parameter governing the search process is \( \nu \), the constant per-second turnover hazard. The full parameter vector describing a single team type’s search process is given by \( \theta = [f_{\pi=2}, M_2, V_2, M_3, V_3, \nu] \).\(^9\)

\(^7\)A team on defense that is trailing near the end of the game may choose to intentionally foul its opponent in order to increase the number of possessions in the remainder of the game. In practice, the expected number of points scored by the offense when it is intentionally granted two free throws far exceeds the expected number of points scored on a typical possession. Two free throws result in an average of nearly 1.4 points, and an average possession results in less than one point. See Tables 3 and 4 for the relevant success rates across point-spread categories.

\(^8\)The family of beta density functions is a convenient choice because the functional form is flexible, parsimonious, and has support confined to the interval \([0, 1]\). Because the random variables being drawn from the densities \( f(p|\pi) \) are themselves probabilities (of particular shots succeeding), any chosen functional forms for \( f(p|2) \) and \( f(p|3) \) must have support confined to \([0, 1]\). A common parameterization of the beta density is given by \( f_X(x) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} \, dx} \) for \( x \in [0,1] \). I use the re-parameterization \( M = \frac{\alpha}{\alpha + \beta} \) and \( V = \alpha + \beta \). \( M \) is the mean of the random variable and \( V \) is inversely related to the variable’s dispersion.

\(^9\)For numerical stability in the estimation routines, I estimate continuous transformations of parameters that have unbounded support instead of directly estimating those parameters. The likelihood-maximizing parameter vector is invariant to these transformations, and the transformations ensure that an intermediate iteration of the hill-climbing algorithm does not step outside of a parameter’s support.
Maximum likelihood estimation requires an expression of the conditional probability of observed outcomes in terms of model parameters. In available data, I observe the duration of each possession and a record of which of the five mutually exclusive events caused the possession to end. Terminal events include turnovers and successful and unsuccessful two- or three-point shots. I observe the score differential $X$ and time $t$ when each possession begins, but I do not observe the sequence of state variables $p$ and $\pi$.

Given a reservation rule $R_0(s) \equiv R_0(s; \phi = 0)$ defined as in equation (10), the conditional probabilities of each event for each value of the shot clock can be calculated by taking appropriate integrals over values of $(p, \pi)$ with respect to the parameterized joint density function $f$. I let $F$ represent the distribution function corresponding to density function $f$ and $S$ represent the survivor function corresponding to density function $f$. The conditional probabilities of the various events $e$ are given by,

$$
P(e|s) = \begin{cases} 
    v & \text{for } e = \text{turnover} \\
    (1 - v) f_\pi(2) S\left(\frac{R(s)}{2} \big| \pi = 2\right) E_F\left[p | s, \pi = 2, 2p > R(s)\right] & \text{for } e = \text{successful 2} \\
    (1 - v) f_\pi(2) S\left(\frac{R(s)}{2} \big| \pi = 2\right) E_F\left[1 - p | s, \pi = 2, 2p > R(s)\right] & \text{for } e = \text{unsuccessful 2} \\
    (1 - v) f_\pi(3) S\left(\frac{R(s)}{3} \big| \pi = 3\right) E_F\left[p | s, \pi = 3, 3p > R(s)\right] & \text{for } e = \text{successful 3} \\
    (1 - v) f_\pi(3) S\left(\frac{R(s)}{3} \big| \pi = 3\right) E_F\left[1 - p | s, \pi = 3, 3p > R(s)\right] & \text{for } e = \text{unsuccessful 3} \\
    (1 - v) \left[ f_\pi(2) F\left(\frac{R(s)}{2} \big| \pi = 2\right) + f_\pi(3) F\left(\frac{R(s)}{3} \big| \pi = 3\right) \right] & \text{for } e = \text{continued search} 
\end{cases} \tag{11}
$$

I use these expressions for the conditional probabilities of discrete events to construct the likelihood function that is the basis for estimation.

I next construct a likelihood function. For each possession, I observe the time elapsed from the shot clock when the possession ended ($s^*$) and the event that caused the possession to end ($e^*$). Five of the six possible events listed in the piecewise definition of equation (11) are terminal events (turnovers and successful and unsuccessful two-point and three-point attempts), and, hence, are directly observed. All non-terminal events fall in the sixth category, continued search. The full sequence of events in any possession is a sequence of choices for continued search followed by a terminal event. By CIA, the probability of observing a possession described by the pair $(s^*, e^*)$ is given by:

$$
Pr(s^*, e^*) = Pr(e^*|s^*) \prod_{s=0}^{s^*-1} Pr(\text{continued search}|s) \tag{12}
$$

CIA further implies that the probability of observing a sample containing possessions $j = 1..N$ described
by the pairs \( \{(s_j^*, e_j^*)\}_{j=1}^N \) is given by:

\[
Pr\left( \left\{ (s_j^*, e_j^*) \right\}_{j=1}^N \right) = \prod_{j=1}^N Pr(s_j^*, e_j^*)
\] (13)

The right-hand side of equation (13) makes use of the expression defined in equation (12), and the right-hand side of equation (12) makes use of the expressions defined in equation (11). Finally I define the log-likelihood function,

\[
l(\theta) = \ln \left( Pr\left( \left\{ (s_j^*, e_j^*) \right\}_{j=1}^N \right| \theta) \right),
\] (14)

where \( \theta = [f_{\pi=2}, M_2, V_2, M_3, V_3, \nu]' \).

During estimation an inner loop computes the log-likelihood function at each candidate parameter vector using the numerical solution to the optimal reservation rule, and an outer loop searches the parameter space for the likelihood maximizing parameter vector. To reduce the chances that a set of parameter estimates represent a local maximum to the likelihood but not a global maximum, I repeat the estimation routine from several different starting points in the parameter space.

4 Data

I construct the dataset used for estimation from two sources. The first data source is a compilation of detailed play-by-play records for a subset of the regular-season basketball games played between November 2003 and March 2008 downloaded from the website statsheet.com. Appendix 2 describes the process of constructing possession-level data from the raw event data, and the procedure for coding possessions that did not strictly fit in to one of the outcomes included in the model. The second data source is a set of point spreads for regular-season games played during the same time period for which a point spread was available. These data come from the website covers.com. The final dataset contains 5,258 games that appear in both data sources.

Table 1 describes the distribution of games across seasons and across point spreads. More recent seasons are more heavily represented in the dataset, reflecting the increasing availability of detailed play-by-play game data. The lowest point-spread categories are most common, and a smaller fraction of games fall in each larger point-spread category.

Table 2 presents regression estimates of the home team’s winning margin (negative if the home team loses) on the amount by which the home team is favored on the point spread (negative if the home team was an underdog). I estimate an OLS regression to fit a conditional mean and a minimum absolute deviation regression to fit a conditional median. The point spread appears to provide an excellent forecast of the final-score differential. Consistent with efficiency in the point-spread betting market, neither estimated constant is statistically different from zero, and neither estimated slope coefficient is statistically different from one.

\[\text{My estimation routines converge to the same parameter vectors regardless of the initial guess.}\]
Table 3 provides descriptive statistics at the possession level. I report these values separately for favorites and underdogs in each point-spread category. Because the estimation routine restricts attention to possessions from the first halves of games, I provide one set of descriptive statistics for all possessions and another restricted to first-half possessions. To accommodate estimation, I code all possessions meeting my sample-selection criteria to one of the outcomes that is explicitly modeled. As expected, possessions of favored teams end more frequently with made shots and less frequently with missed shots and turnovers than the possessions of underdogs, and the disparity between the outcomes of favorites and underdogs tends to grow larger in the higher point-spread categories.

Figure 4 illustrates the skewness patterns that are the focus of the study. Panels 4-a and 4-b plot kernel density estimates of the favored team’s winning margin relative to the point spread. Panels 4-c and 4-d plot kernel density estimates of the difference between the favorite’s lead at halftime and the predicted value of that quantity. Consistent with the findings of Wolfers (2006), the distribution of favorites’ winning margins is approximately symmetric in games with a low point spread and is right skewed in games with a high point spread. Panels 4-c and 4-d demonstrate that any skewness in winning margins arises during the second half of games, as the distributions of halftime leads are nearly symmetric.

5 Structural Parameter Estimates

To recover estimates of the model parameters, I apply the estimation routine described in Section 3 to the data described in Section 4. Table 4 provides estimates of the model’s parameters by point-spread category. Table 5 assesses the fit of the model to the first half data. I compare empirical average possession duration to the average duration predicted by the model, and I compare the fractions of possessions ending in each of the five modeled terminal events to the fractions predicted by the model. The predicted moments closely resemble the empirical moments both within and across point-spread categories. Figure 5 allows for a visual inspection of the model’s fit to the observed dynamics of the offensive possessions during the first halves of games. To reduce clutter, the figures restrict attention to one low point-spread category [0,4] and one high point-spread category (16,20]. The figure presents the predicted reservation values and predicted points per attempted shot by possession duration, along with actual mean points per shot attempt over the course of the 35-second shot clock. Consistent with the model, all observed points per shot attempt fall above the predicted reservation value, and average points per attempt tend to fall as time passes and the reservation value falls.

To understand how the average drift in the relative score should vary across game states during the second half of games, figure 6 plots the average change in the favorite’s lead predicted to result from the play of one offensive possession for each team as a function of $\phi(\omega)$, the marginal rate of substitution between time and points for the favorite. In all six point-spread categories the average drift in the favorite’s lead decreases

---

11The favorite’s predicted halftime lead is estimated with a regression of that quantity on a constant and the point spread
12For a given $\phi$ I compute each team type’s optimal reservation policy according to Equation 10, taking expectations with respect to each type’s estimated search parameters, and I calculate the average points per possession associated with these policies. The average scoring drift at a given $\phi$ for a particular point spread is the difference between the average points per possession for the favorite and the underdog. I construct the curves in figure 6 by repeating this procedure for a range of $\phi$. 

13
with $\phi(\omega)$ and is convex in $\phi$ when $\phi$ is high. Given this drift pattern one would expect the distribution of the relative score to become right skewed in the higher point-spread categories, because the relative score commonly evolves through the high-$\phi$ portion of the state space when team $A$ is significantly stronger than team $B$. The next section uses simulations to examine the magnitude of this induced skewness.

6 Dynamic Simulation

Using the estimated model, I conduct a series of simulation experiments to assess the skewness patterns that result from teams’ optimally chosen strategies. I conduct the simulations separately for each point-spread category. I initialize each simulation by drawing a halftime score difference from a discretized normal distribution that matches the first two moments of the empirical distribution of favorite’s halftime leads. This approach guarantees that any skewness that is found in the simulated distributions is generated by the model’s predicted strategies. Table 6 reports the mean and standard deviation of the favorite’s halftime lead in each point-spread category. Then I use the model to simulate the distribution of the favorite’s winning margin.\(^{13}\) I then compute analogs of the two quantities required to construct the skewness based test for point shaving. For each point-spread category, I compute the fraction of the simulated favorite’s winning margins that fall between zero and the median of the winning-margin distribution. I also compute the fraction of the simulated favorite’s winning margins that fall between the median of the winning-margin distribution and twice the median.\(^{14}\)

I compute three separate versions of the simulations in order to isolate the impact of several model features. In the first set of simulations, I impose that each offense simply maximizes expected points per possession in all game states. This is not an optimal policy, but the exercise provides a reference to which more nearly optimal policies can be compared. In a second set of simulations, the offensive team optimally solves the model, but I do not allow the defense to foul intentionally. In a third set of simulations, both the offense and defense play optimal strategies.

Figure 7 plots the fraction of the favorite’s winning margins falling between zero and the median winning margin (lower region) and between the median winning margin and twice the median (higher region) for each of the three simulation scenarios and for the empirical distribution. Within each panel, I plot these two quantities against the midpoint of the point spread-category from which it was computed.

Panel 7-a reports the results of the first set of simulations in which the offensive team always maximizes its expected points per possession. The simulations finds that the distribution of winning margins is nearly symmetric, with about the same fraction of winning margins falling in the lower region as in the upper region. Panel 7-b reports the results of the second set of simulations in which the offense solves the full model and the defensive team never fouls. Under that scenario, the fraction of winning margins falling in the lower region exceeds the fraction of games falling in the higher region in all but the lowest point-spread

\(^{13}\)See Appendix 3 for details on computing the simulated distributions.

\(^{14}\)I use the median of the distribution instead of a particular point spread for two reasons. First, each point-spread category contains many individual point spreads, so the comparison to a single quantity is convenient. Second, the medians of the predicted distributions fall systematically below each category’s mean point spread. Presumably this is attributable to an unmodelled dimension by which the strengths of favorites and underdogs differ, rebounding for instance. Using the median preserves the interpretation of a difference in the two proportions as a departure from symmetry.
category. This result suggests that optimal offensive strategies introduce right skewness into the favorite’s winning margin, and that the degree of right skewness grows with the strength of the favorite.

Panel 7c reports the results of the third set of simulations in which the offense and the defense both adopt the optimal strategies predicted by the model. The simulations find that the fraction of simulated winning margins falling in the lower and higher is almost identical to the fractions of winning margins falling in those regions in actual games. That finding is the key result of the study. The empirical patterns are graphed in panel 7d. In games with a large favorite (the two highest point-spread categories), the fraction of games falling in the lower region exceeds the fraction falling in the higher region in simulations of the full model and in actual games. In games with a small favorite, similar fractions of games fall in the lower region and higher region in simulations of the full model and in actual games.

Applying the skewness test for point shaving to these simulated data, the conclusion is that the favorite intends to shave points in roughly 7 percent of games in the two highest point-spread categories. That predicted point-shaving prevalence is statistically indistinguishable from that computed from the empirical winning-margin distribution. Because panel 7c follows from innocent optimizing play, the simulation exercise suggests that the patterns previously attributed to point shaving are actually indistinguishable from the patterns expected under a null hypothesis of no point shaving.

7 Corroborating Evidence

Finally, I provide direct evidence that NCAA basketball teams employ the strategic stalling, hurrying, and intentional fouling strategies that account for the right skewed winning margin distributions in the model.

Table 7 shows that during the second halves of games, the length of offensive possessions is significantly positively associated with offensive team’s current lead, and the association becomes increasingly strong as the end of the game approaches. Specifically, I regress possession duration on the offensive team’s current lead and a set of fixed effects for combinations of point spread categories and indicators for the favorite/underdog being on offense (results in column 1). I then estimate a similar specification that interacts the favorite’s current lead with categories indicating how much time is left in the game (results in column 2). On average, a one point increase in the favorite’s lead increases the average length of a second half possession by about .05 seconds. During the first five minutes of the second half the effect is less than .02 seconds per point, and during the last five minutes of the second half the effect is nearly .09 seconds per point.

Table 8 reports the results of linear probability models comparing the frequency of likely intentional fouls (defined as fouls in the first 15 seconds of a possession) in game states where the estimated model finds intentional fouling to be optimal and game states where the estimated model does not. Figure 8 illustrates the states in which intentional fouling is an optimal strategy in the estimated model. Broadly speaking, it is optimal for a team on defense to foul when the team faces a small deficit is faced near the end of the game. For very small deficits, fouling is not optimal until very near the end of the game. For larger deficits, fouling is optimal with more time remaining in the game. Column 1 reports the results of regressing a dummy for a possession ending with a quick defensive foul on an indicator that the game state was in the model’s optimal
fouling region. The estimates find that quick fouls occur about four times more often in optimal fouling states than in non-optimal fouling states. Column 2 includes separate dummies for the possession occurring on the interior of the optimal fouling region, the edge of the optimal fouling region, and just outside the optimal fouling region. The results find that quick defensive fouls occur at similarly elevated rates on both the interior and edge of the optimal fouling region, and occur at only a slightly elevated rate just outside of the optimal fouling region. In sum, these results suggest that the strategic adjustments in the model that induce skewness to the distribution of favorites’ winning margins also occur in actual games.

8 Conclusion

This paper finds that the inference of widespread point shaving from skewness in the distribution of final score differentials is ill-founded. While the skewness-based test for point shaving relies on an assumption that winning margins are symmetric in the absence of point shaving, the model considered in this paper finds that teams adopt end-of-game strategies that do not in general lead to symmetric distributions. Calibrated with parameters estimated from first-half play-by-play data, the model of innocent dynamic competition predicts skewness patterns that are statistically indistinguishable from the empirical patterns. While we know from the historical record (Porter, 2002; Rosen 2001) that the true prevalence of point shaving is not zero, this finding suggests that the skewness-based test is likely to drastically overstate the true prevalence.

A possible extension to this study might formally model the behavior of a player or team engaged in point shaving. Wolfers (2006) proposes an exercise of that sort as a potential extension. Developing a credible model of a game in which one team is point shaving poses several obstacles that are not a problem for this study. A realistic model of point shaving requires a departure from the perfect-information framework. A more complex information structure would recognize that a team is probably never certain that its opponent is point shaving. The team that is point shaving must then consider the beliefs of its opponent regarding its objective, beliefs regarding those beliefs, and so on. An important second complication is the need for a point shaver to avoid detection. A simple but flawed model of point shaving might begin with the model used in this paper and replace the favorite’s objective function with one providing a reward to not covering the point spread. In that model, the optimal strategy for a large favorite who is shaving points is to play normally until near the end of the game and then, if necessary, deploy the strategy that most rapidly reduces its lead. Using that strategy, the favorite would win with the same frequency as when not point shaving and would almost never cover the point spread. But in practice, a casual spectator would recognize that the corrupt team was not simply trying to win. Because point shaving is illegal, a realistic model of point shaving must include some penalty for strategies that are easily detected.

The findings of this study cast some doubt on forensic economic studies that rely on unmodelled assumptions about innocent behavior. In the case considered in this study, a theoretical model is sufficient to raise the possibility that the skewness-based test for point shaving leads to biased estimates. Theory alone

\[15\] For instance, the favorite might receive a payoff of one if its winning margin fell between zero and the point spread and receive a payoff of zero otherwise.

\[16\] In the model considered in this paper, that strategy is to intentionally foul one’s opponent when on defense and to attempt the first available shot that provides a low success probability when on offense.
is insufficient to predict the direction or magnitude of any bias, and therefore a calibration exercise proves informative. These findings suggest that the indirect inference techniques that are common in forensic economic studies can be sensitive to seemingly minor institutional features of the environment in which the behavior of interest takes place, and highlight the promise of structural estimation as a tool for validating forensic economic methodology.

References


Appendix I

Proof of Proposition 2: Under the proposition’s premise, it follows immediately from the first order condition in equation \( \frac{\partial}{\partial X_1} \) that \((\sigma_{A2} + \sigma_{B2})\) does not vary with \(X_1\). The strategic adjustments of A and B exactly offset.

Consider how \(E(X_2|X_1)\) varies with \(X_1\). Totally differentiating the first order condition in Equation \[3\] and rearranging terms finds,

\[
\frac{\partial \sigma_{A2}}{\partial X_1} = \frac{1}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{A2})} \quad \text{and} \quad \frac{\partial \sigma_{B2}}{\partial X_1} = \frac{-1}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{B2})}
\]

Using these expressions, one can then express,

\[
\frac{\partial E(X_2|X_1)}{\partial X_1} = \frac{\partial \mu(\sigma_{A2})}{\partial \sigma_{A2}} \frac{\partial \sigma_{A2}}{\partial X_1} - \frac{\partial \mu(\sigma_{B2})}{\partial \sigma_{B2}} \frac{\partial \sigma_{B2}}{\partial X_1} = \frac{\mu'(\sigma_{A2})}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{A2})} + \frac{\mu'(\sigma_{B2})}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{B2})}
\]

\[
= \frac{\mu'(\sigma_{A2})}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{A2})} - \frac{1}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{B2})} \left( \frac{\mu'(\sigma_{A2})}{(\sigma_{A2} + \sigma_{B2})\mu''(\sigma_{A2})} \right)
\]

Under the symmetry assumption, the bracketed term is equal to zero, and the proposition follows.

Proof of Proposition 3: Again, make use of the expression,
The data used in this study come from two sources. The first data source is a compilation of detailed play-by-play records for 8,102 regular-season basketball games played between November, 2003 and March, 2008 downloaded from the website statsheet.com. The raw play-by-play data consists of a list of game events for each team (i.e. shot attempts, rebounds, and turnovers) and, for each event, the time elapsed from the shot clock when the event occurred. The second data source is a set of gambling point spreads for 24,868 regular-season games played during the same time period. This data is publicly available from the website covers.com. Merging these two data sources by the two team names and the game’s date shows that the intersection contains 5,258 games. Data from those 5,258 games are used to construct a data set for analysis.

From these raw data, I construct a data set containing one observation per possession. The set of information for each possession includes the team that was on offense, the time that elapsed from the shot clock during the possession, and the event that caused the possession to end. The key piece of information that must be computed is the time remaining in the game when the event occurred. For each possession, I compute the time elapsed from the shot clock at the conclusion of each possession, and the possession’s duration. The possession’s duration is the time elapsed from the start of the possession to the end of the possession. For possessions including multiple events, maintain the record of the event sequence that occurred during the possession.

1. Classify each play-by-play observation in one of the categories; made two point shot, missed two point shot, made three point shot, missed three point shot, turnover, rebound, foul, free throw attempt, assist, blocked shot, substitution, or timeout using keyword searches of the record’s raw text string.

2. Drop events that are not relevant to the model’s estimation and do not impact the shot clock. These include assists, blocked shots, substitutions, and timeouts. The events that are not dropped retain all information relevant to estimation.

3. Flag events that reset the shot clock. These include made shots, fouls, turnovers, and rebounds.

4. For each event, compute the time elapsed from the shot clock when the event occurred as the difference between the time at which the event occurred and the game time at which the shot clock was most recently reset.

5. Reshape the dataset from one containing one observation per event to one containing one observation per possession. For possessions including multiple events, maintain the record of the event sequence that occurred during the possession, and the possession’s duration. The possession’s duration is the time elapsed from the shot clock at the time of the possession’s final event.

6. Code each possession to one of the model’s terminal events. For possessions ending with a turnover, a made or missed two point shot, or a made or missed three point shot, this coding is straightforward. In other cases, use the following rules. Possessions in which a defensive foul was committed during a two point shot attempt are coded as successful two-point shots. Possessions in which a defensive foul was committed during a three-point shot attempt are coded as successful three-point shots. Possessions that end with an offensive foul are coded as turnovers. Shot clock violations are coded as turnovers occurring in the final second of the shot clock.

In particular, some games not selected for the sample do not contain a record of turnovers.

For possessions following a made shot by the opponent, I set the time at which the shot clock was reset to four seconds after the game time at which the shot was made. Following a made shot, the next shot clock does not begin counting until the team to be on offense carries the ball out of bounds and passes the ball in. This procedure typically takes several seconds.
7. Drop possessions for which this procedure computes a possession duration exceeding the maximum possible duration (by the shot clock rule) of 35 seconds.

8. Exclude possessions from the analysis if they have duration of five seconds or less.

9. Exclude possessions from the analysis if they end with a defensive foul for which no free throws are awarded.

I create a second dataset containing one observation per observed free throw. Each observation in this dataset describes the point spread of the game in which the free throw occurred, the team that attempted the free throw, the game time at which the free throw occurred, and the result of the free throw.

Appendix 3: Simulations

I simulate the distribution of winning margins in three steps. First, I numerically solve the full model calibrated by the estimated structural parameters to obtain teams optimal policy functions. Second, I used the optimal policy function to compute a transition matrix describing the probability of achieving each possible winning margin conditional on each possible half-time score differential. Third, I apply these transition probabilities to a smoothened version of the empirical halftime distribution. The smoothed distribution is a discretized normal distribution with the same first and second moments as the empirical distribution. This procedure ensures that any observed asymmetry in the simulated winning margin distribution comes from simulated second-half play.

I numerically compute the solution to the full value function $V: \Omega \rightarrow [0, 1]$, mapping each element of the full state space to a probability of victory using backward induction according to equation (5), and I obtain the corresponding optimal policy function $P: \Omega \rightarrow A^d \times A^o$. Backward induction is feasible if the payoff functions $U^i$ provide a fixed boundary condition at the game’s conclusion: time $T$. However, in NCAA basketball games, a five-minute overtime period is played whenever the game is tied at the end of play. To facilitate backward induction, I impose that each team receives a payoff of 0.5 when the game ends with a tie when I compute the numerical solution to the full model using backward induction for the entire second half, which contains 1200 seconds.. Then, in the simulations I allow five additional minutes (using teams’ policies for the final five minutes) to be played when the score is tied at the end of the game and repeat until the game ends with a winner and a loser.

I handle several details of the simulation process as follows. I discretize the shot opportunity distributions, $F^o$, using ten points of support, so that I can use discrete dynamic programming methods. In states in which one team has a very large lead and therefore the winning team is not in doubt, $V = 0$ or $1$. I imposed that the offensive team reverts to maximizing expected points per possession and the defensive team chooses not to foul. The model does not make a unique prediction in these states, because many policies yield the same expected payout when the winner is not in doubt.

---

19This procedure coded 1.21 percent of observations to durations exceeding 35 seconds.
20The winning team is no longer in doubt in the context of the model when insufficient time remains for the trailing team to come from behind even given the most fortuitous possible string of events.
21One could imagine other possible policies adopted by teams when the game’s winner is no longer in doubt. These policies are consistent with rational play if, for very small $\epsilon > 0$, each team placed a weight of $(1 - \epsilon)$ on the discrete outcome of winning and a weight of $\epsilon$ on the margin of victory/defeat.
Table 1: Descriptive Statistics - Games

<table>
<thead>
<tr>
<th>Category</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-2004 Season</td>
<td>12</td>
</tr>
<tr>
<td>2004-2005 Season</td>
<td>17</td>
</tr>
<tr>
<td>2005-2006 Season</td>
<td>21</td>
</tr>
<tr>
<td>2006-2007 Season</td>
<td>24</td>
</tr>
<tr>
<td>2007-2008 Season</td>
<td>26</td>
</tr>
<tr>
<td>Point Spread ∈ [0, 4]</td>
<td>29</td>
</tr>
<tr>
<td>Point Spread ∈ (4, 8]</td>
<td>27</td>
</tr>
<tr>
<td>Point Spread ∈ (8, 12]</td>
<td>18</td>
</tr>
<tr>
<td>Point Spread ∈ (12, 16]</td>
<td>13</td>
</tr>
<tr>
<td>Point Spread ∈ (16, 20]</td>
<td>7</td>
</tr>
<tr>
<td>Point Spread ∈ (20, ∞]</td>
<td>6</td>
</tr>
<tr>
<td>Observations</td>
<td>5,258</td>
</tr>
</tbody>
</table>

Note: The sample comprises play-by-play game data from statsheet.com merged to point-spread data from covers.com. A game is included in the sample if it appears in both data sources and the game’s play-by-play data contains enough detail to perform the analysis conducted in this study (see Data Appendix 1 for details). Source: author’s calculations.

Table 2: Regression Analysis of the Point Spread’s Predictive Accuracy

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>MAD</td>
</tr>
<tr>
<td>Amount by which home team is favored</td>
<td>1.01</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Observations</td>
<td>5,258</td>
<td>5,258</td>
</tr>
</tbody>
</table>

Standard errors in parentheses.
Note: This table displays regression estimates of the conditional mean and median of the home team’s winning margin (negative if the home team loses). Column (1) reports coefficient estimates for an OLS regression of the home team’s winning margin on a constant and the amount by which the home team is favored (negative if the home team is the underdog). Column (2) reports coefficient estimates for minimum absolute deviation (median) regression of the home team’s winning margin on the amount by which the home team is favored. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Table 3: Descriptive Statistics - Possessions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Possessions</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
</tr>
<tr>
<td>Duration of Possession (sec)</td>
<td>16.11</td>
<td>16.32</td>
<td>16.26</td>
<td>16.27</td>
<td>16.17</td>
<td>15.78</td>
<td>16.18</td>
</tr>
<tr>
<td>Points</td>
<td>0.84</td>
<td>0.85</td>
<td>0.82</td>
<td>0.88</td>
<td>0.81</td>
<td>0.90</td>
<td>0.78</td>
</tr>
<tr>
<td>Made two point shot</td>
<td>0.30</td>
<td>0.31</td>
<td>0.30</td>
<td>0.31</td>
<td>0.29</td>
<td>0.33</td>
<td>0.28</td>
</tr>
<tr>
<td>Missed two point shot</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.24</td>
<td>0.26</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>Made three point shot</td>
<td>0.10</td>
<td>0.10</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>Missed three point shot</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
<td>0.17</td>
<td>0.19</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>Observations</td>
<td>641,447</td>
<td>93,547</td>
<td>93,760</td>
<td>87,139</td>
<td>88,005</td>
<td>39,879</td>
<td>41,005</td>
</tr>
</tbody>
</table>

|                   | All Possessions        | F            | U            | F            | U            | F            | U       |
| Duration of Possession (sec) | 16.11                  | 16.18        | 16.29        | 15.96        | 16.37        | 15.71        | 16.59   |
| Points            | 0.84                   | 0.82         | 0.79         | 0.83         | 0.77         | 0.87         | 0.75    |
| Made two point shot | 0.27                   | 0.28         | 0.27         | 0.28         | 0.26         | 0.30         | 0.25    |
| Missed two point shot | 0.26                   | 0.25         | 0.26         | 0.25         | 0.27         | 0.24         | 0.27    |
| Made three point shot | 0.10                   | 0.10         | 0.10         | 0.10         | 0.09         | 0.11         | 0.09    |
| Missed three point shot | 0.17                   | 0.17         | 0.17         | 0.18         | 0.17         | 0.18         | 0.17    |
| Turnover          | 0.20                   | 0.19         | 0.20         | 0.19         | 0.20         | 0.18         | 0.22    |
| Observations      | 320,064                | 46,531       | 46,697       | 43,845       | 43,889       | 28,719       | 28,928  |

Standard errors in parentheses.
Note: The sample excludes possessions with a duration of five seconds or less and possessions that end with a defensive foul that does not result in free throws. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
<td>F</td>
<td>U</td>
</tr>
<tr>
<td>Prob. shot opportunity is worth two points</td>
<td>0.533</td>
<td>0.536</td>
<td>0.603</td>
<td>0.563</td>
<td>0.568</td>
<td>0.520</td>
<td>0.607</td>
<td>0.551</td>
<td>0.797</td>
<td>0.515</td>
<td>0.823</td>
<td>0.489</td>
</tr>
<tr>
<td>Mean of success prob. distribution - two pointers</td>
<td>0.139</td>
<td>0.133</td>
<td>0.134</td>
<td>0.127</td>
<td>0.157</td>
<td>0.118</td>
<td>0.163</td>
<td>0.112</td>
<td>0.144</td>
<td>0.102</td>
<td>0.151</td>
<td>0.121</td>
</tr>
<tr>
<td>Mean of success prob. distribution - three pointers</td>
<td>0.068</td>
<td>0.058</td>
<td>0.089</td>
<td>0.056</td>
<td>0.095</td>
<td>0.047</td>
<td>0.109</td>
<td>0.043</td>
<td>0.177</td>
<td>0.038</td>
<td>0.223</td>
<td>0.030</td>
</tr>
<tr>
<td>Turnover hazard</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.015</td>
<td>0.017</td>
<td>0.015</td>
<td>0.018</td>
<td>0.016</td>
<td>0.018</td>
<td>0.017</td>
<td>0.020</td>
</tr>
<tr>
<td>Observations</td>
<td>46,531</td>
<td>46,697</td>
<td>43,845</td>
<td>43,889</td>
<td>28,719</td>
<td>28,928</td>
<td>19,945</td>
<td>20,277</td>
<td>10,644</td>
<td>10,814</td>
<td>9,736</td>
<td>10,039</td>
</tr>
<tr>
<td>Free throw success probability</td>
<td>0.691</td>
<td>0.68</td>
<td>0.694</td>
<td>0.686</td>
<td>0.697</td>
<td>0.68</td>
<td>0.692</td>
<td>0.667</td>
<td>0.695</td>
<td>0.678</td>
<td>0.702</td>
<td>0.652</td>
</tr>
<tr>
<td>Observations</td>
<td>10,294</td>
<td>9,913</td>
<td>11,187</td>
<td>9,709</td>
<td>8,084</td>
<td>6,516</td>
<td>5,813</td>
<td>4,324</td>
<td>3,426</td>
<td>2,352</td>
<td>3,475</td>
<td>2,102</td>
</tr>
</tbody>
</table>

Standard errors in parentheses.
Note: Estimation is by maximum likelihood using data from the first halves of games. Free throw success probabilities are estimated by the sample mean success rate for free throws. A nested fixed point algorithm is used to compute all remaining parameters. The parameters are estimated separately for the favorite and underdog within each of the six point-spread categories. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Table 5: Comparison of Empirical and Predicted Moments

<table>
<thead>
<tr>
<th></th>
<th>$PS \in [0,4]$</th>
<th></th>
<th>$PS \in (4,8]$</th>
<th></th>
<th>$PS \in (8,12]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$U$</td>
<td>$F$</td>
<td>$U$</td>
<td>$F$</td>
</tr>
<tr>
<td>Duration</td>
<td>$m$</td>
<td>$\hat{m}$</td>
<td>$m$</td>
<td>$\hat{m}$</td>
<td>$m$</td>
</tr>
<tr>
<td></td>
<td>16.18</td>
<td>15.85</td>
<td>16.29</td>
<td>15.94</td>
<td>15.96</td>
</tr>
<tr>
<td>Made 2</td>
<td>0.278</td>
<td>0.279</td>
<td>0.270</td>
<td>0.268</td>
<td>0.280</td>
</tr>
<tr>
<td>Missed 2</td>
<td>0.254</td>
<td>0.265</td>
<td>0.256</td>
<td>0.271</td>
<td>0.249</td>
</tr>
<tr>
<td>Made 3</td>
<td>0.0993</td>
<td>0.0969</td>
<td>0.0962</td>
<td>0.0974</td>
<td>0.102</td>
</tr>
<tr>
<td>Missed 3</td>
<td>0.175</td>
<td>0.188</td>
<td>0.175</td>
<td>0.185</td>
<td>0.182</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.194</td>
<td>0.171</td>
<td>0.203</td>
<td>0.179</td>
<td>0.187</td>
</tr>
<tr>
<td>Observations</td>
<td>46,531</td>
<td>46,697</td>
<td>43,845</td>
<td>43,889</td>
<td>28,719</td>
</tr>
<tr>
<td>$t$ (equality of mean duration)</td>
<td>10.00</td>
<td>10.29</td>
<td>7.94</td>
<td>10.57</td>
<td>5.48</td>
</tr>
<tr>
<td>p-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pearson’s $\chi^2$ (equality of proportions)</td>
<td>209.95</td>
<td>215.66</td>
<td>154.82</td>
<td>218.56</td>
<td>93.21</td>
</tr>
<tr>
<td>p-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$PS \in [12,16]$</th>
<th></th>
<th>$PS \in (16,20]$</th>
<th></th>
<th>$PS &gt; 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$U$</td>
<td>$F$</td>
<td>$U$</td>
<td>$F$</td>
</tr>
<tr>
<td>Duration</td>
<td>$m$</td>
<td>$\hat{m}$</td>
<td>$m$</td>
<td>$\hat{m}$</td>
<td>$m$</td>
</tr>
<tr>
<td></td>
<td>15.46</td>
<td>15.28</td>
<td>16.61</td>
<td>16.17</td>
<td>14.97</td>
</tr>
<tr>
<td>Made 2</td>
<td>0.303</td>
<td>0.303</td>
<td>0.248</td>
<td>0.245</td>
<td>0.305</td>
</tr>
<tr>
<td>Missed 2</td>
<td>0.236</td>
<td>0.244</td>
<td>0.274</td>
<td>0.292</td>
<td>0.239</td>
</tr>
<tr>
<td>Made 3</td>
<td>0.110</td>
<td>0.109</td>
<td>0.0867</td>
<td>0.0886</td>
<td>0.111</td>
</tr>
<tr>
<td>Missed 3</td>
<td>0.178</td>
<td>0.186</td>
<td>0.167</td>
<td>0.179</td>
<td>0.173</td>
</tr>
<tr>
<td>Turnover</td>
<td>0.173</td>
<td>0.158</td>
<td>0.225</td>
<td>0.196</td>
<td>0.172</td>
</tr>
<tr>
<td>Observations</td>
<td>19,945</td>
<td>20,277</td>
<td>10,644</td>
<td>10,814</td>
<td>9,736</td>
</tr>
<tr>
<td>$t$ (equality of mean duration)</td>
<td>3.60</td>
<td>8.63</td>
<td>1.52</td>
<td>7.18</td>
<td>1.52</td>
</tr>
<tr>
<td>p-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pearson’s $\chi^2$ (equality of proportions)</td>
<td>40.68</td>
<td>127.39</td>
<td>22.88</td>
<td>75.80</td>
<td>12.67</td>
</tr>
<tr>
<td>p-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Note: Within each point-spread category, sample moments ($m$) and predicted moments ($\hat{m}$) are provided separately for favorites and underdogs. Empirical moments are sample means. Predicted moments are the equivalent means predicted by the model when each team searches optimally given estimated structural parameters with the objective of maximizing expected points per possession. Two test statistics and corresponding p-values are provided for each combination of point-spread category and favorite/underdog. A two-sided $t$-test is performed to test the null that the empirical mean possession duration is equal to the predicted mean duration. A Pearson’s chi-squared test is performed to test the null that the five empirical proportions are equal to the predicted proportions. Under the null, the Pearson’s statistic is distributed chi-squared with $J - 1 = 4$ degrees of freedom (where $J$ is the number of categories). Source: authors calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Table 6: Mean and Standard Deviation of Favorite’s Halftime Lead by Point-Spread Category

<table>
<thead>
<tr>
<th>Point Spread</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[0, 4]$</td>
<td>1.16</td>
<td>8.59</td>
</tr>
<tr>
<td>$(4, 12]$</td>
<td>3.04</td>
<td>8.39</td>
</tr>
<tr>
<td>$(8, 12]$</td>
<td>5.57</td>
<td>8.83</td>
</tr>
<tr>
<td>$(12, 16]$</td>
<td>8.19</td>
<td>8.56</td>
</tr>
<tr>
<td>$(16, 20]$</td>
<td>10.17</td>
<td>8.82</td>
</tr>
<tr>
<td>$&gt; 20$</td>
<td>13.82</td>
<td>9.40</td>
</tr>
</tbody>
</table>

Note: The simulations described in this paper assume that half-time score differentials follow a (discretized) normal distribution with the empirical means and standard deviations contained in this table. The simulated final score distributions, then, reveal the extent to which optimal second-half play induces asymmetries. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.

Table 7: Second Half Possession Durations by Time Remaining and Score Differential

<table>
<thead>
<tr>
<th>(1) Duration</th>
<th>(2) Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offensive team’s lead</td>
<td>0.047*** (0.0018)</td>
</tr>
<tr>
<td>Offensive team’s lead × 1(15-20 minutes left in 2nd half)</td>
<td>0.017*** (0.004)</td>
</tr>
<tr>
<td>Offensive team’s lead × 1(10-15 minutes left in 2nd half)</td>
<td>0.029*** (0.004)</td>
</tr>
<tr>
<td>Offensive team’s lead × 1(5-10 minutes left in 2nd half)</td>
<td>0.051*** (0.004)</td>
</tr>
<tr>
<td>Offensive team’s lead × 1(&lt;5 minutes left in 2nd half)</td>
<td>0.085*** (0.004)</td>
</tr>
<tr>
<td>15-20 minutes left in 2nd half</td>
<td>–</td>
</tr>
<tr>
<td>10-15 minutes left in 2nd half</td>
<td>1.078* (0.473)</td>
</tr>
<tr>
<td>5-10 minutes left in 2nd half</td>
<td>0.726 (0.512)</td>
</tr>
<tr>
<td>&lt;5 minutes left in 2nd half</td>
<td>0.479 (0.491)</td>
</tr>
<tr>
<td>Constant</td>
<td>13.65*** (0.014)</td>
</tr>
</tbody>
</table>

Observations: 337,243

FEs for point spread × favorite/dog on offense: yes

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Note: The dependent variable in each regression is the possession duration in seconds. The sample includes all second half possessions that begin with the score differential in $[-20, 20]$. Column (1) computes an average hurrying/stalling response across all second half possessions, and column (2) computes effects separately for each 5-minute period within the second half. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Table 8: Regression Analysis of Defensive Fouling Policy

<table>
<thead>
<tr>
<th></th>
<th>(1) foul</th>
<th>(2) foul</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted foul region</td>
<td>0.312**</td>
<td>0.363***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Predicted foul region - interior</td>
<td>0.206**</td>
<td>0.059***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Close to predicted foul region</td>
<td>0.059***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.110***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>820,672</td>
<td>820,672</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *p < 0.05, **p < 0.01, ***p < 0.001.

Note: All possessions played during the first and second halves are included. The dependent variable in each regression is an indicator that a defensive foul occurred during the possession. A possession is coded as “near the edge” of the predicted foul region if the favorite’s lead is within two points of a lead at which fouling is not optimal. A possession is coded as on the “interior” of the predicted foul region if the possession falls inside the predicted foul region and is not coded as “near the edge.” A possession is coded as “close” to the predicted foul region if the possession falls outside of the predicted foul region but the favorite’s lead is within two points of a lead at which fouling is optimal. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.

Figure 1: The Optimal Choice of a Mean-Variance Pair During Stage 2 of the Stylized Model

Note: Taking B’s strategy as given, A faces a choice among combinations of the mean and the variance of the final score. Two example choice sets are provided, one corresponding to A trailing after stage 1 and the other corresponding to A leading after stage 1. Iso-expected payoff curves are represented by dashed lines. Because $X_2$, the score during stage 2, is normally distributed, A’s probability of winning depends on the ratio of the expected value of the final score to the standard deviation of $X_2$. As such, the iso-expected payoff curves are rays away from the origin.
Figure 2: State Dependent Strategies Can Skew Winning Margins in Either Direction

(a) Drift when low-variance strategy (stalling) is less costly
(b) Drift when high-variance strategy (hurrying) is less costly

(c) Drift being convex in score → right skewness
(d) Drift being concave in score → left skewness

Note: Panels (a) and (b) illustrate the average change in the relative score in the second half as a function of the halftime score in the illustrative two stage model. Panel (a) gives an example second-half scoring drift pattern where the drift declines with the favorite’s halftime lead, a pattern that occurs in the illustrative two stage model when adjusting to a low variance strategy incurs a lower marginal cost than adjusting to a high variance strategy. Panel (b) gives an example second-half scoring drift pattern where the drift increases with the favorite’s halftime lead, a pattern that occurs when adjusting to a low variance strategy incurs a higher marginal cost than adjusting to a high variance strategy. Panels (c) and (d) illustrates the effect of second half play on the shape of scoring distribution in these two cases. Figure (c) shows that when second half scoring drift is a convex function of the halftime lead over the support of the halftime lead distribution, as in the right portion of figure (a), second half play introduces right skewness to the winning margin distribution. Figure (d) shows that when second half scoring drift is a concave function of the halftime lead, as in the right portion of figure (b), second half play introduces left skewness to the winning margin distribution.
Figure 3: Predicted Reservation Policies Across Game States

Note: Shot opportunities are characterized by a success probability and a point value. The graphed reservation values depict for a single team type (teams favored by 0 to 4 points) the expected success probability times point value above which it is optimal to accept the opportunity. The optimal policy depends on the game state. The top, middle, and bottom lines correspond to game states in which the marginal rate of substitution between time and points ($\phi$) for the favorite is positive (0.15), zero, and negative (−0.15). When the partial effect of the passing time on a team’s expected payoff is higher, the team adopts a higher reservation policy that results in longer average possessions. Source: author’s calculations.

Figure 4: Game Outcomes Relative to Point Spread Prediction

Note: Plots (a) and (b) provide kernel-density estimates of the difference between the favorite’s winning margin and the point spread. Plots (c) and (d) provide kernel-density estimates of the difference between the favorite’s halftime lead and the predicted value of that quantity from a linear regression of the halftime lead on a constant and the point spread. As a reference, each plot is overlaid with a normal density function with the same first two moments as the estimated density. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Figure 5: Predicted Reservation Values and Average Points by Time Elapsed from Shot Clock

(a) Favorite (0 ≤ Point Spread < 4)
(b) Underdog (0 ≤ Point Spread < 4)
(c) Favorite (16.5 ≤ Point Spread < 20)
(d) Underdog (16.5 ≤ Point Spread < 20)

Note: Shot opportunities are characterized by a success probability and a point value. A reservation policy is an expected point value (success probability times point value) above which an optimizing offense is predicted to attempt an available shot. The solid line on each figure depicts the optimal reservation policy by possession duration consistent with the estimated structural parameters. The dashed line on each plot depicts the predicted average point value of attempted shots (shots with expected point values exceeding the reservation level). The scatter plot depicts the empirical average points per attempted shot against possession duration for first-half possessions included in the estimation sample. The points marked with x’s depict two-point attempts and the points marked with o’s depict three-point attempts. Source: author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Figure 6: Predicted Scoring Drift Across Game States

Note: The plotted curves depict the predicted average drift in the favorite’s lead over a pair of possessions (one for the favorite one for the underdog) across states. Game states are characterized by $\phi$, the marginal rate of substitution between time and the favorite’s points. Source: author’s calculations using estimated model parameters.
Figure 7: False Experiments - Skewness Based Test for Point Shaving Applied to Simulated Outcomes

(a) Simulations Assuming Offense Maximizes Points in all Possessions
(b) Simulations Assuming Offense Plays Optimal End-of-Game Strategy
(c) Simulations Assuming Offense and Defense Play Optimal End-of-Game Strategy
(d) Empirical Pattern

Note: The solid line on each plot provides the predicted probability that the favorite’s winning margin falls between zero and the median of the winning margin distribution. The dashed line on each plot provides the predicted probability that the favorite’s winning margin falls between the median of the winning margin distribution and twice the median of the winning margin distribution. Source: (a-c) author’s simulations calibrated with estimated model parameters, and (d) author’s calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.
Figure 8: Predicted Optimal Defensive Fouling Policies

(a) Point Spreads 0 to 4
(b) Point Spreads 4.5 to 8
(c) Point Spreads 8.5 to 12
(d) Point Spreads 12.5 to 16
(e) Point Spreads 16.5 to 20
(f) Point Spreads 20.5 and Above

Note: The region outlined by a solid line on each plot depicts the portion of the state space in which the favored team is predicted to foul intentionally when on defense. The region outlined by a dashed line on each plot depicts the portion of the state space in which the favored team is predicted to foul intentionally when on defense. Source: author’s calculation of the numerical solution to the model calibrated with estimated model parameters.