The Hot Hand: A New Approach to an Old "Fallacy"

Andrew Bocskocsky, John Ezekowitz, and Carolyn Stein

Harvard University

September 21, 2013

Table of Contents

What is the Hot Hand?

In the context of basketball, the *Hot Hand* is the belief that a player who has made several of his past shots is more likely to make his next shot.

In other words, shots are not independent events - the likelihood of make depends on the outcome of past shots.

The Hot Hand is a Fallacy...

The Hot Hand has been disproven in academic literature:

- Most famously by Gilovich, Vallone, and Tversky in 1985 ("The Hot Hand in Basketball: On the Misperception of Random Sequences")
  - Authors show that $P(Hit)$ does not vary conditional on the number of consecutive hits or misses.
  - The length of "streaks" observed is consistent with the expected length under the assumption of independence.
  - There is no streakiness in free-throw shooting.
- Subsequent studies have provided additional evidence that these results hold.

The Hot Hand is a Fallacy...

And is accepted as a fallacy by the academic community:

- This past winter, Larry Summers addressed the Harvard basketball team after one of their practices.
  - He asked them if they believed in the Hot Hand. After they nodded, he told them they were wrong. "People apply patterns to random data" he explained.
- In February, David Brooks wrote a NYT op-ed about the philosophy of big data. In it, he uses the Hot Hand as an example of when our intuition leads us astray.
  - "When a player has hit six shots in a row, we imagine that he has tapped into some elevated performance groove. In fact, its just random statistical noise, like having a coin flip come up tails repeatedly. Each individual shots success rate will still devolve back to the players career shooting percentage."

...Or is it?

The famous Gilovich, Vallone, and Tversky results hinge on one critical assumption:

> *"It may seem unreasonable to compare basketball shooting to coin tossing because a player's chances of hitting a basket are not the same on every shot. Lay-ups are easier than 3-point field goals and slam dunks have a higher hit rate than turnaround jumpers. Nevertheless, the simple binomial model is equivalent to a more complicated process with the following characteristics: Each player has an ensemble of shots that vary in difficulty (depending, for, example, on the distance from the basket and on defensive pressure), and each shot is randomly selected from this ensemble."*

...Or is it?

This is the assumption we challenge in our paper.

If players believe in the Hot Hand, their perception of heat may affect the difficulty of the shots they select. If hot players select more difficult shots, this could offset the effect of heat.

Therefore, the question we seek to answer is twofold:

1. Do players attempt more difficult shots as they become hotter?
2. Are hot players more likely to make their next shot, *controlling for the difficulty of that shot?*

Table of Contents

Raw Data

We were able to obtain and merge data from four different sources:

1. **NBA Roster:** Player traits, including height, weight, and position
2. **NBA Expanded Play-by-Play:** Lists all major events in the game, with additional data such as the time and player(s) associated with the event
3. **SportVU Optical Tracking:** Spatial data that includes x, y, z coordinates for each player and the ball in $1/25$ of second increments
4. **SportVU Play-by-Play Optical:** This dataset has a unique sequence number that matches events in the NBA Play-by-Play data to the SportVU Optical Tracking data

Data by the Numbers

- 6 cameras on each court
- 15 arenas with cameras (50% of total)
- 30 teams all with partial data
- 474 players with shots taken
- ~3,500 data events per game
- 83,000 shots attempts in 2012-2013 season
- ~1 million optical observations per game.
- 600+ million optical observations in our dataset.

Shot Log

We used this data to compile a shot log.
For every shot attempted, we had data on:

- Shot Conditions: Shot location, Shot Type (13 mutually exclusive categories), etc.
- Game Conditions: Score differential, time remaining, quarter, etc.
- Defensive Conditions: Locations of all defenders, defender angle, etc.

From these datapoints, we can further extrapolate other variables, such as measures of defensive pressure.

Table of Contents

Measuring Shot Difficulty

To measure shot difficulty, we came up with a model that predicts the probability that player i makes shot s:

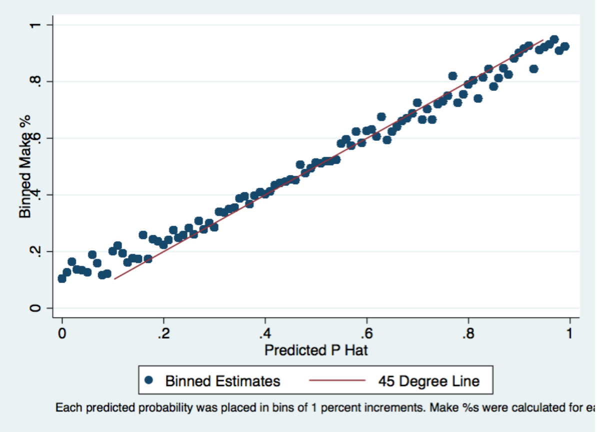
$$\hat{P}_{is} = \alpha + \beta * (\mathit{Game\ Condition\ Controls}_{is}) + \gamma * (\mathit{Shot\ Controls}_{is}) \\ + \delta * (\mathit{Defensive\ Controls}_{is}) + \theta * (\mathit{Player\ Fixed\ Effects}_{i})$$

$\hat{P}$ gives us a single number that is easily interpretable and encapsulates the overall difficulty of a given shot.

Testing the Model

To test the accuracy of the model, we ran it on a randomized training set. We then applied the model on the remaining data to predict the $\hat{P}$'s.

The figure below compares the predicted and actual make percentages:



Measuring Heat

How do we define a player's heat? We use two distinct methods:

1. **Simple Heat$_n$** measures a player's shooting percentage over his past n shots.
   - *Example:* If a player made 3 out of his past 4 shots, then Simple $\text{Heat}_4 = \frac{3}{4} = 0.75$
2. **Complex Heat$_n$** measures the difference between a player's actual and expected shooting percentage over the past n shots, based on the $\hat{P}$ values of those shots.
   - *Example:* If a player made 3 out of his past 4 shots, and those shots had $\hat{P}$ of 0.1, 0.5, 0.8, and 0.6 then Complex $\text{Heat}_4 = \frac{3}{4} - \left(\frac{0.1+0.5+0.8+0.6}{4}\right) = 0.75 - 0.5 = 0.25$

Complex Heat is the Better Measure

Though perhaps less intuitive, we argue that Complex Heat is a better measure of heat.

- It measures *true* overperformance - a player who goes 2 for 3 from the 3-point line is "hotter" than the player who makes 2 out of 3 layups.
- It controls for serial correlation betweens shots:

$$\textit{Complex Heat} = \textit{Actual Pct.} - \textit{Expected Pct.} = \textit{Simple Heat} - \underbrace{\textit{Expected Pct.}}_{\text{difficulty of past shots}}$$

Example: A player is covered by a short defender.

Table of Contents

Do Players Believe in the Hot Hand?

If players believe in the Hot Hand, they will adjust their play accordingly.

- Do hot players take shots from further away?
- Do defenders cover hot players more closely?
- Are hot players more likely to take their team's next shot?
- Does overall shot difficulty increase with heat?

If players do believe in the Hot Hand, we would expect that the answers to these questions are yes.

Empirical Strategy

$$\textit{Shot Distance}_{is} = \alpha + \beta * (\textit{Heat}_{is}) + \gamma * (\textit{Controls}_{is}) + \theta * (\textit{Player Fixed Effects}_i)$$

$$\textit{Defender Distance}_{is} = \alpha + \beta * (\textit{Heat}_{is}) + \gamma * (\textit{Controls}_{is}) + \theta * (\textit{Player Fixed Effects}_i)$$

$$P(\textit{Same}_{is}) = \Phi(\alpha + \beta * (\textit{Heat}_{is}) + \gamma * (\textit{Controls}_{is}) + \theta * (\textit{Player Fixed Effects}_i))$$

$$\hat{P}_{is} = \alpha + \beta * (\textit{Heat}_{is})$$

Results

The results suggest that players do believe in the Hot Hand, and alter their play to reflect these beliefs.

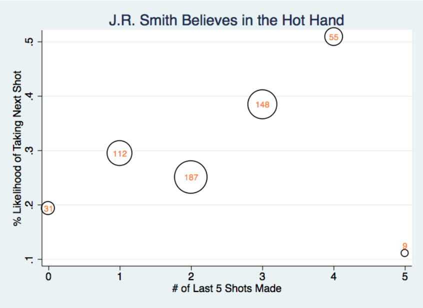
VARIABLES	(1) Distance	(2) Distance	(3) Defender Distance	(4) P(Same)	(5) P(Same)
Simple Heat (4)	2.385***			0.0578***	
	(0.186)			(0.00871)	
Complex Heat (4)		2.240***	-0.151***		0.0637***
		(0.185)	(0.0397)		(0.00909)
Constant	7.249***	8.387***	4.123***		
	(0.289)	(0.266)	(0.135)		
Observations	45,123	45,047	45,047	45,115	45,039
R^2	0.290	0.289	0.161		

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Response	Raw Effect Size	% Effect Size
Shot Distance	7.0 in	4.6%
Defender Distance	0.5 in	1.0%
P(Take Next Shot)	1.4%	7.2%

These effect sizes reflect a player making an additional shot of his last four attempts.

Results - Individual Player Evidence



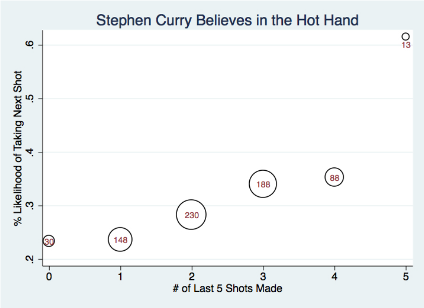


Table of Contents

Testing for the Hot Hand

We have established that heat affects shot selection. This proves that critical Tversky assumption of "random shot selection" does not hold.

Next, we want to see if after controlling for shot difficulty, a Hot Hand effect does emerge.

Empirical Strategy

Baseline specification (no difficulty controls):

$$P(Make_{is}) = \alpha + \beta * (Heat_{is}) + \theta * (Player\ Fixed\ Effects_i)$$

Difficulty-controlled specification:

$$P(Make_{is}) = \alpha + \beta * (Heat_{is}) + \gamma * \hat{P}_{is}$$

Results - Simple Heat

When we do not control for difficulty, we do not find evidence of the Hot Hand.

VARIABLES	(1) P(Make)	(2) P(Make)
Fitted values	1.026***	1.031***
	(0.0143)	(0.0159)
Simple Heat (1)	-0.00570	
	(0.00369)	
Simple Heat (3)		-0.0159**
		(0.00794)
Constant	-0.0204***	-0.0336***
	(0.00590)	(0.00777)
Observations	64,432	51,542
R^2	0.157	0.154

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: $\hat{P}$ represents the predicted probability that a shot goes in. Both specifications include Player Fixed Effects.

Simple Heat does not control for within-player variation.

Results - Complex Heat

When we control for difficulty and use Complex Heat, a positive and significant Hot Hand effect emerges.

VARIABLES	(1) P(Make)	(2) P(Make)	(3) P(Make)
$\hat{P}$	1.015***	1.018***	1.019***
	(0.00902)	(0.00972)	(0.0106)
Complex Heat (3)	0.0239***		
	(0.00757)		
Complex Heat (4)		0.0254***	
		(0.00930)	
Complex Heat (5)			0.0317***
			(0.0112)
Constant	-0.00901*	-0.0108**	-0.0117**
	(0.00464)	(0.00497)	(0.00537)
Observations	51,572	45,123	38,674
R^2	0.148	0.147	0.146

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: $\hat{P}$ is the predicted probability that a given shot goes in - a proxy for overall difficulty. Simple/Complex Heat (n) represents simple/Complex Heat over the past n shots.

Table of Contents

Conclusions

1. Players believe in the Hot Hand and act accordingly.
    - They take shots from further away
    - Defenders cover them closer
    - They are more likely to take their team's next shot
    - They take more difficult shots (controlling for the difficulty of their prior shots)
2. Players who are hot (outperforming given the difficulty of the shots they have taken) are more likely to make their next shot, controlling for their next shot's difficulty.
    - Said differently, the Hot Hand emerges when we control for both past and current shot difficulty