HOT SHOTS: AN ANALYSIS OF THE 'HOT HAND' IN NBA FIELD GOAL AND
FREE THROW SHOOTING

Robert M. Lantis
Erik T. Nesson

Hot Shots: An Analysis of the 'Hot Hand' in NBA Field Goal and Free Throw Shooting
Robert M. Lantis and Erik T. Nesson
NBER Working Paper No. 26510
November 2019, Revised March 2020
JEL No. D91

ABSTRACT

We investigate the hot hand hypothesis using detailed data on free throws and field goal attempts for the 2004-2005 through 2015-2016 NBA regular seasons. Free throws represent a more controlled setting, allowing a closer examination of the potential physiological mechanisms behind success in repeated motions, while field goal attempts represent the setting most observers have in mind when commenting on a player's repeated shooting success. We examine these two settings together, within the same players in the same games, permitting a more comprehensive analysis of the hot hand. We find a small hot hand effect for free throws, which more than doubles for longer streaks of consecutively made free throws. However, if a player makes a field goal, he is no more or less likely to make his next field goal attempt, and longer streaks of consecutively made field goals reduce the probability that a player makes his next field goal attempt. These results are robust to a large number of controls, including controlling for the characteristics of the previous shots.

Robert M. Lantis
Ball State University
2000 West University Ave.
Muncie, IN 47306
rmlantis@bsu.edu

Erik T. Nesson
Department of Economics
Ball State University
Muncie, IN 47306
and NBER
etnesson@bsu.edu

1 Introduction

The question at the center of research into the hot hand concerns the common belief that previous shot success predicts future shot success. This research began with Gilovich et al. (1985), who state their research question in the opening sentences of their paper as follows, "In describing an outstanding performance by a basketball player, reporters and spectators commonly use expressions such as 'Larry Bird has the hot hand' or 'Andrew Toney is a streak shooter.' These phrases express a belief that the performance of a player during a particular period is significantly better than expected on the basis of the player's overall record." Gilovich et al. (1985) find no existence of the hot hand in professional free throw and field goal shooting, or in collegiate field goal shooting in a controlled experiment, and thus, they conclude that the widespread belief in the hot hand is the result of a cognitive bias. Gilovich et al. (1985) spawned a large interest in the hot hand across a variety of settings and sports.

Recent research into the hot hand often focuses on controlled settings such as shooting experiments, the NBA 3-point contest, and in-game free-throw shooting (e.g. Arkes, 2010; Aharoni and Sarig, 2012; Miller and Sanjurjo, 2014, 2015, 2018; Yaari and Eisenmann, 2011). In a series of papers, Joshua Miller and Adam Sanjurjo find evidence of shooting streaks in controlled and semi-controlled settings, including controlled shooting experiments and the NBA 3-Point Contest (Miller and Sanjurjo, 2014, 2015, 2018). Importantly, Miller and Sanjurjo (2018) demonstrate that the type of conditional probability analysis commonly used in controlled studies of the hot hand is biased against finding a hot hand effect. Correcting this bias produces strong evidence of a hot hand in controlled shooting experiments, including in the data originally used by Gilovich et al. (1985). Results from other movement patterns, such as pitching horseshoes, tennis, and bowling, also suggest that success in repeated movement patterns can increase the probability of success in future movements (Smith, 2003; Klaassen and Magnus, 2001; Dorsey-Palmateer and Smith, 2004).

However, the original question posed by Gilovich et al. (1985) is not whether a physiological mechanism exists for success in repeated movement patterns, but rather whether the widespread belief in the hot hand is correct. As referenced in the quote above from Gilovich et al. (1985), the most relevant setting for testing these beliefs is the run of play. Gilovich et al. (1985) conduct surveys suggesting that players and coaches believe the hand exists in the run of play, and general

basketball fans also exhibit gambling habits that reflect beliefs in the hot hand (Brown and Sauer, 1993). It is unclear whether the results from controlled shooting situations will translate to game situations, as players' shots in games can be many minutes apart, may be from very different locations, and possibly in very different game situations. An additional challenge in testing for the hot hand effect in game situations are endogenous responses by both offenses and defenses in response to made shots. Previous papers find that both offenses and defenses react to previously made shots (Aharoni and Sarig, 2012; Rao, 2009; Csapo and Raab, 2014; Csapo et al., 2015; Attali, 2013; Bocskocsky et al., 2014), although the few papers that do examine the hot hand in field goals have been mixed in their findings, perhaps because of the difficulty in accounting for the endogenous responses to made field goals (e.g. Rao, 2009; Bühren and Krabel, 2015; Csapo and Raab, 2014; Bocskocsky et al., 2014).

In this paper we analyze the hot hand in the NBA for both free throws, a more controlled shooting environment, and field goals in the run-of-play, which while not a controlled environment, are more relevant to evaluating spectators' beliefs in the hot hand. We compile detailed data regarding the universe of free throws and field goal attempts for the 2004-2005 through 2015-2016 NBA regular seasons, representing over 500,000 free throws and two million field goal attempts, and combine them with a large set of player, game and shot characteristics. We introduce into the analysis of the hot hand in basketball new methods of controlling for players' underlying shooting ability, new methods of accounting for previous shot difficulty, and we examine multiple measures of previous shot success, including whether a player made his previous shots and also whether the player made streaks of shots. Finally, we use two seasons for which we have additional shot and defensive player characteristics available to examine whether additional shot and defense characteristics affect our results.

We find a small hot hand effect in free throws, similar to Arkes (2010), although smaller in magnitude. If a player makes his previous free throw, he is about two percentage points more likely to make his next free throw, and the effect is larger for the second and third free throws in a set. These effects grow further for players who take more free throws in a game, and are even more pronounced when players make streaks of free throws. If a player makes four free throws in a row, he is 4.5 percentage points more likely to make his next free throw.

In the run of play, we find that both offenses and defenses respond to field goals as if the hot

hand effect exists, but we do not find any evidence of a hot hand effect. We find a negative and statistically significant relationship between a player making his last field goal attempt and the probability that he makes his next field goal attempt, although once we account for previous shot characteristics, making a field goal does not change the probability that a player makes his next field goal attempt, and we can bound the coefficient to extremely small magnitudes. However, we do find statistically significant negative coefficients when examining longer streaks of consecutively made field goals, even when accounting for previous shot characteristics. When a player makes three field goals in a row, the probability he make his next field goal decreases by 0.6 percentage points. Our results are not affected by including the additional controls in the 2013-2014 and 2014-2015 seasons; if anything, these extra controls increase the magnitude of the negative effect of making a previous field goal on the probability of making the next field goal.

Our paper makes a number of contributions to the literature on the hot hand. First, we examine roughly 516,000 free throw attempts and 2,074,000 field goal attempts taken by the same players within the same games. This allows a more direct comparison of the existence and magnitude of the hot hand effect in more controlled and less controlled settings. We also bring a larger set of controls than used in previous literature, and we demonstrate the importance of controls, even in fairly controlled settings such as free throw shooting. We further advance the literature by including controls for both current and previous shot characteristics. Finally, we examine the importance of potential omitted variable bias by adapting a methodology developed in Oster (2019) to create bounds for our coefficients of interest under different assumptions regarding omitted variable bias.

Second, we also introduce a new methodology to control for differences in players' shooting ability, adapted from Chen et al. (2016) and Green and Zwiebel (2017), but to our knowledge not previously used in studying the hot hand in basketball. Previous work has included player fixed effects to account for shooting ability and other time-invariant player characteristics (e.g. Arkes, 2010; Bocskocsky et al., 2014), although including fixed effects in a lagged-dependent variable model will lead to biased coefficients (Nickell, 1981; Nerlove, 1967, 1971). Instead, we include two measures of player's shooting percentage which do not include the current game, following a strategy employed by Chen et al. (2016) in testing the gambler's fallacy in three settings, refugee asylum court decisions, reviews of loan applications, and calls made by Major League Baseball umpires, and by Green and Zwiebel (2017) to determine whether the hot hand that exists in at

bats in baseball.

Third, we introduce new methods of testing for streaks of shots. We first test for the hot hand using a straightforward regression framework where we regress an indicator for a player making his current shot on an indicator for making his previous shot, similar to Arkes (2010)'s examination of free throws and some specifications in Chen et al. (2016). We also introduce a new test for success in streaks of previous shots. Here we examine the interactions of indicator variables for success in each of the previous four shots to estimate the effect of different lengths of streaks of success. Our strategy again builds upon Chen et al. (2016), and we are, to our knowledge, the first to apply it to the analysis of the hot hand. Although other papers have analyzed shooting success over more than just the previous shot in a regression framework (e.g. Bocskocsky et al., 2014; Green and Zwiebel, 2017), our strategy allows us to examine different streak lengths separately as opposed to analyzing the average or shooting percentage or batting average over a period of shots or at-bats.

Although previous studies have used a multivariate framework to approach the hot hand in basketball, as we outline above, our approach offers advances over previous work in this area. Some previous work has focused only on a few specific players or game situations. For example, Rao (2009) find no evidence of a hot hand when employing a random effects probit model for a set of players on the Los Angeles Lakers, while Bühren and Krabel (2015) use OLS and tobit models to estimate the impact of making a shot which forced overtime on the field goal shooting of that player in overtime. Arkes (2010) estimates fixed effects logit models for free throws taken during the 2005-2006 NBA regular season, and he finds that if a player makes his previous free throw in a set, he is about three percentage points more likely to make his next free throw. Green and Zwiebel (2017) examine at-bats in baseball using a methodology to control for players' batting averages which is similar to our approach. They argue that the allocation of offensive and defensive resources is less fluid in baseball than in basketball, thus endogenous responses are less of a concern. Green and Zwiebel (2017) find a hot hand effect in at bats for various measures of previous batting success.

Finally Bocskocsky et al. (2014) control for other factors which may affect shot success, and they find mixed evidence of whether the hot hand exists, depending on which of their measures of previous success is used. Bocskocsky et al. (2014) develop two measures of previous shot success, "simple heat" and "complex heat." Simple heat is the player's shooting percentage over the previous four shots within each game. Complex heat adjusts for previous shot difficulty. Bocskocsky et al.

(2014) first predict the difficulty of each shot. They then adjust actual shooting percentage over the previous four shots with the predicted shooting percentage to construct complex heat. In a similar manner, Bocskocsky et al. (2014) include a predicted shot difficulty measure to control for current shot difficulty. However, including a predicted value in a regression in the manner of Bocskocsky et al. (2014), instead of directly controlling for the characteristics of shots, likely generates biased coefficients.[1]

The rest of this paper is organized as follows. Section 2 overviews our play-by-play shot data and statistical methods, Section 3 discusses our results, and Section 4 concludes.

2   Data and Methods

2.1   Data

We use three data sources for our analysis. Our main data source is play by play data for the 2004-2005 through 2015-2016 NBA regular seasons from `http://www.bigdataball.com`.[2] This data set provides the date and teams for every game and detailed information for every event in each game, including field goal attempts, free throws, rebounds, fouls, timeouts, etc., as well as the players who were on the court at the time of each event. For field goal attempts, this data contains the time of the field goal attempt, whether the field goal attempt went in, the shot type, and the shot location in x-y coordinates, which we use to calculate the shot distance from the basket. We use the `bbr` package in R, available here: `https://github.com/mbjoseph/bbr`, to obtain and merge in player characteristics for the offensive player, including position, height, and weight, and age, from `http://www.basketball-reference.com`. Additionally, we add game characteristics, also from `http://www.basketball-reference.com`, including which team was the home team and away team and the crowd size.

In addition to this primary data set, we take advantage of additional shot characteristics and defensive player characteristics which are available for the 2013-2014 and 2014-2015 seasons. We

---

[1]We reproduced many of the results in Bocskocsky et al. (2014) using our data, including their finding of a hot hand. Our reproduction suggests that their finding of a hot hand is driven by the inclusion of predicted shot difficulty instead of directly controlling for shot characteristics. When predicted shot difficulty is replaced with controls for shot difficulty, we find that previous shot success is negatively correlated with current shooting success. Details regarding our reproduction are available upon request.

[2]See `https://www.bigdataball.com/nba-playbyplay-stats-multipleseasons/` for more information on this data set.

start with our primary data set and merge in these additional field goal attempt characteristics, which we collected from `http://www.nbasavant.com`.[3] Representing the universe of field goal attempts taken during these seasons, this data contains the number of dribbles taken before shooting and number of seconds on the ball before shooting. Importantly, the data also contain information on the defense, including the name of the nearest defender and distance from the nearest defender to the shooting player, which allows for additional analysis of defensive responses to made field goals.

2.2   Methods

2.2.1   The Determinants of Made Shots

To examine the effects of previous shot success on current shot success, we estimate variations on the following regression model:

$$s_{p,i,g,t} = \alpha_0 + \alpha_s s_{p,i-1,g} + \boldsymbol{v_{p,i,g}}\boldsymbol{\alpha_v} + \bar{s}_{p,-g} + \epsilon_{p,i,g}, \tag{1}$$

where the dependent variable, $s_{p,i,g}$, is an indicator variable equal to 1 if player p makes shot i in game g in season t, and the independent variable of interest, $s_{p,i-1,g}$ is an indicator variable for whether the same player made the previous shot in the same game. If α_s is positive and statistically significant, it provides evidence for the hot hand effect, and if α_s is negative and statistically significant, it not only provides evidence that a hot hand does not exist, but that making the previous shot actually decreases the probability of making the current shot.

We begin these analyses by estimating naive models using the lagged shot indicator as the only independent variable. We add in additional controls in $\boldsymbol{v_{p,i,g}}$. For both free throws and field goals, we include controls for the game situation, including indicators for month, day of the week, team by season, and opponent by season, an indicator for the player being on the home team interacted with game attendance, and the period interacted with the seconds remaining in that period. We also include controls for player characteristics, including a quadratic function of the player's age, a quadratic function of the player's height and weight, an interaction between height and weight, and

---

[3]NBA Savant collects the additional characteristics from ESPN's Shot Tracker and `stats.nba.com`. See `http://www.nbasavant.com/shot_search.php` for more information about the data from NBA Savant.

indicators for player position. We additionally control for free throw or field goal characteristics. For free throws, we include indicators for the second and third free throws in a set, whether the free throw resulted from a personal, shooting or other foul (with flagrant foul as the omitted category), whether there was a substitution before the free throw, and whether there was a timeout before the free throw. We also include the point gap before the current set of free throws interacted with indicators for the home team and with the period.

For field goals, we control for the point gap before the current shot interacted with the home team and with the period, the number of points the opposing team scored on their last possession, the play length in seconds, indicators for whether the player took his team's previous shot, whether the player was fouled on his previous shot, the time between the player's shots, the distance between the player's previous shot and current shot, shot distance (and its square), indicators for and the interaction between shot types and shot distance, an interaction between shot type and season, and indicators for and the interaction between the angle of the shot to the rim and whether it was on the left side of the court. Finally, we analyze whether the difficulty of previous field goals affects the estimation of the hot hand, following Bocskocsky et al. (2014). The intuition is that making a more difficult shot, like a long-range jump shot, is a stronger indication of shooting success than making an easier shot, like a layup. To this end, we additionally estimate specifications which include lagged time-varying controls in addition to our contemporaneous controls.[4]

Finally, it is important that we control for a player's underlying probability of making a shot and Simpson's Paradox (Wardrop, 1995). To this end, we adapt the methodology employed by Chen et al. (2016) and Green and Zwiebel (2017) and control for the player's "leave-out" field goal or free throw percentage, denoted by $\bar{s}_{p,-g}$. We include two measures, the player's field goal or free throw percentage in the relevant season, excluding the current game, and a four week moving average of the player's field goal or free throw percentage, including two weeks on either side of the current game, again excluding the current game. One other strategy for controlling for player ability could be to replace the observable player characteristics with player fixed effects, which would control for all time-invariant player characteristics. Including fixed effects with a lagged dependent variable

---

[4]The lagged game and shot controls already correspond to a player's previous shot. We do not lag shot measures which incorporate the second lagged shot, as this would limit our sample to only including third and higher shots in a game. Thus, we do not include lagged indicators for whether the player was fouled on his previous shot, took the team's previous shot, the lagged time between shots and the lagged distance between shots.

would bias our coefficients downward (Nickell, 1981; Arellano and Bond, 1991; Nerlove, 1967, 1971). We show in Appendix Tables B1 and B2 that including fixed effects in our regressions instead of the leave-out field goal or free throw percentages does bias our results significantly downward. In all shooting models, we cluster our standard errors at the player by season level.

2.2.2 Changes in Behavior Following Made Field Goals

When examining field goals, we additionally examine whether making field goals induces changes in offensive and defensive behaviors. We estimate fixed effects panel data regression models as follows,

$$y_{p,i,g,t} = \beta_0 + \beta_s s_{p,i-1,g} + \boldsymbol{w_{p,i-1,g}}\boldsymbol{\beta_w} + \boldsymbol{x_{p,i,g}}\boldsymbol{\beta_x} + \psi_{pt} + \epsilon_{p,i,g}, \tag{2}$$

where $y_{p,i,g}$ is a shot-related outcome for player p's shot i in game g in season t. We examine shot-related outcomes for both offensive and defensive players. We create three variables measuring the confidence of offensive players: an indicator for whether the same player who took the previous shot takes the next shot, the time before taking the next shot, and the shot distance in feet. Relatedly, we measure the distance in feet between a player's shots. We also create an indicator variable for whether the player has an assist in the next play to measure whether, if players are guarded more closely by the defense, they respond by passing to increasingly open teammates. We create two variables measuring possible defensive adjustments: whether the defending team substitutes a player in the next two minutes and whether the defending team calls a timeout in the next two minutes. We extend this analysis to two seasons where we can measure two additional defensive adjustments, assigning a new defender to a player on his next shot and the defender distance from the shooter in Section 3.5.

The independent variable of interest is again $s_{p,i-1,g}$, the indicator for player p making shot $i-1$ in the same game. The vector $\boldsymbol{w_{p,i-1,g}}$ includes controls for the game characteristics and shot characteristics of the lagged shot as well.[5] In all specifications, $\psi_{p,t}$ represents player by season fixed effects for the player who took shot $i-1$, which account for time-invariant observable and unobservable differences across players and seasons, allowing us to examine changes *within* players

---

[5]The game and shot controls already correspond to a player's previous shot, so we do not include lagged shot measures which incorporate the second lagged shot as this would limit our sample to only including third and higher shots in a game. Thus, we do not include lagged indicators for whether the player was fouled on his previous shot, took the team's previous shot, the time between shots and the distance between shots.

and seasons *across* shots. Including fixed effects will not bias our coefficients here, as we do not have a lagged dependent variable. In all offensive and defensive response models, we cluster our standard errors at the shooting player by season level.

2.2.3 Extensions to Baseline Specifications

We expand our definition of previous shooting success to examine streaks of shooting success for both free throws and field goal attempts, as success from previous shots further back in a shooter's history may impact the probability of making the current shot as well as the responses to made field goals. To this end, we modify equation 1 as follows,

$$s_{p,i,g} = \gamma_0 + \sum_{j=-4}^{-1} \gamma_{sj} s_{p,i-j,g} + \sum_{j=-4}^{-1} \sum_{\substack{k=-4, \\ k>j}}^{-1} \gamma_{sjk} s_{p,i-j,g} \times s_{p,i-k,g} + \boldsymbol{v_{p,i,g}} \boldsymbol{\gamma_v} + \bar{s}_{p,-g} + \epsilon_{p,i,g}. \quad (3)$$

Here, our independent variables of interest are the first through fourth lagged shot indicators and their interactions. We calculate linear combinations of the lagged shot coefficients to measure the additional impact that comes from making several shots consecutively. Thus, we can separately estimate the effect of one shot, two shot, three shot and four shot streaks on the probability of making the next shot.[6] We also modify equation 2 in a similar manner to examine how offenses and defenses respond to streaks of made field goals. Our strategy here follows from Chen et al. (2016)'s analysis of the gambler's fallacy, and to our knowledge, it has not been adopted to examine the hot hand hypothesis. We show these results in Section 3.4.

2.3 Summary Statistics

The free throw data covers 697,238 free throws during the 2004-2005 through 2015-2016 NBA regular seasons. To be in our analysis, free throws cannot be a player's first free throw of a game, which removes 180,536 free throws. Thus, our final sample for the free throw analysis is 516,699 free throws. Table 1 shows summary statistics for our free throw sample. Some of our results are previewed by the difference between free throw success following a made and a missed free throw. The percent of free throws made following a made free throw is approximately 78 percent but only

---

[6]For example, the effect of making two shots in a row is the sum of the coefficient on the indicator for making the previous shot, the twice lagged shot, and the interaction of the previous and twice lagged shot.

72 percent following a missed free throw. Many of the other observable characteristics, however, are similar between a previously made and missed free throw.

The field goal data covers 2,375,873 field goals during the 2004-2005 and 2015-2016 NBA regular seasons. To be in our analysis, field goal attempts cannot be a player's first field goal attempt of the game, which removes 301,372 field goal attempts. Our final sample contains 2,074,501 field goal attempts across all 12 seasons.[7] Table 2 shows summary statistics for our field goal sample. Of our sample of field goal attempts, about 46 percent came after that player made his previous field goal in the same game. Here again, the differences between field goals following a made and a missed field goal preview some of our main results. First, the field goal percentage following a made field goal is about 45 percent, lower than the field goal percentage after a missed field goal of 46 percent.[8] Following a made field goal, players shoot from nearly eight tenths of a foot farther away, take their next field goal about 20 seconds sooner, are 4.5 percentage points more likely to take the next field goal on their team, and the opposing team is about two percentage points more likely to substitute in the next two minutes and six percentage points more likely to call a timeout in the next two minutes.

Table 2 also illustrates how different consecutive shots often are for players. Players have a 22 percent chance of taking consecutive shots, an average of over three and a half minutes elapse between a player's shots, and a player's consecutive shots are nearly 16 feet away from each other. These characteristics again call into question how generalizable the results from shooting experiments and semi-controlled settings like free throws are to the analysis of the hot hand in field goal shooting.

3 Results

3.1 The Hot Hand Hypothesis in Free Throws

We begin by examining free throws. Results from estimating equation 1 for our sample of free throws are shown in Table 3. The first six columns show regressions estimated over all free throws. We add in different levels of controls, culminating in the addition of the leave-out free throw percentages

---

[7]Importantly, field goal attempts are not recorded if there is a shooting foul and the field goal attempt is missed. Field goal attempts are counted if there is a shooting foul and the shot is made.

[8]Since these are averages, the bias identified by Miller and Sanjurjo (2018) may affect these averages.

in Column (5) and lagged controls in Column (6). In Column (7) we exclude the first free throw in any set, keeping all the same controls as in Column (6). Thus, we are only examining free throws immediately following another free throw by the same player to focus on the physiological aspects of repeated motions.

In Columns (1) and (2), making the previous free throw increases the likelihood of making the next free throw by about six percentage points. However, once we control for player characteristics in Column (3) the effect is lowered to 4.0 percentage points. Including the leave-out free throw percentages decreases the size of the effect to 2.1 percentage points, which illustrates the importance of accounting for Simpson's Paradox. Column (6) demonstrates that the hot hand effect is not meaningfully impacted by including lagged free throw controls. In Column (7), the effect size increases slightly from 2.1 to 2.4 percentage points, consistent with the hypothesis of physiological mechanisms leading to changes in the success of repeated motions. In the Appendix we highlight the importance of using a player's leave out free throw percentage as compared to player fixed effects. Table B1 provides the results of using fixed effects and shows significantly lower in magnitude as a result of the bias caused by including fixed effects with a lagged dependent variable as a regressor.

3.2   Offensive and Defensive Adjustments to Field Goals

We next examine potential offensive and defensive adjustments to field goals. Table 4 shows results from these regressions. We examine five outcomes for offensive responses, whether the player that took the previous field goal attempt takes the next field goal attempt, the seconds before the player takes his next field goal attempt, the shot distance of the next field goal attempt, the distance between the current and previous shot, and whether the player who took the previous field goal attempt creates an assist on the next play. We examine two outcomes for defensive responses, whether the defending team substitutes a player in the next two minutes, and whether the defending team calls a timeout in the next two minutes.

Starting with offensive adjustments, if a player makes his previous field goal attempt, he is about five percentage points more likely to take the next field goal attempt on his team, the time before he takes his next field goal attempt decreases by almost 20 seconds, and he takes his next field goal attempt about 1.5 feet further away from the basket, although the distance between shot locations is not affected by a large amount, and he is 0.9 percentage points more likely to make an

assist on the next play. Combining this result with the increased probability of taking the next field goal attempt, a player is more likely to be involved in his team's next scoring attempt if he makes his previous field goal attempt. Defenses also react to made field goals, where they are about two percentage points more likely to make a substitution and five percentage points more likely to call a timeout in the next two minutes.

Below the coefficients and standard errors, we report the means of the dependent variables to help put the coefficient sizes into context. Roughly translating our marginal effects into semi-elasticities, the size of the effects of field goals on the offensive adjustments are substantial. Players are about 24 percent more likely to take the next field goal attempt, decrease the time between their field goal attempts by about nine percent, increase their shooting distance by about twelve percent, and 18 percent more likely to record an assist. However, the distance from their previous shot only increases by 0.3 percent, evidence they are not attempting shots from different locations of where the last shot was made. The size of the defensive adjustments are smaller, although still notable. Teams are 4.4 percent more likely to make a substitution and 23 percent more likely to call a timeout.

3.3   The Hot Hand Hypothesis in Field Goals

Next, we examine the hot hand hypothesis for field goals in Table 5. As with free throws, we begin our analysis of the hot hand with a pared back specification where we only include whether or not the player in question made his previous field goal attempt as the independent variable. Moving from left to right, we add in additional controls to create a more apples-to-apples comparison between field goal attempts. In Columns (2) through (4), we sequentially add in controls for the game situation, player characteristics, and shot characteristics. In Column (5) we add controls for the player's leave-out field goal percentages, and finally in Column (6) we add in lagged shot characteristics to further control for the difficulty of the previous shot.

We find a negative and statistically significant relationship between making the previous field goal attempt and making the next field goal attempt across a majority of the specifications. In Columns (1) through (3), the negative effect of making a shot on the probability of making the next shot grows in magnitude from -0.6 to -1.1 percentage points. Controlling for shot characteristics reduces the effect of a previous field goal on the probability of making the current field goal attempt,

while accounting for the player's leave-out field goal percentages increases the magnitude and statistical significance of the negative correlation between field goal success. Finally, including lagged controls reduces the estimated effect of making a previous field goal on making the next field goal to nearly zero, with an extremely tight confidence interval.

We investigate which lagged controls are the most important variables for decreasing the magnitude and statistical significance of the results between Columns (5) and (6) in Table 5. In Table 6, we start with the estimate in Column (5) of Table 5, and successively add in lagged controls to finally arrive at the estimate in Column (5) of Table 5 in the last column. We find that lagged game controls (period indicators fully interacted with the seconds remaining in the period and the point gap and its interactions with indicators for the home team and the period) and possession in order to determine what best captures the difficulty of the prior shot) and lagged possession controls (points scored by the opposing team on the last possession and the play length) do not markedly affect the results. The lagged shot type influences the estimates the most, providing estimates that are no longer statistically significant and negative, but not statistically different than zero.

3.4   Streaks of Previous Success

Next, we expand our definition of previous shot success to include streaks of free throws and field goals as described in Section 2.2.3. Tables 7, 8, and 9 present estimates for the effect of making the previous shot, the previous two shots, the previous three shots, and finally making the previous four shots.[9] Comparing these estimates will show whether the effect of previous success on the current shot varies with the the length of consecutive previous successes.

Table 7 shows results estimating the effects of streaks in past free throw shooting on current free throw shooting success. Column (1) of Table 7 replicates the results from Column (6) of Table 3, and Columns (2) through (4) show estimates for players who took two, three, and four previous free throws within the same game. We find that when only including success in the previous two free throws that there is no significant difference in our estimate of the effect of making only the

---

[9]Specifically, we present linear combinations of the coefficients which represent the effect of making a specific number of consecutively made shots. For example, the effect of making the previous three shots in Column (3) of Tables 7 and 9 is calculated by $\gamma_{L1} + \gamma_{L2} + \gamma_{L3} + \gamma_{L1,L2} + \gamma_{L1,L3} + \gamma_{L2,L3} + \gamma_{L1,L2,L3}$, where γ_{L1}, γ_{L2}, and γ_{L3} refer to the coefficient of the indicator for making the first second and third lagged shots, respectively, and $\gamma_{L1,L2}$, $\gamma_{L1,L3}$, $\gamma_{L2,L3}$, and $\gamma_{L1,L2,L3}$ represent interactions. Point estimates for any combination of makes and misses in the previous four shots can be calculated, but are not presented.

previous free throw. However, Column (2) shows that the effect of making two consecutive free throws increases the chances of making a current shot by about 3.5 percentage points. There is a stronger effect still if the player has made two or three consecutive free throws in a row, as shown in Column (4). When limiting the analysis to players who took at least five free throws in a game, the estimates for longer streaks of success grow in magnitude, but are not statistically different from each other.[10]

In Table 8 we repeat the analysis for the offensive and defensive adjustments in Table 4. Each panel shows the estimates for different lengths of consecutively made field goals that are included in the regression. For the offensive adjustments of taking the next field goal, time in between field goal attempts, and shot distance, longer streaks of made field goals have increasingly larger effects on these adjustments. Most notably, making the previous field goal increases the distance of the next shot by approximately 1.5 feet while making the previous four field goals increases the distance of the current shot by over three feet. The difference in shot location from the previous shot also grows in magnitude for longer streaks of success, however, the effect of just making the previous shot becomes statistically no different from zero once four lagged shot outcomes are included. The offensive adjustment of recording an assist on the next play is larger for longer streaks of success, however, very similar in magnitude.

Similar to the offensive adjustments, the probability a defense makes a substitution or calls a timeout in the next two minutes increases for longer streaks of made field goals. The effect of making just the previous shot increases the probability the defensive team makes a substitution in the next two minutes by 2.2 percentage points, while the effect of making the previous four field goals increases that probability to 3.4 percentage points. Similarly, the effect of making just the previous field goal increases the probability the defensive team calls a timeout in the next two minutes by five percentage points, while making the previous three field goals increases it to six percentage points.

In Table 9 we repeat the analysis for field goal attempts. Column (1) of Table 9 replicates the estimates from Column (6) in Table 5, and the additional columns show the results from

---

[10]One potential issue is how quickly our sample size falls in Columns (3) and (4) due to many players not having the required number of free throws in a game. The selection of players that may exist into these sub-samples may be a result of defenses' fouling behavior responding to the successes in prior free throw shooting or in offenses' behavior in getting the ball into the hands of players having free throw success. Our detailed game, player and season/moving average free throw shooting percentage controls should mitigate these concerns.

progressively adding in additional lags of previously made field goals and their interactions terms. The effect of making just the previous field goal or previous two field goals on the probability of making the current field goal is extremely small, with a very small 95% confidence interval, no matter how many lagged shots are included. However, we find a statistically significant and negative effect from making three or four consecutive field goals in a row, which both decrease the probability of making the current shot by 0.6 percentage points.[11]

3.5 Additional Shot and Player Characteristics in 2013-2014 and 2014-2015 Seasons

As a further check of the robustness of our estimates we limit our sample to the 2013-2014 and 2014-2015 seasons, for which have additional shot and defensive characteristics. Representing the universe of field goal attempts taken during these seasons, this data additionally contains the number of dribbles taken before shooting and number of seconds on the ball before shooting, as well as the name of the nearest defender and distance from the nearest defender to the shooting player, which allows for additional analysis of defensive responses to made field goals.[12]

We first identify the effect of making a field goal on the offensive and defensive adjustments in Table 4, as well as the defensive adjustments of assigning a new defender and defender distance from the shooter. Table 10 shows results from these regressions. We find similar estimates for all offensive and defensive adjustments as in Table 4, except for the change in shot location now becomes statistically no different from zero. The effect of making the previous field goal increases the probability a new defender is assigned to that player by 0.7 percentage points, while the distance of the nearest defender decreases by 0.07 feet. These translate to changes that are about one and 1.8 percent of the dependent variable's mean, respectively.

We also repeat the analysis of Table 5 with the sub sample of data that includes additional shot and defensive characteristics. These results are presented in Table 11. Once again we present estimates of the effect of making the previous field goal on the probability of making the current

---

[11]In Appendix Tables A3 and A4 we re-estimate Tables 7 and 9 without including the lagged shot characteristics. While our results for free throws are very similar, we do find negative and statistically significant results for field goals. Specifically, the small negative effect of making a field goal on the probability of making the next field goal is still statistically no different than zero, however, grows in magnitude for longer streaks. Here, making the previous 3 or 4 field goals in a row decreases the probability of making the current field goal by 0.9 percentage points.

[12]Summary statistics for this subset of our primary data set can be found in Appendix Table A5.

field goal, including growing levels of controls. We add opponent controls in Column (4), which include a quadratic function for the defensive player's age, indicators for the defensive player's position and for a new defender covering the shooting player, a quadratic function of defender distance from the shooting player, and the height and weight difference between the shooting and defensive player and their interactions with defender distance. We also replace the time-invariant opponent controls with defender fixed effects in Column (7). As in Table 5, we find a persistent and statistically significant negative relationship between making a previous field goal and the probability of making the next field goal across the majority of the models, although the majority of the coefficients are greater in magnitude than in Table 5. Additionally, as in Table 5, the most important controls are shot controls and controls for players' shooting ability as well as accounting for the characteristics of the previous shot. Neither the inclusion of opponent controls in Column (4) nor defender fixed effects in Column (7) change the magnitude or statistical significance of the estimate. The final estimate, including lagged shot controls, is now negative and statistically significant at the 10 percent confidence level.

We also reproduce the models in Table 9 with the sub sample, including the additional shot and defender controls from Column (7) in Table 11. We show these results in Table 12. Here again, we find results that are similar to Table 11, although in general slightly larger in magnitude. However, the standard errors in these models are in general twice as large as those in Table 9, and this reduces the statistical significance of many of the coefficients. We do find that making the previous four field goals has a large, negative impact on the probability of making the current field goal, decreasing the probability of making the current shot by 1.1 percentage points (p¡.1).

3.6   Robustness Checks and Heterogeneity

To ensure the robustness of our estimates of the hot hand and examine any further heterogeneity in our findings, we examine a large number of robustness checks for both our free throw and field goal samples. For free throws, we impose additional restrictions on the four specifications found in Columns (5) through (7) of Table 3. We estimate these specifications separately for each of the following restrictions: home team only, away team only, excluding free throws from overtime, excluding free throws within the last two minutes of the game, excluding free throws with substitutions made in between them, using only free throws from shooting fouls, and using only free

17

throws from personal fouls. We show these results in Appendix Table A1. Across the different sub samples, we find similar coefficient sizes and statistical significance levels as in the baseline results in Table 3. In general, making the previous free throw increases a player's probability of making his next free throw by about two percentage points, and the effects are larger for second and third free throws in a set.

Next, we impose several additional restrictions on the field goal sample. We estimate equations 1 and 2 separately for each of the following restrictions; home team only, away team only, shots within the same half, shots within the same quarter, no overtime shots, no shots within five seconds of each other, no shots within the last two minutes of the fourth quarter of the game, and shots only within two minutes of each other. We show these regression results in Appendix Table A2 and briefly describe them here. The effect of making the previous shot on the likelihood of that player taking their team's next shot and the following shot's distance from the rim are similar across the different sub samples, each with a few exceptions. Excluding shots within five seconds of each other increases the effect of making the previous shot on the probability of taking the next shot by about eight percentage points, as many shots within five seconds of each other are rebounds. We find the distance of a player's shot increases to an even greater degree under the restrictions of only using shots within the same half, only shots within the same quarter and only shots within two minutes of each other. Defenses also consistently adjust for made shots across the different sub samples. The effect of prior field goal success on the probability of making the current field goal is relatively constant across all sub samples, and in most sub-samples it is not statistically different from zero and of small magnitude. However, the effect becomes negative and statistically significant when excluding shots taken over two minutes from each other, excluding shots taken within five seconds of each other, or examining shots within the same quarter.

To examine omitted variable bias, we follow a method outlined in Oster (2019) to create a bounding set for the true effect of making the previous shot on current shot success for both free throws and field goals after taking into account potential omitted variable bias. The methodology uses three pieces of information to create a bounding set for the true value of a coefficient under potential omitted variable bias: (1) coefficients on the independent variable of interest with and with other controls; (2) the R^2 values from these two regressions; and (3) a value of the proportional

selection between the independent variable of interest and the omitted variable, called δ.[13]

For free throws, we find that the bounded set for the effect of making the previous free throw on the probability of making the current free throw using the regression found in Column (6) of Table 3 is $[0.007; 0.021]$, which suggests that unobservable variables may be biasing our estimated hot hand effect for free throws downward. For field goals, the bounded set for the effect of making the previous field goal attempt on the current probability of making the next field goal attempt using the regression found in Column (8) of Table 5 is $[-0.0005; 0.003]$.[14] Therefore, we can say that even if we are missing some important unobservable variables the true effect is at the most slightly positive and very close to zero. Furthermore, we find that the estimate would still be negative as long as the unobservables were 17% as important as the observables we included in the regression. The extensive set of controls we are able to include in our regression alleviate concerns that there are unobservables which would be able to explain 17% of what our observables can. Furthermore, the bounded set for our subset of data which includes additional controls for current and previous shot difficulty is $[-0.0029; 0.0008]$, and the estimate would remain negative as long as the unobservables are less than 79% as important as this more extensive set of controls. Thus, we believe that omitted variable bias is unlikely to explain why our results differ from those in controlled shooting experiments.

We additionally re-estimate our results in Tables 3, 5, 7 and 9 using player fixed effects instead of controlling for the leave-out free throw or field goal percentage. These results can be found in Appendix B. In general, these results show the size of the negative bias that exists when including a lagged dependent variable as a control with the use of fixed effects.

One further potential issue is whether our analyses have enough statistical power to detect a hot hand effect under different potential data generating processes (DGPs). To further examine this issue, we conduct simulations described in Appendix C. We adapt DGPs from Wetzels et al. (2016) and Stone (2012), and the results of these exercises suggest that our models have sufficient power to detect hot hand effects, should they exist in the true data generating process. Finally, in results available upon request, we demonstrate that our results for models with a dichotomous

---

[13]More formally, Oster (2019) considers a model of the form $Y = \beta X + W_1 + W_2$, where W_1 is a linear combination of a vector of observed controls and W_2 is a is a linear combination of a vector of unobserved controls. Then δ is defined as $\frac{\sigma_{2X}}{\sigma_{22}}$, where σ_{2X} is the covariance between W_2 and X and σ_{22} is the variance of W_2.

[14]For both free throws and field goal attempts, we use the method for the selection of R_{Max} suggested by Oster (2019) which uses the R^2 found from the controlled model multiplied by 1.3.

dependent variable are robust to a probit specification (Spector and Mazzeo, 1980).

4   Conclusion

Recent hot hand research has focused on experimental settings due to their ability to remove external factors.[15] However, the original question posed by Gilovich et al. (1985) regards observers' perceptions of shooting success *in game situations*. Thus, although controlled experiments can uncover whether a physiological mechanism for repeated shot success exists, external validity concerns mean that controlled experiments cannot directly speak to players' performance in game situations. We use detailed data on the universe of free throw shots and field goal attempts during the 2004-2005 and 2015-2016 NBA regular seasons to investigate the hot hand hypothesis. Free throws are a more controlled shooting environment, allowing a closer examination of the possible physiological mechanisms at play in repeated shot motions. But field goals are the setting that observers most often refer to when describing a player's shooting performance as "hot" or "streaky," although changes in shot characteristics and endogenous responses by the offense and defense to made shots make it more difficult to examine the hot hand in field goals. We examine both free throws and field goal attempts together within the same players in the same games to more directly compare the results of hot hand analyses in each setting. We examine a wide variety of controls for game, team, player and shot characteristics to reduce the possibility of omitted variable bias, and we additionally use the methodology devised by Oster (2019) to bound the effect that omitted variable bias may have on our results.

We find a small, persistent hot hand effect in free throws. If a player makes his previous free throw, he is about two percentage points more likely to make his next free throw, and the effect of making the previous free throw grows to 2.4 percentage points for free throws within a set. Our free throw results are roughly in line with Arkes (2010), who finds that making the previous free free throw in a set increases the probability of making the next free throw in the set by about three percentage points. However, the free throw hot hand effect grows for streaks of free throws. Making three or four free throws in a row increases the probability a player makes their current shot by about 4.5 percentage points. Even these large hot hand effects are much smaller than effects from

---

[15]See Miller and Sanjurjo (2014) for more on the argument that controlled shooting experiments are preferred to in-game data.

experimental studies. For example, Miller and Sanjurjo (2018) reevaluate the controlled shooting experiment in Gilovich et al. (1985) and find that a player who has made three shots in a row is 13 percentage points more likely to make his or her next shot than if the player missed three shots in a row.

In stark contrast to free throws, we find no relationship between making a previous field goal and the probability that a player makes his next field goal in our full sample, although when we introduce additional controls for a subset of two seasons we find a negative association between a player making a field goal and making his next field goal. Additionally, we show that introducing controls for the game situation and difficulty of the current and previous shots does not produce a hot hand effect, as previous research suggests (Bocskocsky et al., 2014). Most surprisingly, we find that making consecutive field goals reduces the probability that a player makes his next field goal. Specifically, we estimate that if a player makes three field goals in a row, he is 0.6 percentage points less likely to make his next field goal attempt. Both offenses and defenses respond to made field goals as if the hot hand effect exists. Following a made shot, players are more likely to make an assist on their team's next made shot, take their team's next shot, and take that shot quicker and from a further distance. Defenses are more likely to take a time out, make a substitution, assign a new defender to that player, and defend that player closer. Put together, our results reinforce the suggestion that, while the physiological mechanisms behind repeated shot success found in experimental settings do exist in game situations, they are easily overwhelmed by other factors.

We believe the largest concern in our results is controlling for endogenous responses and the omitted variable bias that may exist from insufficient controls. While omitted variable bias is always a concern in applied microeconomics, we take a number of steps to minimize potential omitted variable bias. First, while our results show the importance of controls in both free throw and field goal settings, many of our results are fairly stable when adding additional controls. For example, in Table 5 our results are consistent once controlling for shooting player behaviors. Additional controls for individual shooting ability, shot controls, defensive player controls, and even defensive player fixed effects do not markedly change our results. The one exception is the inclusion of lagged shot controls, which better capture the difficulty of the previous shot. Regardless, the coefficients across Table 5 suggest an extremely small effect of making a previous field goal on the probability a player makes his next field goal. Second, when we introduce additional controls in the subset of

21

years for which they are available, these additional controls increase the magnitude of the negative relationship between field goal shooting. Third, we adapt the methodology introduced in Oster (2019) to bound the effects that omitted variables may play. Our results suggest that omitted variable bias is unlikely to explain our results, and in fact omitted variable bias is likely to generate at most an extremely small hot hand effect.

Another concern that has been raised by previous research is measurement error in both the dependent and independent variables. As previous research points out, making a shot is a noisy indicator of changes in the underlying probability that a player makes a shot (e.g. Stone, 2012; Arkes, 2013). This measurement error may affect both our dependent variable, whether a player makes his current free throw attempt or field goal attempt, and also our independent variable of interest, whether a player made his previous free throw attempt(s) or field goal attempt(s). However, we believe that this concern is misplaced. First, if making the current shot is a noisy but unbiased measure of the underlying probability of making a shot, our coefficients will be unbiased but our standard errors will likely increase. Second, we do not attempt to estimate changes in the underlying probability that players make previous shots. Instead, we focus on what is observable, and thus more relevant to spectators' beliefs about the hot hand: whether a player makes a shot and the characteristics surrounding that shot.

More broadly, it is unclear how analyses of the hot hand should control for shot difficulty or examine the underlying probability of making a shot versus the observed outcome of a made or missed shot. The original question posed by Gilovich et al. (1985) is whether the widespread belief in the hot hand among players, coaches, announcers and fans is correct. Answering this question necessitates a correct interpretation of what these beliefs are. It is unclear to what extent players, coaches, announcers and fans "control" for external factors when evaluating the hot hand effect in game situations and whether players, coaches, announcers and fans evaluate the probability of making a shot versus the observed outcome of whether a shot is made. In this paper, we focus on observable outcomes, as we believe this is the most relevant for examining perceptions regarding the hot hand. We believe that our paper calls for more work connecting results from controlled settings to the run of play and for connecting spectators' beliefs with shooting success. For example, controlled experiments could be conducted which more closely mimic the run of play, varying the time in between shooting attempts the types of shots, adding in defenders. While some experiments

have been conducted along these lines (e.g. Miller and Sanjurjo, 2018), we believe more work is called for in this area to further examine the conditions under which the hot hand effect emerges.

An important question is where this leaves the original hypothesis in Gilovich et al. (1985). Our results suggest that the original hypothesis posed by Gilovich et al. (1985) is largely correct for field goals, but not for free throws. Our findings would align with the hypothesis of a hot hand existing due to muscle memory of repeated physical movements. But this muscle memory appears to dissipate quickly, as evidenced by the hot hand effect in free throws being driven by repeated free throws in a set and by players who take more free throws in a game. Given the size of the hot hand effect, a two percentage point increase in the probability of making the next free throw, it is unclear whether casual fans can discern an effect of this magnitude. But we also find that the size of the hot hand effect grows for players taking more free throws in a game, and it is plausible that observers may be able to discern a 4.6 percentage point increase in the probability of making the next free throw. However, we do not find a hot hand in the most important setting, the run of play. In fact, we find that making a streak of field goals decreases the probability that a player makes his next field goal. Thus, our results suggest that the widespread belief in the hot hand in the run of play is incorrect, supporting the original hypothesis in Gilovich et al. (1985).

References

Aharoni, G. and O. H. Sarig (2012). Hot hands and equilibrium. Applied Economics 44(18), 2309–2320.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. The Review of Economic Studies 58(2), 277–297.

Arkes, J. (2010). Revisiting the hot hand theory with free throw data in a multivariate framework. Journal of Quantitative Analysis in Sports 6(1).

Arkes, J. (2013). Misses in "hot hand" research. Journal of Sports Economics 14(4), 401–410.

Attali, Y. (2013). Perceived hotness affects behavior of basketball players and coaches. Psychological Science 24(7), 1151–1156.

Bocskocsky, A., J. Ezekowitz, and C. Stein (2014). Heat check: New evidence on the hot hand in basketball. Working Paper, 1–30.

Brown, W. O. and R. D. Sauer (1993). Does the basketball market believe in the 'hot hand'? comment. The American Economic Review 83(5), 1377–1386.

Bühren, C. and S. Krabel (2015). Individual performance after success and failure: A natural experiment. Working Paper, 1–24.

Chen, D. L., T. J. Moskowitz, and K. Shue (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. The Quarterly Journal of Economics 131(3), 1181–1242.

Csapo, P., S. Avugos, M. Raab, and M. Bar-Eli (2015). The effect of perceived streakiness on the shot-taking behaviour of basketball players. European Journal of Sport Science 15(7), 647–654.

Csapo, P. and M. Raab (2014). "Hand down, man down." analysis of defensive adjustments in response to the hot hand in basketball using novel defense metrics. PloS One 9(12), e114184.

Dorsey-Palmateer, R. and G. Smith (2004). Bowlers' hot hands. The American Statistician 58(1), 38–45.

Gilovich, T., R. Vallone, and A. Tversky (1985). The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology 17(3), 295–314.

Green, B. and J. Zwiebel (2017). The hot-hand fallacy: Cognitive mistakes or equilibrium adjustments? Evidence from Major League Baseball. Management Science.

Klaassen, F. J. and J. R. Magnus (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. Journal of the American Statistical Association 96(454), 500–509.

Miller, J. B. and A. Sanjurjo (2014). A cold shower for the hot hand fallacy. Working Paper.

Miller, J. B. and A. Sanjurjo (2015). Is it a fallacy to believe in the hot hand in the NBA three-point contest? Working Paper, 1–24.

Miller, J. B. and A. Sanjurjo (2018). Surprised by the gambler's and hot hand fallacies? A truth in the law of small numbers. Econometrica 86(6), 2019–2047.

Nerlove, M. (1967). Experimental evidence on the estimation of dynamic economic relations from a time series of cross-section. The Economic Studies Quarterly (Tokyo. 1950) 18(3), 42–74.

Nerlove, M. (1971). Further evidence on the estimation of dynamic economic relations from a time series of cross sections. Econometrica 39(2), 359.

Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica, 1417–1426.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37(2), 187–204.

Rao, J. M. (2009). Experts' perceptions of autocorrelation: The hot hand fallacy among professional basketball players. Working Paper.

Smith, G. (2003). Horseshoe pitchers' hot hands. Psychonomic Bulletin & Review 10(3), 753–758.

Spector, L. C. and M. Mazzeo (1980). Probit analysis and economic education. The Journal of Economic Education 11(2), 37–44.

Stone, D. F. (2012). Measurement error and the hot hand. The American Statistician 66(1), 61–66.

Wardrop, R. L. (1995). Simpson's paradox and the hot hand in basketball. The American Statistician 49(1), 24–28.

Wetzels, R., D. Tutschkow, C. Dolan, S. van der Sluis, G. Dutilh, and E.-J. Wagenmakers (2016). A bayesian test for the hot hand phenomenon. Journal of Mathematical Psychology 72, 200–209.

Yaari, G. and S. Eisenmann (2011). The hot (invisible?) hand: can time sequence patterns of success/failure in sports be modeled as repeated random independent trials? PloS One 6(10), e24532.

Table 1: Summary Statistics for Free Throw Sample

Variable	All Free Throws		Made Prev. Free Throws		Missed Prev. Free Throws	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Made Prev. Free Throw	0.756	0.429				
Free Throw Made (Y/N)	0.770	0.421	0.785	0.411	0.724	0.447
Second FT	0.582	0.493	0.567	0.496	0.631	0.482
Third FT	0.011	0.103	0.012	0.108	0.007	0.085
Personal Foul FT	0.203	0.402	0.206	0.404	0.193	0.395
Shooting Foul FT	0.734	0.442	0.730	0.444	0.748	0.434
Other Foul FT	0.040	0.195	0.039	0.194	0.041	0.199
Sub b/t Free Throw	0.176	0.381	0.172	0.378	0.187	0.390
Timeout b/t Free Throw	0.003	0.053	0.003	0.053	0.003	0.053
And One Free Throw	0.000	0.002	0.000	0.002	0.000	0.003
Time b/t Player FTs (Sec.)	237.581	455.118	245.352	459.558	213.473	440.191
N	516,699		390,742		125,957	

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons.

Table 2: Summary Statistics for Field Goals Sample

Variable	All Shots Mean	Std.Dev	Made Prev. FG Mean	Std.Dev	Missed Prev. FG Mean	Std.Dev
Previous Shot Made (Y/N)	0.463	0.499				
Shot Made (Y/N)	0.457	0.498	0.453	0.498	0.460	0.498
Shot Distance (Ft.)	12.559	10.008	12.981	9.915	12.194	10.073
Distance b/t Shots (Ft.)	15.917	10.804	14.990	10.747	16.717	10.789
Time b/t Player Shots (Sec.)	218.782	286.925	208.014	272.410	228.074	298.574
Player Took Prev. Shot (Y/N)	0.215	0.411	0.239	0.427	0.194	0.395
Assist on Next Shot	0.050	0.218	0.054	0.226	0.047	0.211
Sub w/in Two Min	0.499	0.500	0.512	0.500	0.487	0.500
Timeout w/in Two Min	0.219	0.413	0.249	0.433	0.193	0.394
Play Length (Sec.)	12.654	6.747	12.961	6.559	12.389	6.894
N	2,074,501		960,979		1,113,522	

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons.

Table 3: Effects of Making a Previous Free Throw on Probability of Making Next Free Throw

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Made Prev. Free Throw	0.061***	0.055***	0.040***	0.041***	0.021***	0.021***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Dep. Var. Mean	0.770	0.770	0.770	0.770	0.770	0.770	0.781
Adjusted R-Squared	0.004	0.009	0.024	0.026	0.046	0.046	0.046
N	516,699	516,699	516,699	516,699	516,699	516,699	306,520
Game Controls		X	X	X	X	X	X
Player Controls			X	X	X	X	X
Shot Controls				X	X	X	X
Leave-Out FT%					X	X	X
Lag Controls						X	
2nd & 3rd FTs Only							X

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is an indicator equal to one for a made free throw and zero for a missed free throw. Standard errors, clustered at the player by year level, are in parentheses. Game controls include indicators for month, day of the week, team by season, and opponent by season, an indicator for the player being on the home team interacted with game attendance, the period interacted with the seconds remaining in that period, and the point gap before the current set of free throws interacted with the home team and with the period. Player controls include indicators for player position, a quadratic function of the player's age, a quadratic function of the player's height and weight, and an interaction between height and weight. Shot controls include indicators for the second and third free throws in a set, whether the free throw resulted from a personal, shooting or other foul (with flagrant foul as the omitted category), whether there was a substitution before the free throw, and whether there was a timeout before the free throw. Leave-Out FT% controls include the player's free throw percentage in the relevant season, excluding the current game, and a four week moving average of the player's free throw percentage, including two weeks on either side of the current game, again excluding the current game. Lagged controls include all time-varying controls for the previous free throw, excluding controls which are calculated based on the second lagged free throw. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 4: Offensive and Defensive Adjustments to Field Goals

	Offensive Adjustments					Defensive Adjustments	
	Take Next Field Goal Attempt	Time b/t Field Goal Attempts	Shot Distance	Shot Location Difference	Assist on Next Play	Sub w/in Two Mins.	Time Out w/in Two Mins.
Previous Shot Made (Y/N)	0.050***	-19.635***	1.549***	0.053***	0.009***	0.022***	0.051***
	(0.001)	(0.553)	(0.021)	(0.016)	(0.000)	(0.001)	(0.001)
Dep. Var. Mean	0.213	218.782	12.559	15.917	0.050	0.499	0.219
Adjusted R-Squared	0.026	0.102	0.181	0.300	0.019	0.110	0.047
N	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is given by the column title. Standard errors, clustered at the player by year level, are in parentheses. All models control for the game characteristics and shot characteristics of the field goal attempt, as well as player by season fixed effects for the player who took the field goal attempt. Game controls include indicators for month, day of the week, team by season, and opponent by season, an indicator for the player being on the home team interacted with game attendance, the period interacted with the seconds remaining in that period, and the point gap before the current shot interacted with the home team and with the period. Shot controls include the number of points the opposing team scored on their last possession, the play length in seconds, shot distance (and its square), indicators for and the interaction between shot types and shot distance, an interaction between shot type and season, and indicators for and the interaction between the angle of the shot to the rim and whether it was on the left side of the court. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 5: Effects of a Previous Field Goal on Probability of Making Next Field Goal Attempt

	(1)	(2)	(3)	(4)	(5)	(6)
Previous Shot Made (Y/N)	-0.006***	-0.008***	-0.011***	-0.001	-0.002**	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dep. Var. Mean	0.457	0.457	0.457	0.457	0.457	0.457
Adjusted R-Squared	0.000	0.003	0.006	0.079	0.080	0.080
N	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501
Game Controls		X	X	X	X	X
Player Controls			X	X	X	X
Shot Controls				X	X	X
Leave-Out FG%					X	X
Lag Controls						X

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is an indicator equal to one for a made field goal and zero for a missed field goal. Standard errors, clustered at the player by year level, are in parentheses. Game controls include indicators for month, day of the week, team by season, and opponent by season, an indicator for the player being on the home team interacted with game attendance, the period interacted with the seconds remaining in that period, and the point gap before the current shot interacted with the home team and with the period. Player controls include indicators for player position, a quadratic function of the player's age, a quadratic function of the player's height and weight, and an interaction between height and weight. Shot controls include the number of points the opposing team scored on their last possession, the play length in seconds, an indicator for if the shooting player also took his team's previous shot, the time between the player's shots, the distance between the player's previous shot and current shot, whether the player was fouled on his last shot, shot distance (and its square), indicators for and the interaction between shot types and shot distance, an interaction between shot type and season, and indicators for and the interaction between the angle of the shot to the rim and whether it was on the left side of the court. Leave-Out FG% controls include the player's field goal percentage in the relevant season, excluding the current game, and a four week moving average of the player's field goal percentage, including two weeks on either side of the current game, again excluding the current game. Lagged controls include all time-varying controls for the previous field goal, except for controls which are calculated using the second-lagged field goal. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 6: Effects of a Previous Field Goal on Probability of Making Next Field Goal Attempt
Investigating Lagged Controls

	(1)	(2)	(3)	(4)	(5)	(6)
Previous Shot Made (Y/N)	-0.002**	-0.002***	-0.002**	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dep. Var. Mean	0.457	0.457	0.457	0.457	0.457	0.457
Adjusted R-Squared	0.080	0.080	0.080	0.080	0.080	0.080
N	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501
Lag Game Controls		X	X	X	X	X
Lag Poss. Controls			X	X	X	X
Lag Shot Type				X	X	X
Lag Shot Dist.					X	X
Lag Shot Dist Int.						X

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. Columns (1) and (6) replicate Columns (5) and (6), respectively, from Table 5. Column (2) adds in lagged game situation controls (period indicators fully interacted with the seconds remaining in the period and the point gap and its interactions with indicators for the home team and the period). Column (3) additionally adds lagged play controls, including the points scored by the opposing team on the last possession and the play length. Column (4) additionally adds lagged shot type by season indicators. Column (5) additionally adds lagged shot distance and shot distance squared, and Column (6) finally adds in the lagged shot type by shot distance interactions. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 7: Effects of Making Consecutive Free Throws on Probability of Making Next Free Throw

	(1)	(2)	(3)	(4)
Made Prev. FT (Y/N)	0.021***	0.021***	0.037***	0.022
	(0.001)	(0.003)	(0.007)	(0.015)
Made Prev. Two FTs (Y/N)		0.035***	0.048***	0.036***
		(0.003)	(0.007)	(0.013)
Made Prev. Three FTs (Y/N)			0.053***	0.044***
			(0.007)	(0.012)
Made Prev. Four FTs (Y/N)				0.045***
				(0.012)
Dep. Var. Mean	0.770	0.771	0.777	0.777
Adjusted R-Squared	0.046	0.047	0.048	0.049
N	516,699	358,201	261,181	179,782

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is an indicator equal to one for a made free throw and zero for a missed free throw. Standard errors, clustered at the player by year level, are in parentheses. The independent variables of interest are indicators for whether or not a player made the lagged, twice lagged, third lagged and fourth lagged free throw, as well as all the possible interactions between these indicators. We create linear combinations of the relevant coefficients to calculate the effect of only making the previous free throw, only the previous two free throws, only the previous three free throws and making all of the previous four free throws. The controls included are the same as in Column (6) of Table 3, with the addition of time-varying controls for the relevant lagged free throws. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 8: Offensive and Defensive Adjustments to Consecutively Made Field Goals

	Offensive Adjustments					Defensive Adjustments	
	Take Next FG Attempt	Time b/t FG Attempts	Shot Dist	Shot Location Diff	Assist on Next Play	Sub w/in 2 Mins.	T.O. w/in 2 Mins.
A. One Lag Shot (N= 2,074,501)							
Made Prev. FG (Y/N)	0.050***	-19.635***	1.549***	0.053***	0.009***	0.022***	0.051***
	(0.001)	(0.553)	(0.021)	(0.016)	(0.000)	(0.001)	(0.001)
Dep. Var. Mean	0.213	218.782	12.559	15.917	0.050	0.499	0.219
Adjusted R-Squared	0.026	0.102	0.181	0.300	0.019	0.110	0.047
B. Two Lagged Shots (N= 1,811,757)							
Made Prev. FG (Y/N)	0.050***	-21.841***	1.522***	0.108***	0.010***	0.022***	0.047***
	(0.001)	(0.674)	(0.025)	(0.022)	(0.000)	(0.001)	(0.001)
Made Prev. Two FGs (Y/N)	0.062***	-29.490***	2.234***	0.316***	0.011***	0.031***	0.060***
	(0.002)	(0.815)	(0.032)	(0.025)	(0.001)	(0.001)	(0.001)
Dep. Var. Mean	0.217	214.552	12.589	15.910	0.050	0.520	0.230
Adjusted R-Squared	0.027	0.117	0.176	0.297	0.019	0.092	0.050
C. Three Lagged Shots (N= 1,568,672)							
Made Prev. FG (Y/N)	0.054***	-21.861***	1.471***	0.029	0.010***	0.022***	0.045***
	(0.002)	(0.846)	(0.033)	(0.030)	(0.001)	(0.001)	(0.001)
Made Prev. Two FGs (Y/N)	0.067***	-30.516***	2.199***	0.292***	0.012***	0.030***	0.058***
	(0.002)	(0.963)	(0.039)	(0.033)	(0.001)	(0.002)	(0.001)
Made Prev. Three FGs (Y/N)	0.069***	-32.395***	2.731***	0.503***	0.012***	0.033***	0.058***
	(0.002)	(1.065)	(0.045)	(0.036)	(0.001)	(0.002)	(0.002)
Dep. Var. Mean	0.222	204.771	12.637	15.935	0.050	0.530	0.236
Adjusted R-Squared	0.028	0.125	0.171	0.293	0.019	0.088	0.055
D. Four Lagged Shots (N= 1,346,178)							
Made Prev. FG (Y/N)	0.062***	-22.519***	1.549***	0.036	0.009***	0.020***	0.047***
	(0.002)	(1.083)	(0.047)	(0.045)	(0.001)	(0.002)	(0.002)
Made Prev. Two FGs (Y/N)	0.074***	-30.358***	2.200***	0.267***	0.012***	0.028***	0.056***
	(0.002)	(1.173)	(0.052)	(0.047)	(0.001)	(0.002)	(0.002)
Made Prev. Three FGs (Y/N)	0.077***	-33.453***	2.809***	0.488***	0.011***	0.033***	0.060***
	(0.003)	(1.258)	(0.059)	(0.051)	(0.001)	(0.003)	(0.002)
Made Prev. Four FGs (Y/N)	0.073***	-32.408***	3.196***	0.701***	0.013***	0.035***	0.055***
	(0.003)	(1.383)	(0.066)	(0.055)	(0.001)	(0.003)	(0.002)
Dep. Var. Mean	0.227	192.568	12.696	15.968	0.051	0.528	0.238
Adjusted R-Squared	0.029	0.124	0.167	0.289	0.018	0.093	0.061

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is given by the column title. Standard errors, clustered at the player by year level, are in parentheses. The independent variables of interest are indicators for whether or not a player made the lagged, twice lagged, third lagged and fourth lagged field goals, as well as all the possible interactions between these indicators. We create linear combinations of the relevant coefficients to calculate the effect of only making the previous field goal, only the previous two field goals, only the previous three field goals and making all of the previous four field goals. The controls included are the same as in Table 4, with the addition of time-varying controls for the relevant lagged field goals. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 9: Effects of Consecutive Field Goals on Probability of Making Next Field Goal Attempt

	(1)	(2)	(3)	(4)
Made Prev. FG (Y/N)	-0.000	0.001	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.002)
Made Prev. Two FGs (Y/N)		-0.002	-0.002	-0.001
		(0.001)	(0.002)	(0.002)
Made Prev. Three FGs (Y/N)			-0.005***	-0.006***
			(0.002)	(0.002)
Made Prev. Four FGs (Y/N)				-0.006**
				(0.003)
Dep. Var. Mean	0.457	0.456	0.456	0.455
Adjusted R-Squared	0.080	0.080	0.080	0.079
N	2,074,501	1,811,757	1,568,672	1,346,178

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The dependent variable is an indicator equal to one for a made field goal and zero for a missed field goal. Standard errors, clustered at the player by year level, are in parentheses. The independent variables of interest are indicators for whether or not a player made the lagged, twice lagged, third lagged and fourth lagged field goals, as well as all the possible interactions between these indicators. We create linear combinations of the relevant coefficients to calculate the effect of only making the previous field goal, only the previous two field goals, only the previous three field goals and making all of the previous four field goals. The controls included are the same as in Column (6) of Table 5, with the addition of time-varying controls for the relevant lagged field goals. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 10: Offensive and Defensive Adjustments to Field Goals
2013-2014 and 2014-2015 Seasons with Additional Shot Characteristics

	Offensive Adjustments					Defensive Adjustments			
	Take Next Field Goal Attempt	Time b/t Field Goal Attempts	Shot Distance	Shot Location Difference	Assist on Next Play	Defender Distance	New Defender	Sub w/in Two Mins.	Time Out w/in Two Mins.
Previous Shot Made (Y/N)	0.055***	-21.569***	1.356***	-0.052	0.010***	-0.076***	0.007***	0.021***	0.055***
	(0.004)	(1.547)	(0.056)	(0.038)	(0.001)	(0.010)	(0.002)	(0.002)	(0.002)
Dep. Var. Mean	0.211	222.418	12.742	15.857	0.051	4.141	0.729	0.506	0.235
Adjusted R-Squared	0.025	0.100	0.185	0.305	0.019	0.265	0.029	0.098	0.034
N	353,926	353,926	353,926	353,926	353,926	353,926	353,926	353,926	353,926

Notes: Data from http://www.bigdataball.com, http://www.nbasavant.com, and http://www.basketball-reference.com for the 2013-2014 and 2014-2015 regular seasons. With a few exceptions, the specifications are the same as those in Table 4. First, we add two additional dependent variables regarding defender responses: whether the defense assigns a new defender and the distance between the defender and shooter. Second, given the limited number of seasons, we replace player by season and opponent by season fixed effects with player, season and opponent fixed effects, and shot type by season indicators with shot type indicators. Third, the models also include additional shot characteristics for the previous shot (the number of dribbles the player took and touch time prior to the shot) and opposing defender (defender distance from the shooting player in feet (and its square), and the height and weight difference between the shooting and defensive player and their interactions with defender distance). Fourth, we include the shot distance and its square as independent variables when the dependent variable is either the defender distance or assigning a new defender. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 11: Effects of a Previous Field Goal on Probability of Making Next Field Goal Attempt
2013-2014 and 2014-2015 Seasons with Additional Shot Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Previous Shot Made (Y/N)	-0.009***	-0.011***	-0.015***	-0.015***	-0.004**	-0.004**	-0.004***	-0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dep. Var. Mean	0.453	0.453	0.453	0.453	0.453	0.453	0.453	0.453
Adjusted R-Squared	0.000	0.002	0.007	0.008	0.096	0.097	0.098	0.098
N	353,940	353,940	353,940	353,940	353,940	353,940	353,940	353,713
Game Controls		X	X	X	X	X	X	X
Player Controls			X	X	X	X	X	X
Opponent Controls				X	X	X	X‡	X‡
Shot Controls					X	X	X	X
Leave-Out FG%						X	X	X
Def. Player FE							X	X
Lag Controls								X

Notes: Data from `http://www.bigdataball.com`, `http://www.nbasavant.com`, and `http://www.basketball-reference.com` for the 2013-2014 and 2014-2015 regular seasons. The specifications are the same as those in Table 5 with a few exceptions. First, given the limited number of seasons, we replace team by season and opponent by season fixed effects with team, season and opponent fixed effects, and shot type by season with shot type indicators. Second, we include additional controls as follows. Shot characteristics additionally include the number of dribbles the player took and touch time prior to the shot. Opponent controls include a quadratic function for the defensive player's age, indicators for the defensive player's position and for a new defender covering the shooting player, defender distance from the shooting player in feet (and its square), and the height and weight difference between the shooting and defensive player and their interactions with defender distance. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%. ‡ When including defensive player fixed effects we do not include any time-invariant player characteristics for the defensive player (e.g. position).

Table 12: Effects of Consecutive Field Goals on Probability of Making Next Field Goal Attempt
2013-2014 and 2014-2015 Seasons

	(1)	(2)	(3)	(4)
Made Prev. FG (Y/N)	-0.003*	-0.002	-0.003	-0.004
	(0.002)	(0.002)	(0.004)	(0.005)
Made Prev. Two FGs (Y/N)		-0.002	0.000	0.000
		(0.003)	(0.004)	(0.006)
Made Prev. Three FGs (Y/N)			-0.007*	-0.005
			(0.004)	(0.006)
Made Prev. Four FGs (Y/N)				-0.011*
				(0.006)
Dep. Var. Mean	0.453	0.453	0.452	0.451
Adjusted R-Squared	0.098	0.099	0.098	0.097
N	353,713	308,326	266,204	227,553

Notes: Data from `http://www.bigdataball.com`, `http://www.nbasavant.com`, and `http://www.basketball-reference.com` for the 2013-2014 and 2014-2015 regular seasons. The specifications are the same as those in Table 9 but include the additional controls described above in Table 11 and the relevant lagged shot controls. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Appendices

A Appendix Tables

Table A1: Robustness Checks: Free Throws

	All Free Throws w/o Lagged Controls	All Free Throws w/ Lagged Controls	Only 2nd and 3rd Free Throws
Full Sample	0.021***	0.021***	0.024***
	(0.001)	(0.001)	(0.002)
	[0.770]	[0.770]	[0.781]
	N=516,699	N=516,699	N=306,520
Home Team Only	0.020***	0.020***	0.023***
	(0.002)	(0.002)	(0.003)
	[0.770]	[0.770]	[0.781]
	N=265,815	N=265,815	N=156,955
Away Team Only	0.019***	0.019***	0.024***
	(0.002)	(0.002)	(0.003)
	[0.770]	[0.770]	[0.781]
	N=250,884	N=250,884	N=149,565
No Overtime Free Throws	0.021***	0.021***	0.024***
	(0.001)	(0.001)	(0.002)
	[0.770]	[0.770]	[0.781]
	N=509,944	N=509,944	N=303,088
No Last 2 Minutes of End of Game	0.021***	0.020***	0.024***
	(0.001)	(0.002)	(0.002)
	[0.769]	[0.769]	[0.779]
	N=459,870	N=459,870	N=277,770
No Sub. Between Free Throws	0.020***	0.020***	0.024***
	(0.002)	(0.002)	(0.002)
	[0.768]	[0.768]	[0.783]
	N=424,968	N=424,968	N=214,789
Shooting Fouls Only	0.021***	0.021***	0.025***
	(0.002)	(0.002)	(0.002)
	[0.767]	[0.767]	[0.779]
	N=379,271	N=379,271	N=230,590
Personal Fouls Only	0.016***	0.016***	0.017***
	(0.003)	(0.003)	(0.004)
	[0.781]	[0.781]	[0.795]
	N=104,856	N=104,856	N=62,530

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic Columns (5) through (7) of Table 3, except as described in each row. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table A2: Robustness Checks: Field Goals

	Offensive Adjustments					Defensive Adjustments		Prob. Make Next Shot
	Take Next Shot	Time b/t Shots	Shot Distance	Change in Shot Location	Assist Next Play	Sub w/in Two Mins.	Time Out w/in Two Mins.	
Full Sample (N=2074501)	0.050***	-19.635***	1.549***	0.053***	0.009***	0.022***	0.051***	-0.000
	(0.001)	(0.553)	(0.021)	(0.016)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.213]	[218.782]	[12.559]	[15.917]	[0.050]	[0.499]	[0.219]	[0.457]
Home Team Only (N=1037353)	0.049***	-18.455***	1.563***	0.073***	0.010***	0.022***	0.054***	-0.000
	(0.001)	(0.690)	(0.025)	(0.021)	(0.001)	(0.001)	(0.001)	(0.001)
	[0.213]	[218.403]	[12.417]	[15.817]	[0.051]	[0.502]	[0.221]	[0.463]
Away Team Only (N=1037148)	0.051***	-20.500***	1.557***	0.045**	0.008***	0.020***	0.047***	-0.001
	(0.001)	(0.683)	(0.025)	(0.021)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.214]	[219.161]	[12.700]	[16.017]	[0.049]	[0.495]	[0.217]	[0.451]
Shots w/in Same Half (N=1839690)	0.052***	-9.583***	1.700***	0.074***	0.009***	0.022***	0.052***	-0.001
	(0.001)	(0.421)	(0.022)	(0.017)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.236]	[163.723]	[12.483]	[15.853]	[0.050]	[0.495]	[0.220]	[0.459]
Shots w/in Same Quarter (N=1536755)	0.054***	-1.201***	1.881***	0.110***	0.008***	0.023***	0.055***	-0.002**
	(0.001)	(0.250)	(0.024)	(0.018)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.274]	[111.264]	[12.458]	[15.758]	[0.050]	[0.489]	[0.230]	[0.460]
No Overtime Shots (N=2056805)	0.049***	-19.480***	1.553***	0.056***	0.009***	0.021***	0.051***	-0.001
	(0.001)	(0.554)	(0.021)	(0.016)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.213]	[218.915]	[12.550]	[15.914]	[0.050]	[0.499]	[0.217]	[0.457]
No Shots w/in 5 Seconds (N=2034630)	0.081***	-28.401***	1.295***	-0.205***	0.008***	0.022***	0.052***	-0.002***
	(0.001)	(0.535)	(0.020)	(0.015)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.199]	[223.015]	[12.721]	[16.104]	[0.051]	[0.499]	[0.219]	[0.456]
No Last 2 Minutes of 4th Qtr. (N=1970689)	0.048***	-18.672***	1.558***	0.068***	0.009***	0.022***	0.052***	-0.001
	(0.001)	(0.555)	(0.021)	(0.016)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.213]	[218.714]	[12.488]	[15.878]	[0.050]	[0.500]	[0.212]	[0.459]
No Shots over 2 Minutes (N=1098415)	0.072***	5.247***	2.240***	0.193***	0.005***	0.021***	0.048***	-0.004***
	(0.002)	(0.099)	(0.029)	(0.022)	(0.000)	(0.001)	(0.001)	(0.001)
	[0.388]	[57.656]	[12.181]	[15.614]	[0.040]	[0.493]	[0.230]	[0.463]

Notes: Data from http://www.bigdataball.com and http://www.basketball-reference.com for the 2005-2006 through 2016-2017 regular seasons. Columns (1) - (7) use the same specifications as Table 4, while Column (8) uses the specification from Column (6) of Table 5. All controls are the same as in the relevant specifications in Table 4 or Table 5, except as indicated by the row description. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table A3: Effects of Making Consecutive Free Throws on Probability of Making Next Free Throw Not Controlling for Previous Shot Characteristics

	(1)	(2)	(3)	(4)
Made Prev. FT (Y/N)	0.021***	0.022***	0.039***	0.024
	(0.001)	(0.003)	(0.007)	(0.015)
Made Prev. Two FTs (Y/N)		0.037***	0.050***	0.040***
		(0.003)	(0.007)	(0.013)
Made Prev. Three FTs (Y/N)			0.056***	0.049***
			(0.007)	(0.012)
Made Prev. Four FTs (Y/N)				0.050***
				(0.012)
Dep. Var. Mean	0.770	0.771	0.777	0.777
Adjusted R-Squared	0.046	0.047	0.048	0.049
N	516,749	358,202	261,181	179,782

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic those in Table 7, with the exception that they do not control for lagged time-varying controls. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table A4: Effects of Consecutive Field Goals on Probability of Making Next Field Goal Attempt Not Controlling for Previous Shot Characteristics

	(1)	(2)	(3)	(4)
Made Prev. FG (Y/N)	-0.002**	0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)
Made Prev. Two FGs (Y/N)		-0.003***	-0.003**	-0.002
		(0.001)	(0.002)	(0.002)
Made Prev. Three FGs (Y/N)			-0.007***	-0.008***
			(0.002)	(0.002)
Made Prev. Four FGs (Y/N)				-0.009***
				(0.002)
Dep. Var. Mean	0.457	0.456	0.456	0.455
Adjusted R-Squared	0.080	0.080	0.080	0.079
N	2,074,501	1,812,115	1,569,280	1,346,943

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic those in Table 9, with the exception that they do not control for lagged time-varying controls. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table A5: Summary Statistics for Field Goals Sample
2013-2014 and 2014-2015 Seasons with Additional Shot Characteristics

	All Shots		Made Prev. FG		Missed Prev. FG	
Variable	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Previous Shot Made (Y/N)	0.459	0.498				
Shot Made (Y/N)	0.453	0.498	0.449	0.497	0.458	0.498
Shot Distance (Ft.)	12.742	9.932	13.130	9.854	12.413	9.986
Defender Distance (Ft.)	4.141	2.760	4.169	2.699	4.118	2.812
Time b/t Player Shots (Sec.)	222.402	292.235	211.103	277.650	231.994	303.739
Player Took Prev. Shot (Y/N)	0.213	0.409	0.240	0.427	0.190	0.392
Assist on Next Shot	0.051	0.220	0.055	0.228	0.047	0.213
New Defender (Y/N	0.729	0.445	0.737	0.440	0.722	0.448
Sub w/in Two Min	0.506	0.500	0.519	0.500	0.494	0.500
Timeout w/in Two Min	0.235	0.424	0.267	0.443	0.207	0.405
Play Length (Sec.)	12.558	6.673	12.839	6.493	12.320	6.813
Num. Dribbles	2.022	3.493	2.067	3.558	1.984	3.436
Touch Time (Sec.)	2.726	3.025	2.792	3.076	2.670	2.979
N	353,948		162,517		191,431	

Notes: Data from `http://www.bigdataball.com`, `http://www.nbasavant.com`, and `http://www.basketball-reference.com` for the 2013-2014 and 2014-2015 regular seasons.

B   Results Using Player Fixed Effects

Table B1: Effects of Making Previous Free Throw on Probability of Making Next Free Throw

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Made Prev. Free Throw	0.061***	0.055***	0.040***	0.041***	0.008***	0.008***	0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Dep. Var. Mean	0.770	0.770	0.770	0.770	0.770	0.770	0.781
Adjusted R-Squared	0.004	0.009	0.024	0.026	0.054	0.054	0.056
N	516,699	516,699	516,699	516,699	516,699	516,699	306,511
Game Controls		X	X	X	X	X	X
Player Controls			X	X			
Shot Controls				X	X	X	X
Player FE					X	X	X
Lag Controls						X	
2nd & 3rd FTs Only							X

Notes: Data from http://www.bigdataball.com and http://www.basketball-reference.com for the 2005-2006 through 2016-2017 regular seasons. The models here mimic the models in Table 3, except that we replace player characteristics and the leave-out free throw percentages with player by season fixed effects. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table B2: Effects of a Previous Field Goal on Probability of Making Next Field Goal Attempt

	(1)	(2)	(3)	(4)	(5)	(6)
Previous Shot Made (Y/N)	-0.006***	-0.008***	-0.011***	-0.001	-0.036***	-0.038***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dep. Var. Mean	0.457	0.457	0.457	0.457	0.457	0.457
Adjusted R-Squared	0.000	0.003	0.006	0.079	0.113	0.113
N	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501	2,074,501
Game Controls		X	X	X	X	X
Player Controls			X	X		
Shot Controls				X	X	X
Player/Season FE					X	X
Lag Controls						X

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic the models in Table 5, except that we replace player characteristics and the leave-out field goal percentages with player by season fixed effects. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table B3: Effects of Making Consecutive Free Throws on Probability of Making Next Free Throw

	(1)	(2)	(3)	(4)
Made Prev. FT (Y/N)	0.008***	0.003	0.018**	-0.003
	(0.001)	(0.003)	(0.007)	(0.015)
Made Prev. Two FTs (Y/N)		0.008**	0.016**	-0.005
		(0.003)	(0.007)	(0.013)
Made Prev. Three FTs (Y/N)			0.016**	-0.005
			(0.007)	(0.013)
Made Prev. Four FTs (Y/N)				-0.009
				(0.012)
Dep. Var. Mean	0.770	0.771	0.777	0.777
Adjusted R-Squared	0.054	0.054	0.055	0.057
N	516,699	358,083	260,779	179,536

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic the models in Table 7, except that we replace player characteristics and the leave-out free throw percentages with player by season fixed effects. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table B4: Effects of Consecutive Field Goals on Probability of Making Next Field Goal Attempt

	(1)	(2)	(3)	(4)
Made Prev. FG (Y/N)	-0.038***	-0.041***	-0.046***	-0.051***
	(0.001)	(0.002)	(0.002)	(0.003)
Made Prev. Two FGs (Y/N)		-0.082***	-0.091***	-0.099***
		(0.003)	(0.004)	(0.004)
Made Prev. Three FGs (Y/N)			-0.135***	-0.150***
			(0.006)	(0.006)
Made Prev. Four FGs (Y/N)				-0.193***
				(0.008)
Dep. Var. Mean	0.457	0.456	0.456	0.455
Adjusted R-Squared	0.113	0.116	0.118	0.121
N	2,074,501	1,811,757	1,568,672	1,346,178

Notes: Data from `http://www.bigdataball.com` and `http://www.basketball-reference.com` for the 2005-2006 through 2016-2017 regular seasons. The models here mimic the models in Table 9, except that we replace player characteristics and the leave-out field goal percentages with player by season fixed effects. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

C   Description and Results of Simulations

In this Appendix, we create simulated data to test the statistical power and accuracy of our estimation strategies. We construct data generating processes (DGP) of three models of shot success and show how our models perform in estimating each DGP. We calculate average coefficients across the simulations, the percent of the coefficients that are statistically significant to examine the statistical power of our estimation strategy, and the percent of the coefficients that are negative to examine how likely the results in our paper would be for each DGP. The first DGP comes from a Markov switching model similar to Wetzels et al. (2016) where both a hot and cold state exist, however, we simplify the model so that the state a player is in is solely dependent on the outcome of the previous shot.[16] The second model expands on the first, where entering the hot state is dependent on the outcome of the previous three shots. The third model is taken directly from Stone (2012), however, we report results for fewer values of the model's parameters.

Our simulation results suggest that, if in fact a hot hand effect exists, our estimation strategy would likely be able to identify it. Our estimation strategy is able to pick up even relatively low magnitudes of the hot hand, and it is unlikely to find a negative effect if in fact a positive relationship exists. Our coefficients are unbiased for the first DGP, although if instead the true DGP is one of the second two DGPs, our strategy results in attenuated coefficients. However, it performs better than simply controlling on observables or using fixed effects, as has been done in previous work. We also show that even if one of the second two DGPs is in fact the true DGP, we would still be unlikely to find the negative coefficients from our in game field goal data. We explain each data generating process and present the results of our estimation strategy below.

C.1   Simulated Data

We begin by creating a simulated data set of 100 players with different underlying probabilities of shot success, coming from a normal distribution with a mean of 0.73 and a standard deviation of 0.105, corresponding to the mean and standard deviation field goal percentage from our field goal data. In order to show bias that could exist from omitted variables we also generate two additional variables, W and Z, each with their own mean, standard deviation, and covariances with the underlying probability of shot success for players.[17]

Next, we generate a different number of shots per game for each player for 82 games, the number of shots are generated from a distribution matching that found in the field goal data, with a mean of 5.312, standard deviation of 8.361, skewness of 0.824 and kurtosis of 3.553. The above process results in 8,200 unique player game combinations and 66,069 total shots. Each of our three data generating processes for shot success start with this generated data.

We run 1,000 iterations of the above data. For each iteration, we create three DGPs of shot success and then estimate statistical models on each of these DGPs as described below.

C.2   Regime Shift Model 1

The first model of shot success uses a regime shift model, where a player enters a hot state if he makes his previous shot. The probability of success on the first shot of each game is equal to the player's underlying probability of shot success, μ_p. For each subsequent shot the probability of success, $P(s_{t,p} = 1)$, is the player's underlying probability of shot success, μ_p, plus the hot hand parameter, δ if the previous shot was made. Therefore, the probability of success for player p current shot t is as follows,

$$P(s_{p,t} = 1) = \mu_p + \delta * \mathbb{1}\{s_{p,t-1} = 1\}, \tag{4}$$

where $\mathbb{1}\{s_{p,t-1} = 1\}$ is an indicator function for whether player p made his previous shot $t - 1$. We repeat this process for five values of the hot hand parameter, $\delta = $ (-0.10, -.01, 0.00, 0.01 , 0.10). These values correspond to no hot hand existing, and an increase (decrease) in a player's probability of shot success of approximately $1/10^{th}$ of a standard deviation, and an increase (decrease) of approximately one standard deviation.

---

[16]We use the model outlined in Wetzels et al. (2016) with α=1 to make this simplification.

[17]The mean, standard deviation and covariance of W were set to 5, 2 and 0.25, while the mean, standard deviation and covariance of Z were set at 50, 20 and 0.75.

C.3 Regime Shift Model 2

The second model of shot success uses a regime shift model where the probability of entering the hot state is 1 only if the previous three shots were made, which reflects a common theme in the hot hand literature that the search for the hot hand should not solely extend to success from the previous shot but success in several of the previous shots. Therefore, we alter Model 1 such that the probability of success on the current shot is as follows:

$$P(s_{p,t} = 1) = \mu_p + \delta * \mathbb{1}\{s_{p,t-1} = s_{p,t-2} = s_{p,t-3} = 1\}, \tag{5}$$

where $\mathbb{1}\{s_{p,t-1} = s_{p,t-2} = s_{p,t-3} = 1\}$ is an indicator that takes on the value of one only if the previous three shots were all successful. We then repeat the simulation process used for Model 1, but replace the probability of shot success on the current shot with the above equation. We also use the same three values of δ that were used in Model 1.

C.4 Model 3

Finally, our third model uses a data generating process that follows a data generating process outlined by Stone (2012), in which the probability of success on every shot depends not on whether the previous shot was made, but is generated by the following AR(1) process:

$$P(s_{p,t} = 1) = \rho P(s_{p,t-1} = 1) + (1 - \rho)\mu_p + \epsilon_t, \tag{6}$$

where $\epsilon_t | P(s_{p,t} = 1) \sim U[-\alpha\delta_t, \alpha\delta_t]$ and $\delta_t = min\{\rho P(s_{p,t-1} = 1) + (1 - \rho)\mu_p, 1 - (\rho P(s_{p,t-1} = 1) + (1 - \rho)\mu_p)\}$. We use the same underlying player's probability of shot success as the previous two models, and set $\alpha = 0.5$.[18] We use the same simulation process as Model 1 and 2, however, we replace the probability of shot success on the current shot with equation 6. We run Model 3 for five values for $\rho =$: 0, 0.1, 0.5, 0.9, and 1. As the values of ρ increase, the probability of making the current shot increasingly depends on the probability that the player made the previous shot instead of the player's underlying probability of shot success, μ_p.

C.5 Regression Models

For each of the three models discussed above we estimate the following regression for each iteration of the 1,000 simulations, and for each value of δ in Model 1 and 2, and each value of ρ in Model 3:

$$s_{p,i,t} = \alpha + \beta s_{p,t-1}\gamma W + \chi Z + \phi \bar{s}_{p,-i} + \epsilon, \tag{7}$$

where $s_{p,t}$ is an indicator for player p making his shot t in game i. We begin by estimating a model only including the lagged shot indicator, $s_{p,t-1}$. We then estimate a second model where we add in the control variables W and Z and a third model which includes these controls as well as the player's field goal percentage excluding the current game, $\bar{s}_{p,-i}$. To further mimic our regression results, we cluster our standard errors at the player level. Finally, to highlight the differences between a player's leave out field goal percentage and fixed effects models, we estimate a fourth regression which includes the two control variables, W and Z, as well as player fixed effects.

C.6 Simulation Results

We estimate these four models across each value of either δ or ρ for 1,000 simulated data sets. We store the coefficient estimate and associated p-value for each iteration. We average the coefficients across all 1,000 simulations and present them in Table C1. We also report the number of statistically significant coefficients and the number of coefficients that were negative of the 1,000 simulations.

For the first regime shift model the estimate of β should be very close to the value of δ in the model. Our results show that our estimation strategy of controlling for observables and the leave out field goal percentage performs best for this data generating process. Additionally, it highlights the negative bias that exists when using player fixed effects instead of the leave out field goal percentage. If the parameter δ is in fact zero, reflecting no existence of a hot hand, the leave out field goal percentage model results in an average β coefficient of 0.001, and 388 of the 1,000 iterations were negative. Additionally, if Regime Shift Model 1 is the true DGP, we can accurately estimate both relatively small and large hot hand effects. If $\delta = -0.01$, representing an extremely small negative autocorrelation in shot success, 979 of the 1,000 iterations result in negative coefficients

---

[18]For more values of α and ρ see Stone (2012)

and 517 coefficients are statistically significant. If instead there is a small positive hot hand effect and $\delta = 0.01$, only four of the 1,000 iterations are negative and not statistically significant, and 714 of the coefficients are statistically significant and positive. Importantly, this demonstrates that under this DGP, our negative and statistically significant coefficients would be extremely unlikely. For larger magnitudes of δ, every single coefficient is statistically significant. Put together, these results suggest that the leave-out field goal percentage performs very well both in terms of generating unbiased coefficients and sufficient statistical power.

In the second regime shift model, we expect the average estimate of β to be lower in magnitude compared to the actual value of δ. Because the hot hand depends on three shots made in succession, and the regression model only examines the last shot, this measurement error in the independent variable of interest will attenuate the β coefficients. Once again if there is in fact no hot hand, when $\delta = 0$, our estimation strategy of controlling on observables and including the leave out field goal percentage results in an average estimate of approximately zero. As in Regime Shift Model 1, just under 400 of the estimates are negative, and only 15 are statistically significant. The attenuated coefficients here do affect the performance of the leave-out field goal percentage model for extremely small positive and negative values of δ. However, once again, for larger absolute values of δ, every single iteration resulted in statistically significant coefficients.

Finally, we present the estimates for the data generating process outlined by Stone (2012). Here again, the leave-out field goal percentage regression specification performs relatively well. Even when 90% of the probability of current shot success is determined by a player's underlying shooting percentage, μ_p, we only find 133 negative coefficients, and just 2 of them are statistically significant, while 182 coefficients are statistically significant and positive. Thus, even here it is unlikely to find a consistently negative and statistically significant coefficient if this is the true DGP. Additionally, our estimates from the field goal data suggest that higher values of ρ are even less likely to be the true parameter values if in fact this is the correct DGP, because we find no negative coefficient estimates from any of the simulations for the higher values of ρ. Finally, these results suggest that our results found for free throws may be understating the true autocorrelation in the probabilities of shot success, since for higher values of ρ our strategy has a negative bias.

Table C1: Simulated Estimates of the Hot Hand

	No Controls			Control Variables			Leave Out FG %			Player FE		
	Avg. Coef.	# Stat. Sig.	# Neg.	Avg. Coef.	# Stat. Sig.	# Neg.	Avg. Coef.	# Stat. Sig.	# Neg.	Avg. Coef.	# Stat. Sig.	# Neg.
A. Regime Shift Model 1												
$\delta = $ -0.10	-0.054	1000	1000	-0.084	1000	1000	-0.099	1000	1000	-0.101	1000	1000
$\delta = $ -0.01	0.045	1000	0	0.010	449	16	-0.009	517	979	-0.012	773	997
$\delta = $ 0.00	0.057	1000	0	0.020	988	0	0.001	72	388	-0.002	72	666
$\delta = $ 0.01	0.068	1000	0	0.030	1000	0	0.011	714	4	0.008	466	31
$\delta = $ 0.10	0.173	1000	0	0.126	1000	0	0.101	1000	0	0.098	1000	0
B. Regime Shift Model 2												
$\delta = $ -0.10	0.012	269	2	-0.018	970	1000	-0.033	1000	1000	-0.035	1000	1000
$\delta = $ -0.01	0.052	1000	0	0.016	912	0	-0.002	68	709	-0.005	233	909
$\delta = $ 0.00	0.057	1000	0	0.020	982	0	0.001	65	388	-0.002	70	647
$\delta = $ 0.01	0.062	1000	0	0.025	1000	0	0.005	201	109	0.002	82	295
$\delta = $ 0.10	0.115	1000	0	0.069	1000	0	0.044	1000	0	0.040	1000	0
C. Stone DGP												
$\rho = 0.0$	0.057	1000	0	0.020	989	0	0.002	63	360	-0.001	63	623
$\rho = 0.1$	0.060	1000	0	0.023	994	0	0.005	184	133	0.002	79	362
$\rho = 0.5$	0.076	1000	0	0.040	1000	0	0.022	1000	0	0.019	992	0
$\rho = 0.9$	0.133	1000	0	0.100	1000	0	0.082	1000	0	0.079	1000	0
$\rho = 1.0$	0.197	1000	0	0.166	1000	0	0.150	1000	0	0.146	1000	0

Notes: Each panel represents results from a different data generating process. Each individual row of a panel represents results for the specified value of the relevant parameter of that model for the four different regression equations. Each different regression specification shows the average coefficient, number of statistically significant coefficients, and number of negative coefficients across 1,000 simulations of the data generating process named. The Leave Out FG % and Player FE columns also include the two control variables. Panel A an Panel B correspond to a modified model from Wetzels et al. (2016). Panel C corresponds to the model outlined by Stone (2012).