Do Basketball Scoring Patterns Reflect Illegal Point Shaving or Optimal In-Game Adjustments?: Online Appendix

Jesse Gregory

University of Wisconsin, NBER

jmgregory@ssc.wisc.edu

Current Version: August, 2017

Appendix 2: Data (Online – Not for Publication)

The data used in this study come from two sources. The first data source is a compilation of detailed play-by-play records for 8,102 regular season basketball games played between November, 2003 and March, 2008 downloaded from the website statsheet.com. The raw play-by-play data consists of a list of game events for each team (i.e. shot attempts, rebounds, and turnovers) and, for each event, the time remaining in the game when the event occurred. These 8,102 games exclude games for which the play-by-play record is insufficiently detailed to allow an accurate calculation of the time elapsed from the shot clock at the conclusion of each possession[1]. The second data source is a set of gambling point spreads for 24,868 regular-season games played during the same time period. This data is publicly available from the website covers.com. Merging these two data sources by the two team names and the game's date shows that the intersection contains $5{,}258$ games. Data from those $5{,}258$ games are used to construct a data set for analysis.

From these raw data, I construct a data set containing one observation per possession. The set of information for each possession includes the team that was on offense, the time that elapsed from the shot clock during the possession, and the event that caused the possession to end. The key piece of information that must be computed is the time elapsed from the shot clock. Beginning with the raw play-by-play records, I use the following steps.

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

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

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

4. For each event, compute the time elapsed from the shot clock when the event occurred as the difference between the game time at which the event occurred and the game time at which the shot clock was most recently reset[2].

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

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

[2] For possessions following a made shot by the opponent, I set the time at which the shot clock was reset to four seconds after the game time at which the shot was made. Following a made shot, the next shot clock does not begin counting until the team to be on offense carries the ball out of bounds and passes the ball in. This procedure typically takes several seconds.

6. Code each possession to one of the model's terminal events. For possessions ending with a turnover, a made or missed two point shot, or a made or missed three point shot, this coding is straightforward. In other cases, use the following rules. Possessions in which a defensive foul was committed during a two point shot attempt are coded as successful two-point shots. Possessions in which a defensive foul was committed during a three-point shot attempt are coded as successful three-point shots.[3] Possessions that end with an offensive foul are coded as turnovers. Shot clock violations are coded as turnovers occurring in the final second of the shot clock.

7. Drop possessions for which this procedure computes a possession duration exceeding the maximum possible duration (by the shot clock rule) of 35 seconds.[4]

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

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

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

Appendix 3: Marginal rate of substitution between time and points across game states (Online – Not for Publication)

A3.1: Theoretical Evidence The strategy for estimating the model's structural parameters described in Section 3 hinges on the assumption that the marginal rate of substitution between time and points is approximately zero in the first halves of games. To better understand why one might expect the marginal rate of substitution between time and points (ϕ) is close to zero early in games, first consider one special case of the fully dynamic game in which the value function, and hence ϕ, can be approximated with a closed form expression. Specifically, consider a game between two evenly matched teams in a setting where "stalling" and "hurrying" are equally costly in terms of foregone expected points per possession. In this case, teams' state-specific adjustments in pairs of possessions will offset one another, and the game's score will evolve approximately as a random walk. By the central limit theorem, the value function (probability of team A winning) with a current score differential of X at t seconds into the game can be approximated by,

$$V^A(X,t) \approx \Phi\left(\frac{X}{\sigma\sqrt{T-t}}\right) \tag{1}$$

[3]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.

[4]This procedure coded 1.21 percent of observations to durations exceeding 35 seconds.

where $\Phi(.)$ is the normal CDF, σ is the average standard deviation of the score differential per second, and $T-t$ is the number of seconds remaining in the game. Making use of this expression, the maginal rate of substitution between time and points is approximately,

$$\begin{aligned}
\phi(X,t) &\approx \frac{\partial\Phi\left(\frac{X}{\sigma\sqrt{T-t}}\right)/\partial t}{\partial\Phi\left(\frac{X}{\sigma\sqrt{T-t}}\right)/\partial X} \\
&= \frac{\phi\left(\frac{X}{\sigma\sqrt{T-t}}\right)\times\left(\frac{\frac{1}{2}\times X}{\sigma(T-t)^{3/2}}\right)}{\phi\left(\frac{X}{\sigma\sqrt{T-t}}\right)\times\left(\frac{1}{\sigma\sqrt{T-t}}\right)} \\
&= \frac{X}{2\,(T-t)}
\end{aligned} \tag{2}$$

This expression suggests two reasons to expect the ϕ to be small in absolute value during the first half of games. First, point differentials X are on average smaller in absolute value during the first halves of games than later in games. Second, and probably more importantly, the absolute value of ϕ declines geometrically in the amount of time remaining in the game for any given score differential, since $T-t$ appears in the denominator. Since in the first half one possession is only a small fraction of the total time remaining, adjusting possession duration has a small impact on win probabilities compared to the impact of accumulating a larger scoring lead. This expression suggests that for a given score differential, teams' incentives to hurry or stall may be an order of magnitude smaller during the first halves of games than at the end of games. For example, if team A leads by 10 points with 1 minute left in the game the approximation yields $\phi(10, 39\times 60)\approx .083$, but if team A leads by 10 points midway through the first half of the game the approximation yields $\phi(10, 10\times 60)\approx .003$.

A3.2: Empirical Evidence While this calculation seems to support the assumption that the marginal rate of substitution between time and points is small during the first half of games, it is important to verify that the assumption holds empirically, since, as emphasized throughout the paper, the acutal basketball scoring environment is not one in which stalling and hurrying are equally costly. To do this, I compute nonparametric estimates of ϕ across all game states and point spread categories. I first nonparametrically estimate the probability of the favorite winning conditional on the time remaining in the game, the favorite's current lead, and the pre-game point spread. The estimating equation is,

$$1(favwin_i) = g(X_{ij}, t_{ij}, ps_{ij}) + e_{ij} \tag{3}$$

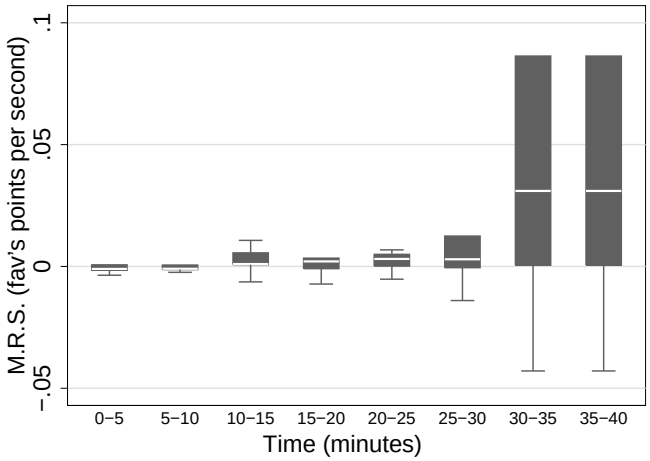
with i indexing games and j indexing individual possessions. Specifically, I estimate local linear regressions with respect to X_{ij} and t_{ij} across a grid in $X\times t$ seperately by point spread category. I then compute a

predicted marginal rate of substitution for each possession using,

$$\widehat{\phi}_{ij}(X_{ij}, t_{ij}, ps_{ij}) = \frac{\partial \widehat{g}(X_{ij}, t_{ij}, ps_{ij})/\partial t}{\partial \widehat{g}(X_{ij}, t_{ij}, ps_{ij})/\partial X} \tag{4}$$

Figure 1 presents box plots illustrating the distribution of these predicted $\widehat{\phi}$'s by 5-minute game intervals. Consistent with the identifying assumption of a zero marginal rate of substitution between time and points during the first half, the distribution of estimated $\widehat{\phi}_{ij}(X_{ij}, t_{ij}, ps_{ij})$ values (across point spread categories and score differentials) is clustered close to zero throughout games' first twenty minutes. Consistent with the intuition suggested by equation A–2, the marginal rate of substitution begins to vary substantially across game states later in the game, particularly in the final ten minutes.

Figure 1: Empirical Marginal Rates of Substitution Between Time and Points by Time Played



Source: author's calculations using play-by-play data from statsheet.com merged to point-spread data from covers.com.

Appendix 4: Simulations (Online – Not for Publication)

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

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

I handle several details of the simulation process as follows. I discretize the shot opportunity distributions, F^o, using ten points of support, so that I can use discrete dynamic programming methods. In states in which one team has a very large lead and therefore the winning team is not in doubt, $V = 0$ or 1[5], I imposed that the offensive team reverts to maximizing expected points per possession and the defensive team chooses not to foul[6]. The model does not make a unique prediction in these states, because many policies yield the same expected payout when the winner is not in doubt.

[5]The winning team is no longer in doubt in the context of the model when insufficient time remains for the trailing team to come from behind even given the most fortuitous possible string of events.

[6]One could imagine other possible policies adopted by teams when the game's winner is no longer in doubt. These policies are consistent with rational play if, for very small $\epsilon > 0$, each team placed a weight of $(1 - \epsilon)$ on the discrete outcome of winning and a weight of ϵ on the margin of victory/defeat.