⊗ ‖‖‖ inpredictable   🔍

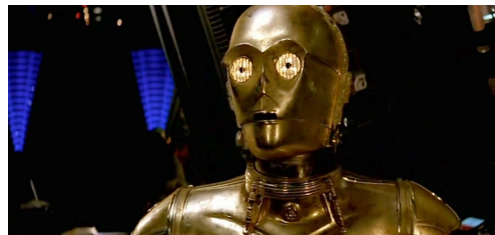HOME   NBA ⌄   WNBA ⌄   SPORTS ⌄   RANKINGS ⌄   TOOLS ⌄   ABOUT ⤬

Updated NBA Win Probability Calculator

FEBRUARY 06, 2015

Last month, I rolled out a new version of my NBA Win Probability Graphs and Box Scores (new link | old link). In addition to adding some new features, such as the option of displaying real time along the horizontal axis, the underlying win probability model was rebuilt as well.

The caption reads: The odds of winning a game when down by 6 with 18 seconds left are approximately 250 to 1.

The dataset was updated and expanded, model parameters were further optimized, and handling of late game situations was improved, particularly in the final seconds.

Until now, that new model was only used to generate the graphs. The interactive Win Probability Calculator was still using the old model. The calculator tool has now been updated with the new, improved model. I have also removed the "Beta" tag that had been there since its inception.

But how do I know the model is improved, and not just new? One way to assess a probability model's accuracy is by measuring log-likelihood. Likelihood, in this context, signifies the probability the model assigned to any specific game outcome. For example, if the model says that the win probability for a team is 15%, and the team actually goes on to win the game, the likelihood is 0.15. If the team lost, the likelihood was 85%. We can do this calculation for all game situations in which the model estimates a win probability. The total likelihood is just the product of all of those individual likelihoods. As a mathematical convenience, one often takes the natural logarithm of that product.

A higher likelihood signifies a better model. A model that routinely assigns 1% win probability to teams that ultimately win will generate low likelihood scores. A model that routinely assigns 95% win probability to teams that actually win will have high likelihood scores. It rewards confident, accurate predictions, while penalizing overconfident, inaccurate predictions.

To apply this test, I took all games from the current NBA season (through January 7, 2015) and had each model assign a likelihood score to each play. The old model was developed from 2004-2011 season data. The new model is based on games from 2000-2012. The 2014 season data is out of sample for both, so this should be a fair, unbiased test. Log likelihood has its own version of an R-squared measure (I'm using the McFadden version). The old model had
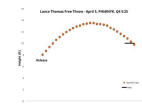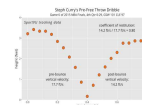
Popular Posts
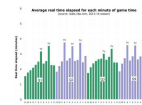
How long is each minute of NBA gametime?

Super Bowl Squares - Every Score Pays Out

Introducing ShArc: Shot Arc Analysis

A Deflategate Analysis for the NBA

Clutch Shooting Isn't Easy (unless it's a free throw)

Updated NBA Win Probability Calculator

Tweets

an R-squared of 0.281 against the 2014 season data. The new model showed modest improvement, with an R-squared of 0.285.

The graph below shows average log likelihood improvement from the new model by minute of game time.

new model - improvement in log likelihood
2014-15 NBA Games thru Jan 7, 2015
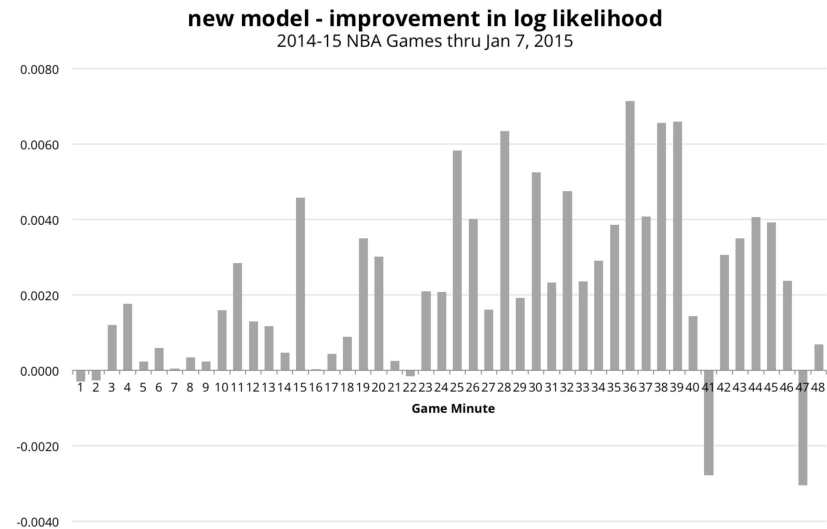
As you can see, there was improvement across the board, with just a few steps back. I think (hope) the dips shown for minutes 41 and 47 are just noise.

Methodology

> **Data**: Play by play data for nearly all NBA games from the 2000 to 2012 seasons. I merged that data with point spreads from sportsdatabase.com

> **Model inputs**: The win probability is a function of game time, point differential, possession, and the Vegas point spread. I use the Vegas point spread to control for differences in team strength that could bias the results. The calculator tool returns probability assuming the two teams are evenly matched (a 0 point spread). While not available in the calculator tool, the point spread adjusted probabilities are available as an option in the new win probability graphs.

> **Modeling approach**: Locally weighted logistic regression (with the assistance of R's locfit package). It is an extension of the more common LOESS methodology to logistic regression. Logistic regression is more appropriate for modeling probabilities. The smoothing window was calibrated via cross validation. The optimal smoothing window shrank as time remaining in the game approached zero. For the final few seconds of game time, I abandoned regression entirely and built a simple decision tree to calculate the probabilities.

> **Non-possession states**: With basketball, there are game situations that don't qualify as "pure" possession states. For example, after a missed shot, but before the rebound. Or if a team has two free throws to shoot off of a personal foul. To calculate those probabilities, rather than building separate regression models, I derive them from the base "pure" possession model, and applied some simple assumptions. For example, on average, missed shots are rebounded by the defense 69.5% of the time. So, the win probability after a missed shot is just 0.695 x the win

Mike Beuoy
@inpredict

Replying to @inpredict

Same type of chart, but broken down by quarter instead.

The Bucks just step on your throat through all four quarters.

OKC is actually a below 0.500 team through 3 quarters. And then they turn it loose in the 4th.

Raptors are like the Warriors of old, winning games in the 3rd.

Embed　　　　　　　　View on Twitter

Blog Archive

probability of the team with possession plus 0.305 times the win probability if the team's opponent has possession. See below for the parameters used to derived win probabilities for these interstitial states:

> Probability of:

>> Defensive rebound (off of field goal): 69.5%

>> Offensive rebound (off of field goal): 30.5%

>> Defensive rebound (off of free throw): 86.6%

>> Offensive rebound (off of free throw): 13.4%

>> Made free throw: 75.6%

>> Missed free throw: 24.4%

NBA   WIN PROBABILITY

RELATED POSTS

NEW NBA FEATURE: DOPPLEGAMERS

THE HACKULATOR: A WIN PROBABILITY BASED GUIDE FOR WHEN TO FOUL

JUDGING WIN PROBABILITY MODELS

LEAVE A COMMENT

Related Posts

Create a Link

Anthony Davis's Torrid MVP Pace

Updated NBA Win Probability Calculator

Squares Probabilities at FiveThirtyEight

► **January** (5)

► **2014** (74)

► **2013** (78)

► **2012** (84)

► **2011** (1)

Tags

AIR BALLS   BANK SHOTS   BETTING

BOOKIES HATE HIM   COLTS

COVER PROBABILITY   DEFLATEGATE

ESPN   EVERY SCORE PAYS

FIVETHIRTYEIGHT   FOUR FACTORS

FREE THROW DEEP DIVES

FREE THROWS   GAME LENGTH

GUESS THE LINES   HERO BALL

HOME COURT   HORSE RACING   MLB

NBA   NCAA BRACKET   NCAAB

NCAAF   NEW FEATURE   NFL   NHL

PACE   PHYSICS   PLAYER ANALYSIS

PLAYOFFS   POINTLESSNESS

RANKINGS   ROBOCAP   SHARC

SITE NEWS   SOCCER   SPORTVU

SQUARES PROBABILITY

STEPHEN CURRY   STRATEGY

TEAM ANALYSIS   TENNIS

TURNOVER INDEX   VISUALIZATION

WIN PROBABILITY   WNBA

Contact Form

Name

Email *

Message *

Send

Search Inpredictable

Search

Recent Posts

New Feature: Measuring Pace Based on Time to Shot

Handicapping the Handicappers

New NBA feature: Dopplegamers

Random Posts

Guess the Lines - Final Results

Turnover Index Week 16

Guess the Lines Week 16

Popular Posts

Super Bowl Squares - Every Score Pays Out

How long is each minute of NBA gametime?

The Turnover Index (possibly overfit edition)