NBER WORKING PAPER SERIES

THE HOT HAND IN THE NBA 3-POINT CONTEST: THE IMPORTANCE OF LOCATION, LOCATION, LOCATION

Robert M. Lantis
Erik T. Nesson

Working Paper 29468
http://www.nber.org/papers/w29468

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2021

We are grateful to Joshua Miller and Adam Sanjurjo for making their data and code publicly available. We also thank Chad Cotti, Phil DeCicca, Jeremiah Harris, and Matthew Harris for helpful comments. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Hot Hand in the NBA 3-Point Contest: The Importance of Location, Location, Location
Robert M. Lantis and Erik T. Nesson
NBER Working Paper No. 29468
November 2021
JEL No. D81,D91,Z20,Z29

ABSTRACT

Do basketball players exhibit a hot hand? Results from controlled shooting situations suggest the answer is yes, while results from in-game shooting are mixed. Are the differing results because a hot hand is only present in similar shots or because testing for the hot hand in game situations is difficult? Combining repeated shots in a location and shots across locations, the NBA 3-Point Contests mimics game situations without many of the confounding factors. Using data on the 1986-2019 contests, we find a hot hand, but only within shot locations. Shooting streaks increase a hot hand only if a player makes his previous shot and only within locations. Even making three shots in a row has no effect on making the next shot if a player moves locations. Our results suggest that any hot hand in basketball is only present in extremely similar shooting situations and likely not in the run-of-play.

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

If a basketball player makes one, two, or three shots in a row, are they more likely than otherwise expected to make their next shot? Research suggests that many players, coaches, and fans believe the answer is 'yes' (Gilovich et al., 1985; Brown and Sauer, 1993). Despite the widespread belief in the hot hand, consistent empirical proof of a hot hand in basketball has been elusive. Controlled and semi-controlled shooting settings often find evidence of a hot hand (e.g. Arkes, 2010; Yaari and Eisenmann, 2011; Aharoni and Sarig, 2012; Miller and Sanjurjo, 2014, 2018; Lantis and Nesson, 2021), while evidence from the run of play, the most relevant setting to the belief in the hot hand, has been mixed (e.g. Rao, 2009; Bühren and Krabel, 2015; Csapo and Raab, 2014; Bocskocsky et al., 2014; Lantis and Nesson, 2021). Why are results from these two settings different? It could be that the physiological mechanisms behind the hot hand are sensitive to changes in shot location or shot conditions, or it could be that measurement error and omitted variable bias conceal a hot hand effect in field goal shooting.

Blending characteristics of controlled shooting experiments and the run of play, the NBA 3-Point Contest is an attractive setting to resolve these conflicting sets of estimates. In the contest, players shoot 3-point shots from five locations around the 3-point line, generally shooting five shots in each location before moving to the next location. The combination of repeated shots within the same location and shots spanning locations in a relatively controlled setting allows a close examination of whether a hot hand exists and, if so, whether it is affected by a player moving shot locations.

In this paper, we examine the hot hand in the NBA 3-Point Contest. We use data collected by Miller and Sanjurjo (2021) spanning nearly all of the contests from the first NBA 3-Point Contest in 1986 to the 2019 contest, and we merge in additional player characteristics and NBA 3-Point Contest characteristics. We first estimate linear probability models examining whether making the previous shot affects the probability a player makes his current shot. We then investigate the robustness of the hot hand to different specifications of previous shooting success using a flexible framework similar to Chen et al. (2016) and Lantis and Nesson (2021). We examine the hot hand across all shots as well as separately for shots within shooting locations, where at least one previous shot was in the same location, and the first shot in a location, where all previous shots were in a different location.

We find that a hot hand exists in the contest, but only for shots within shooting locations. When a player moves shooting locations, no combination of makes and misses from the previous shot location affects the probability of making the first shot in the next location. For example, making the fifth shot in the previous location decreases the probability that a player makes his next shot, the first shot in the next location, by two percentage points, which is not statistically significant. If a player makes the last two or three shots in the previous location, the probability that he makes his next shot in the new location is essentially the same as if he missed those same shots. The lack of a hot hand across shooting locations even extends to location changes where shots across locations are the same distance away from the basket, for example moving from the wing to the center of the court.

Within a shooting location, making the previous shot increases the probability of making the next shot by between 4.9 and 9.5 percentage points, depending on the shot number within the location. Making streaks of consecutive shots generates a larger hot hand effect than just making the previous shot, but the streak must include the previous shot. For example, missing a shot and then making two shots in a row increases the probability of making the next shot by nearly eight percentage points, and if a player misses two shots in a row, making the next shot still increases the probability he makes the shot after that by 5.6 percentage points compared to if he missed three shots in a row. However, if a player makes two shots in a row and then misses a shot, he is no more likely to make his next shot than if he had missed three shots in a row.

Our results provide context for why results from studies of the hot hand in basketball using controlled shooting experiments differ from those using data from the run-of-play. Many recent studies of the hot hand in basketball often focus on controlled and semi-controlled settings to limit confounding factors that may generate omitted variable bias (e.g. Arkes, 2010; Yaari and Eisenmann, 2011; Aharoni and Sarig, 2012; Miller and Sanjurjo, 2014, 2018). Consecutive field goals are often taken from different locations, and additionally, endogenous responses in response to previously made shots may generate omitted variable bias (Rao, 2009; Aharoni and Sarig, 2012; Attali, 2013; Csapo and Raab, 2014; Csapo et al., 2015; Bocskocsky et al., 2014; Lantis and Nesson, 2021). To summarize one paper's findings, Lantis and Nesson (2021) find that if a NBA player makes a field goal, he is more likely to take his team's next shot and take that shot quicker and from a further distance, while the defense is more likely to take a time out, make a substitution, assign a new defender to that player, and defend that player closer on his next shot.

While previous studies examining the hot hand in the run of play attempt to control for many of these factors, it remains a question whether the lack of a consistent hot hand in the run of play is due to bias from these omitted factors or whether the hot hand does not exist in the run of play. We find that, while the hot hand exists in the NBA 3-Point Contest, moving shooting positions in a fairly controlled setting, when the shot difficulty is roughly the same between locations, is enough to disrupt the mechanisms behind the hot hand among the best shooters in the NBA. This suggests that results from the run of play which fail to find a hot hand, or even find some evidence of a negative autocorrelation between shooting success, like Lantis and Nesson (2021), are not solely driven by omitted variable bias but rather accurately reflect the absence of a hot hand in game situations.

Results from neuroscience experiments also support our findings. Verstynen and Sabes (2011) find that motor skills become less variable the more an individual repeats the same task, supporting our finding that the third and fourth shots in each location have the largest hot hand effects. Relatedly, Zheng et al. (2020) find that success in motor skills tasks increase in repetitions, but only when there is a visual confirmation of success. Scheidt et al. (2001) show that motor skill adaptations rely on the most recent performances, supporting our finding that success in the most recently taken shot has the greatest effect on success in the current shot. The lack of a hot hand effect when players moved shot locations aligns with findings that higher variability in the required movement for success decreases motor learning (Cardis et al., 2018). Hot hand research in other sports also show a similar pattern to our findings. Hot hands are typically found in sports with repeated movement patterns from similar locations, such as bowling (Dorsey-Palmateer and Smith, 2004), horseshoe pitching (Smith, 2003), and tennis (Klaassen and Magnus, 2001). Interestingly, studies which aggregate movements together which are similar and slightly different may be masking heterogeneity in the hot hand and thus understating the hot hand effect.

We next provide background to the NBA 3-Point Contest and research into the hot hand in Section 2. Then we discuss the data and statistical models in Section 3. Section 4 discusses the results, and Section 5 concludes.

2 Background and Literature Review

2.1 NBA 3-Point Contest Background

The NBA 3-Point Contest began in 1986 as part of the NBA All-Star Weekend and has occurred during every All-Star Weekend since then.[1] Each contestant takes five shots from five different locations beyond the three point line and has 60 seconds to shoot all 25 shots (this format changed slightly in 2020, as we describe below). Figure 1 shows the general locations of the shots in the NBA 3-Point Contest. The five locations are the same for every shooter. Shooters begin in one corner of the court, then move to the wing, then shoot from directly in line with the basket, then from the other wing, and finally shoot from the corner on the opposite side of the court. Each contestant may start from either the left or right corner location. The corner shots are 22 feet from the basket, while the wing locations and the shot directly in the middle of the court are 23.75 feet from the basket.

[Figure 1 About Here]

[1] There was no contest in 1999 due to the NBA lockout.

The first four shots of each location are worth one point if the basket is made, while the fifth shot of each location is worth two points if made, resulting in a maximum score of 30 points. The contest began with three rounds from 1986-1998 and two rounds since 2000. The shooters who have the highest scores in each round advance to the next round. The shooter with the highest score in the final round wins the competition, which includes a monetary prize.[2] If there is a tie in either round, the shooters who are tied shoot again and the shooter with the highest score advances.[3]

Throughout the years the number of shooters in the contest has varied, with 21 contests including eight players and 12 contests including six players, as well as a single year which included nine players and one year which included 10 players. In 2014 the NBA added a "Money Rack", which consisted of five balls all worth two points if the shot was made, and allowed shooters to select which location they wanted the Money Rack to be placed. In 2020 the NBA also extended the time limit to 70 seconds, and added two "Mountain Dew" shots, each worth three points. The Mountain Dew shots are located between the shots from wings and the location directly in line with the basket, but are located further back, at 29.9 feet away from the basket.

2.2 Research into the Hot Hand in Basketball Shooting

The question of whether or not a hot hand in basketball shooting exists extends back to Gilovich et al. (1985). Gilovich et al. (1985) describe their research question 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) collect data on NBA field goals and free throws, as well as a controlled shooting experiment involving collegiate players, and conclude that the belief

[2]The prize money began at $10,000 and has increased in recent years to $50,000.

[3]The time the shooters have in tie-break rounds varies across years and is either 24, 30 or 60 seconds.

that a hot hand exists in basketball shooting is not supported in any of these settings.

Miller and Sanjurjo (2018) demonstrate that the results in Gilovich et al. (1985) are biased downward. After correcting for this bias, the original data in Gilovich et al. (1985) support the existence of a hot hand. However, Ritzwoller and Romano (2021) develop new tests for evaluating the proportion of successes following consecutive successes with either the overall proportion of successes or the proportion of successes following consecutive failures. Applying their tests to the Gilovich et al. (1985) controlled shooting experiment data, they find that the hot hand found in Miller and Sanjurjo (2018) is driven by only one shooter.

Other analyses of the hot hand in basketball focus on free throws or field goals in game situations. The majority of analyses of free throws find a hot hand effect. Two recent examples are Arkes (2010) and Lantis and Nesson (2021). Arkes (2010) collects all free throws from the 2005-2006 NBA regular season and finds that making the previous free throw in a set increases the probability of making the next free throw in the set by about three percentage points. Lantis and Nesson (2021) examine free throws for the 2004-2005 through 2015-2016 NBA regular seasons. They find that making a free throw raises the probability that a player makes his next free throw by about two percentage points, which grows to 4.5 percentage points for players making streaks of three to four free throws in a row.

However, the widespread belief in the hot hand refers to the run of play, not controlled shooting experiments or free throws. Testing for a hot hand in the run of play is more difficult because players' shots may be from different locations, may be minutes apart, and may be affected by endogenous responses to previously made shots (Rao, 2009; Aharoni and Sarig, 2012; Attali, 2013; Csapo and Raab, 2014; Csapo et al., 2015; Bocskocsky et al., 2014; Lantis and Nesson, 2021). The few papers that do examine the hot hand in field goals have been mixed in their findings (e.g. Rao, 2009; Bühren and Krabel, 2015; Csapo and Raab, 2014; Bocskocsky et al., 2014), although the most recent evidence fails to find evidence for a hot hand in the run of play (Lantis and Nesson, 2021).

The NBA 3-Point Contest offers one setting that blends characteristics of field goal shooting without many of the confounds of game situations. The NBA 3-Point Contest features many of the same elite basketball players commonly thought to have a hot hand in game situations competing in a high-stakes competition. Perhaps most importantly, these players shoot 3-point shots within the same location but also move to different locations around the 3-point line. Only a few papers have tested for a hot hand in the NBA 3-Point Contest, notably Koehler and Conley (2003) and Miller and Sanjurjo (2021). Koehler and Conley (2003) study the 1994-1997 NBA 3-Point Contests and compare the number of shooting runs, the probability that a player makes his next shot after hitting or missing three shots in a row compared to the player's base shooting percentage, and examine a player's performance after "announcers' spontaneous temperature outbursts." In all these tests, Koehler and Conley (2003) fail to find evidence of a hot hand.

Miller and Sanjurjo (2021) collect a larger data set, spanning the 1986 to 2020 NBA 3-Point Contests, and correct for a downward bias in Koehler and Conley (2003)'s analyses. They find consistent evidence for a hot hand in the NBA 3-Point Contest across their different measures of the hot hand. Our paper builds on the data collected by Miller and Sanjurjo (2021), and our methods are similar to some of their methods, so we summarize their results in more depth here and then quickly summarize the differences in our approaches. In an Appendix (Section 7.3), we describe a replication of main results from Miller and Sanjurjo (2021) and detail how our results differ. Miller and Sanjurjo (2021) first focus on five measures of the hot hand: (1) H_F, the relative frequency of shots taken immediately following three or more consecutive hits; (2) H_M: the relative frequency of hits immediately following three or more consecutive hits, or shooter's momentum; (3) H_L: the longest streak of consecutive hits; (4) H_L: a composite measure of the first three measures, calculated as the first principal component of the three measures; and (5) R: The runs statistic, where each streak of consecutive runs or misses is counted as a run, and compared to the expected number of runs. Miller and Sanjurjo (2021) also calculate the above statistics for misses to test for a "cold hand." They find consistent evidence for a hot hand across these five measures, but no evidence for a cold hand.

Miller and Sanjurjo (2021) then estimate fixed effects linear probability regressions with different combinations of fixed effects, using an indicator for the player hitting their three previous shots as the measure of previous shooting success (see Figure 9 in Miller and Sanjurjo (2021)). They then estimate specifications for each shot number within shot locations (see Figure 10 in Miller and Sanjurjo (2021)). Figure 10 suggests that in their main specification, the first and second shots in each location have the largest hot hand effect. One reason for this finding is the choice of how to code previous shot success. Miller and Sanjurjo (2021) code their measure of previous shot success to be zero for the first three shots in each round, as a player has not yet taken enough shots to hit his three previous shots. Since players' shooting percentage is lower at the beginning of each round, this coding choice increases the coefficient on the first, second, and third shots in each location above what it would be had the first three shots in each round been coded to missing. The effect of this choice can be seen in Figure 10, as when the first three shots in each round are dropped from the analysis, the estimated hot hand for the first shot in each location drops from almost 0.09 to about 0.05. In our regressions, we do not code previous shot success to be zero in situations when players have not yet taken all the required previous shots.

Our analyses differ in a few other important ways from Miller and Sanjurjo (2021). First, while we also use linear probability models, we do not include fixed effects as this introduces bias in a lagged dependent variable model. Rather we follow previous research into the hot hand and gambler's fallacy and include a leave-out 3-point average and other player characteristics to control for the underlying expected probability that a player will make a particular shot (Chen et al., 2016; Green and Zwiebel, 2018; Lantis and Nesson, 2021). As we describe in Section 4.4.2, our results are robust to estimation with lagged dependent variable models. Second, we examine shot success in the previous one, two, and three shots in a more flexible framework than an indicator for making the previous three shots. An indicator for making the previous three shots compares success in making the previous three shots against all other combinations of hits and misses in the previous three shots and can mask the existence of a hot hand effect. We follow Chen et al. (2016) and Lantis and Nesson (2021) by instead comparing makes and misses against a streak of misses in each number of shots. This framework also allows an examination of which shots in the previous three shots are most influential in determining shooting success in the next shot.

3 Data and Methods

3.1 Data Collection

We use data on the 1986 to 2020 NBA 3-Point Contests created by Miller and Sanjurjo (2021). Miller and Sanjurjo (2021) collect video records of every NBA 3-Point Contest except some rounds in 1986 and 1993 and code each shooter's makes, misses, and the time in between each shot. We direct readers to Miller and Sanjurjo (2021) for more information about their data. Importantly for our analysis, we know the shooting location for each shot (the two corners, the two wings, or the center of the court), and we also know the shot number within each location. Because of the addition of the "Mountain Dew" shots and time extension in 2020, we drop 2020 from our analysis and examine the 1986 to 2019 contests.

We supplement the data collected by Miller and Sanjurjo (2021) with two additional data sources. First, we merge data on player characteristics collected from `http://www.basketball-reference.com` using the the `bbr` package in R.[4]. We collect the season 3-point percentage as well as the total number of 3-point attempts and the player's age. Beginning in the 2014 season players were able to choose which location they would like to make their "Money Rack", where every ball on that rack would count as two points. We build a data set from viewing YouTube videos for each NBA 3-Point Contest, identifying for every player which shot location was their Money Rack of all money balls. We use this data

[4]The `bbr` package is available here: `https://github.com/mbjoseph/bbr`.

to additionally control for whether the shot belonged to the Money Rack post-2014.

3.2 Methods

We base our analysis of the hot hand in the NBA 3-point contest on linear probability models, following previous research into the hot hand (e.g. Arkes, 2010; Bocskocsky et al., 2014; Green and Zwiebel, 2018; Lantis and Nesson, 2021), instead of probability measures used in the main analysis in previous research into the NBA 3-Point Contest (Koehler and Conley, 2003; Miller and Sanjurjo, 2021). First, regression allows us to examine the importance of other factors in shot success by including these factors as additional controls in the regression. These other factors are important as they may generate omitted variable bias, even in quasi-controlled shooting situations such as the NBA 3-Point Contest. For example, suppose that it is practice, rather than previous success, that matters for future shooting success. In that case, omitting controls for the number of shots taken may generate a spurious correlation between past and future shot success. Second, regression allows us to naturally examine subsamples of shots to determine if a hot hand effect is concentrated in certain shot locations or shot numbers. Third, regression allows a flexible specification for measuring multiple measures of streaks of success over the previous two and three shots.

In our first results, we examine the effects of making the previous 3-point shot on the probability of making the current 3-point shot by estimating variations of the following linear probability model:

$$s_{p,y,r,i} = \alpha_0 + \alpha_s s_{p,y,r,i-1} + \alpha_{\bar{s}} \bar{s}_{p,y,r,-l} + \boldsymbol{v_{p,y,r,i} \alpha_v} + \alpha_i i + \alpha_t t + \epsilon_{p,y,r,i}, \tag{1}$$

where the dependent variable, $s_{p,y,r,i}$, is an indicator variable equal to 1 if player p in year y and round r makes shot i. The independent variable of interest, $s_{p,y,r,i-1}$, is an indicator variable for whether the same player made his previous shot in the same round. If α_s is positive and statistically significant, it provides evidence of the hot hand effect. The other independent variables include controls for the player's 3-point shooting average ($\bar{s}_{p,y,r,-l}$), shot setting controls ($\boldsymbol{v_{p,y,r,i}}$), the shot number within each round (i), and the time between shots (t).

We start with a naive model using the lagged shot indicator as the only independent variable along with controls for the player's 3-point contest shooting average, $\bar{s}_{p,y,r,-l}$. Controlling for the player's underlying probability of making a shot alleviates concerns about Simpson's Paradox, which suggests that even though individual players may not exhibit evidence of a hot hand, pooling players together may lead to a spurious hot hand effect (Wardrop, 1995). Therefore, we use a methodology similar to Chen et al. (2016), Green and Zwiebel (2018) and Lantis and Nesson (2021) and control for the player's "leave-out" three point contest percentage. More specifically, we calculate each player's shooting percentage for each year of the NBA 3-Point Contest, excluding the five shots before the current shot, five shots after the current shot, and the current shot. We additionally include the player's season 3-point percentage, the season total number of 3-point shots taken, and the player's age. We use this strategy as opposed to player fixed effects because using fixed effects with a lagged dependent variable generates downward-biased coefficients (Nickell, 1981; Arellano and Bond, 1991; Nerlove, 1967, 1971). We show that our results in Equation 1 are robust to estimating dynamic panel data models in Section 4.4.2. In all models we cluster our standard errors at the player by year level.

We then add in controls for the shot setting, denoted by $\boldsymbol{v_{p,y,r,i}}$, including a quadratic year trend, Money Rack identifier, round fixed effects, indicators for whether the shots were from the corner or center (with the wing location as the omitted category), and fixed effects for the shot number within each shot location.[5] Next, we add in the shot number within each round to control for any practice effects that may generate omitted variable bias, denoted by i. Finally, we control for the time between shots, denoted by t.

[5]Some years have few enough shooters to generate a small lagged dependent variable bias with year fixed effects, especially in some subsamples. Thus, in our main specifications we include a quadratic year function instead of fixed effects. Results using year fixed effects, available upon request, are very similar to our main results except that the coefficients are generally a bit smaller.

We then modify Equation 1 to examine heterogeneity within shot numbers in each shot location. To this end, we interact indicators for the shot number within each location with the indicator for making the previous shot:

$$\begin{aligned} s_{p,y,r,i} &= \beta_0 + \beta_s s_{p,y,r,i-1} + \sum_{b=2}^{5} \beta_b \mathbb{1}[n=b] + \sum_{b=2}^{5} \beta_{sb} \mathbb{1}[n=b] \times s_{p,y,r,i-1} \\ &+ \beta_{\bar{s}} \bar{s}_{p,y,r,-l} + \boldsymbol{v_{p,y,r,i}} \boldsymbol{\beta_v} + \beta_i i + \beta_t t + \epsilon_{p,y,r,i}. \end{aligned} \tag{2}$$

Here, $\mathbb{1}[n=b]$ is an indicator function equal to one if the shot is number b in each location, with $b = \{2,3,4,5\}$, and $\mathbb{1}[n=b] \times s_{p,y,r,i-1}$ represents the interaction between the indicator for making the previous shot and the indicators for the shot number within each shot location. The independent variables of interest are the effects of making the lagged shot for each of the shot numbers within each shot location, with the first shot in each location being the omitted category. Thus, the effect of making the previous shot on the probability of making the first shot in each location is simply β_s, and the probability for the other shots are β_s + β_{sb}, with $b = \{2,3,4,5\}$. The coefficients β_{sb} also show the difference in the effect of making the last shot on the probability of making the current shot between the first shot in each location and shots two through five, and a statistically significant coefficient indicates a difference between shot number b and the first shot.

To further investigate the effects of moving shot locations, we estimate Equation 2 separately for each shot location. As Figure 1 shows, the two corner locations are 22 feet away from the basket, while the two wing locations and the center location are 23.75 feet away from the basket. Thus, shooting motions between either of the wings and the center should involve similar physical motions while the two corner locations should involve a different shooting motion. There are two movements between location in which shooting motions should thus be similar: the wing to the center and then the center to the other wing. Moving from the corner to the wing or the wing to the corner should involve changes in shooting motions. Extending the analysis to each individual location allows a closer examination to whether the hot hand is sensitive to movements in location, even if the shooting motions are very similar between locations.

3.3 Streaks of Success

Shooting success further back than the previous shot may impact the probability of making the current shot. Streaks of made shots may generate a larger hot hand effect than simply making the previous shot, or shooting success further back may work independently of the previous shot. To examine the effects of shooting success over the past three shots in a flexible framework, we modify Equation 1 as follows:

$$\begin{aligned} s_{p,y,r,i} = \gamma_0 + \sum_{j=-3}^{-1} \gamma_{sj} s_{p,y,r,i+j} + \sum_{j=-3}^{-1} \sum_{\substack{k=-3,\\ k>j}}^{-1} \gamma_{sjk} s_{p,y,r,i+j} \times s_{p,y,r,i+k} \\ + \gamma_{\bar{s}} \bar{s}_{p,y,r,-l} + \boldsymbol{v_{p,y,r,l,i} \gamma_v} + \gamma_i i + \gamma_t t + \epsilon_{p,y,r,i}. \end{aligned} \tag{3}$$

Our independent variables of interest are the first through third lagged shot indicators and their interactions. Equation 3 allows a more flexible framework for analyzing previous shot success than creating an indicator for whether a player has made his previous two or previous three shots. In Equation 3, the baseline category is a player having made none of the previous three shots, and we can compute the effect of making any combination of the previous three shots against this baseline. The baseline for an indicator of making all three previous shots would instead include any other combination of makes and misses in the previous three shots, which would possibly mask any hot hand effect.

We first calculate linear combinations of the lagged shot coefficients to estimate the effect of streaks of two and three made shots on the probability of making the next shot.[6] We then further leverage the flexibility of Equation 3 to estimate the effect of all possible combinations of makes and misses over the past three shots on the probability of making the current shot.

[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.

Finally, we estimate these specifications on only the first shot in each location and then on shots two through five in each location to again estimate how hot hand effects differ within and across shooting locations.

3.4 Summary Statistics

The three point contest data covers 9,557 shots taken across the 1986 to 2019 contests. To be in our analysis, the three point shots cannot be a player's first three point shot of each round, which removes 397 shots, leaving us with a final sample of 9,160 shots. Of these, 4,779 shots follow a made shot while 4,381 follow a missed shot. To preview our later results, Table 1 shows summary statistics for our sample, before and after a made or missed three point shot. Following a made three point shot the percent of three point shots made is approximately 56%, while the percentage of three points shots made following a missed shot is 49%, suggesting the presence of a hot hand even though this difference is biased downward by the bias discovered in Miller and Sanjurjo (2018). Table 1 also shows interesting differences between shots following a made and missed shots. Shots taken following a made shot are more likely to be in later rounds, later in each round, less likely to be in the corner, and the player takes longer to shoot the next shot following a made shot.

[**Table 1 About Here**]

4 Results

4.1 Analysis Using One Lagged Shot

We begin by examining a simple formulation of the hot hand: whether a player who makes his last shot is more likely than expected to make his next shot. Table 2 shows results from these regressions. In Column (1) we show a parsimonious specification controlling only for a player's shooting percentage and age to control for Simpson's Paradox. In Column (2), we add controls for the shot setting, in Column (3), we include the shot number within each round, and in Column (4), we additionally control for the time between shots. The results across specifications are remarkably consistent. Players who make their previous 3-point shot are between 5.2 and 5.9 percentage points more likely to make their next 3-point shot, and the coefficients are significant at the one percent level in all specifications. The effect is sizable and translates to an increase of roughly 10 percent off the mean 3-point percentage of 53 percent.

[**Table 2 About Here**]

Next, we explore heterogeneity across shot numbers within each shot location. To this end, we expand our above model to interact the lagged shot indicator with indicators for the shot number from the location of the current shot as described in Equation 2. Our results, displayed in Table 3, show that the hot hand effect found in Table 2 is driven exclusively by the second through fifth shots in each location. Making the last shot in the previous location decreases the probability of making the first shot in the new location by about two percentage points, not statistically significant in any specification. In general, the hot hand effect grows between the second and third shots in the round, holds steady between the third and fourth shots, and falls on the fifth shot, which is the "Moneyball Shot." For the third and fourth shots in each location, the hot hand effect is extremely large: making the previous shot increases the probability of making the next shot by 9.5 and 8.0 percentage points, for the third and fourth shot, respectively, in Column (4). This is an increase of about 15 to 18 percent off the mean.

[**Table 3 About Here**]

Finally, we change Equation 2 to examine shot locations separately. Movements from locations one to two and four to five involve changes in the distance from the basket, while movements from locations two to three and three to four represent movements without a change in shooting distance. The physical shot motions between these locations should be more similar than between location changes with a change in distance. Table 4 shows results from these models, allowing an examination of how sensitive the hot hand is to changes in location. As the sample size is smaller, the coefficients are less-precisely estimated than in Table 3, but the results suggest that the hot hand is extremely sensitive to changes in location. All coefficients for the first shot in each location are negative, except for when shooters move from location four to five, where the coefficient is very small. There is no evidence that moving from location two to three and three to four generate a hot hand effect for the first shot in that location, even though the mechanics of the last shot in the previous location are similar to the first shot in the new location. The large hot hand effects for the third and fourth shots in each location from Table 2 are driven by shot locations later in each round.

[Table 4 About Here]

4.2 Analysis of Time between Shots

In Table 1, the time between shots was different between previously made and missed shots. We also find the the time between shots has a negative effect on the probability of making a shot in Table 2, although we do not report these coefficients. Therefore, we now explore this in more detail in Table 5. Table 5 shows regressions where the dependent variable is the number of seconds between each shot and the independent variable is again an indicator for making the previous shot. We include player by year fixed effects, as these regressions do not contain lagged dependent variables, and we still cluster our standard errors at the player by year level. Additionally, we include round, shooting location, and shot number within location fixed effects, along with the overall shot number in each round. Column (1) of Table 5 shows results for all shots and Columns (2) through (6) show results for each ball. Overall, if a player makes his previous 3-point shot, he takes about 2/10ths of a second longer to take his next shot. These results are very consistent across shot numbers two through five in each location, except that the effect for the third shot is slightly larger at 3/10ths of a second. For the first shot in each location, the effect is very small and not statistically significant for the first shot. The size of the effect is especially interesting as the seconds between shots are about twice as long for the first shot in each location, since player must move between locations. Put together, these results suggest that changes in shooting behavior, which may generate a hot hand effect, are also limited to shots taken from the same location.

[Table 5 About Here]

4.3 Analysis Using Multiple Lagged Shots

We now expand our analysis of the hot hand to examine shooting across the previous two and three shots, reporting results from estimating Equation 3. As our previous results suggest the hot hand effect is absent for the first shot within each location, we conduct a similar analysis here. The columns of Table 6 report results from estimating Equation 3 across one, two and three previous shots. We estimate these models for all shots in Panel A, first shots in each location in Panel B, and then shots two through five in each location in Panel C. In each column, the coefficients and standard errors are relative to the player making no shots in the lagged shots considered, which is one shot for the Column (1), two shots for the Column (2), and three shots for Column (3).

[Table 6 About Here]

As in Table 3, the hot hand effects in Table 6 for all shots in Panel A are completely driven by shots two through five in each shot location. In Panels A and C, the effect of previous shot success increases somewhat for longer streaks of made shots. In Panel A, the effect grows from a 5.4 percentage point increase in making the next shot if a player makes his previous shot to a 6.5 percentage point increase in making the next shot if a player makes his three previous shots. In Panel C, if a player makes his previous shot he is seven percentage points more likely to make his next shot while if he makes three shots in a row he is 8.1 percentage points more likely to make his next shot. However, in Panel B there is no relationship between previous shot success and future shot success across any of the streak lengths. Specifically, if a player makes the last three shots in a location, the probability that he makes the first shot in the next location decreases by 0.6 percentage points compared to if he missed those three shots.

Equation 3 allows a more detailed analysis into which lagged shots are the most important in streaks of shots. In the previous three shots, there are seven possible combinations of making at least one shot: HMM, MHM, MMH, HHM, HMH, MHH, and HHH, where H denotes hitting a shot and M denotes missing a shot. In Table 7 we disaggregate the results for each of the seven possible combinations outlined above.[7] In Column (1) we show results for all shots, Column (2) shows results for first shots in each location only, and Column (3) shows results for the second through fifth shots in each location. Each row shows results for different combinations of made shots, and the last row corresponds to the streak of three makes in Table 6.

[**Table 7 About Here**]

In Table 7, Column (2), no combinations of makes or misses over the previous three shots in a shot location affects the probability of making the first shot in a new location. In fact, the largest magnitude coefficients are negative. In the second through fifth shots in Column (3), which drive the results across all shots in Column (1), the previous shot is the most important shot for predicting future shot success. In Column (2), just making the previous shot while missing the previous two shots generates a sizable hot hand effect of 5.6 percentage points compared to if the player missed all three shots. The effect grows to 7.9 percentage points when missing a shot then making the previous two shots, and grows further to the 8.1 percentage point coefficient in Table 6 for making all three previous shots. No combinations of making the second or third lagged shots have an effect on a player making his current shot unless these shots are combined with making the previous shot.

[7] The estimates shown for different combinations of makes in the previous three shots are the relevant linear combinations of the related coefficients from interactions of the indicators of lagged shot success.

4.4 Robustness Checks

4.4.1 Heterogeneity

We examine the heterogeneity of our results by estimating the models in Tables 2, 3, and 6 across multiple subsamples. First, following Miller and Sanjurjo (2021), we examine players with at least 75, 100, and 150 shots across contests. Second, we examine whether results differ across different rounds in the NBA 3-Point Contest. Third, we examine different years of the contests. We group the years 1986-1998 together, as the contest was relatively stable in format through this time. The years 2000 to 2014 represent a time after the NBA lockout, and finally the years 2014 to 2019 represent a change in format where the Money Rack was introduced. Finally, in the last row of the first heterogeneity analysis we show that our results are robust to estimation with a Probit model instead of a linear probability model.

Table A1 shows heterogeneity analyses for Tables 2 and Table 3. Column (1) shows results across all shots, comparable to Table 2, and Columns (2) through (6) show results for individual shots, comparable to Table 3. For comparison, we repeat the results from the main specifications in the first row. Across the different subsamples we continue to find a hot hand effect of between 4.0 and 6.6 percentage points. Notably, there is no hot hand effect found for the first shot in a location for any subsample, and a few of the sub samples actually report a negative and statistically significant relationship between making the last shot in a location and the first shot in the next location. The third and fourth shots in a location also display the largest hot hand effects consistently across the subsamples, while success in the fifth shot in each location is in general not as related to success in the previous shot.

In Table A2 we repeat the heterogeneity analyses but for streaks of shots. Streaks of two shots are consistently related to success in the next shot, although streaks of three shots do not persist in all subsamples. Again, no subsamples suggest that making the first shot in each location is positively related to success in the previous one, two or three shots, while non-first shots in each round display a much larger relationship to streaks of previous shooting success.

4.4.2 Dynamic Panel Data Models

As noted in Section 4, our method of utilizing the leave-out 3-point percentage may not fully account for Simpson's Paradox or other unobservable player characteristics which may lead to an upward-biased estimate of the hot hand effect. Another method to account for these unobservables is to include player or player × year fixed effects, thus purging anything constant about a player's performance in a given NBA 3-Point Contest. We would modify Equation 1 to

$$s_{p,y,r,i} = \delta_0 + \delta_s s_{p,y,r,i-1} + \boldsymbol{v_{p,y,r,i}} \boldsymbol{\alpha_v} + \alpha_i i + \alpha_t t + \sigma_{p,y} + \epsilon_{p,y,r,i}, \tag{4}$$

where $s_{p,y,r,i}$ and $s_{p,y,r,i-1}$ are again indicators for a player making his current and previous shot, $\boldsymbol{v_{p,y,r,i}}$ represents shot setting controls, excluding the quadratic year trend, i represents the shot number within each round, t represents the time between shots, and $\sigma_{p,y}$ represents player by year fixed effects.

As Nickell (1981) points out, when variables are effectively transformed to be deviations from their mean in a fixed-effects model, the lagged-dependent variable is correlated with the error term. Thus, Equation 4 may generate biased and inconsistent coefficients, and this inconsistency will translate to regressions including independent variables that are functions of the lagged-dependent variable. Nickell (1981) derives an approximation for the bias in a simple setting, which depends on ρ, the coefficient on the lagged dependent variable, and T, the number of time-periods within each panel. For reasonably large values of T, $\text{plim}_{N\to\infty}(\hat{\rho} - \rho) \cong \frac{-(1+\rho)}{T-1}$. This equation suggests that the bias is nearly always negative, is not eliminated when ρ is zero, and is larger when T is small.

A way to account for these issues is to estimate dynamic panel data models (e.g. Anderson and Hsiao, 1981; Arellano and Bond, 1991; Arellano and Bover, 1995), and Klaassen and Magnus (2001) utilize a similar approach in research into the hot hand in tennis, estimating dynamic panel data models with random effects. One of the first papers to propose solutions to the bias from incorporating lagged dependent variables in panel data, Anderson and Hsiao (1981), suggests removing the fixed effects through first differencing, then instrumenting for the lagged dependent variable with further lags of the dependent variable. Arellano and Bover (1995) suggest using "forward orthogonal deviations," where each observation is subtracted from the average of all available values of that variable, and this is the regression we run.[8]. Table 8 shows results from these models. Panel A shows results using all shots, Panel B shows results over the first shot from each shooting location, and Panel C shows results using shots two through five from each shooting location.

[Table 8 About Here]

The results here are similar to those in Table 2, although generally a bit smaller. Again, there is a consistent hot hand effect measured across all shots. The coefficients here suggest that if a player makes his previous shot, he is between 3.0 and 4.0 percentage points more likely to make his next shot. As in our earlier results, the hot hand effect is not found for the first shot in each shooting location. Panel B shows no statistically significant relationship between a player making his previous shot and making his current shot. In fact, most of the coefficients are negative. Panel C shows that the coefficients when the results are restricted to the second through fifth shots in each round increase to between 4.6 and 5.9 percentage points.

A common concern with models derived from Arellano and Bond (1991) is overidentification. In the main specifications, all possible lags of the dependent variable are used as instruments which generates 2,092 instruments. In the first four columns, we undertake various strategies to reduce the number of instruments, highlighted in Roodman (2009a). First, we limit the number of lags to one, instead of using all available lags. This reduces

[8]See Arellano and Bover (1995) and Roodman (2009a) for more information

the instruments to 24 (one for each lag for each of the eight months we estimate our regressions). Next, we use principal components to reduce the instruments, following Roodman (2009a), Bai and Ng (2010), and Kapetanios and Marcellino (2010). Finally, we collapse the instrument matrix (Roodman, 2009b). We show these results in Appendix Table A3. Across the specifications, we see similar results to the main dynamic panel data models above.

5 Conclusion

The NBA 3-Point Contest provides an attractive location to study the hot hand hypothesis, the belief that making previous shots improves a player's current shooting percentage above what would be otherwise expected. In the NBA 3-Point Contest, some of the best 3-point shooters in the NBA compete for a cash prize, shooting five shots each from five locations around the 3-point line. The combination of repeated shots from the same location combined with shots in which a player moves locations allows an examination of the hot hand both within shot locations, where the shooting mechanism is very similar across shots, and across shot locations, where the shooting mechanisms change. Results from an examination of whether and where the hot hand exists in the NBA 3-Point Contest can thus help reconcile results from previous hot hand studies which consistently find a hot hand in controlled shooting situations but do not consistently find a hot hand in field goal shooting in game situations. If a robust hot hand effect is found in the NBA 3-Point Contest, both within and especially between shooting locations, this would indicate that the lack of a hot hand effect seen in many studies of in-game field goal shooting (e.g. Lantis and Nesson, 2021) is possibly due to omitted variable bias or measurement error. If instead results from the NBA 3-Point Contest indicate that the hot hand is easily disrupted by changes in location or other factors, this would indicate that a hot hand does not exist in game situations, where changes in locations are common between consecutive shots.

We examine the existence and robustness of the hot hand in the NBA 3-Point Contest using data provided by Miller and Sanjurjo (2021) merged with additional information on player characteristics and NBA 3-Point Contest characteristics. Like Miller and Sanjurjo (2021), we find a hot hand effect in the NBA 3-Point Contest. Players who make their previous shot are about five percentage points more likely to make their next shot. Examining shot success over the previous three shots, we find that the previous shot is the most important shot in predicting future shot success. Streaks of shots are only related to future shooting success if that streak includes the previous shot. For example, if a player makes his third and second lagged shot, but misses the previous shot, he is no more likely to make his next shot than if he had missed all three previous shots.

However, the hot hand effect is extremely sensitive to changes in shooting location. The hot hand effects exist only when a player is not moving shooting locations, in the second through fifth shots in each shooting location. Making one, two or even three shots in a row in the previous shot location has no statistically significant effect on making the first shot in the next shot location, and the largest coefficients in magnitude are negative. Moreover, we find that location changes which do not affect shooting distance, and thus should still have similar shooting mechanisms, still display no evidence of a hot hand. The lack of a hot hand when players are moving shot locations, and the persistent hot hand effect within shot locations, are extremely robust across different measures of previous shooting success, sub-samples, and other robustness checks.

One possible concern is that our analysis is underpowered, and that if we had a larger sample, we would detect a hot hand effect for the first shot in each shooting location. We do not believe our results are underpowered for a few reasons. First, our statistically significant results for all shots and shots two through five in each shot location suggest that we do have sufficient power to detect hot hand effects in these circumstances. Moreover, the point estimates across measures of previous shot success are generally very small when examining the first shot in each location, or negative, compared to the point estimates that are larger in magnitude and positive for the second through fifth shot in each location. Finally, Table

3 demonstrates that we find statistically significant differences in the effects of making the previous shot in shots two through five compared to shot one in each location.

Another possible concern is measurement error. Making a previous shot may be a noisy indicator of changes in the underlying probability that a player makes his next shot (e.g. Stone, 2012; Arkes, 2013). These noisy measures attenuate the observed relationship between past and present shooting, and thus our measures here may represent a lower-bound measurement of the hot hand. To address this concern, we test multiple measures of previous shot success. All measures suggest the same conclusion: that the hot hand found in the NBA 3-Point Contest is only found when the previous shot is from the same location as the current shot. Moreover, Table 7 suggests that, among all combinations of shot success in the previous three shots, the previous shot is the most important shot in creating a hot hand effect. A streak of two or three made shots only affects the probability of making the next shot only if that streak includes the previous shot. Thus, the simple formulation of the hot hand presented in Tables 2 and 3 may be a sufficiently accurate measure of underlying changes in the probability that a player makes his next shot so as not create large attenuation bias.

Our results tie together multiple themes from previous research into the hot hand hypothesis in basketball. First, they confirm that the results from controlled shooting experiments which find a robust hot hand, such as Miller and Sanjurjo (2018)'s paradigm-shifting article on the hot hand and gambler's fallacy, extend outside of controlled experiments into semi-controlled settings. Our finding that the existence of a hot hand for shots within the same location also provides support for previous papers which find a hot hand in NBA free throw shooting (e.g. Arkes, 2010; Lantis and Nesson, 2021). Finally, our results highlight the lack of a hot hand when shot location changes, providing support for recent research which suggests a hot hand does not exist in NBA field goals in the run of play, (Lantis and Nesson, 2021).

The important characteristic which ties these results together is shot location. Our results suggest that the physiological mechanisms behind a hot hand are extremely sensitive to changes in shot location, and they may be sensitive to changes in other characteristics as well. Controlled shooting situations, especially those where shots are repeatedly taken from the same location, will likely exhibit a hot hand effect. However, if players move locations, even around the 3-point arc, this movement disrupts the physiological mechanisms which generate a hot hand effect. Put together, our results suggest that the widespread belief in the hot hand in field goal shooting, where players are often moving locations and other characteristics are also changing, is likely a fallacy, supporting the original conjecture in Gilovich et al. (1985).

Our results highlight the importance of ensuring that characteristics that change the physical or mental process of what is required for success, in our study shot location, are consistent across trials or controlled for when identifying whether a hot hand exists. Previous research investigating the hot hand fallacy in areas like gambling behavior, or decision making may be disguising or underestimating a hot hand effect if characteristics that alter the process required for success changes between the previous and current trial. Future research could re-investigate the hot hand in these areas with this insight in mind.

References

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

Anderson, T. W. and C. Hsiao (1981). Estimation of dynamic models with error components. Journal of the American Statistical Association 76(375), 598–606.

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.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68(1), 29–51.

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.

Bai, J. and S. Ng (2010). Instrumental variable estimation in a data rich environment. Econometric Theory 26(6), 1577–1606.

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.

Cardis, M., M. Casadio, and R. Ranganathan (2018). High variability impairs motor learning regardless of whether it affects task performance. Journal of Neurophysiology 119(1), 39–48.

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 (2018). The hot-hand fallacy: Cognitive mistakes or equilibrium adjustments? Evidence from Major League Baseball. Management Science 64(11), 4967–5460.

Kapetanios, G. and M. Marcellino (2010). Factor-GMM estimation with large sets of possibly weak instruments. Computational Statistics & Data Analysis 54(11), 2655–2675.

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.

Koehler, J. J. and C. A. Conley (2003). The "hot hand" myth in professional basketball. Journal of Sport and Exercise Psychology 25(2), 253–259.

Lantis, R. and E. Nesson (2021). Hot shots: An analysis of the "hot hand" in NBA field goal and free throw shooting. Journal of Sports Economics 22(6), 639–677.

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

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.

Miller, J. B. and A. Sanjurjo (2021). Is it a fallacy to believe in the hot hand in the NBA three-point contest? European Economic Review 138(9), 1–23.

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.

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

Ritzwoller, D. M. and J. P. Romano (2021). Uncertainty in the hot hand fallacy: Detecting streaky alternatives to random bernoulli sequences. The Review of Economic Studies, 1–43.

Roodman, D. (2009a). How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal 9, 86–136.

Roodman, D. (2009b). A note on the theme of too many instruments. Oxford Bulletin of Economics and Statistics 71(1), 135–158.

Scheidt, R. A., J. B. Dingwell, and F. A. Mussa-Ivaldi (2001). Learning to move amid uncertainty. Journal of Neurophysiology 86(2), 971–985.

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

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

Verstynen, T. and P. N. Sabes (2011). How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. Journal of Neuroscience 31(27), 10050–10059.

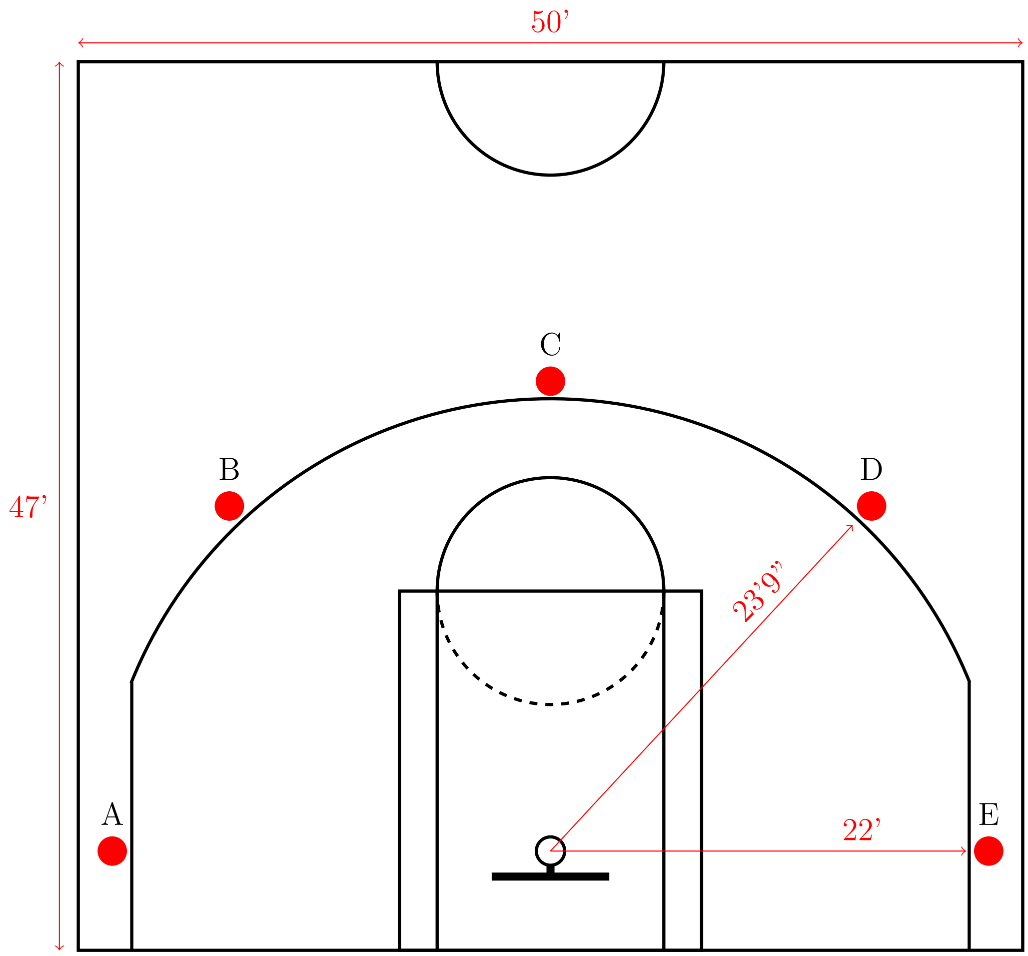
Wardrop, R. L. (1995). Simpson's paradox and the hot hand in basketball. The American Statistician 49(1), 24–28.

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.

Zheng, Y., Y.-R. Mao, T.-F. Yuan, D.-S. Xu, and L.-M. Cheng (2020). Multimodal treatment for spinal cord injury: a sword of neuroregeneration upon neuromodulation. Neural Regeneration Research 15(8), 1437.

6 Figures and Tables

Figure 1: NBA 3-Point Contest Shooting Locations



Notes: This figure shows one half of a basketball court along with the five shooting locations in the NBA 3-Point Contest. Players may start at either location A or location E, then work their way around the 3-Point line to the other corner. Players shoot five shots from each location before moving to the next location.

Table 1: Summary Statistics For Sample

Variable	All Shots		Made Previous Shot		Missed Previous Shot		
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev	T-Test
Made Shot	0.528	0.499	0.559	0.497	0.494	0.500	0.000
Made Prev. Shot	0.522	0.500					
Round	1.546	0.765	1.581	0.779	1.508	0.747	0.000
Shot Number	13.245	6.889	13.677	6.738	12.773	7.021	0.000
Shot Location: Center	0.209	0.406	0.220	0.414	0.197	0.398	0.007
Shot Location: Wing	0.421	0.494	0.432	0.495	0.408	0.492	0.020
Shot Location: Corner	0.371	0.483	0.348	0.476	0.395	0.489	0.000
Shot Number in Location	3.070	1.379	3.135	1.396	2.998	1.357	0.000
Money Rack	0.038	0.192	0.040	0.196	0.036	0.186	0.304
Time between Shots	2.299	0.766	2.314	0.758	2.282	0.775	0.052
3Pt. Attempts (Season)	335.504	163.714	338.127	167.416	332.642	159.548	0.108
3Pt. Percent (Season)	0.407	0.045	0.410	0.045	0.405	0.046	0.000
Player Age	27.244	3.559	27.259	3.554	27.227	3.566	
N	9,160		4,779		4,381		

Notes: Data from Miller and Sanjurjo (2021), `http://www.basketball-reference.com`, and data collected by the authors for the 1986 through 2019 NBA 3-Point Contests. There was no contest in 1999 due to the NBA lockout. The last column shows the p-value for a t-test for differences in shots following a previously made shot and shots following a previously missed shot.

Table 2: Effects of Making a 3-Point Shot on the Probability of Making the Next 3-Point Shot

	(1)	(2)	(3)	(4)
Made Prev. Shot	0.059***	0.055***	0.052***	0.054***
	(0.012)	(0.012)	(0.012)	(0.012)
Dep. Var. Mean	0.528	0.528	0.528	0.528
Adjusted R-Squared	0.010	0.014	0.015	0.020
N	9,160	9,160	9,160	9,160
Leave-Out 3pt. %	X	X	X	X
Shot Controls		X	X	X
Shot Number			X	X
Time b/t Shots				X

Notes: The data sources are described in Table 1. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. In Column (1), we include the controls for the player's shooting percentage (leave-out 3-point percentage, season 3-point shooting percentage and 3-point shots taken, and player age). In Column (2), we additionally control for a quadratic year function, round, shot location, an indicator for whether the shot location was the "Money Rack" post-2014, and shot number fixed effects. In Column (3), we add a control for the shot number within each round. In Column (4), we add a control for the time between each shot. Standard errors are clustered at the player by season level. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 3: Effects of Making a 3-Point Shot on the Probability of Making the Next 3-Point Shot: Results by Shot Number within Shot Location

	(1)	(2)	(3)	(4)
Made Prev. Shot x Shot 1	-0.021	-0.021	-0.020	-0.020
	(0.028)	(0.028)	(0.028)	(0.028)
	[n/a]	[n/a]	[n/a]	[n/a]
Made Prev. Shot x Shot 2	0.056**	0.051**	0.047*	0.049**
	(0.024)	(0.024)	(0.024)	(0.024)
	[0.030]	[0.042]	[0.056]	[0.052]
Made Prev. Shot x Shot 3	0.100***	0.096***	0.092***	0.095***
	(0.022)	(0.022)	(0.022)	(0.022)
	[0.001]	[0.001]	[0.002]	[0.001]
Made Prev. Shot x Shot 4	0.084***	0.080***	0.078***	0.080***
	(0.023)	(0.023)	(0.024)	(0.023)
	[0.003]	[0.004]	[0.005]	[0.005]
Made Prev. Shot x Shot 5	0.056**	0.053**	0.052**	0.052**
	(0.025)	(0.025)	(0.025)	(0.025)
	[0.037]	[0.043]	[0.049]	[0.049]
Dep. Var. Mean	0.528	0.528	0.528	0.528
Adjusted R-Squared	0.014	0.015	0.016	0.021
N	9,160	9,160	9,160	9,160
Leave-Out 3pt. %	X	X	X	X
Shot Controls		X	X	X
Shot Number			X	X
Time b/t Shots				X

Notes: The data sources are described in Table 1. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. The four columns have the same progression as Table 2, and results in each column represent one regression. The coefficients and standard errors are linear combinations of an indicator for making the previous shot and the interaction between making the previous shot and the shot number in each location. P-values for t-tests of the difference in the effect of making each particular shot and the first shot are shown in brackets. Standard errors are clustered at the player by season level. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 4: Effects of Making a 3-Point Shot on the Probability of Making the Next 3-Point Shot: Results by Shot Number and Shot Location

	Shot Location Number				
	1	2	3	4	5
Made Prev. Shot x Ball 1	n/a	-0.041	-0.005	-0.029	0.009
	(n/a)	(0.051)	(0.050)	(0.051)	(0.052)
Made Prev. Shot x Ball 2	0.033	0.022	0.069	0.064	0.013
	(0.052)	(0.053)	(0.051)	(0.056)	(0.051)
Made Prev. Shot x Ball 3	0.095*	0.048	0.078	0.110**	0.103**
	(0.051)	(0.047)	(0.050)	(0.052)	(0.049)
Made Prev. Shot x Ball 4	0.053	0.067	0.001	0.158***	0.125**
	(0.049)	(0.050)	(0.048)	(0.053)	(0.055)
Made Prev. Shot x Ball 5	0.068	0.046	0.071	0.024	0.045
	(0.050)	(0.052)	(0.052)	(0.051)	(0.052)
Dep. Var. Mean	0.489	0.522	0.547	0.538	0.539
Adjusted R-Squared	0.016	0.021	0.017	0.028	0.015
N	1,588	1,985	1,911	1,870	1,806

Notes: The data sources are described in Table 1. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. The five columns each conform to the shot location for each shooter. Locations 1 and 5 are the corners, 2 and 4 are the wings, and 3 is the center. The controls are the same as Table 3, with the exception that we do not control for shot location since each regression is only estimated for a single shot location. The coefficients and standard errors are linear combinations of an indicator for making the previous shot and the interaction between making the previous shot and the shot number in each location. Standard errors are clustered at the player by season level. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 5: Effects of Making a 3-Point Shot on the Time between Shots

		Specific Shot Number in Location				
	All Shots	1	2	3	4	5
Made Prev. Shot	0.023***	0.011	0.023***	0.032***	0.020**	0.021*
	(0.006)	(0.019)	(0.008)	(0.008)	(0.009)	(0.012)
Dep. Var. Mean	2.299	3.727	1.897	1.947	2.010	2.191
Adjusted R-Squared	0.883	0.707	0.719	0.704	0.725	0.623
N	9,160	1,540	1,927	1,918	1,906	1,869

Notes: The data sources are described in Table 1. The dependent variable in each regression is the number of seconds between 3-point shots. All regressions restrict the analysis to previous shots taken within the same round of the current shot. The first column shows results for all shots, and Columns (2) through (6) show results for each shot within shot location. All regressions include controls for player by year fixed effects, round fixed effects, indicators for shot location, the shot number within each round, an indicator for the Money Rack, and the first column additionally uses fixed effects for shot number within each shot location. Standard errors are clustered at the player by season level. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 6: Effects of Making Consecutive 3-Point Shots on the Probability of Making the Next 3-Point Shot

	Streak of One Hit	Streak of Two Hits	Streak of Three Hits
A. All Shots			
Made Prev. Number Shots	0.054***	0.063***	0.065***
	(0.012)	(0.017)	(0.021)
Dep. Var. Mean	0.528	0.534	0.537
Adjusted R-Squared	0.020	0.019	0.019
N	9,160	8,763	8,366
B. First Shot in Shot Location			
Made Prev. Number Shots	-0.018	0.004	-0.006
	(0.028)	(0.035)	(0.051)
Dep. Var. Mean	0.478	0.478	0.478
Adjusted R-Squared	0.014	0.014	0.014
N	1,540	1,540	1,540
C. Non-First Shots in Shot Location			
Made Prev. Number Shots	0.070***	0.077***	0.081***
	(0.012)	(0.018)	(0.023)
Dep. Var. Mean	0.538	0.546	0.550
Adjusted R-Squared	0.021	0.019	0.019
N	7,620	7,223	6,826

Notes: The data sources are described in Table 1. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. Panel B restricts shots to the first shot in each location, such that all lagged shots included in the model are taken in a different location. Panel C restricts shots to shots two through five from each location, such that at least one lagged shot included in the model is taken in the same location. The controls included are those in Column (4) of Table 2. Standard errors are clustered at the player by season level. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 7: Effects of Making Combinations of 3-Point Shots on the Probability of Making the Next 3-Point Shot

	All Shots	First Shot in Shot Location	Non-First Shots in Shot Location
MM**H**	0.046**	-0.003	0.056**
	(0.021)	(0.058)	(0.024)
M**H**M	0.008	0.034	-0.004
	(0.022)	(0.059)	(0.024)
HMM	-0.003	-0.062	0.007
	(0.024)	(0.057)	(0.026)
M**HH**	0.055**	-0.061	0.079***
	(0.022)	(0.056)	(0.023)
HM**H**	0.030	-0.055	0.046*
	(0.023)	(0.057)	(0.025)
HHM	-0.006	-0.012	-0.010
	(0.025)	(0.058)	(0.027)
HHH	0.065***	-0.006	0.081***
	(0.021)	(0.051)	(0.023)
Dep. Var. Mean	0.537	0.478	0.550
Adjusted R-Squared	0.019	0.014	0.019
N	8,366	1,540	6,826

Notes: The data sources are described in Table 1, and the table structure is similar to Table 6. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. Each row shows coefficients corresponding to different combinations of made and missed shots over the past three shots, compared to missing all three shots. The shots move from left to right, so the first row (MM**H**) shows the effect of missing shots $i-3$ and $i-2$ and then making shot $i-1$ on the probability of making shot i. The coefficients provided for different combinations of previously made shots are the linear combinations of the coefficients for the relevant interactions of the indicators for making each of the previous three shots. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

Table 8: Effects of Making a 3-Point Shot on the Probability of Making the Next 3-Point Shot: Dynamic Panel Data Models

	(1)	(2)	(3)	(4)
A. All Shots				
Made Prev. Shot	0.035***	0.031**	0.030**	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)
Dep. Var. Mean	0.529	0.529	0.529	0.531
N	9,085	9,085	9,085	8,924
AR Test 1	0.000	0.000	0.000	0.000
AR Test 2	0.338	0.314	0.319	0.442
AR Test 3	0.329	0.406	0.419	0.472
B. First Shot in Shot Locations				
Made Prev. Shot	-0.039	-0.026	-0.020	-0.009
	(0.029)	(0.029)	(0.029)	(0.030)
Dep. Var. Mean	0.535	0.535	0.535	0.550
N	1,695	1,695	1,695	1,534
AR Test 1	0.000	0.102	0.085	0.111
AR Test 2	0.141	0.952	0.801	0.913
AR Test 3	0.269	0.901	0.683	0.826
C. Shots Within Same Shot Location				
Made Prev. Shot	0.052***	0.047***	0.046***	0.057***
	(0.013)	(0.013)	(0.013)	(0.013)
Dep. Var. Mean	0.528	0.528	0.528	0.528
N	7,390	7,390	7,390	7,390
AR Test 1	0.000	0.000	0.000	0.000
AR Test 2	0.112	0.157	0.158	0.131
AR Test 3	0.344	0.349	0.359	0.666
Player x Year FE	X	X	X	X
Shot Controls		X	X	X
Shot Number			X	X
Time b/t Shots				X

Notes: The data sources are described in Table 1. The dependent variable in each regression is an indicator for a player making a 3-point shot. All regressions restrict the analysis to previous shots taken within the same round of the current shot. The regressions follow the same progression of controls as Table 2, with the exception that the leave-out 3-point average and quadratic year trend is replaced by player by year fixed effects. All regressions are estimated using an Arellano and Bover (1995) style specification. Standard errors are robust, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

7 Appendix

7.1 Heterogeneity Analysis

Table A1: Effects of Making a 3-Point Shot on the Probability
Of Making the Next 3-Point Shot: Heterogeneity Analysis

		Specific Shot Number in Location				
	All Shots	1	2	3	4	5
All Shots (N=9160)	0.054***	-0.020	0.049**	0.095***	0.080***	0.052**
	(0.012)	(0.028)	(0.024)	(0.022)	(0.023)	(0.025)
Players with $\geq$ 75 shots (N=6624)	0.048***	-0.031	0.057**	0.081***	0.093***	0.026
	(0.014)	(0.032)	(0.028)	(0.026)	(0.027)	(0.030)
Players with $\geq$ 100 shots (N=5922)	0.049***	-0.043	0.059**	0.086***	0.108***	0.015
	(0.014)	(0.033)	(0.030)	(0.027)	(0.028)	(0.031)
Players with $\geq$ 150 shots (N=4021)	0.040***	-0.103**	0.052	0.127***	0.094***	-0.002
	(0.015)	(0.042)	(0.037)	(0.031)	(0.031)	(0.033)
First Round Only (N=5488)	0.055***	-0.021	0.067**	0.079***	0.075**	0.060**
	(0.013)	(0.034)	(0.029)	(0.028)	(0.030)	(0.030)
Excluding First Round (N=3672)	0.051***	-0.017	0.019	0.117***	0.087**	0.036
	(0.019)	(0.048)	(0.038)	(0.035)	(0.036)	(0.040)
1986-1998 Contests (N=4119)	0.044**	-0.090*	0.006	0.120***	0.077**	0.084**
	(0.019)	(0.046)	(0.039)	(0.033)	(0.037)	(0.036)
2000 to 2014 Contests (N=3307)	0.055***	0.020	0.088**	0.074*	0.074*	0.011
	(0.017)	(0.047)	(0.035)	(0.037)	(0.038)	(0.041)
2014-2019 Contests (N=1734)	0.066**	0.044	0.092	0.067	0.089*	0.034
	(0.028)	(0.054)	(0.056)	(0.044)	(0.045)	(0.064)
Probit Model (N=9160)	0.054***	-0.019	0.049**	0.094***	0.080***	0.052**
	(0.012)	(0.027)	(0.024)	(0.022)	(0.023)	(0.025)

Notes: The data sources are described in Table 1. This table reproduces Column (4) in Table 2 (for Column (1)) and Table 3 (for Columns (2)-(6)), with the exception that all models are run over the sample specified.

Table A2: Effects of Making Streaks of 3-Point Shots on the Probability of Making the Next 3-Point Shot
Heterogeneity Analysis

	All Shots			First Shot in Shot Location			Non-First Shots in Shot Location		
	One Make	Two Makes	Three Makes	One Make	Two Makes	Three Makes	One Make	Two Makes	Three Makes
All Shots	0.054***	0.063***	0.065***	-0.018	0.004	-0.006	0.070***	0.077***	0.081***
	(0.012)	(0.017)	(0.021)	(0.028)	(0.035)	(0.051)	(0.012)	(0.018)	(0.023)
Players with $\geq$ 75 shots	0.048***	0.052***	0.044*	-0.035	-0.004	-0.026	0.066***	0.063***	0.055**
	(0.014)	(0.020)	(0.024)	(0.033)	(0.042)	(0.064)	(0.014)	(0.021)	(0.026)
Players with $\geq$ 100 shots	0.049***	0.052**	0.042	-0.043	-0.007	-0.043	0.070***	0.064***	0.055**
	(0.014)	(0.021)	(0.026)	(0.034)	(0.046)	(0.069)	(0.015)	(0.022)	(0.028)
Players with $\geq$ 150 shots	0.040***	0.048*	0.014	-0.102**	-0.063	-0.114	0.071***	0.069***	0.035
	(0.015)	(0.025)	(0.032)	(0.043)	(0.061)	(0.081)	(0.015)	(0.025)	(0.034)
First Round Only	0.055***	0.059***	0.070**	-0.020	-0.030	-0.065	0.071***	0.080***	0.105***
	(0.013)	(0.019)	(0.027)	(0.035)	(0.047)	(0.061)	(0.015)	(0.022)	(0.031)
Excluding First Round	0.051***	0.068**	0.046	-0.011	0.063	0.072	0.066***	0.069**	0.037
	(0.019)	(0.027)	(0.034)	(0.048)	(0.055)	(0.088)	(0.019)	(0.029)	(0.036)
1986-1998 Contests	0.044**	0.049*	0.020	-0.085*	-0.038	-0.110	0.072***	0.069***	0.045
	(0.019)	(0.026)	(0.034)	(0.047)	(0.054)	(0.075)	(0.019)	(0.026)	(0.037)
2000 to 2014 Contests	0.055***	0.044*	0.072**	0.025	0.010	0.039	0.064***	0.049*	0.080**
	(0.017)	(0.026)	(0.030)	(0.047)	(0.062)	(0.085)	(0.019)	(0.029)	(0.035)
2014-2019 Contests	0.066**	0.104***	0.109**	0.050	0.084	0.116	0.072**	0.110***	0.099*
	(0.028)	(0.036)	(0.048)	(0.056)	(0.085)	(0.125)	(0.029)	(0.038)	(0.057)

Notes: The data sources are described in Table 1. This table reproduces the results in Table 6, with the exception that all models are run over the sample specified.

7.2 Dynamic Panel Data Models

Table A3: Effects of Making a 3-Point Shot on the Probability of Making the Next 3-Point Shot: Dynamic Panel Data Models Robustness Checks

	Limit Number of Instruments	Use PCA to Reduce Instruments	Collapse Instruments
A. All Shots			
Made Prev. Shot	0.049***	0.049***	0.050***
	(0.013)	(0.015)	(0.013)
Dep. Var. Mean	0.531	0.531	0.531
N	8,924	8,924	8,924
AR Test 1	0.000	0.000	0.000
AR Test 2	0.269	0.319	0.252
AR Test 3	0.476	0.476	0.477
B. First Shot in Shot Locations			
Made Prev. Shot	-0.015	-0.012	-0.015
	(0.030)	(0.037)	(0.030)
Dep. Var. Mean	0.550	0.550	0.550
N	1,534	1,534	1,534
AR Test 1	0.112	0.110	0.112
AR Test 2	0.927	0.920	0.926
AR Test 3	0.835	0.831	0.835
C. Shots Within Same Shot Location			
Made Prev. Shot	0.070***	0.073***	0.072***
	(0.013)	(0.016)	(0.013)
Dep. Var. Mean	0.528	0.528	0.528
N	7,390	7,390	7,390
AR Test 1	0.000	0.000	0.000
AR Test 2	0.050	0.056	0.044
AR Test 3	0.668	0.669	0.668

Notes: The data sources are described in Table 1, and the regressions have similar controls to Column (4) of Table 8. Changes in specifications are described by each column title. Standard errors are robust. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

7.3 Reconciliation with Miller & Sanjurjo (2021)

In this Appendix, we describe the difference in methodologies between the regression results in Miller and Sanjurjo (2021) and our results. As noted above, Miller and Sanjurjo (2021) focus on five measures of the hot hand: (1) H_F, the relative frequency of shots taken immediately following three or more consecutive hits; (2) H_M: the relative frequency of hits immediately following three or more consecutive hits, or shooter's momentum; (3) H_L: the longest streak of consecutive hits; (4) H_L: a composite measure of the first three measures, calculated as the first principal component of the three measures; and (5) R: The runs statistic, where each streak of consecutive runs or misses is counted as a run, and compared to the expected number of runs. Miller and Sanjurjo (2021) also calculate the above statistics for misses to test for a "cold hand."

Miller and Sanjurjo (2021) also estimate fixed effects linear probability models generally of the form,

$$s_{p,i} = \delta_0 + \delta_s s3_{p,i} + \rho_p + \epsilon_{p,i}, \tag{5}$$

where $s_{p,i}$ is an indicator for whether player p hit shot i, $s3_{p,i}$ is an indicator for whether player p hit his previous three shots before shot i, and ρ_p represents player fixed effects. In some specifications, Miller and Sanjurjo (2021) also interact the player fixed effects with session (round by year), and also include shooting location and ball number fixed effects.

For specifications including all shots, the authors simulate regression coefficients as if there was no hot hand using each player's shooting percentage and use this to correct for dynamic panel data bias (Nickell, 1981; Arellano and Bond, 1991; Nerlove, 1967, 1971). Miller and Sanjurjo (2021) do not run this correction when examining individual shot numbers within each location. This bias will likely not exist when examining individual shot numbers, as the lagged shots will not include the previous shot number. For example, the first shot in each location will include shots three, four, and five from the previous location in $s3_{p,i}$, but not the first shot from the previous location.

As we describe above, the amount of dynamic panel data bias depends both on the number of time periods within each panel and on the correlation between the lagged and current levels of the dependent variable (Nickell, 1981). Using simulated data to correct for bias may correct for the level of bias which depends on the number of time periods, but may not correctly account for bias that depends on the correlation between the lagged dependent variable and dependent variable.

Another difference between our analysis and the analysis in Miller and Sanjurjo (2021) is the indicator variable for hitting the three previous shots. As the authors' indicator looks backward over the three previous shots, the first three shots for each player in each session should have missing values for the indicator for hitting three previous shots. However, this indicator is coded as "0" for the first three shots in each session. Since the first three shots have a much lower shooting percentage than later shots in each round (0.41 vs. 0.54), this coding creates a bias toward finding a hot hand effect. This bias is especially large when examining the first few shots in each location. A final important difference is Miller and Sanjurjo (2021) restrict many of their regressions, in particular those examining individual shots in each shot location, to players who shoot at least 100 shots in the NBA 3-Point Contests. This primarily drops players who appear in a small number of contests and did not make it past the first round of the contest.

In Table A4, we recreate the main specifications from Figures 9 and 10 in Miller and Sanjurjo (2021). We then show the effects of different specifications, mapping changes from Miller and Sanjurjo (2021) to our results. In the first row, we replicate the main specification in Miller and Sanjurjo (2021) Figures 9 and 10, which include player fixed effects and restrict the sample to players taking at least 100 shots.[9] In the next row, we switch from the fixed effects model and the bias correction to instead using the leave-out 3-point percentage and quadratic year function, as well as fixed effects for round, location, and in the first column,

[9]Miller and Sanjurjo (2021) adjust the coefficients from Figure 9 to account for dynamic panel data bias but do not follow the same procedure for the shots from the specific shot locations, as lagged dependent variable bias is not a concern in individual shot numbers. We follow their methodology in the first row.

shot number within each location. This does not have a large impact the results, although some of the coefficients for individual shot locations change slightly.

In row three, we change the coding of the lagged shots to not include zeros when players have not yet taken three previous shots. This reduces the magnitude of the overall hot hand effect by just over one percentage point in aggregate. However, this has a very large effect for the first shot in each location, reducing the magnitude of the effect by three percentage points from 8.7 percentage points to 5.8 percentage points. The magnitude of the effects for the second and third shots in each location are also reduced by 2.5 percentage points and 1.5 percentage points, respectively. In row four, we retain the original shot streak coding from rows one and two, but we remove the restriction on players who shoot more than 100 shots. This does not have a large effect on the overall estimate of the hot hand effect, but it does change some of the results for the individual shots. In particular, including all players again reduces the magnitude of the hot hand effect for the first shot in each location by over 50 percent, from 8.7 percentage points in row two to 4.0 percentage points. The effects for later shots within a location grow in magnitude while having smaller standard errors.

In row five, we combine the changes from rows three and four together. We see an overall hot hand effect across all shots, and the results roughly fall between rows four and five for the second and third shots, which are affected by the changes in both rows three and four. However, the coefficient for the first shot in each location is now extremely small at 1.8 percentage points, more in line with our results in Table 3. In row six, we add additional controls for the time in between each shot and the overall shot number in each round, which do not markedly change the results from row five. Finally, in row seven we change from the indicator for hitting three previous shots, which implicitly uses any other combination of hits and misses in the previous three shots as the control group, to the interaction model described in Equation 3. We show only the results for hitting three shots in a row compared to hitting no shots to compare the results to the previous rows. The coefficients for all shots and shot one are identical results in Table 6, Panel A and B. In general, the effect of using the interactions is to increase the statistical power of the results. The hot hand effect is strong for the third and fourth shots in each location, while the coefficient for the first ball in each location is extremely small.

The results here suggest that, while the overall hot hand effect is robust to different specification choices, the finding of a large hot hand effect for the first shot in each location is very sensitive to two factors in particular: (1) the choice of coding shot streaks for the first three shots as zero instead of missing; and (2) the choice to restrict the sample to players shooting more than 100 shots. Either of these choices by itself removes the statistical significance of the hot hand effect for first shots in each location, but the combination of these two effects reduces the size of the effect of making the previous shot on the probability of making the current shot from 8.7 percentage points to 1.8 percentage points while keeping the standard errors relatively constant. Furthermore, using indicators for every combination of makes and misses in the previous three shots instead of just using an indicator for making the previous three shots reduces the hot hand effect to -0.6 percentage points. It is worth noting that the baseline for our comparison is a player missing his previous three shots instead of any other combination of makes and misses in the previous three shots aside from a streak of three hits. We would expect that the change in the interpretation of the coefficient would increase the estimate for the hot hand effect since our control group no longer includes players who made only the previous shot, which we found to have a significant positive effect. This makes the -0.6 percentage point effect even more surprising when compared to the 1.8 percentage point effect.

Table A4: Replication of Miller& Sanjurjo (2021)

	All Shots	Specific Shot Number in Location				
		1	2	3	4	5
1. Miller & Sanjurjo Specifications	0.065***	0.086*	0.084*	0.068	0.062*	0.039
	(0.021)	(0.044)	(0.049)	(0.044)	(0.035)	(0.037)
2. Use LO Average & Drop 2020	0.065***	0.087**	0.076*	0.062	0.072**	0.041
	(0.020)	(0.040)	(0.045)	(0.042)	(0.036)	(0.034)
3. Change Coding of Shot Streak	0.052***	0.058	0.051	0.047	0.072**	0.041
	(0.020)	(0.041)	(0.045)	(0.043)	(0.036)	(0.034)
4. All Players (with Original Coding)	0.063***	0.040	0.078**	0.074**	0.064**	0.069**
	(0.017)	(0.033)	(0.038)	(0.035)	(0.031)	(0.028)
5. Change Coding of Shot Streak & All Players	0.052***	0.018	0.056	0.056	0.064**	0.069**
	(0.017)	(0.034)	(0.038)	(0.036)	(0.031)	(0.028)
6. Add Other Controls	0.049***	0.015	0.055	0.052	0.053*	0.070**
	(0.017)	(0.034)	(0.038)	(0.036)	(0.031)	(0.028)
7. Use Interaction Streak Specification	0.065***	-0.006	0.038	0.138***	0.109**	0.044
	(0.021)	(0.051)	(0.051)	(0.047)	(0.044)	(0.044)

Notes: The data sources are described in Table 1. The first row replicates the result indicated by a "+" in Miller and Sanjurjo (2021) Figure 9 for all shots and the open circles in Miller and Sanjurjo (2021) Figure 10 for individual shots within each location. Rows two through seven change the specifications as described in Section 7.3. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.