Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Is it a Fallacy to Believe in the Hot Hand in the NBA Three-Point Contest?

Joshua B. Miller and Adam Sanjurjo

Working Paper n. 548

This Version: June 11, 2015

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Is it a *Fallacy* to Believe in the Hot Hand in the NBA Three-Point Contest?

Joshua B. Miller[a] and Adam Sanjurjo[b] *†

June 11, 2015

Abstract

The hot hand fallacy refers to a belief in the atypical clustering of successes in sequential outcomes when there is *none*. It has long been considered a massive and widespread cognitive illusion with important implications in economics and finance. The strongest evidence in support of the fallacy remains that from the canonical domain of basketball, where the widespread belief in the existence of hot hand shooting, among expert players and coaches, has been found to have no evidential basis (Gilovich, Vallone, and Tversky 1985). A prominent exhibit of the fallacy is Koehler and Conley (2003)'s study of the NBA Three-Point Contest (1994-1997), a setting which is viewed as ideal for a test of the hot hand (Thaler and Sunstein 2008). In this setting, despite the well-known beliefs of players, coaches, and fans alike, Koehler and Conley find no evidence of hot hand shooting. In the present study, we collect 29 years of shooting data from television broadcasts of the NBA Three-Point Contest (1986-2015), and apply a statistical approach developed in Miller and Sanjurjo (2014), which is more powered, contains an improved set of statistical measures, and corrects for a substantial downward bias in previous estimates of the hot hand effect. In contrast with previous studies, but consistent with Miller and Sanjurjo (2014)'s recent finding of substantial hot hand shooting in all previous controlled shooting studies (including that from the original study of Gilovich, Vallone, and Tversky), we find substantial evidence of hot hand shooting in the NBA Three-Point Contest. This leaves little doubt that the hot hand not only exists, but actually occurs regularly. Thus, belief in the hot hand, in principle, is not a fallacy.

JEL Classification Numbers: C12; C14; C91; C93; D03.
Keywords: Hot Hand Fallacy; Hot Hand Effect.

*a: Department of Decision Sciences and IGIER, Bocconi University, b: Fundamentos del Análisis Económico, Universidad de Alicante. Financial support from the Department of Decision Sciences at Bocconi University, the Spanish Ministerio de Ciencia y Tecnología and Feder Funds (SEJ-2007-62656), and the Spanish Ministry of Economics and Competition (ECO2012-34928) is gratefully acknowledged.

†Both authors contributed equally.

1 Introduction

On January 23, 2015, professional basketball player Klay Thompson, of the NBA's Golden State Warriors, hit 13 consecutive shots, 9 of which were taken from long distance ("3-point" range). Just three weeks later, his teammate, Stephen Curry, hit 13 consecutive shots in the all-star weekend 3-point shooting contest. In each case, teammates (and other observers) universally described the player as being "hot," i.e. in a state of elevated performance. Consistent with these beliefs, the evidence suggests that players are more likely to allocate the ball to a teammate who they think has a hot hand.[1] As an example, in a game that occurred within weeks of Thompson's performance, LeBron James, whom many consider the world's best player, spontaneously changed the design of plays during a time stoppage in order to continue targeting his teammate Kevin Love; when asked about this after the game, James replied: "He had the hot hand, I wanted to keep going to him."

The academic consensus, however, following nearly 30 years of research, is that the beliefs of these players is clear evidence of a "massive and widespread cognitive illusion," by which people perceive an atypical clustering of successes when there is *none* (Kahneman 2011).[2] Accordingly, the *hot hand fallacy* has been considered, more generally, as a candidate explanation for various puzzles and behavioral anomalies identified in a variety of domains, including financial markets, sports wagering, casino gambling, and lotteries.[3]

[1] See Aharoni and Sarig (2011); Attali (2013); Bocskocsky, Ezekowitz, and Stein (2014); Cao (2011); Neiman and Loewenstein (2011); Rao (2009).

[2] The original evidence against the existence of a hot hand in basketball involved NBA field goal and free throw data, and a controlled shooting study with collegiate players (Gilovich et al. 1985). Many researchers have criticized the statistical methods used in the original study (Albert 1993; Albert and Williamson 2001; Arkes 2013; Dorsey-Palmateer and Smith 2004; Hooke 1989; Korb and Stillwell 2003; Miyoshi 2000; Stern and Morris 1993; Stone 2012; Swartz 1990; Wardrop 1999). Typically, these criticisms have highlighted the improper use of field goal data, due to its lack of control. The use of free-throw data has also been criticized for not being representative of hot-hand shooting (Koehler and Conley 2003). Recent studies have revisited field goal and free throw data. While there are unavoidable limitations when studying the field goal and free throw data (see Miller and Sanjurjo (2014)), these studies are important because their results are the opposite of what was found in the original study. Bocskocsky et al. (2014) use the most extensive set of controls available yet for NBA field goal data, and find an average effect consistent with hot hand shooting. For free throw data there is evidence that players shoot better on their second shot in a pair, for pairs in which the first shot is a success (Arkes 2010; Wardrop 1995; Yaari and Eisenmann 2011). There is a similar literature in baseball, but a recent novel and careful analysis shows that hot streaks exist (Green and Zwiebel 2013).

[3] The work in financial markets includes Barberis and Thaler (2003); De Bondt (1993); De Long, Shleifer, Summers, and Waldmann (1991); Kahneman and Riepe (1998); Loh and Warachka (2012); Malkiel (2011); Rabin and Vayanos (2010), in sports wagering Arkes (2011); Avery and Chevalier (1999); Brown and Sauer (1993); Camerer (1989); Durham, Hertzel, and Martin (2005); Lee and Smith (2002); Paul and Weinbach (2005); Sinkey and Logan (2013), in casino gambling Croson and Sundali (2005); Narayanan and Manchanda (2012); Smith, Levere, and Kurtzman (2009); Sundali and Croson (2006); Xu and Harvey (2014), and in lotteries Galbo-Jørgensen, Suetens, and Tyran (2013); Guryan and Kearney (2008); Yuan, Sun, and Siu (2014)

Recent studies by Miller and Sanjurjo (2014, 2015), however, identify critical limitations in the analysis of the canonical study, which call its conclusions into question. The limitations include an estimation procedure which substantially biases estimates of the hot hand effect downward, as well as an inability to control for confounding factors in its study of game data. As a consequence, Miller and Sanjurjo (2014) develop an unbiased (and more powered) statistical testing procedure which, when applied to the least confounded data from the original study—that from the controlled shooting experiment—finds strong evidence of the hot hand, in contrast with the canonical results. Further, Miller and Sanjurjo (2014) conduct a controlled shooting field experiment with semi-professional shooters which (1) identifies hot hand shooters, (2) finds that hot hand shooters can be predicted out-of-sample, based only on prior performance, and (3) reveals that teammates, with no knowledge of a player's performance in the task, can predict which players have a greater tendency to exhibit hot hand shooting.

While Miller and Sanjurjo (2014)'s results suggest that hot hand shooting should be expected to occur in games, direct identification in that environment may not be feasible given the many potential confounds present.[4] Though this conclusion is perhaps disappointing, fortunately, there does exist a shooting environment in which (1) the type of shooting found in NBA games is approximated, and (2) several of the crucial features of a controlled shooting experiment are satisfied, so that many of the confounds present in games are eliminated. The environment of interest is the annual NBA all-star weekend three point shooting contest, which has been described by Thaler and Sunstein (2008) as "an ideal situation in which to study the hot hand." In the shootout, a group of the NBA's best shooters are selected to compete against one another in multiple rounds. In the first round each shooter is given 60 seconds to shoot 25 balls—5 at each of 5 racks which span the three point line, and are roughly equally spaced. More successful shooters then advance to compete in additional rounds, in which the task is the same as in the first. As in games, players shoot on an NBA court, under high stakes, and in front of a large crowd and television audience. Unlike in games, each shot is taken from (roughly) the same distance, in close temporal proximity, and without many of the strategic confounds present in games (variable defensive pressure, offensive and defensive play calls, etc.). Koehler and Conley (2003) study four years of NBA shootout contest data and, despite the well-known beliefs of players, coaches, and fans, they find no evidence of a

[4] For a thorough discussion, which includes examples with game data, and supporting literature, see Miller and Sanjurjo (2014).

hot hand in performance, which is consistent with the canonical results of Gilovich et al. (1985).

In this study we collect 29 years of NBA three point shootout data from television broadcasts, and apply the novel empirical approach of Miller and Sanjurjo (2014), which is more powered, contains an improved set of statistical measures, and corrects for a strong downward bias in estimates of the hot hand effect that Miller and Sanjurjo (2015) uncovered in previous analyses. In contrast with the results of Koehler and Conley (2003), we find substantial evidence of the hot hand: many individual NBA shooters have a statistically significant (and substantial) hot hand effect, and the number is far more than would be expected by chance. In particular, 28 of the 33 players shoot better than usual immediately following three or more hits in a row, with many players' shooting performance increasing by double digits (in percentage points) when on a hit streak (see Table 1 in the results section). Further, 8 of these 33 players have hot hand statistics that are significant at the 5 percent level, a highly significant finding itself ($p < .001$, binomial test). Further, despite heterogeneity in the size (and sometimes even the sign) of the hot hand effect across individual players, in a pooled analysis of all players, on average, players shoot significantly better when on a streak of hits, by around 6 percentage points, controlling for the inherent bias in this measure.[5] These effect sizes are substantial, as the difference between the median and very best three point shooter in the 2013-2014 NBA season was 10 percentage points. Finally, by also testing for the cold hand, as in Miller and Sanjurjo (2014), we can identify whether observed streak shooting is due to the hot hand or the cold hand. Indeed, there is substantial evidence of hot hand shooting, and little evidence of cold hand shooting.

Thus, strong evidence of hot hand shooting has now been discovered within and across individual shooters, and on the pooled level in (1) all extant controlled shooting experiments (Miller and Sanjurjo 2014), and (2) an intermediate shooting environment which lies somewhere between a controlled shooting experiment and an NBA game, despite previous studies having found no evidence of the hot hand in these domains. Further, while recent analyses of in-game shooting data are subject to the aforementioned confounding factors, and limited to pooled analyses, their finding of a slight positive pooled hot hand effect is qualitatively in line with what we have found in the controlled shooting experiments and the NBA three point contest (Arkes 2010; Bocskocsky et al.

[5] To illustrate the magnitude of the bias, if we were to apply the measure of the hot hand effect used by Gilovich et al. (1985), and later Koehler and Conley (2003), which has a stronger bias, to a sample of 25 shots, from a 50 percent shooter, we would underestimate the hot hand effect by 26 percentage points, on average, and for a sample of 100 shots, by 8 percentage points, relative to the unbiased measure.

2014; Yaari and Eisenmann 2011), and presuming a degree of heterogeneity in the size (and sign) of the hot hand effect across individuals that is similar to what we observe in the NBA shootout, and in controlled shooting studies (Miller and Sanjurjo 2014), any average effect will understate the size of the hot hand effect present in particular individuals. The hot hand effect has thus been found to be robust across differing shooting environments and degrees of expertise—from college players, to semi-professionals, to professionals. When taken together, this body of evidence leaves little doubt that the hot hand not only exists, but actually occurs regularly, and thus belief in the hot hand, in principle, is not a fallacy.

A secondary implication of these results is that the term *hot hand bias* should perhaps be introduced to the literature, in order to allow for the possibility that a bias in perception exists that is less extreme than the one represented by the hot hand fallacy. For example, one might be correct in believing that the hot hand exists, but overestimate either its frequency of occurrence, or its magnitude when it occurs. Whether such a bias exists, and to what extent it can be found in experts, versus amateurs, are empirical questions to be addressed in future work, and which we discuss a bit further in Section 5. The primary implications of the results presented in this paper, and Miller and Sanjurjo (2014), are that an extreme version of this bias—the hot hand fallacy—in which one strongly believes that the hot hand is sometimes present, when it in fact does not exist, is resoundingly false.

Section 2 describes our data, Section 3 our empirical strategy, Section 4 our results, and Appendix A reports robustness checks on the analysis presented in Section 4.

2 Data

NBA Three point shooting contest

The annual NBA all-star weekend three point shooting contest began in 1986, with the following format. Eight of the NBA's best three point shooters are selected to participate in a tournament consisting of three rounds. In each round, each participant has the opportunity to shoot 25 times—5 times from each of 5 rack locations which span the three point line, and are roughly equally spaced. The three point line is 22 feet (6.7 meters) from the basket in the corners of the court, a distance that increases along a 14 foot (4.3 m) line that runs parallel to the sidelines on either side.

At the 14 foot mark the three point line becomes an arc 23.75 feet (7.2 m) from the basket. The shooter has 60 seconds to complete as many of these 25 shots as possible. All misses are worth zero points, while hits (makes) are worth either one point (first four balls on each rack) or two points (fifth ball on each rack). The four players with the highest point totals in round one advance to round two.[6] Those two players with the highest point totals in round two advance to the finals, and the player with the highest point total in the finals wins the contest ($10,000)—along with the indirect benefits that are likely to result from the strengthening of one's reputation as an elite shooter.

From the first contest in 1986 to the most recent, in 2015, the format of the contest has experienced some minor changes. The number of participants was 6 during the contests of 2003-2013 rather than the usual 8, with the number of rounds initially three (1986-1998), and later two (2000-2015). The distance of the three point line changed for the 1995-1997 contests to 22 feet (6.7 m) along the entire arc, the number of two-point balls changed from 5 to 9 for the two most recent years (2014, 2015), and the cash prize of the winner is now $35,000, with $22,500 to 2nd place, $15,000 to 3rd place, and $4,500 for 4th to 6th place.[7]

Discussion

Because the hot hand fallacy refers to the purportedly fallacious beliefs held by basketball practitioners (and observers) regarding expert shooters' performance, it might seem at first glance that the most obvious approach is to test for the hot hand in shooting data from professional games. Nevertheless, as explained in Miller and Sanjurjo (2014), the many potential confounds present in game situations may make it impossible to identify, or rule out, the hot hand with this data. On the other hand, one can consider a controlled shooting experiment, as the one conducted in Miller and Sanjurjo (2014), in which expert shooters take each shot under conditions that are free of strategic considerations, and otherwise as close to identical as possible, in terms of location, difficulty, temporal proximity, incentives, physical state, etc.

Relative to NBA games and controlled shooting experiments, the NBA three point shootout

[6] In the case of a tie between two shooters in contention to move to the next round, 24 second elimination rounds, with the rules otherwise the same, are conducted until one participant has a strictly lower score than the other, in which case he is eliminated.

[7] Also, the amount of seconds given to shooters in tie-breaking elimination rounds has varied from 24, to 30, to 60 seconds.

can be viewed as allowing identification of the hot hand to an intermediate degree, as it offers more control than in-game shooting, but less than controlled shooting experiments. In particular, the NBA shootout is free of most of the strategic considerations present in games (variable defensive pressure, offensive and defensive play calls, etc.), while at the same time providing a relatively large volume of similar shots, taken in close temporal proximity. On the other hand, the shootout offers relatively less control than shooting experiments in certain respects. For example, incentives are arguably not constant across shots, given the tournament structure and varied value of balls, shooters sometimes shoot in the same location as the previous shot, but other times not, the 25 shots taken in each shooter round are few compared to what is obtained in experiments, and sessions are separated across significant periods of time (even up to years for some return shooters).

Meanwhile, the NBA three point contest shares some features with NBA games that are absent in a controlled shooting experiment. In particular, the stakes are substantial, and feedback from the crowd and player cohort may affect the players' performance.

Data Collection

We were able to collect video records from every annual NBA three point shootout telecast since the first in 1986, up to the most recent in January 2015, though two of the telecasts were incomplete.[8] This gives us shots from 29 years of the contest, and 105 players, with an average of 78 shots per player (33 players take at least 100 shots).

Data for each shooting session was coded from the original televised video footage, and consisted of hits (1), misses (0), and the time elapsed between shots. Shots that were taken after the time in a session had expired were not included. In a few rare cases, players stepped on the line, but were typically unaware of this until after the session had ended. We included these shots, though the contest did not, as we judged the shots to be taken under near identical conditions (a few centimeters difference) as shots in which the player was just slightly behind the line.[9] The summary statistics of the derived data were cross-validated with Table 1 of Koehler and Conley

[8] In particular, we were unable to obtain complete video records for the first two rounds of the first contest (1986; 14 sessions lost) given that the producers cut out most of the shots from the telecast. Also, we were unable to obtain six first round sessions from the 1993 contest. There was no contest in 1999.

[9] An even rarer occurrence was when the video footage did not clearly display the outcome of a shot, as in the case of a momentary obstruction in front of the camera. In these few cases the combination of the announcers' audio data, the player's final point total, and the televised scoresheet (in the years that this was available) was always enough to confidently infer the outcome of the shot.

(2003), which summarizes individual player performance in four contests between 1994-1997, with the table successfully reproduced up to minor discrepancies.[10]

3 Empirical Strategy

We first briefly describe the empirical approach used by Koehler and Conley (2003), and some of its important limitations. Then we explain our empirical approach and how it improves on those of previous studies.

Koehler and Conley (2003)'s analysis consists of two types of tests, each with two variations. The first type involves intuitive measures of the hot hand effect size, comparing each player's shooting performance (relative frequency of hits) when on a streak of three or more hits, in test (1) to his base rate shooting performance, and in test (2) to his shooting performance when on a streak of three or more *misses*, in both cases stitching together all of a player's shooting sessions and treating the composite sample as if it came from a single session. The second type is the common test used for the degree of clustering in shot outcomes, the *runs* test, in which a run is defined as a streak of consecutive hits, or misses, and the test compares the total number of observed (sample) runs to what is expected under the null hypothesis of a player who shoots with a constant probability of success (i.e. a Bernoulli shooter). The hot hand (or cold hand) hypothesis predicts fewer total runs than expected.[11] In test (3) the authors conduct the runs test once for each shooter, stitching together all of a shooter's 25 shot sessions, and in test (4) the authors conduct a runs test on each 25 shot session, treating all sessions (across shooters) as if they were generated by independent shooters.

The most serious concern regarding Koehler and Conley (2003)'s approach is that their tests of conditional shooting performance, tests (1) and (2), which represent the typical measures of hot hand effect size used in the literature, severely bias estimates of the hot hand effect downward.[12]

[10] The data from Koehler and Conley (2003) no longer exists. We thank Jonathan J. Koehler for searching his records for us. The discrepancies we find are (1) we record a round from Steve Kerr which was not recorded by Koehler and Conley, (2) Koehler and Conley record a round from Dennis Scott (1996 round 3) which we do not record, (3) For 4 other players (Miller, Rice, Legler, Scott) our records differed by 1 or 2 shots, which could be attributed to different coding rules for shots taken while a player's foot is on the line.

[11] For example, in the 10 shot sequence 1101000111 there are five runs: "11," "0," "1," "000," and "111," whereas in the "streakier" shot sequence 1111110000 there are only two: "111111" and "0000."

[12] Miller and Sanjurjo (2015) discovered this problem in regards to (2), which was the intuitive effect size estimate used first by Gilovich et al. (1985).

The intuition behind the bias is easier to see for test (1), in which conditioning on a streak of three or more hits creates a selection bias in which these hits are removed from the sample, leaving a smaller fraction of hits, thus driving conditional performance on the subsequent shot below the base rate. While the fact that Miller and Sanjurjo (2015) first detected this bias after it existed unnoticed in the literature for nearly 30 years, suggests that it is difficult to detect, the magnitude of the bias is substantial; for example, in a 25 shot session, a 50 percent shooter would be expected to have a difference in conditional hit rates, when on a streak of hits vs. misses (test 2), of negative 26 percentage points (instead of zero); in a 100 shot session, negative 8 percentage points.[13] Other concerns, also first recognized by Miller and Sanjurjo (2014), are that tests (2), (3), and (4) allow false negatives and false positives of hot hand performance, as among other issues, they provide no way of separating hot hand from cold hand (shooting worse after several consecutive misses) shooting, and (2) cannot separate a Bernoulli shooter from one that performs equally better following several hits *and* following several misses.

By contrast, the empirical approach of Miller and Sanjurjo (2014) is able to measure the effect of hit streaks on players' conditional shooting performance without bias, has been shown to increase statistical power relative to those of previous studies, and is the first to separate hot hand from cold hand shooting. We follow Miller and Sanjurjo (2014)'s empirical approach here, which also adds the benefit of allowing a direct comparison of the shooting performance of NBA shootout contestants with that of expert shooters in all extant controlled shooting studies.

We use a set of five statistics: (1) the *hit streak frequency* (H_F) measures the relative frequency of shots taken immediately following three or more consecutive hits (made shots), i.e. "how often a shooter is on a streak," (2) the *hit streak momentum* (H_M) measures the relative frequency of hits on shots taken immediately following three or more consecutive hits (made shots), i.e. "how well a shooter performs when on a streak," (3) the *hit streak length* (H_L) measures the length of the longest streak of consecutive hits, which was previously used by Wardrop (1999), (4) a composite hit streak statistic which is the first principal component of all three hit streak statistics (H_C), and (5) the standard runs statistic (R). For each of our hit streak statistics H_F, H_M, H_L, and H_C we have an analogous miss streak statistic M_F, M_M, M_L, and M_C, which we can use to test for a simultaneous presence of hot hand and absence of cold hand, which allows us to test with the

[13] The bias in measure (2) is roughly twice that of (1). In Section 4 we compare our unbiased measure to (1), the less biased measure used by Koehler and Conley (2003).

possibility of complete identification of the hot hand.

Our statistical test procedure, the details of which can be found in Miller and Sanjurjo (2014), starts with the standard null hypothesis (H_0) that a player does not get hot (or cold), and thus the player's shooting performance is a sequence of iid Bernoulli trials with a fixed probability of success. Therefore, while a player's true success rate is unknown to the experimenter, under H_0, and conditional on the number of observed successes, the shot outcomes are *exchangeable*, i.e. all orderings of the shot outcomes are equally likely. This means that for a single player's realization of a sequence of shots, an exact (discrete) distribution exists for each statistic outlined above, under H_0, and by exchangeability; enumerating each permutation of the player's shots, and calculating the value of the test statistic gives this distribution. With this, it becomes clear why our test of a shooter's performance when on a hit streak is unbiased, unlike the analogous tests of Gilovich et al. (1985) and Koehler and Conley (2003), as in our approach the selection bias "cancels out" by being present both in observed H_M as well as the calculated value of the test statistic for each permutation that makes up the distribution. Throughout our analysis we report one-sided p-values for all of our tests, as the alternative hypothesis of hot hand shooting establishes a clear ex ante directional prediction for the statistical measures we define.

Due to the few number of shots in each shooting session (25), and the well-documented power issues concerning small samples of binary data (Miller and Sanjurjo 2014), in our primary analysis we take a similar approach to Koehler and Conley (2003) and measure each statistic across an individual shooter's multiple shooting sessions, permuting at the player level.[14] Further, in order to reduce the noisiness of our estimates, and allow at least a minimal number of instances in which a player has hit three or more shots in a row, in our primary analysis we restrict the sample of shooters to those who took at least 100 shots—the number at which GVT's analysis of controlled shooting sessions is well-known to be considerably underpowered (Miller and Sanjurjo 2014)—which leaves us with 33 of the original 72 shooters.[15,16] Also, like Koehler and Conley (2003), we treat all

[14]Because permuting at the player level, across sessions, can create a selection bias if there are idiosyncratic session effects in performance, in Appendix A we repeat our analysis, but instead permuting on the player, year and round level.

[15]Koehler and Conley (2003) do something similar in their analysis of shooting performance on streaks, by only considering shooters who had at least five three-hit sequences. By taking 100 shots as our cutoff we reduce the relative noisiness of estimates considerably, as a 50% shooter is expected to generate 12 three-hit sequences in every 100 shots. This cutoff also facilitates comparison of effect sizes with those of the Gilovich et al. (1985) data presented in Miller and Sanjurjo (2014).

[16]Players who have taken more shots have generally been more successful in the shootout, which means that they

one point and two point shots as taken under the same conditions.

4 Results

We begin with a basic description of overall shooting performance, then perform a detailed analysis of shooting performance on the individual player level. Our individual-level analysis follows two separate approaches: first, we estimate the magnitude of the hot hand (and cold hand) effect in each player; second, we test whether the number of individual shooters who exhibit hot hand (cold hand) performance is greater than the number expected by chance. Finally, we conduct a pooled analysis of the player-clustered average for each of our statistics. In all cases we find strong evidence of hot hand shooting.[17]

Overall Performance

The (player-clustered) average shooting percentage among the players with at least 100 attempts is 54 percent, with a minimum of 43 percent (Danny Ainge) and a maximum of 66 percent (Jason Kopono).[18] The average number of shots taken is 166, with a minimum of 100 (multiple players) and a maximum of 454 (Craig Hodges). As is the case in all extant controlled studies (Miller and Sanjurjo 2014), as well as free throw shooting (Arkes 2010; Yaari and Eisenmann 2011), players perform relatively (and significantly) worse on their first two shot attempts, shooting 26% on their first shot, 39% on the second, and 56% percent thereafter. On average, players do not shoot significantly better or worse from any of the five rack locations. For the fifth shot from each rack, which is worth two points (1986-2013), players take significantly more time to shoot the ball (2.21 sec vs. 2.01 sec), but their average performance is no different (56% for both). Further, shooting performance does not systematically change over time: (player-clustered) average shooting performance taken in the first half of shots in a round (55%) does not significantly differ from that taken in the second (57%), and there are no significant trends across rounds and years (.02 average

typically have higher relative frequencies of hits, which, under the null that each player shoots with a constant probability of success, and our standardized test procedure, does not create any selection bias.

[17] Appendix A.1 presents the following alternative analyses of the data: (1) shots are permuted with stratification occurring on the player, year, and round level, rather than the player level, and (2) shooters who took fewer than 100 shots are included. For (1) results are similar, and remain highly significant. For (2) results are similar, but slightly attenuated.

[18] For the 72 players with fewer than 100 shots, whom we analyze in Appendix A.1 Table 4, the average shooting percentage was 44 percent.

serial correlation).

Because of the "warm-up" effect observed in the first two shots of each round, we exclude these shots, as in (Miller and Sanjurjo 2014). Once this exclusion is made, the otherwise comparable levels of average performance across the variables mentioned above give no reason to suspect that the approach of Koehler and Conley (2003), which treats all shots from each shooter as if they were taken under the same conditions, is problematic. Thus we follow this approach, which allows us to use the majority of the shot data collected. We report all tests as one-sided, as the hot hand hypothesis makes a clear directional prediction.

Individual Performance

Statistical tests on small samples of binary outcomes are considerably underpowered, as demonstrated in the numerous power analyses conducted in Miller and Sanjurjo (2014). While the average of 166 shots taken in the NBA's 3-point contest is dwarfed by the average of 975 shots taken in the controlled shooting task of Miller and Sanjurjo (2014), the analysis developed in that study has been applied to the data of the original hot hand study of Gilovich et al. (1985), which had just 100 shot attempts from each player, with the outcome being that Miller and Sanjurjo (2014) found significantly more players with extreme hot hand performance than would be expected if players were to have a constant probability of shooting success, thus overturning the results of the original study.

In the third column of Table 1, we report the raw estimate of the hot hand effect size for each player, which indicates the change in the player's hit rate (in percentage points) on those shots which immediately follow a streak of three or more hits, as compared to any other recent shot history.[19] Given that the effect size is expected to be negative for a shooter with a constant probability of success (an expected average difference of -2.6 percentage points), the raw effect size results, alone, are striking: 25 out of 33 shooters have a positive hot hand effect, before correcting for this bias, and the average uncorrected difference (clustered by player) is +3.7 percentage points (pooled statistical tests are reported below).[20] Once the bias is corrected for, the average difference increases to +6.3

[19] This comparison with any other recent shot history, rather than only recent shot histories consisting of streaks of three or more misses, as performed in Koehler and Conley (2003), uses relatively more data, and makes significant positive effects less likely to be attributable to a miss streak rather than a hit streak (see Miller and Sanjurjo (2014) for an in-depth discussion).

[20] A similar analysis of whether shooters perform relatively worse than usual on those shots which immediately follow a streak of three or more *misses* does not yield analogous evidence of cold hand shooting.

Table 1: *Reported for each player are the number of shots taken, the player's overall hit rate, the change in hit rate after hitting three or more shots in a row (estimated and bias-corrected), and the composite statistic for hit streaks and miss streaks. Players are ordered by the* p-value *of their composite hit streak statistic (measured in standard deviation units from the mean, with significance stars corresponding to a player-stratified permutation test).*

			Change in hit rate after hitting 3 or more		Composite Statistic	
player	# shots	hit rate	estimate	bias corrected	hit streaks	miss streaks
Davis, Hubert	125	.53	.31	.34	3.34^{***}	.04
Scott, Dennis	100	.57	.11	.14	2.31^{***}	.51
Richardson, Quintin	100	.46	.17	.22	2.09^{**}	1.16^{*}
Allen, Ray	247	.56	.11	.12	1.95^{**}	−.76
Curry, Stephen	150	.61	.11	.12	1.90^{**}	.92
Hodges, Craig	454	.56	.05	.05	2.02^{**}	.94
Lewis, Rashard	123	.46	.16	.20	1.68^{**}	1.80^{**}
Pierce, Paul	125	.49	.08	.11	1.66^{**}	1.11
Hornacek, Jeff	160	.51	.16	.18	1.50^{*}	1.15
Billups, Chauncey	100	.47	.06	.11	1.34^{*}	.40
Korver, Kyle	125	.54	.12	.15	1.38^{*}	−.06
Price, Mark	199	.60	.06	.07	1.02	.74
Cook, Daequan	100	.56	.02	.05	.95	.01
Durant, Kevin	100	.48	.18	.22	.97	.17
Kopono, Jason	150	.66	.02	.03	.91	−.38
Legler, Tim	150	.63	.14	.16	.91	.15
Ellis, Dale	320	.51	.00	.02	.85	.33
Rice, Glen	197	.49	.01	.03	.77	.69
Porter, Terry	160	.49	.03	.06	.76	−1.25
Bird, Larry	211	.57	.11	.12	.66	−.73
Irving, Kyrie	124	.64	−.05	−.03	.50	1.42^{*}
Kerr, Steve	208	.56	−.00	.01	.40	−1.43
Person, Wesley	136	.60	.02	.04	.30	−.40
Schrempf, Detlef	124	.48	.00	.04	.09	.78
Belinelli, Marco	100	.57	.03	.06	.09	−.55
Jones, James	100	.57	.00	.03	−.02	−.59
Novitski, Dirk	248	.50	−.00	.01	−.20	−.44
Miller, Reggie	235	.53	.01	.02	−.34	−.59
Arenas, Gilbert	100	.56	−.17	−.14	−.37	.49
Barros, Dana	184	.49	−.17	−.15	−.37	−.65
Ainge, Danny	148	.43	−.03	.01	−.72	−1.01
Lenard, Voshon	100	.53	−.14	−.10	−.93	−.86
Stojakovic, Peja	287	.62	−.22	−.22	−1.80	.99
Average	166	.54	.04	.06	.78	.12

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (one-sided, right) percentage points.[21] This is a substantial pooled effect size given that heterogeneity in the size (and sometimes even the sign) of the hot hand effect across individual shooters could, in principle, easily cancel out the presence of hot hand shooters in the average. To further put the magnitude of this *average* hot hand effect into perspective, the difference between the median NBA three point shooter, and the very best, was 10 percentage points, in the 2013-2014 season. In addition, it is important to note that even this number (+6.3) is likely an underestimate of the true average hot hand effect, as the act of hitting three shots in a row is only a noisy signal of a player's underlying performance state (Arkes 2013; Stone 2012).[22]

While the hit streak statistics for frequency, momentum, and length, described in Section 3, could each be calculated and reported for each player, we instead create a composite statistic which is the first principal component of all three hit streak statistics, and report its value for each player in column five of Table 1, and order the players by the *p-value* of this statistic. The composite statistic serves three purposes: (1) it controls for multiple comparisons, (2) it increases power when testing for the hot hand effect in individual shooters, and (3) it allows for the degree of hot hand shooting to be measured by a single statistic. As can be seen in Table 1, the composite hit streak statistic is significant at the 5 percent level for 8 of the 33 shooters, which is a highly statistically significant finding ($p < .001$, binomial test). Column five shows that the analogous composite miss streak statistic is significant at the 5 percent level for only 1 out of 33 shooters. Thus, the hot hand effect is found to be significant even within certain individuals, though the cold hand effect is not. These findings indicate that players have a tendency of getting significantly "hot," and have more of a tendency of getting significantly hot than "cold."

The idea of testing whether more players exhibit hot hand shooting than would be expected if each were to shoot with a constant probability of success can be extended to consider all positive hot hand effects (not just positive and significant effects), and for each of the hit streak statistics, as well as the runs statistic employed by Koehler and Conley (2003). Figure 1 graphically displays these statistics for each of the 33 shooters. In each cell, each shooter's corresponding statistic is plotted against its median under the null (based on the permutation distribution). If the shooter's

[21] The Bias adjustment is computed to three decimal precision.

[22] Sometimes a player will hit three shots in a row purely out of luck, and not because he is momentarily in a superior performance state. This means that when the shots following these streaks of three hits are pooled with shots taken when the player actually is in a superior performance state, the measured effect is diluted with respect to the true hot hand effect.

statistic is on the side of the median (gray line) predicted by the hot hand hypothesis, we will refer to this as the "hot side" of the median; for the three hit streak statistics, H_F, H_M, and H_L, the hot side is above the median, while for the runs statistic, R, it is below. Figure 1 shows that the hit streak frequency (H_F) statistic is on the hot side of the median for 24 of the 33 shooters; this would occur in less than one out of every one hundred studies if the null hypothesis that each player shoots with a constant probability of success were true ($p = .007$, one-sided binomial test). Similarly, the hit streak momentum H_M statistic is on the hot side of the median for 28 of the 33 shooters ($p < .001$, one-sided binomial test), the hit streak length H_L for 23 of the 33 ($p = .018$, one-sided binomial test), and the runs statistic for 29 of the 33 ($p < .001$, one-sided binomial test). By contrast, the analogous miss streak statistics do not have a comparable tendency to cluster on the cold side of the median, as the number of shooters (out of 33) on the cold side of the median for M_F is 19, and for M_M and M_L is 16. Under the null hypothesis that each player shoots with a constant probability of success (i.e. the hot [cold] hand does not exist), the number of cold hand shooters observed in our sample is well within the 95% confidence interval. On the other hand, we would almost never observe as many shooters with directional hot hands as what we observe here.

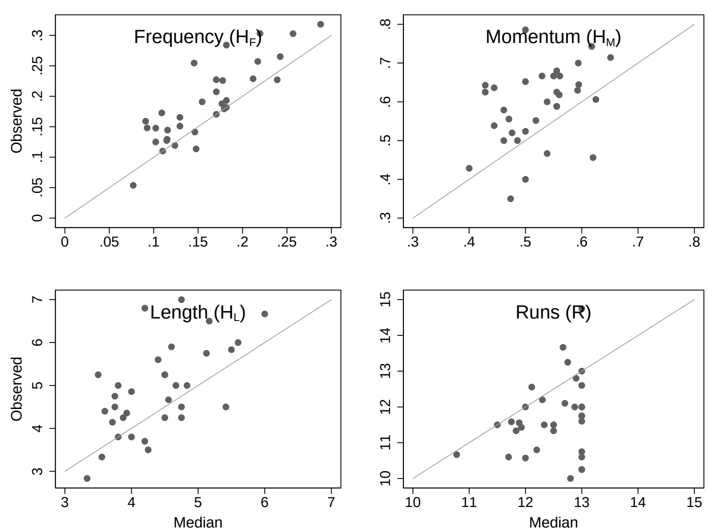

Figure 1: *Observed vs. median hit streak statistics. Permutations are stratified by player.*

Given that each shooting session contained at most 25 shots, and that we do not know precisely how often hot hands occur, or for how long, when they do occur, it is difficult to know whether a shooter performing better in one session than another is because he (1) experienced the hot hand during one session and not the other, or (2) some other reason. As the above analysis was conducted with permutations stratified at the player level, it is possible that specific years and rounds of exceptional performance (due to something other than hot hand shooting) drive the results—an explanation that cannot be ruled out when the null distribution is generated by permuting shots across rounds. As a result, in Appendix A.1 we present Table 3 and Figure 2, which are analogous to this section's Table 1 and Figure 1, respectively, but with permutations instead stratified at the player, year and round level. As can be seen, the results are of comparable strength, which suggests that the substantial evidence of hot hand shooting that we observe in this section is indeed being driven by a hot hand.

Pooled Analysis

In the analysis up to this point we have observed strong evidence of the hot hand, both in terms of the magnitude of the hot hand effect in individual shooters, as well as the number of shooters who exhibit the hot hand. Further, despite considerable heterogeneity in the size (and sometimes even the sign) of the hot hand effect across shooters, we observed a substantial increase of 6 percentage points in average shooting performance on those shots that immediately follow a streak of three or more hits. We now test whether the average effect for each of our hit streak statistics, as well as the runs statistic, are statistically significant. For each statistic, it is important to re-emphasize that even in the presence of strong hot hand effects in certain individual players, in a pooled analysis the presence of shooters with the opposite ("anti-hot hand") effect, or that shoot as if they have a constant probability of success ("consistent hand"), will dilute the average effect with respect to the players with a hot hand, possibly even disguising the presence of hot hand shooting entirely.

In Table 2 we report the hit streak statistics, and runs statistic, averaged across all shooters. Each statistic is reported in standard deviations from the mean, so the value of .92 for H_F, for example, indicates that, on average, each player in our sample finds himself on a hit streak more frequently than would be expected under the assumption that he shoots with a constant probability of success, and that the difference corresponds to nearly a full standard deviation increase in performance under this assumption. As can be seen in the table, all three of the hit streak statistics, as well as the runs statistic, are highly significant.

Table 2: *Pooled test of the hit and miss streak statistics for the NBA's 3 point contest. The average value of the standardized (and player-clustered) statistics, with p-values in parentheses. The test is performed players with at least 100 shot attempts. Permutations stratified by player.*

Statistic	Standard Deviations from the Mean
Hit Streak Frequency (H_F)	.92***
	(.000)
Hit Streak Momentum (H_M)	.62***
	(.000)
Hit Streak Length (H_L)	.78***
	(.000)
Total Runs (R)	−.79***
	(.000)
Miss Streak Frequency (M_F)	.27*
	(.061)
Miss Streak Momentum (M_M)	.02
	(.453)
Miss Streak Length (M_L)	.08
	(.311)

Statistics reported as (player-clustered) average standard deviation from the mean, p-values in parentheses
25,000 Permutations for each player strata.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (one-sided)

What is remarkable about the average hot hand effects is that they are particularly pronounced, despite the considerable heterogeneity in the size (and sometimes even the sign) of the hot hand effect observed in our individual shooters, which as we have mentioned, could have easily allowed for the strong hot hand effect in certain shooters to go undetected in the average performance. Further, it is highly implausible that expert players and coaches (or even the average fans) would believe that all shooters experience the hot hand with the same frequency, and with the same magnitude when it occurs. Therefore, our finding that, *on average*, players exhibit the hot hand (but not the cold hand) is much stronger than the type of claim one might ordinarily expect, such as "sometimes, some players experience a hot hand."[23] Indeed, Miller and Sanjurjo (2014) find, in a detailed questionnaire administered to the expert players who participated in their controlled shooting experiment (and did not observe each others' shooting sessions), that not only do all of

[23]Here, we reproduce evidence of expert players' beliefs in the hot hand from the canonical paper Gilovich et al. (1985), in which the authors interviewed eight players from an NBA team: "Most of the players (six out of eight) reported that they have on occasion felt that after having made a few shots in a row they 'know' they are going to make their next shot—that they 'almost can't miss.' Five players believed that a player 'has a better chance of making a shot after having just made his last two or three shots than he does after having just missed his last two or three shots.' (Two players did not endorse this statement and one did not answer the question.)..."

the eight players believe that many of their teammates generally shoot better when on a streak of hits, but only one player believed that all of them do. Further, both their ratings (-3 to +3), and rankings, of how well their teammates shoot when on a hit streak are highly correlated with the actual shooting performance of the teammates in Miller and Sanjurjo's experiment. When this evidence is taken together with the fact that expert coaches and players have much more information to go by than simply whether a teammate has made or missed the last several shots (e.g. how cleanly shots are entering or missing, the player's shooting mechanics, body language, etc.), what is suggested is that not only can experts identify which shooters have a greater tendency to get the hot hand, but that they may also sometimes be able to tell the difference between a lucky streak of hits and one that is instead due to an elevated state of performance. In sum, the hot hand effects that we measure for each individual shooter are diluted versions of the true effects because we include all shots that follow a three-hit streak in our measure of hot hand shooting, whereas experts' may be able to discriminate, and thus correctly observe a larger hot hand effect. Further, the pooled effects that we observe include shooters who expert players and coaches may (correctly) believe do not get the hot hand at all. Thus the true expected improvement in performance of a player who teammates and coaches identify as having the hot hand could potentially be vastly larger than the lower bound represented by the size of the pooled effects presented here.

Finally, the pooled effect sizes reported here are comparable in size to the also highly significant pooled effect sizes of the same hit streak (and runs) statistics measured in Miller and Sanjurjo's study of all extant controlled shooting data.

5 Discussion

The annual NBA three point shootout has been described as an ideal setting in which to study the hot hand, given that it approximates the conditions of in-game NBA shooting, while ridding of many of the confounds present there. In this environment, in contrast with the results of Koehler and Conley (2003), we find substantial evidence of hot hand shooting, with effect sizes in the double digits (percentage points) for many individual shooters. As in the case of Miller and Sanjurjo (2014)'s study of the canonical Gilovich et al. (1985) controlled shooting experiment, and others, the answer to why we find considerable evidence of the hot hand and the original authors find none is multiple (1) our analysis rids of a substantial downward bias in estimates of the hot hand effect that was present in the analyses of these previous studies, (2) our empirical approach is more powered, and (3) we analyze larger datasets.

The present findings, along with those of Miller and Sanjurjo (2014), indicate that the hot hand effect not only exists in those environments in which the ability to identify it is greatest, but effect sizes are substantial, and in every case it has even been found to be a property of the average shooter. Further, although the difficulties involved with identifying the hot hand in game data means that results in that domain should be interpreted with particular caution, possible hot hand effects are observed there (Arkes 2010; Bocskocsky et al. 2014; Yaari and Eisenmann 2011) which are qualitatively in line with the hot hand effects we observe in the more controlled environments that we study. Further, we find the hot hand effect to be robust across multiple different shooting environments, and degrees of expertise in the shooters—from college level, to semi-professional, to professional.[24]

In light of this body of recent results there is now little doubt that the hot hand not only exists, but actually occurs regularly, and thus that the well-known beliefs of expert players and coaches (as well as common fans) can no longer be considered fallacious. On the other hand, the hot hand fallacy is an extreme bias, given that it requires one to strongly believe that the hot hand sometimes occurs, when in fact, it does not exist. As such, we propose the term *hot hand bias*, to allow for the possibility of a less extreme type of perceptual error, in which, for example, one believes that the hot hand occurs more frequently than it actually does, or in greater magnitude when it does. The degree to which a hot hand bias exists, and if so, to what extent it differs between amateurs and experts (who have greater incentive to eradicate biased beliefs) are empirical questions to be answered in future research.[25]

[24] There are also many studies that make important contributions to the study of streaks and momentum in human performance outside of the basketball setting, but they have had little influence on the hot hand fallacy literature because the beliefs and behavior of decision makers have not been identified (Alter and Oppenheimer 2006). A large literature exists on the presence or absence of the hot hand effect in other sports besides basketball; for a review, see Bar-Eli, Avugos, and Raab (2006); for a meta-analysis, see Avugos, Köppen, Czienskowski, Raab, and Bar-Eli (2013). In addition, in the finance literature there is mixed evidence of a hot hand effect (or "performance persistence") among fund managers (Carhart 1997; Hendricks, Patel, and Zeckhauser 1993; Jagannathan, Malakhov, and Novikov 2010).

[25] While it is reasonable to suspect that some people will over-infer based on the limited information that they receive (Barberis, Shleifer, and Vishny 1998; Burns 2004; Mullainathan 2002; Rabin 2002; Rabin and Vayanos 2010), a substantial level of over-inference would be surprising among highly incentivized experts, given that it would represent a stark deviation from the roughly optimal shooting decisions that have been observed in most situations (Goldman and Rao 2014).

References

Aharoni, G. and O. H. Sarig (2011): "Hot hands and equilibrium," *Applied Economics*, 44, 2309–2320.

Albert, J. (1993): "Comment on "A Statistical Analysis of Hitting Streaks in Baseball" by S. C. Albright," *Journal of the American Statistical Association*, 88, 1184–1188.

Albert, J. and P. Williamson (2001): "Using Model/Data Simulations to Detect Streakiness," *The American Statistician*, 55, 41–50.

Alter, A. L. and D. M. Oppenheimer (2006): "From a fixation on sports to an exploration of mechanism: The past, present, and future of hot hand research," *Thinking & Reasoning*, 12, 431–444.

Arkes, J. (2010): "Revisiting the Hot Hand Theory with Free Throw Data in a Multivariate Framework," *Journal of Quantitative Analysis in Sports*, 6.

——— (2011): "Do Gamblers Correctly Price Momentum in NBA Betting Markets?" *Journal of Prediction Markets*, 5, 31–50.

——— (2013): "Misses in 'Hot Hand' Research," *Journal of Sports Economics*, 14, 401–410.

Attali, Y. (2013): "Perceived Hotness Affects Behavior of Basketball Players and Coaches," *Psychological Science*, forthcoming.

Avery, C. and J. Chevalier (1999): "Identifying Investor Sentiment from Price Paths: The Case of Football Betting," *Journal of Business*, 72, 493–521.

Avugos, S., J. Köppen, U. Czienskowski, M. Raab, and M. Bar-Eli (2013): "The "hot hand" reconsidered: A meta-analytic approach," *Psychology of Sport and Exercise*, 14, 21–27.

Bar-Eli, M., S. Avugos, and M. Raab (2006): "Twenty years of "hot hand" research: Review and critique," *Psychology of Sport and Exercise*, 7, 525–553.

Barberis, N., A. Shleifer, and R. Vishny (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 307–343.

Barberis, N. and R. Thaler (2003): "A survey of behavioral finance," *Handbook of the Economics of Finance*, 1, 1053–1128.

Bocskocsky, A., J. Ezekowitz, and C. Stein (2014): "The Hot Hand: A New Approach to an Old 'Fallacy'," 8th Annual Mit Sloan Sports Analytics Conference.

Brown, W. A. and R. D. Sauer (1993): "Does the Basketball Market Believe in the Hot Hand? Comment," *American Economic Review*, 83, 1377–1386.

Burns, B. D. (2004): "Heuristics as beliefs and as behaviors: The adaptiveness of the "hot hand"," *Cognitive Psychology*, 48, 295–331.

Camerer, C. F. (1989): “Does the Basketball Market Believe in the ’Hot Hand,’?” *American Economic Review*, 79, 1257–1261.

Cao, Z. (2011): “Essays on Behavioral Economics,” Ph.D. thesis, Oregon State University.

Carhart, M. M. (1997): “On Persistence in Mutual Fund Performance,” *Journal of Finance*, 52, 57–82.

Croson, R. and J. Sundali (2005): “The Gamblers Fallacy and the Hot Hand: Empirical Data from Casinos,” *Journal of Risk and Uncertainty*, 30, 195–209.

De Bondt, W. P. (1993): “Betting on trends: Intuitive forecasts of financial risk and return,” *International Journal of Forecasting*, 9, 355–371.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1991): “The Survival of Noise Traders In Financial-markets,” *Journal of Business*, 64, 1–19.

Dorsey-Palmateer, R. and G. Smith (2004): “Bowlers’ Hot Hands,” *The American Statistician*, 58, 38–45.

Durham, G. R., M. G. Hertzel, and J. S. Martin (2005): “The Market Impact of Trends and Sequences in Performance: New Evidence,” *Journal of Finance*, 60, 2551–2569.

Galbo-Jørgensen, C. B., S. Suetens, and J.-R. Tyran (2013): “Predicting Lotto Numbers A natural experiment on the gamblers fallacy and the hot hand fallacy,” Working Paper.

Gilovich, T., R. Vallone, and A. Tversky (1985): “The Hot Hand in Basketball: On the Misperception of Random Sequences,” *Cognitive Psychology*, 17, 295–314.

Goldman, M. and J. M. Rao (2014): “Misperception of Risk and Incentives by Experienced Agents,” Working Paper.

Green, B. S. and J. Zwiebel (2013): “The Hot Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments?” Working Paper.

Guryan, J. and M. S. Kearney (2008): “Gambling at Lucky Stores: Empirical Evidence from State Lottery Sales,” *American Economic Review*, 98, 458–473.

Hendricks, D., J. Patel, and R. Zeckhauser (1993): “Hot hands in mutual funds: Short-run persistence of relative performance,” *Journal of Finance*, 48, 93–130.

Hooke, R. (1989): “Basketball, baseball, and the null hypothesis,” *Chance*, 2, 35–37.

Jagannathan, R., A. Malakhov, and D. Novikov (2010): “Do Hot Hands Exist among Hedge Fund Managers? An Empirical Evaluation,” *Journal of Finance*, 65, 217–255.

Kahneman, D. (2011): *Thinking, Fast and Slow*, Farrar, Straus and Giroux.

Kahneman, D. and M. W. Riepe (1998): “Aspects of Investor Psychology: Beliefs, preferences, and biases investment advisors should know about,” *Journal of Portfolio Management*, 24, 1–21.

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

Korb, K. B. and M. Stillwell (2003): "The Story of The Hot Hand: Powerful Myth or Powerless Critique?" Working Paper.

Lee, M. and G. Smith (2002): "Regression to the mean and football wagers," *Journal of Behavioral Decision Making*, 15, 329–342.

Loh, R. K. and M. Warachka (2012): "Streaks in Earnings Surprises and the Cross-Section of Stock Returns," *Management Science*, 58, 1305–1321.

Malkiel, B. G. (2011): *A random walk down Wall Street: the time-tested strategy for sucessful investing*, New York: W. W. Norton & Company.

Miller, J. B. and A. Sanjurjo (2014): "A Cold Shower for the Hot Hand Fallacy," Working Paper.

——— (2015): "Surprised by the Gambler's and Hot Hand Fallacies? A Proof of the Law of Small Numbers with Implications For Inferring Serial Dependence in Finite Sequential Data." Working Paper.

Miyoshi, H. (2000): "Is the "hot hands" phenomenon a misperception of random events?" *Japanese Psychological Research*, 42, 128–133.

Mullainathan, S. (2002): "Thinking Through Categories," Working Paper.

Narayanan, S. and P. Manchanda (2012): "An empirical analysis of individual level casino gambling behavior," 10, 27–62.

Neiman, T. and Y. Loewenstein (2011): "Reinforcement learning in professional basketball players," *Nature Communications*, 2:569.

Paul, R. J. and A. P. Weinbach (2005): "Bettor Misperceptions in the NBA: The Overbetting of Large Favorites and the 'Hot Hand'," *Journal of Sports Economics*, 6, 390–400.

Rabin, M. (2002): "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 117, 775–816.

Rabin, M. and D. Vayanos (2010): "The Gamblers and Hot-Hand Fallacies: Theory and Applications," *Review of Economic Studies*, 77, 730–778.

Rao, J. M. (2009): "Experts' Perceptions of Autocorrelation: The Hot Hand Fallacy Among Professional Basketball Players," Working Paper.

Sinkey, M. and T. Logan (2013): "Does the Hot Hand Drive the Market?" *Eastern Economic Journal*, Advance online publication, doi:10.1057/eej.2013.33.

Smith, G., M. Levere, and R. Kurtzman (2009): "Poker Player Behavior After Big Wins and Big Losses," *Management Science*, 55, 1547–1555.

Stern, H. S. and C. N. Morris (1993): “Comment on “A Statistical Analysis of Hitting Streaks in Baseball” by S. C. Albright,” *Journal of the American Statistical Association*, 88, 1189–1194.

Stone, D. F. (2012): “Measurement error and the hot hand,” *The American Statistician*, 66, 61–66, working paper.

Sundali, J. and R. Croson (2006): “Biases in casino betting: The hot and the gamblers fallacy,” *Judgement and Decision Making*, 1, 1–12.

Swartz, T. (1990): “Letter to the editor: More on the “hot hand”,” *Chance*, 3, 6–7.

Thaler, R. H. and C. R. Sunstein (2008): *Nudge: Improving Decisions About Health, Wealth, and Happiness*, Yale University Press.

Wardrop, R. L. (1995): “Simpson’s Paradox and the Hot Hand in Basketball,” *The American Statistician*, 49, 24–28.

——— (1999): “Statistical Tests for the Hot-Hand in Basketball in a Controlled Setting,” Working paper, University of Wisconsin - Madison.

Xu, J. and N. Harvey (2014): “Carry on winning: The gambler’s fallacy creates hot hand effects in online gambling,” *Cognition*, 131, 173 – 180.

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, 1–10.

Yuan, J., G.-Z. Sun, and R. Siu (2014): “The Lure of Illusory Luck: How Much Are People Willing to Pay for Random Shocks,” *Journal of Economic Behavior & Organization*, forthcoming.

A Appendix

A.1 Analysis with year and round strata, and all players

On the vertical axis of each cell in Figure 2 we measure the corresponding sample statistic for each player. On the horizontal axis we measure the median value of the null reference distribution, with player, year, and round stratified permutations. In Table 3 we recalculate the values of the composite hit streak statistics, for each player, presented in Table 1, but with the null reference distribution calculated with player, year, and round stratified permutations. In Table 4 we present pooled results analogous to those of Table 2, but with the null reference distribution calculated with player, year, and round stratified permutations.

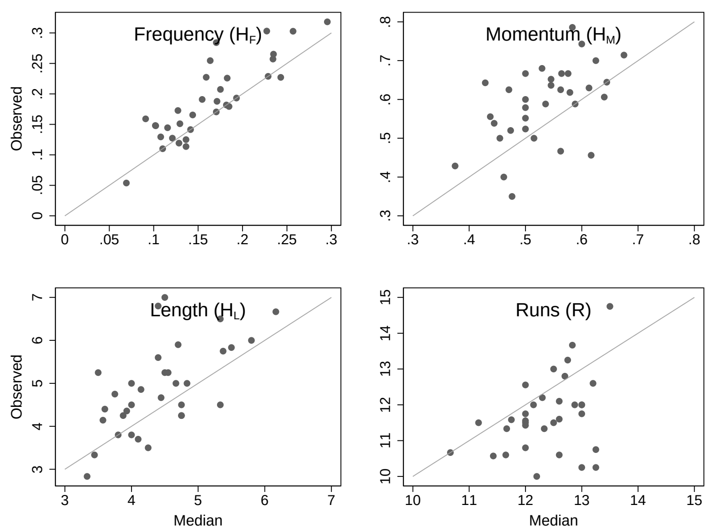

Figure 2: *Observed vs. median hot streak statistics for 3pt, stratified, permutations stratified by player, year, and round.*

Table 3: *Reported for each player are the number of shots taken, the player's overall hit rate, the change in hit rate after hitting three or more shots in a row (estimated and bias-corrected), and the composite statistic for hit streak and miss streaks. Players ordered by the* p-value *of their composite hit streak statistic (measured in standard deviation units from the mean, with significance stars corresponding to a player, year and round stratified permutation test).*

			Change in hit rate after hitting 3 or more		Composite Statistic	
player	# shots	hit rate	estimate	bias corrected	hit streaks	miss streaks
Davis, Hubert	125	.53	.31	.34	3.00***	−.25
Scott, Dennis	100	.57	.11	.14	2.58***	.70
Richardson, Quintin	100	.46	.17	.22	2.15**	1.25*
Hodges, Craig	454	.56	.05	.05	1.87**	.97
Allen, Ray	247	.56	.11	.12	1.86**	−.92
Curry, Stephen	150	.61	.11	.12	1.74**	.85
Korver, Kyle	125	.54	.12	.15	1.59**	.08
Pierce, Paul	125	.49	.08	.11	1.43*	.75
Lewis, Rashard	123	.46	.16	.20	1.35*	1.14
Billups, Chauncey	100	.47	.06	.11	1.35*	.42
Cook, Daequan	100	.56	.02	.05	1.18*	.21
Legler, Tim	150	.63	.14	.16	1.10	.31
Hornacek, Jeff	160	.51	.16	.18	1.08	.58
Porter, Terry	160	.49	.03	.06	1.01	−1.06
Ellis, Dale	320	.51	.00	.02	.93	.35
Rice, Glen	197	.49	.01	.03	.86	.72
Kopono, Jason	150	.66	.02	.03	.70	−.55
Kerr, Steve	208	.56	−.00	.01	.58	−1.35
Bird, Larry	211	.57	.11	.12	.57	−.81
Price, Mark	199	.60	.06	.07	.46	.21
Irving, Kyrie	124	.64	−.05	−.03	.38	1.44*
Durant, Kevin	100	.48	.18	.22	.39	−.92
Person, Wesley	136	.60	.02	.04	.13	−.70
Schrempf, Detlef	124	.48	.00	.04	.10	.66
Belinelli, Marco	100	.57	.03	.06	.03	−.56
Miller, Reggie	235	.53	.01	.02	−.21	−.56
Jones, James	100	.57	.00	.03	−.29	−.85
Barros, Dana	184	.49	−.17	−.15	−.42	−.91
Novitski, Dirk	248	.50	−.00	.01	−.43	−.86
Arenas, Gilbert	100	.56	−.17	−.14	−.50	.50
Ainge, Danny	148	.43	−.03	.01	−.60	−.79
Lenard, Voshon	100	.53	−.14	−.10	−.70	−.67
Stojakovic, Peja	287	.62	−.22	−.22	−1.85	.89
Average	166	.54	.04	.06	.71	.01

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (one-sided, right)

Table 4: *Pooled test of the hit and miss streak statistics for the NBA's 3 point contest. The average value of the standardized (and player-clustered) statistics, with p-values in parentheses.*

	Players with 100+ attempts		All Players	
Statistic	Player Strata	Round Strata	Player Strata	Round Strata
Hit Streak Frequency (H_F)	$.92^{***}$	$.84^{***}$	$.33^{***}$	$.38^{***}$
	(.000)	(.000)	(.001)	(.000)
Hit Streak Momentum (H_M)	$.62^{***}$	$.55^{***}$	$.28^{***}$	$.30^{***}$
	(.000)	(.001)	(.000)	(.000)
Hit Streak Length (H_L)	$.78^{***}$	$.73^{***}$	$.30^{***}$	$.35^{***}$
	(.000)	(.000)	(.002)	(.000)
Total Runs (R)	$-.79^{***}$	$-.70^{***}$	$-.34^{***}$	$-.37^{***}$
	(.000)	(.000)	(.000)	(.000)
Miss Streak Frequency (M_F)	$.27^{*}$	.15	.08	.11
	(.061)	(.200)	(.200)	(.142)
Miss Streak Momentum (M_M)	.02	$-.13$	.06	$-.01$
	(.453)	(.770)	(.297)	(.523)
Miss Streak Length (M_L)	.08	$-.00$	.06	.02
	(.311)	(.502)	(.284)	(.401)

All statistics reported as (player-clustered) average standard deviation from the mean, p-values in parentheses
25,000 Permutations with player, year and round strata.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (one-sided)