The Hot Hand is Dead

An Analysis of the Legitimacy of the Hot Hand Theory in Basketball

Toni Divic, Thomas O'Shea, Kyle Kwong, Chip Weber

Motivation

Statisticians, economists, and data scientists have attempted to assess whether the Hot Hand Theory in basketball is fact or fallacy. Two major papers represent each side of the argument: Gilovich et al. (1985) initially provided convincing evidence that the hot hand was a fallacy, but more recently Bocskocsky et al. (2014) take a different approach to show that the hot hand theory is, in their belief, a fact.

What would the "Hot Hand" look like?

Figure 1: Stephen Curry

A statistical model of hot hand effect might be:

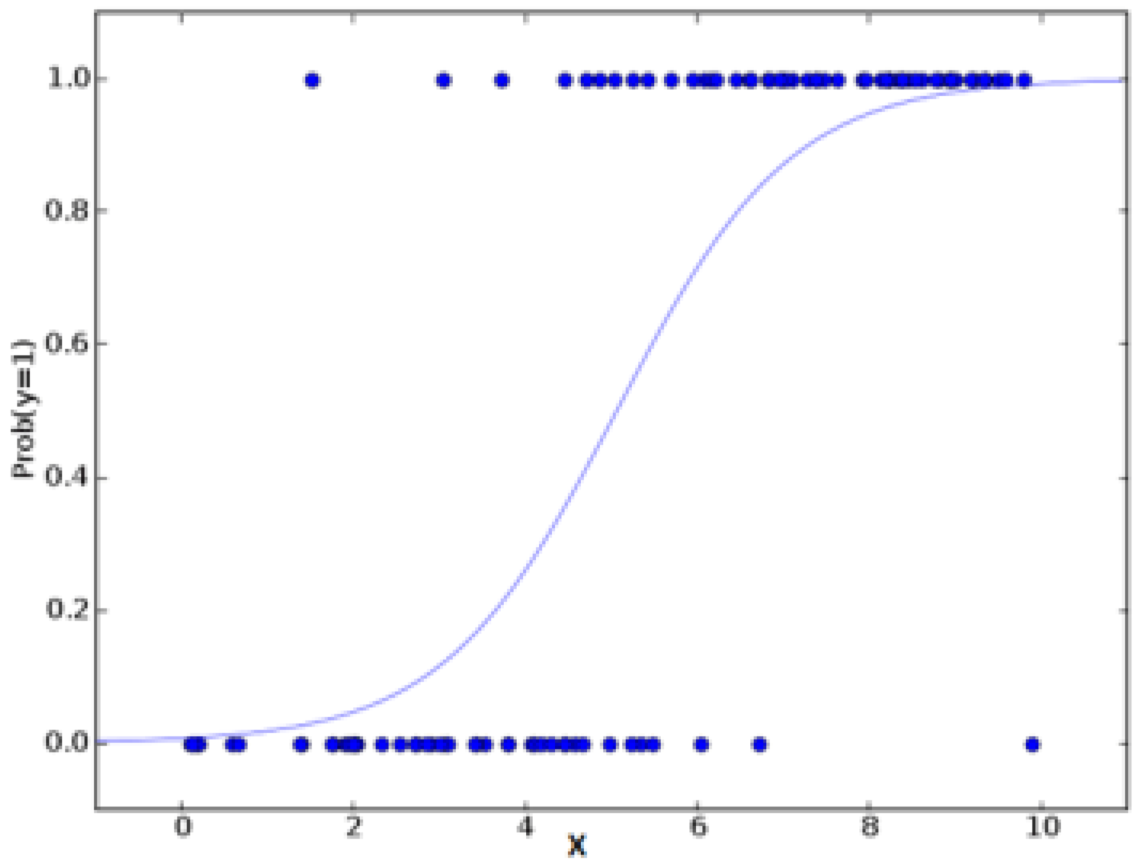

Figure 2: Positive Sloping Logistic Regression Model

The above figure plots a logistic regression classifier that predicts a binary outcome (1 or 0 or in the case of the hot hand, shot made or missed) given some predictor. In the case of the hot hand, that predictor is the shooter's previous *streak*. If the logistic regression predicts that the probability of making a shot increases as the player's streak increases, this indicates the player has the *hot hand*.

Methods

We extract publicly available shot data from the 2013-14 NBA season containing information on the game, player, and other factors in which the shot occurred. From this data, we are able to construct a *previous streak* metric which will serve as our primary predictor in our analysis. We then fit a logistic regression model:

$$ShotMade_s = \beta_1 PreviousStreak_s + \beta \mathbf{X}_s + \epsilon_s \tag{1}$$

where $\mathbf{X}_s$ represents a vector of shot information including shot distance, closest defender's distance, shooter's overall field goal percentage, shot clock, final margin of the game, amongst other predictors. β_1 represents the *hot hand effects*, i.e. the increase in probability of making a shot for each addition to a player's streak.

Main Result

The $PreviousStreak$ variable is generally a poor predictor indicating the hot hand effect is a **fallacy**. In our models, it is marginally significant and either zero or negative. Further, the hot hand effect varies amongst players of interest, but usually is not a significant predictor.

NBA-Wide Results

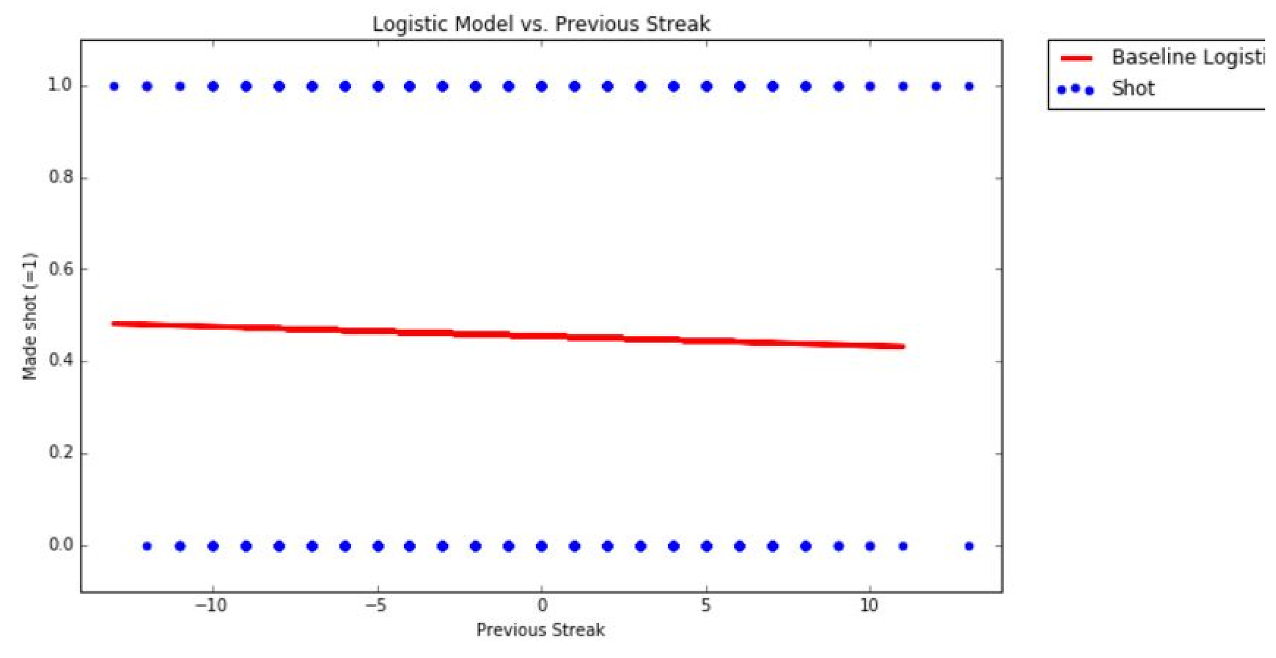

Figure 3: Logit Model with 1 Predictor

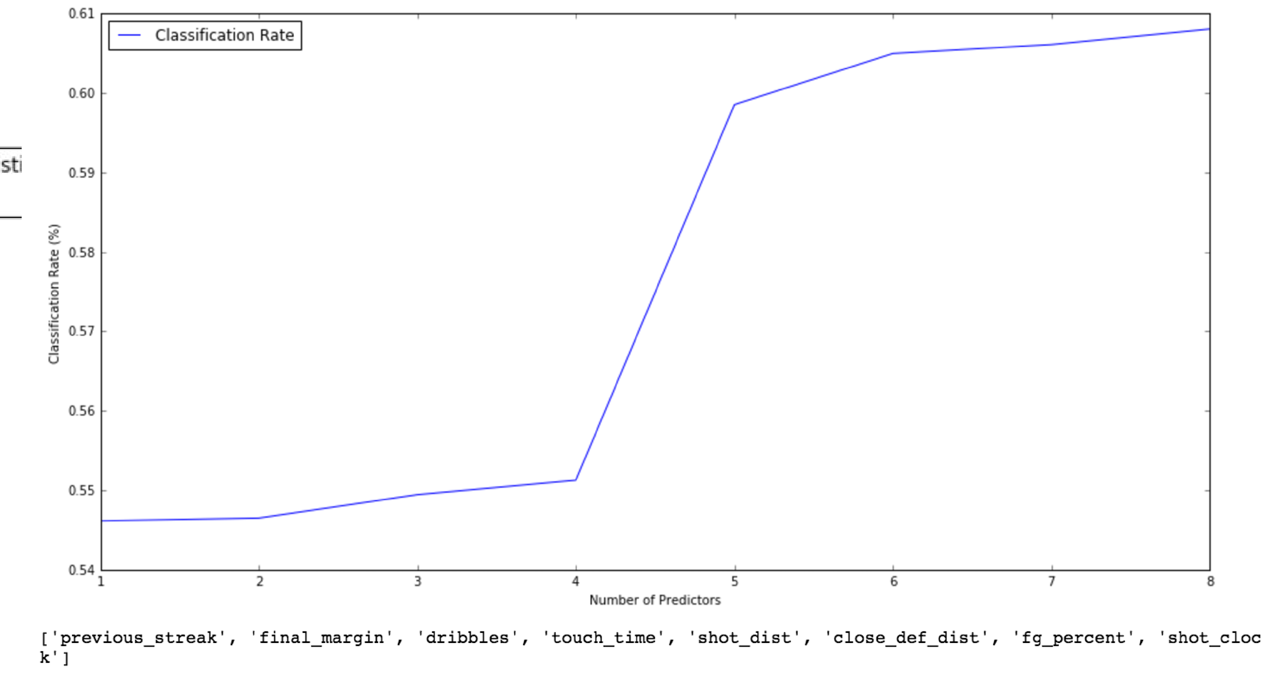

Figure 4: Adding More Predictors

The above illustrates our baseline model using just the previous streak variable. It holds little predictive significance as indicated by the flatness of the line. If anything, the line has a negative slope contrary to our expectations about the hot hand effect. This model is only able to correctly classify **54.6%** of a test set of shots, *barely better than flipping a coin.*

Figure 4 illustrates the model's performance as we add more predictors. With 8 total predictors, the classification rate increases to **60.8%**. The predictor with the greatest predictive significance is *shot distance.*

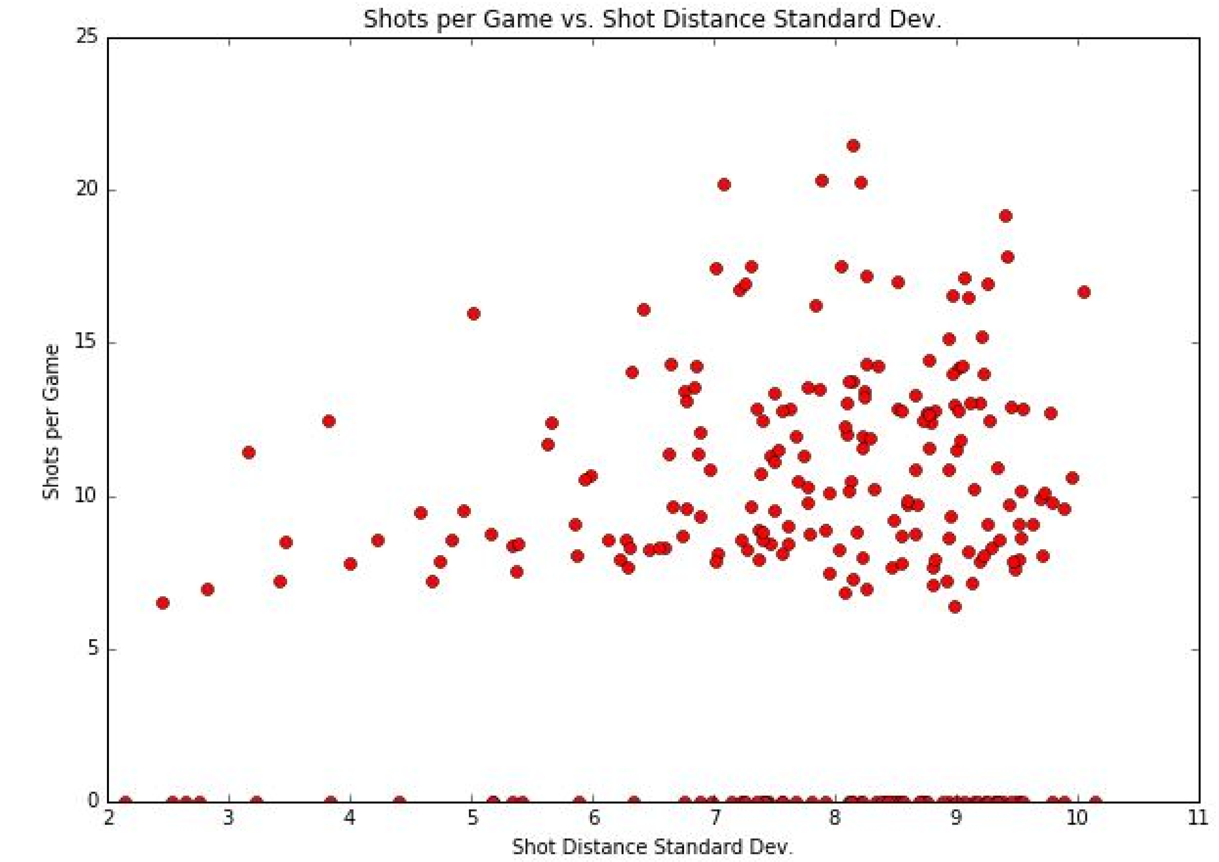

Figure 5: Shots Taken versus Distance Variance

Player by Player Results

Selecting our top players:
We chose to analyze players who take many shots and have a high distance variance (See Fig. 5 plot).

	Player Name	Previous Streak Coefficient	p-Value	Pseudo R^2
0	lebron james	-0.053788	0.116218	0.088236
1	damian lillard	0.029957	0.335786	0.054993
2	russell westbrook	-0.037636	0.252756	0.063786
3	james harden	0.026054	0.350140	0.043948
4	carmelo anthony	-0.052052	0.157199	0.042526
5	stephen curry	-0.075004	0.036462	0.051929
6	kobe bryant	-0.001955	0.954771	0.047502
7	derrick rose	0.051245	0.114370	0.047797
8	kyrie irving	-0.022460	0.485583	0.024724
9	kemba walker	-0.008348	0.820582	0.040961
10	tyreke evans	0.022130	0.500243	0.056303
11	klay thompson	0.016384	0.583960	0.021405
12	mnta ellis	0.033212	0.244191	0.043263
13	lamarcus aldridge	-0.014305	0.635175	0.042875
14	kyle lowry	-0.040356	0.257788	0.059361
15	blake griffin	0.001157	0.970765	0.052497
16	ryan anderson	0.014924	0.699267	0.037514
17	victor oladipo	-0.006465	0.865105	0.082001
18	brandon knight	-0.048439	0.183901	0.033233
19	eric bledsoe	0.002589	0.944338	0.072259

Figure 6: Individual Player Models

The top 20 players are listed above. We fit our robust hot hand model to each and list the model parameters and performance.

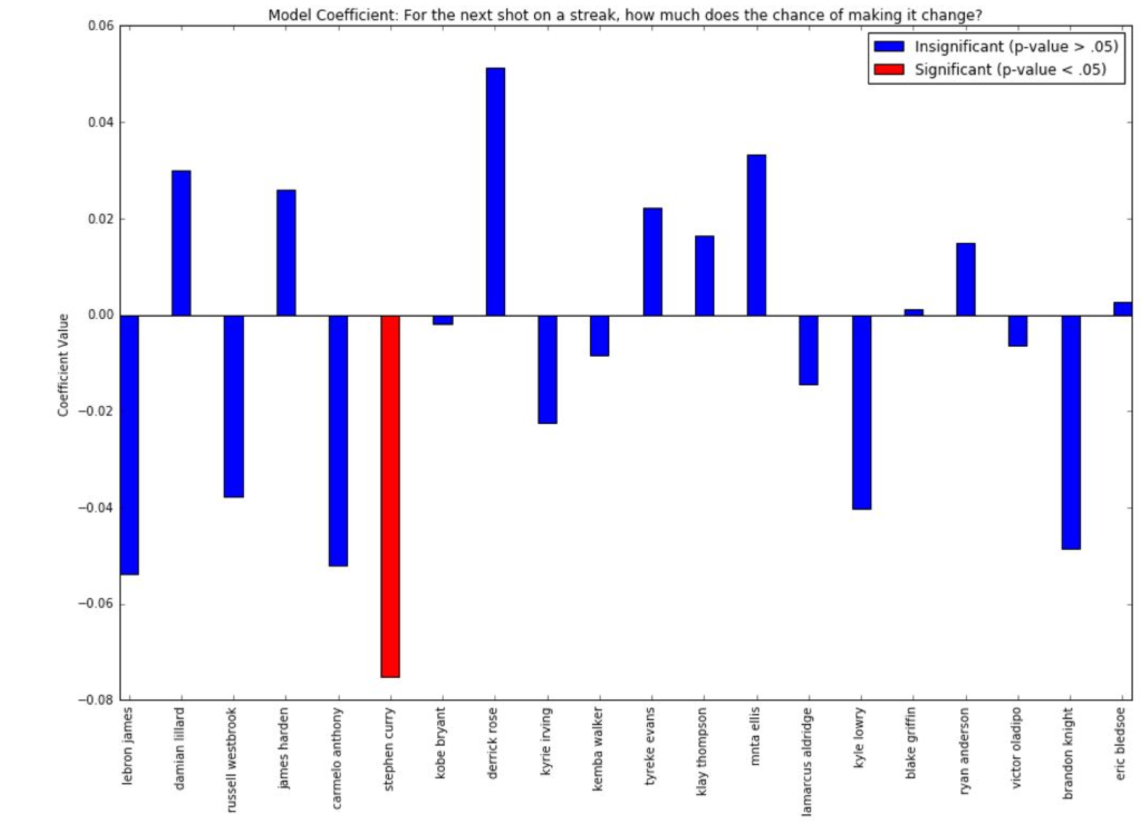

Figure 7: Which Player is the Hottest?

What's notable?

- Curry has the only significant hot hand effect
- Rose has the most positive hot hand effect

Next Steps

- Implement defender, player, game fixed-effects
- If hot hand doesn't exist, does *clutchness* exist?