Do Basketball Scoring Patterns Reflect Illegal Point Shaving or Optimal In-Game Adjustments?*

Jesse Gregory

University of Wisconsin, NBER

jmgregory@ssc.wisc.edu

Current Version: August, 2017

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.[1] 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, occuring in about six percent of games where one team is strongly favored. Wolfers' study garnered significant attention in the popular media[2], 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

[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."

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 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.[3]

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 suggests 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[4]. 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.

[3]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.

[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.

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,

$$X_1 = \mu(\sigma_{A1}) - \mu(\sigma_{B1}) + \Big(\sigma_{A1} + \sigma_{B1}\Big)Z_1 + \Delta \tag{1}$$

$$X_2 = \mu(\sigma_{A2}) - \mu(\sigma_{B2}) + \Big(\sigma_{A2} + \sigma_{B2}\Big)Z_2 + \Delta \tag{2}$$

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 that $\mu()$ is bounded, strictly concave, and reaches an interior maximum on $[0, 1]$. The choice of a very high variance (σ near one) or a very low variance (σ 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 σ approaches 1 and going to ∞ as σ 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. The vertical axis plots the expected final score entering stage 2. The horizontal axis plots the standard deviation of the stage 2 score differential. A's win probability depends on the ratio of the expected final score differential to the standard deviation of stage 2 scoring, so the Iso-payoff curves are rays away from the origin.[5] Proportional increases in the expected final score differential and in the standard deviation of stage 2 scoring do not affect win probabilities. If $X_1 + \Delta = 0$, A and B (depicted by the iso-payoff ray extending horizontally from the origin) each have an 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's optimal strategy is to select 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. These optimal choices are characterized by the tangency of the mean-variance choice set $\mu()$ to the state-dependent iso-payoff curves, with this result occurring because A's iso-payoff curve is downward sloping when its win probability is below $1/2$ and is upward sloping when its win probability is above $1/2$.

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.

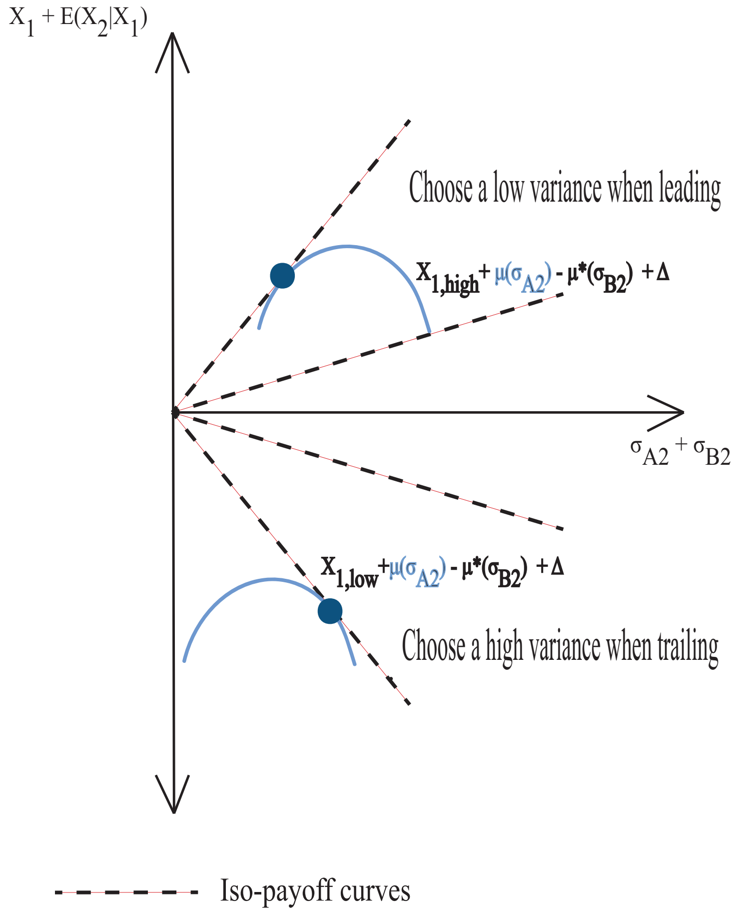
Now consider how the players' optimal strategies influence the shape of the distribution of winning margins. One might expect for the winning margin X to be symmetrically, because $X = X_1 + X_2$ is the sum of two normally distributed random variables. That is not the case in general, because the players' 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.*[6]

[5]Spcifically, A's win probability is $\Phi\Big(\dfrac{X_1 + \mu(\sigma^*_{A2}) - \mu(\sigma^*_{B2}) + \Delta}{\sigma^*_{A2} + \sigma^*_{B2}}\Big)$, where $\Phi()$ is the standard normal CDF.

[6]This result follows directly from the symmetry of the game. Because the solution is unique, switching the names

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.

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.

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 $\mu(.5+c) = \mu(.5-c)$ for any $c \in [0, .5]$. Then, $E(X_1) = E(X_2|X1) = \Delta$, $\sigma_{A1} + \sigma_{B1} = \sigma_{A2} + \sigma_{B2} = 1$, and $X \sim N(2\Delta, 2)$.* Proof provided in Appendix 1.

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 σ^* 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 1.

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 1 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 Δ 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 Δ to be right skewed. Similarly, if $E(X_2|X_1)$

of A and B cannot change the distribution of winning margins.

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 Δ 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.

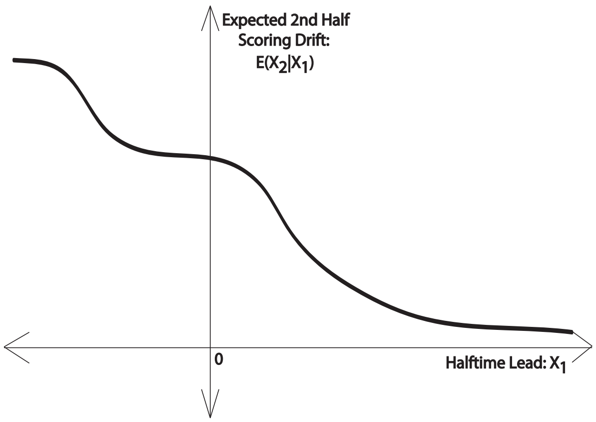
2 Full Model

2.1 Model

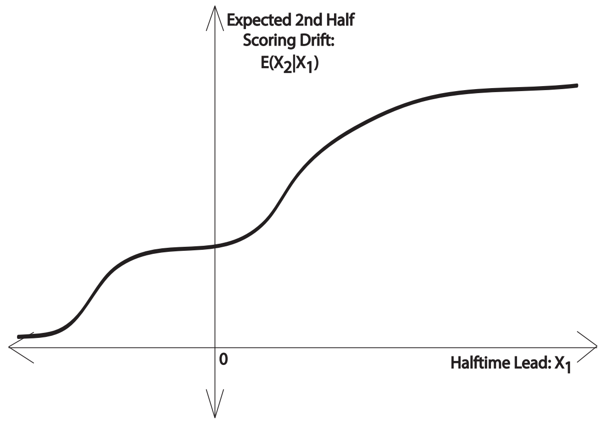
In this section, I propose a more detailed model of the environment in which basketball teams compete. Teams alternate turns as the offensive side, and the team that is on offense faces a sequence of arriving shot opportunities that vary stochastically in quality. I refer to a single turn as the offensive 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

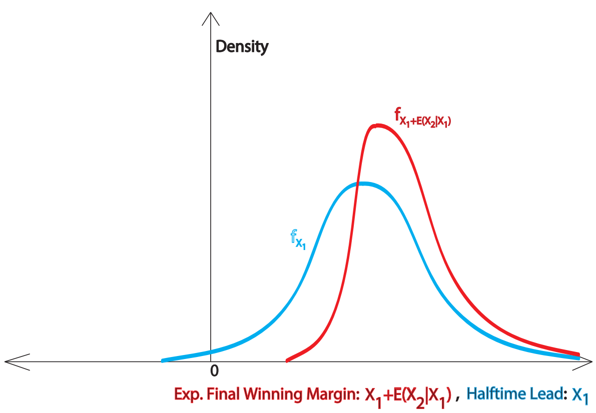
Figure 2: State Dependent Strategies Can Skew Winning Margins in Either Direction



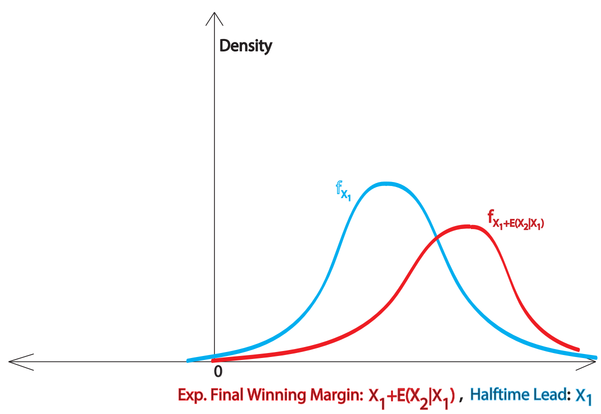
(a) Drift when low-variance strategy (stalling) is less costly



(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.

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 π. 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 π 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[7]. 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 π 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, π) are independent across time. Finally, I assume that turnovers arrive with a constant probability for each team, and I let υ_A and υ_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).[8] 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, π). 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_A^{ft} and p_B^{ft}. If no turnover occurs and the defense chooses not to intentionally foul, a shot opportunity (p, π) is drawn. The offense must choose whether to accept the current opportunity or to continue searching. I denote the choice spaces of the defense 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.[9][10] Because

[7]The 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.

[8]This assumption implies that teams' turnover hazards and arriving shot quality distributions are constant within possessions (over the course of the shot clock) and across possessions (by time-by-score game states). These assumptions may be violated to some extent within actual games if teams adopt tactics that trade off turnover hazard against shot quality (a risk/reward trade-off) or choose to remove star players from the game when the score differential becomes large. I discuss the potential consequences of these sorts of model failures below, arguing that if anything they are likely to exacerbate any right-skewness in the distribution of favored teams' winning margins.

[9]One deviation from this payoff function would be if teams place some utility weight on the final score differential in addition to a discrete preference for winning over losing. This would be rational if, for instance, selection into postseason tournaments is awarded based partially on winning margins in addition to teams' simple win/loss records. If teams do have payoff functions that place weight on score differential, optimal end of game strategies would tend to be closer to those that maximize points per possession than those predicted by this model, and would thus generate less skewed winning margins distributions than those predicted by this model. This is apparent by considering the extreme case where teams only place weight on score differential (and place no additional weight on winning), in which case winning margins will be approximately symmetric as teams would simply maximize expected points per possession throughout the game.

[10]Another deviation from this payoff function would be if teams place disproportionate weight on some of their

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. Denote the vector of state variables with $\omega = \{t, s, X, o, p, \pi\}$, and let $\Omega = \{\omega\}$ denote the state space. Define the value function $V^A(\omega) = E(U^A|\omega)$ to be A's expected payoff from the state ω. 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.

$$V^A(\omega) = \max_{a \in A^A(\omega)} \left\{ \min_{b \in A^B(\omega)} \left\{ E[V^A(\omega')|\omega, a, b] \right\} \right\} \tag{5}$$

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 ω 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. There are several advantages to this approach. Most importantly, the approach ensures that the model's parameters are identified by teams' choices *within* possessions during the portion of games where pace adjustments are not important, as opposed to being identified by teams' end-of-game adjustments, which may or may not be contaminated by point shaving incentives. Also, the approach provides significant computational savings because it avoids having to repeatedly solve the full model with backward induction during estimation.

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

---

games during the season (e.g. conference games or games against rivals) than others. This could be modeled by providing teams with 0 vs. 1 payoffs in some games and, say, 0 vs.$\overline{V}$ (with $\overline{V} > 1$) payoffs in other games. Holding teams' structural search parameters fixed, these payoff changes would have no effect on teams' strategies, since the optimal strategy is still to maximize win percentage. Similarly, even if teams allocate their game-to-game effort over the course of the season disproportionately to these "high-payoff" games, as long as these incentives are reflected in the point spread (as Table 4 suggests is the case) the estimated model should capture these effects because the structural search parameters are allowed to depend on the point spread.

$$EV([t,s,X,o]) = E_{(p',\pi')}\Big[V\Big([t,s,X,o,p',\pi']\Big)\Big] \tag{6}$$

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 $s^\star$ denote the duration of the possession, and let $x^\star$ denote the change in X that occurs as a result of the possession. Then the state of the following the possession is $\omega^\star = [t + s^\star, s = 0, X + x^\star, -o]$, where $-o$ indicates the opponent of team o. I next note that the expected payoff in the state $\omega^\star$ can be approximated linearly using,

$$\begin{aligned}
\widetilde{EV}^A(\omega^\star) &\approx EV^A(\omega_0) + EV_X^A(\omega_0)\, x^\star + EV_t^A(\omega_0)\, s^\star \\
\widetilde{EV}^A(\omega^\star) &\approx EV^A(\omega_0) + EV_X^A(\omega_0)\Big(x^\star + \phi(\omega_0^\star)\, s^\star\Big)
\end{aligned} \tag{7}$$

where $EV_X^A = EV([t,s,X{+}1,o]) {-} EV([t,s,X,o])$, $EV_t^A = EV([t{+}1,s,X,o]) {-} EV([t,s,X,o])$, and $\phi(\omega_0^\star) = EV_t^A(\omega_0^\star)/EV_X^A(\omega_0^\star)$. The term $\phi(\omega_0^\star)$ 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 $\phi(\omega_0^\star)$ point change in the relative score.

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

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,

$$R^A(s;\phi) = \underbrace{E\Big(x^\star \Big| s^\star > s\Big)}_{\text{points from continued search}} + \underbrace{\phi\, E\Big(s^\star - s \Big| s^\star > s\Big)}_{\text{point-value of continued search time}} \tag{8}$$

That is, the reservation expected-point value for a particular shot is equal to the expected number of points scored later in the possession conditional on continued search (the first term) plus the points-value of the expected passage of time later in the possession conditional on continued search (the second term). Early in the game and in other states where $\phi \approx 0$ this is the policy that maximizes expected points per possession. The policy sacrifices expected points to lengthen possessions when $\phi > 0$ (i.e. when a team leads late in a game) and sacrifices expected points to shorten possessions when $\phi < 0$ (i.e. when a team trails late in a game). Given the fixed boundary at $s = 35$, the reservation value can be defined recursively. First, define the auxiliary function $z^A(s;\phi) = E\left[s^\star\phi + x^\star \middle| s^\star \geq s\right]$. I first construct an expression for z^A, and then use z^A to construct R^A.

$$z^A(s;\phi) \;=\; \begin{cases} 35\, \upsilon_A\, \phi \;+\; (1-\upsilon)E\max\Big(p\pi, 0\Big) & \text{if } s = 35 \\ s\, \upsilon_A\, \phi \;+\; (1-\upsilon)E\max\Big(p\pi, z^A(s+1;\phi)\Big) & \text{if } s < 35 \end{cases} \tag{9}$$

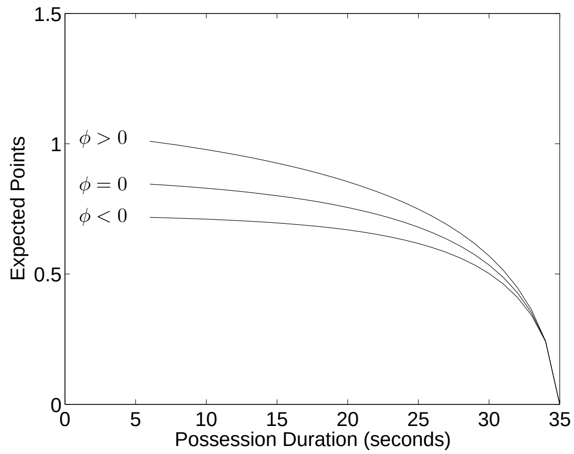
$$R^A(s;\phi) \;=\; \begin{cases} 0 & \text{if } s = 35 \\ (1-\upsilon_A)\Big[E\max\Big(p\pi, z^A(s+1;\phi)\Big) \;-\; s\,\phi\Big] & \text{if } s < 35 \end{cases} \tag{10}$$

Recall that υ 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 $\phi = 0$.

Below I use simulation experiments to examine the extent to which these sorts of optimal pace adjustments introduce skewness to the distribution of winning margins. There are other types of in-game adjustments that teams may make in actual basketball games that are not included in this model. For example, the offensive team and/or defensive team may adjust their tactics to trade off turnover hazard against the quality of arriving shots. If both teams are on the tactical frontier then adjustments (by the offense or defense) that increase the expected turnover hazard should improve the expected quality of arriving shots conditional on the shot opportunity arriving ("high-risk/high-reward"). Tactical adjustments that reduce the turnover hazard should reduce the quality of arriving shot opportunities ("low-risk/low-reward"). Similar to understanding the consequences of pace adjustments on the skewness of winning margins, one would need to understand whether the relative cost in terms of expected points per possession is larger for adjustments towards high-risk or low-risk strategies. If adjustments toward low-risk tactics are less costly than adjustments toward high-risk tactics, one would expect larger adjustments by leading teams, toward those low-

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 (ϕ) 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.

risk strategies, exacerbating the right-skewness of winning margins.[11]

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.

[11]While these other sorts of model features are interesting in their own right, I leave the estimation of the "menu" of such adjustments available to teams for future work.

Table 1: Descriptive Statistics - Possessions

	All	$PS \in [0,4]$		$PS \in (4,8]$		$PS \in (8,12]$		$PS \in (12,16]$		$PS \in (16,20]$		$PS > 20$	
All Possessions		F	U	F	U	F	U	F	U	F	U	F	U
Duration of Possession (sec)	16.11	16.32	16.26	16.27	16.17	16.09	16.23	15.78	16.18	15.28	16.19	14.51	16.14
Points	0.84	0.85	0.82	0.88	0.81	0.90	0.78	0.92	0.75	0.93	0.72	0.94	0.69
Made two point shot	0.30	0.31	0.30	0.31	0.29	0.33	0.28	0.33	0.27	0.33	0.26	0.34	0.24
Missed two point shot	0.25	0.25	0.25	0.24	0.26	0.23	0.26	0.23	0.27	0.24	0.27	0.22	0.28
Made three point shot	0.10	0.10	0.09	0.10	0.09	0.10	0.09	0.10	0.09	0.10	0.08	0.11	0.08
Missed three point shot	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.16	0.17	0.17	0.17
Turnover	0.18	0.18	0.18	0.17	0.19	0.17	0.20	0.17	0.20	0.17	0.22	0.17	0.23
Observations	641,447	93,547	93,760	87,139	88,005	57,293	58,325	39,879	41,005	21,121	21,762	19,384	20,227
First Half Possessions													
Duration of Possession (sec)	16.11	16.18	16.29	15.96	16.37	15.71	16.59	15.46	16.61	14.97	16.76	14.34	16.59
Points	0.80	0.82	0.79	0.83	0.77	0.87	0.75	0.90	0.72	0.90	0.69	0.93	0.65
Made two point shot	0.27	0.28	0.27	0.28	0.26	0.30	0.25	0.30	0.25	0.30	0.24	0.32	0.22
Missed two point shot	0.26	0.25	0.26	0.25	0.27	0.24	0.27	0.24	0.27	0.24	0.27	0.22	0.27
Made three point shot	0.10	0.10	0.10	0.10	0.09	0.11	0.09	0.11	0.09	0.11	0.08	0.11	0.08
Missed three point shot	0.17	0.17	0.17	0.18	0.17	0.18	0.17	0.18	0.17	0.17	0.17	0.18	0.17
Turnover	0.20	0.19	0.20	0.19	0.20	0.18	0.22	0.17	0.22	0.17	0.24	0.17	0.26
Observations	320,064	46,531	46,697	43,845	43,889	28,719	28,928	19,945	20,277	10,644	10,814	9,736	10,039

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.

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.[12] 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.[13]

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 π is discrete, it is convenient to express the joint density as the product of a probability mass function and a conditional density; $f(\pi, p) = f_\pi(\pi)\ f(p|\pi)$. 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, and V_3)[14]. The final parameter governing the search process is υ, the constant per-second turnover hazard. The full parameter vector describing a single team type's search process is given by $\theta = [f_{\pi=2}, M_2, V_2, M_3, V_3, \upsilon]'$[15].

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 π.

[12] 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 1 and 2 for the relevant success rates across point-spread categories.

[13] Appendix 3 (online) provides theoretical and empirical evidence in support of this assumption.

[14] 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 $f(p \mid 2)$ and $f(p \mid 3)$ must have support confined to $[0, 1]$. A common parameterization of the beta density is given by $f_X(x) = \dfrac{x^{\alpha-1}\,(1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1}\,(1-x)^{\beta-1}\,dx}$ for $x \in [0, 1]$. I use the re-parameterization $M = \dfrac{\alpha}{\alpha+\beta}$ and $V = \alpha + \beta$. M is the mean of the random variable and V is inversely related to the variable's dispersion.

[15] 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.

Table 2: Maximum Likelihood Estimates of Structural Parameters by Point-Spread Category

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$PS \in [0,4]$		$PS \in (4,8]$		$PS \in (8,12]$		$PS \in (12,16]$		$PS \in (16,20]$		$PS > 20$	
	F	U	F	U	F	U	F	U	F	U	F	U
Prob. shot opportunity is worth two points	0.533	0.536	0.603	0.563	0.568	0.520	0.607	0.551	0.797	0.515	0.823	0.489
	(0.036)	(0.031)	(0.041)	(0.035)	(0.062)	(0.037)	(0.069)	(0.044)	(0.049)	(0.058)	(0.043)	(0.062)
Mean of success prob. distribution - two pointers	0.139	0.133	0.134	0.127	0.157	0.118	0.163	0.112	0.144	0.102	0.151	0.121
	(0.008)	(0.007)	(0.008)	(0.007)	(0.015)	(0.008)	(0.016)	(0.008)	(0.012)	(0.011)	(0.012)	(0.0144)
Scale of success prob. distribution - two pointers	3.557	3.695	3.543	3.959	3.590	3.579	3.721	3.834	3.633	3.391	3.364	5.436
	(0.138)	(0.143)	(0.144	(0.162)	(0.220)	(0.167)	(0.255)	(0.231)	(0.283)	(0.254)	(0.259)	(0.636)
Mean of success prob. distribution - three pointers	0.068	0.058	0.089	0.056	0.095	0.047	0.109	0.043	0.177	0.038	0.223	0.030
	(0.005)	(0.004)	(0.008)	(0.004)	(0.012)	(0.004)	(0.015)	(0.004)	(0.025)	(0.005)	(0.024)	(0.0034
Scale of success prob. distribution - three pointers	6.215	5.109	7.250	5.233	7.178	4.879	7.652	4.280	11.450	4.637	18.284	2.662
	(0.371)	(0.265)	(0.525)	(0.297)	(0.714)	(0.331)	(0.888)	(0.339)	(2.450)	(0.547)	(4.772)	(0.211)
Turnover hazard	0.016	0.016	0.016	0.016	0.015	0.017	0.015	0.018	0.016	0.018	0.017	0.020
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	46,531	46,697	43,845	43,889	28,719	28,928	19,945	20,277	10,644	10,814	9,736	10,039
Free throw success probability	0.691	0.68	0.694	0.686	0.697	0.68	0.692	0.667	0.695	0.678	0.702	0.652
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.01)	(0.008)	(0.01)
Observations	10,294	9,913	11,187	9,709	8,084	6,516	5,813	4,324	3,426	2,352	3,475	2,102

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.

Given a reservation rule $R^o(s) \equiv R^o(s;\phi = 0)$ 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, π) 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,

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

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 ($s^\star$) and the event that caused the possession to end ($e^\star$). 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 $(s^\star, e^\star)$ is given by:

$$\Pr(s^\star, e^\star) = \Pr(e^\star|s^\star) \prod_{s=6}^{s^\star-1} \Pr(\text{continued search}|s) \tag{12}$$

CIA further implies that the probability of observing a sample containing possessions $j = 1..N$

described by the pairs $\Big\{(s_j^\star, e_j^\star)\Big\}_{j=1}^{N}$ is given by:

$$\Pr\Big(\Big\{(s^\star, e_j^\star)\Big\}_{j=1}^{N}\Big) = \prod_{j=1}^{N} \Pr(s^\star, e_j^\star) \tag{13}$$

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,

$$l\Big(\theta\Big) = \ln\Big(\Pr\Big(\Big\{(s^\star, e_j^\star)\Big\}_{j=1}^{N} \Big| \, \theta\Big)\Big), \tag{14}$$

where $\theta = [f_{\pi=2}, M_2, V_2, M_3, V_3, \upsilon]'$.

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[16].

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 (online) 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.[17] 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 3 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.

[16]My estimation routines converge to the same parameter vectors regardless of the initial guess.

[17]For example the model does not allow for the possibility of possessions ending in defensive fouls in the act of the offense shooting, which result in two or three free throw attempts depending on the value of the attempted shot. These outcomes are classified as “made” attempts. Fewer than six percent of first half possessions end with fouls during the act of shooting. Because these are relatively rare “terminal” events and teams convert nearly 70% of free throw attempts, this simplification is only a small deviation from reality and significantly simplifies the model’s solution.

Table 3: Descriptive Statistics - Games

Category	(1) Percent
2003-2004 Season	12
2004-2005 Season	17
2005-2006 Season	21
2006-2007 Season	24
2007-2008 Season	26
Point Spread $\in [0, 4]$	29
Point Spread $\in (4, 8]$	27
Point Spread $\in (8, 12]$	18
Point Spread $\in (12, 16]$	13
Point Spread $\in (16, 20]$	7
Point Spread $\in (20, \infty]$	6
Observations	5,258

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 2 (online) for details). Source: author's calculations.

Table 4 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.

Table 1 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[18]. Consistent with the findings of Wolfers (2006), the

[18]The favorite's predicted halftime lead is estimated with a regression of that quantity on a constant and the point spread

Table 4: Regression Analysis of the Point Spread's Predictive Accuracy

	(1) OLS	(2) MAD
Amount by which home team is favored	1.01	1.00
	(0.02)	(0.03)
Constant	0.09	0.00
	(0.18)	(0.31)
Observations	5258	5258

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.

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.

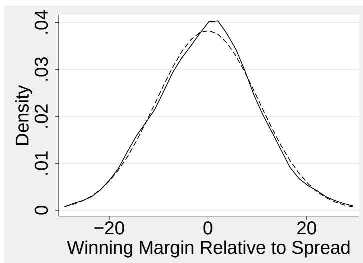
5 Structural Parameter Estimates

To recover estimates of the model parameters, I apply the estimation routine described in Section 3 to the data described in Section 4. Table 2 provides estimates of the model's parameters by point-spread category.

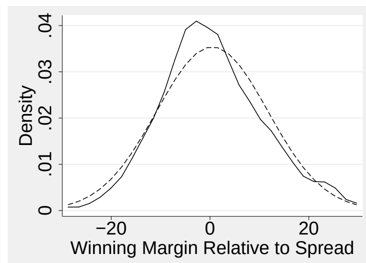
Table 5 assesses the fit of the model to the first half data. I compare empirical average possession duration to the average duration predicted by the model, and I compare the fractions of possessions ending in each of the five modeled terminal events to the fractions predicted by the model. Most of the predicted moments closely resemble the empirical moments both within and across point-spread categories. One exception is that the model slightly underpredicts the fraction of possessions ending in turnovers. Presumably this occurs because the turnover hazard and shot quality distribution are not literally constant over the course of the shot clock within possessions as the model assumes. Allowing these arrival processes to vary over the course of possessions would improve the model's fit. As discussed above, endogenizing these processes by allowing teams to control a tradeoff between the turnover hazard and likely shot quality is one interesting avenue for future work but is not explored in this paper.

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

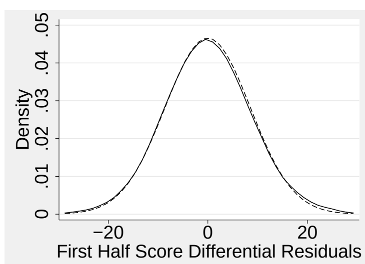
Figure 4: Game Outcomes Relative to Point Spread Prediction



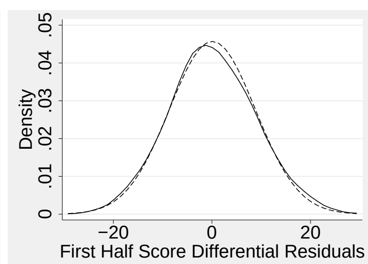
(a) Point Spread < 12



(b) Point Spread $>= 12$



(c) Point Spread < 12



(d) Point Spread $>= 12$

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.

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.

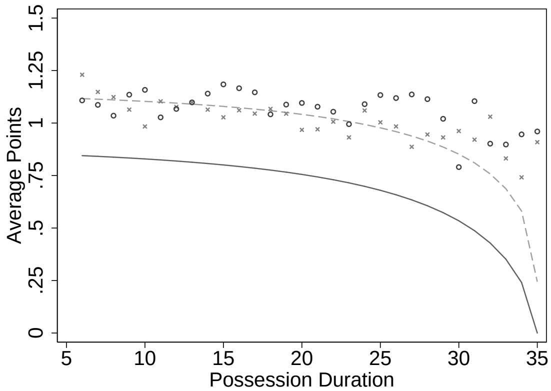
Table 5: Comparison of Empirical and Predicted Moments

	$PS \in [0,4]$				$PS \in (4,8]$				$PS \in (8,12]$			
	F		U		F		U		F		U	
	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$
Duration	16.18	15.85	16.29	15.94	15.96	15.69	16.37	16.00	15.71	15.48	16.59	16.16
Made 2	0.278	0.279	0.270	0.268	0.280	0.281	0.264	0.262	0.296	0.296	0.255	0.254
Missed 2	0.254	0.265	0.256	0.271	0.249	0.260	0.271	0.286	0.237	0.246	0.267	0.282
Made 3	0.0993	0.0969	0.0962	0.0974	0.102	0.100	0.0911	0.0915	0.109	0.107	0.0912	0.0899
Missed 3	0.175	0.188	0.175	0.185	0.182	0.193	0.171	0.182	0.179	0.190	0.170	0.185
Turnover	0.194	0.171	0.203	0.179	0.187	0.167	0.203	0.178	0.180	0.161	0.216	0.189
Observations	46,531		46,697		43,845		43,889		28,719		28,928	
t (equality of mean duration)	10.00		10.29		7.94		10.57		5.48		10.00	
p-value	0.000		0.000		0.000		0.000		0.000		0.000	
Pearson's χ^2 (equality of proportions)	209.95		215.66		154.82		218.56		93.21		170.50	
p-value	0.000		0.000		0.000		0.000		0.000		0.000	

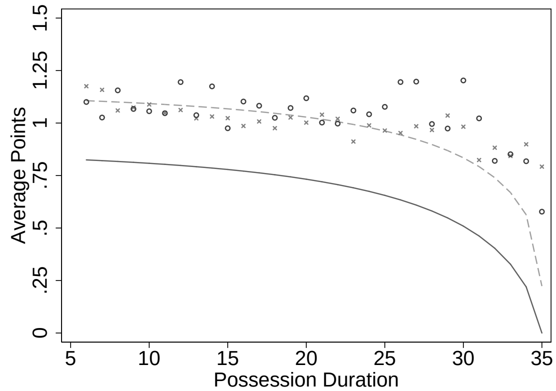
	$PS \in (12,16]$				$PS \in (16,20]$				$PS > 20$			
	F		U		F		U		F		U	
	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$	$\overline{m}$	$\widehat{m}$
Duration	15.46	15.28	16.61	16.17	14.97	14.84	16.76	16.25	14.34	14.24	16.59	16.14
Made 2	0.303	0.303	0.248	0.245	0.305	0.308	0.242	0.239	0.317	0.319	0.218	0.207
Missed 2	0.236	0.244	0.274	0.292	0.239	0.246	0.273	0.292	0.223	0.228	0.274	0.300
Made 3	0.110	0.109	0.0867	0.0886	0.111	0.106	0.0800	0.0809	0.113	0.109	0.0845	0.0959
Missed 3	0.178	0.186	0.167	0.179	0.173	0.182	0.167	0.181	0.180	0.187	0.166	0.171
Turnover	0.173	0.158	0.225	0.196	0.172	0.158	0.238	0.207	0.168	0.157	0.257	0.226
Observations	19,945		20,277		10,644		10,814		9,736,		10,039	
t (equality of mean duration)	3.60		8.63		1.52		7.18		1.52		6.16	
p-value	0.000		0.000		0.065		0.000		0.065		0.000	
Pearson's χ^2 (equality of proportions)	40.68		127.39		22.88		75.80		12.67		86.25	
p-value	0.000		0.000		0.000		0.000		0.013		0.000	

Note: Within each point-spread category, sample moments ($\overline{m}$) and predicted moments ($\widehat{m}$) 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.

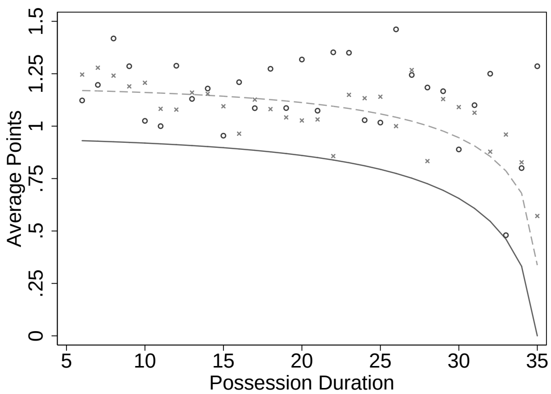
Figure 5: Predicted Reservation Values and Average Points by Time Elapsed from Shot Clock



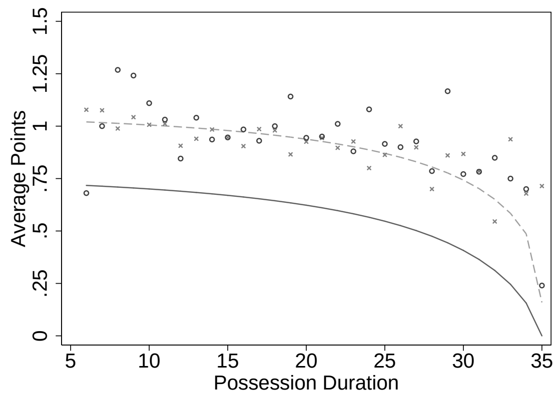
(a) Favorite ($0 \leq$ Point Spread < 4)



(b) Underdog ($0 \leq$ Point Spread < 4)



(c) Favorite ($16.5 \leq$ Point Spread < 20)



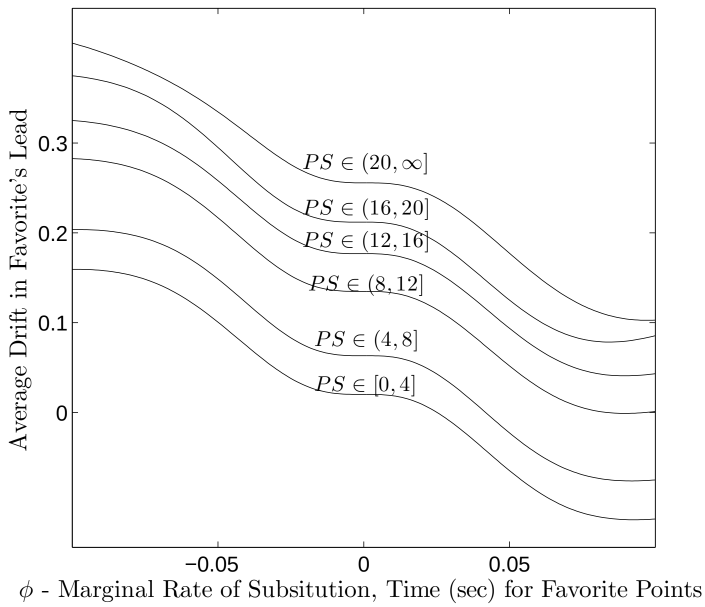
(d) Underdog ($16.5 \leq$ Point Spread < 20)

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.

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 distri-

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 ϕ, the marginal rate of substitution between time and the favorite's points. Source: author's calculations using estimated model parameters.

bution 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.[19] I then compute analogs of the two quantities required to construct the skewness based test for point shaving from the simulated distributions. 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.[20]

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

[19]See Appendix 4 (online) for details on computing the simulated distributions.

[20]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.

Table 6: Mean and Standard Deviation of Favorite's Halftime Lead by Point-Spread Category

	Mean	Standard Deviation
Point Spread $\in [0, 4]$	1.16	8.59
Point Spread $\in (4, 12]$	3.04	8.39
Point Spread $\in (8, 12]$	5.57	8.83
Point Spread $\in (12, 16]$	8.19	8.56
Point Spread $\in (16, 20]$	10.17	8.82
Point Spread > 20	13.82	9.40

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.

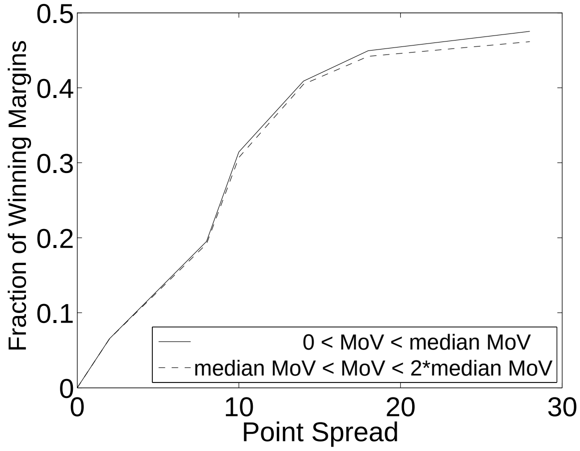
simulations, the offensive team optimally solves the model, but I do not allow the defense to foul intentionally. The difference between the scoring patterns predicted by these two sets of simulations illustrates the impacts of teams' pace adjustments on the direction of skewness in the distribution of favorites' winning margins. In a third set of simulations, both the offense and defense play optimal strategies. This set of simulations shows the extent to which end-of-game fouling by trailing teams exacerbates or lessens this skewness within each point-spread category and provides a benchmark with which the skewness patterns in real games can be compared.

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.

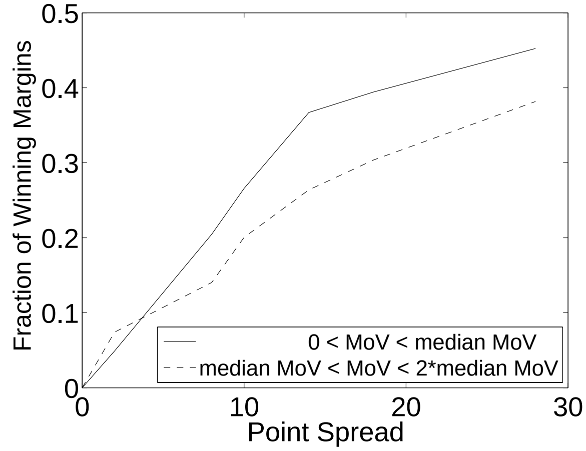
Panel 7-a reports the results of the first set of simulations in which the offensive team always maximizes its expected points per possession. The simulations finds that the distribution of winning margins is nearly symmetric, with about the same fraction of winning margins falling in the lower region 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 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. Note that the predicted right skewness actually exceeds what occurs in actual games (panel 7-d) in games with point spreads less than 12 points.

Panel 7-c 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

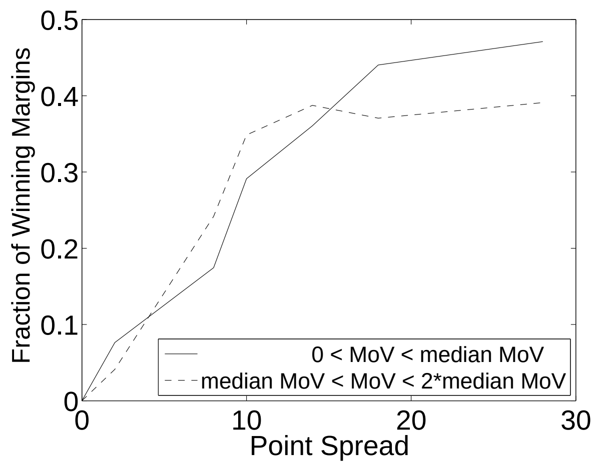
Figure 7: False Experiments - Skewness Based Test for Point Shaving Applied to Simulated Outcomes



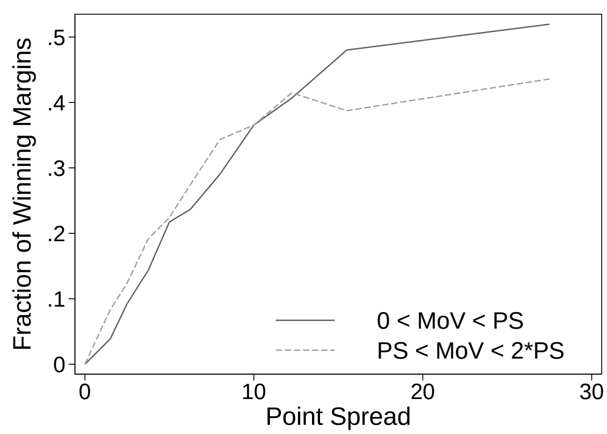
(a) Simulations Assuming Offense Maximizes Points in all Possessions



(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.

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 7-d. 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 7-c 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. The simulation results also suggest that modeling defensive fouling choices is necessary to accurately predict the distribution of winning margins, as the distribution in panel 7-c fit the data substantially better than the distribution in panel 7-b.

7 Corroborating Evidence

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

Table 7 presents evidence on the relationship between the length of possessions and the favorite's lead during different time periods within games. Specifically, I estimate the regression,

$$\begin{aligned}\text{Duration}_j \quad = \quad & a_0 \times \text{FavLead}_j \\ & + a_1 \times \text{FavLead}_j \times 1(\text{mins. 1 to 5 of 2nd half}) \\ & + a_2 \times \text{FavLead}_j \times 1(\text{mins. 6 to 10 of 2nd half}) \\ & + a_3 \times \text{FavLead}_j \times 1(\text{mins. 11 to 15 of 2nd half}) \\ & + a_4 \times \text{FavLead}_j \times 1(\text{mins. 16 to 20 of 2nd half}) + \theta_{ps,t} + e_j\end{aligned} \tag{15}$$

where Duration_j is possession length in seconds, $FavLead_j$ is the favorite's lead *entering* the possession, $\theta_{ps,t}$ is a point-spread by time fixed effect, and the interaction terms allow the impact of the score differential on pace to differ by time. I estimate separate regressions for possessions when the favorite is on offense (column 1) and when the underdog is on offense (column 2). The parameter a_0 describes the impact of the favorite's lead on possession duration during the first halves of games. I estimate a_0 values of 0.014 for favorite possessions and -0.013 for underdog possessions. While these figures are statistically significantly different from zero given the large sample size, they are small and qualitatively consistent with the assumption underlying the structural estimation approach that the marginal rate of substitution between time and points is close to zero in the first halves of games. The point estimates imply that for both the favorite and the underdog average possession duration changes by just above 0.1 seconds for every 10 point change in the score differential during the first half. As expected, the relationship between score differential and possession duration grows steadily stronger as time passes during the second half. The estimates of a_4 reported in the last row (0.215 and -0.280 for the favorite and underdog respectively) imply that average possession duration changes by over two seconds for every 10 point change in score differential during the last five minutes of games.

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

	(1) Favorite Possessions	(2) Underdog Possessions
Favorite's Lead	0.0142***	-0.0132***
	(0.003)	(0.003)
Favorite's Lead × 1(mins. 1 to 5 of 2nd half)	0.0107*	-0.0210***
	(0.005)	(0.005)
Favorite's Lead × 1(mins. 6 to 10 of 2nd half)	0.0255***	-0.0663***
	(0.005)	(0.005)
Favorite's Lead × 1(mins. 11 to 15 of 2nd half)	0.0594***	-0.127***
	(0.005)	(0.006)
Favorite's Lead × 1(mins. 16 to 20 of 2nd half)	0.215***	-0.280***
	(0.006)	(0.006)
Observations	289,736	290,507

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

Standard errors in parentheses. $\star\ p < 0.05$, $\star\star\ p < 0.01$, $\star\star\star\ p < 0.001$.
Note: The dependent variable in each regression is the possession duration in seconds. The sample includes all possessions that begin with the score differential in $[-20, 20]$. Column (1) restricts to possessions where the favored team is on offense, and column (2) restricts to possessions where the underdog is on offense. Source: author's calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.

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 numerically simulated model using estimated parameter values 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 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

Table 8: Regression Analysis of Defensive Fouling Policy

	(1) foul	(2) foul
Predicted foul region	0.312***	
	(0.004)	
Predicted foul region - interior		0.363***
		(0.006)
Predicted foul region - near edge		0.206***
		(0.005)
Close to predicted foul region		0.059***
		(0.004)
Constant	0.110***	0.110***
	(0.000)	(0.000)
Observations	820,672	820,672

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

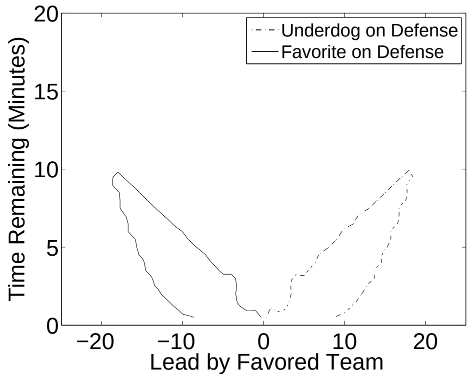
quick defensive fouls occur at similarly elevated rates on both the interior and edge of the optimal fouling region, and occur at only a slightly elevated rate just outside of the optimal fouling region. In sum, these results suggest that the strategic adjustments in the model that induce skewness to the distribution of favorites' winning margins also occur in actual games.

8 Conclusion

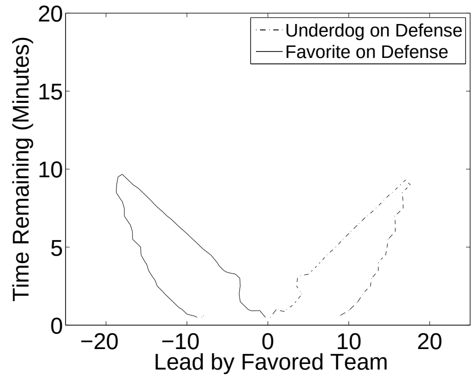
This paper finds that the inference of widespread point shaving from skewness in the distribution of final score differentials is ill-founded. While the skewness-based test for point shaving relies on an assumption that winning margins are symmetric in the absence of point shaving, the model considered 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

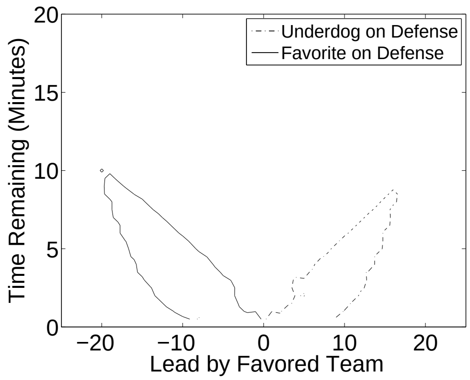
Figure 8: Predicted Optimal Defensive Fouling Policies



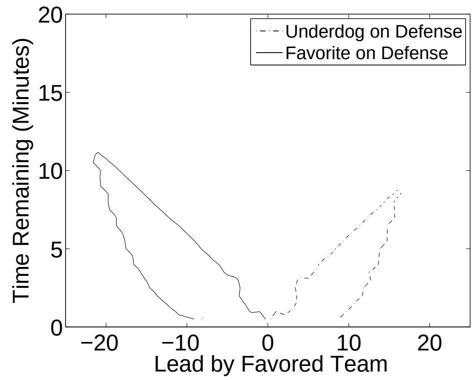
(a) Point Spreads 0 to 4



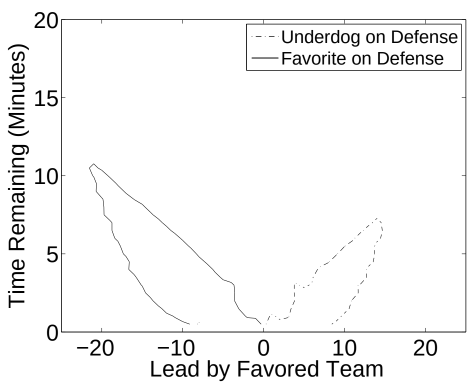
(b) Point Spreads 4.5 to 8



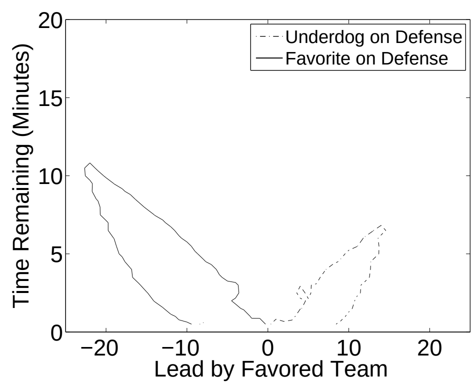
(c) Point Spreads 8.5 to 12



(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.

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[21]. 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[22]. 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 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.

Appendix 1

Proof of Proposition 2: Under the proposition's premise, it follows immediately from the first order condition in equation (3) 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. Totaly 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,

[21]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.

[22]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

$$\begin{aligned}\frac{\partial E(X_2|X_1)}{\partial X_1} &= \frac{\partial \mu(\sigma_{A2})}{\partial \sigma_{A2}}\frac{\partial \sigma_{A2}}{\partial X_1} - \frac{\partial \mu(\sigma_{B2})}{\partial \sigma_{B2}}\frac{\partial \sigma_{B2}}{\partial X_1} \\ &= \frac{\mu'(\sigma_{A2})}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{A2})} + \frac{\mu'(\sigma_{B2})}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{B2})} \\ &= \mu'(\sigma_{A2})\Big(\frac{1}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{A2})} - \frac{1}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{B2})}\Big)\end{aligned}$$

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,

$$\frac{\partial E(X_2|X_1)}{\partial X_1} = \mu'(\sigma_{A2})\Big(\frac{1}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{A2})} - \frac{1}{(\sigma_{A2}+\sigma_{B2})\mu''(\sigma_{B2})}\Big)$$

If for any $\sigma' < \sigma^*$ and $\sigma'' > \sigma^*$ with $u'(\sigma') = -u'(\sigma)$ that $|u''(\sigma')| < |u''(\sigma'')|$ (where σ^* is the action that maximizes $\mu()$), then $\mu'(\sigma_{A2})$ will be opposite in sign from the term in brackets. Conversely if for any $\sigma' < \sigma^*$ and $\sigma'' > \sigma^*$ with $u'(\sigma') = -u'(\sigma)$ that $|u''(\sigma')| > |u''(\sigma'')|$, then $\mu'(\sigma_{A2})$ will be the same sign as the term in brackets. Therefore, the proposition holds.
Proof of Proposition 4: By Zwet (1964) a random variable with distribution function G is more right skewed than another random variable with distribution function H if $G^{-1}(H(x))$ is convex in x. Let H be the CDF of X_1 and let G be the CDF of $X_1 + E(X_2|X_1)t$. Then $G^{-1}(H(x)) = x + E(X_2|X_1 = x)$, and, because the skewness of X_1 is zero, the random variable $X_1 + E(X_2|X_1)$ is right skewed if $E(X_2|X_1 = x)$ is convex in x. By the same reasoning, the random variable $X_1 + E(X_2|X_1)$ is left skewed if $E(X_2|X_1 = x)$ is concave in x.

References

[1] Richard Borghesi. Widespread Corruption in Sports Gambling: Fact or Fiction? *Southern Economic Journal*, 74(4):1063–1069, 2008.

[2] Keith C. Brown, W. V. Harlow, and Laura T. Starks. Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry. *The Journal of Finance*, 51(1):85–110, 1996.

[3] Jeffrey A. Busse. Another Look at Mutual Fund Tournaments. *The Journal of Financial and Quantitative Analysis*, 36(1):53–73, 2001.

[4] William Chan, Pascal Courty, and Li Hao. Suspense: Dynamic Incentives in Sports Contests. *Economic Journal*, 119(534):24–46, 01 2009.

[5] Judith Chevalier and Glenn Ellison. Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 105(6):1167–1200, 1997.

[6] Mark Duggan and Steven D. Levitt. Winning isn't everything: Corruption in sumo wrestling. *The American Economic Review*, 92(5):1594–1605, 2002.

[7] Bernd Frick. Contest Theory and Sport. *Oxf Rev Econ Policy*, 19(4):512–529, 2003.

[8] Alasdair Houston and John McNamara. The choice of two prey types that minimises the probability of starvation. *Behavioral Ecology and Sociobiology*, 17(2):135–141, 1985.

[9] Hans K. Hvide. Tournament Rewards and Risk Taking. *Journal of Labor Economics*, 20(4):877–898, 2002.

[10] Brian A. Jacob and Steven D. Levitt. Rotten Apples: An Investigation of The Prevalence and Predictors of Teacher Cheating*. *Quarterly Journal of Economics*, 118(3):843–877, 2003.

[11] Neal Johnson. NCAA "Point Shaving" as an Artifact of the Regression Effect and the Lack of Tie Games. *Journal of Sports Economics*, 10(1):59–67, 2009.

[12] Edward P. Lazear and Sherwin Rosen. Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89(5):841, 1981.

[13] Steven Levitt. Why are gambling markets organised so differently from financial markets? *Economic Journal*, 114(495):223–246, 2004.

[14] Christopher Parsons, Johan Sulaeman, Michael Yates, and Daniel Hamermesh. Strike Three: Discrimination, Incentives, and Evaluation. *American Economic Review*, 101(4):1410–1435, 2011.

[15] David Porter. *Fixed: How Goodfellas Bought Boston College Basketball*. Taylor Trade, 2002.

[16] Joseph Price and Justin Wolfers. Racial Discrimination Among NBA Referees. *Quarterly Journal of Economics*, 125(4):1859–1887, 2010.

[17] David Romer. Do Firms Maximize? Evidence from Professional Football. *Journal of Political Economy*, 114(2):340–365, 2006.

[18] Charley Rosen. *The Wizard of Odds: How Jack Molinas Almost Destroyed the Game of Basketball*. Seven Stories Press, 2001.

[19] John Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, September 1987.

[20] Justin Wolfers. Point Shaving: Corruption in NCAA Basketball. *American Economic Review*, 96(2):279–283, 2006.

[21] Justin Wolfers and Eric Zitzewitz. Interpreting Prediction Market Prices as Probabilities. *CEPR Discussion Paper No. 5676.*, 2006.

[22] Eric Zitzewitz. Forensic Economics. *Journal of Economic Literature*, Forthcoming.

[23] William Rutger van Zwet. Convex Transoformations of Random Variables. *Mathematish Centrum*, 1964.