The hot-hand fallacy has long been considered a widespread cognitive illusion with important consequences for decision-makers. A sports player may misinterpret a lucky streak of successful shots, and pursue a more aggressive strategy that hurts one’s performance. A physician prone to the hot-hand fallacy may also misinterpret random streaks of poor patient outcomes, and respond by treating the next patient more defensively, incorrectly believing that another poor outcome is likely.

Following Gilovich, Vallone and Tversky (1985), the prevalence of hot-hand fallacy has been studied in a wide range of contexts, including professional sports, gambling, and financial markets. However, there remain few investigations into how this cognitive mistake affects decisions, particularly outside gambling and sports. Evidence suggest that gamblers respond to streaks of outcomes that are evidently random. But when outcomes are truly random, such as which color a roulette ball lands on, these outcomes are not influenced by the player’s decision, and thus a mistaken hot-hand belief is not costly. Put differently, a sophisticated roulette player could very well understand the independent nature of the outcome, and continue to bet on “hot” color just for fun, because it does not hurt to do so. Studies have also found broadly consistent evidence of professional sports players and coaches making strategic adjustment to hot and cold players in team sports such as baseball and basketball.

This paper conducts an empirical investigation into the effects of hot-hand fallacy in the decision-making of two professionals: dart players and physicians. The sport of darts is a compelling environment to study the hot-hand. First, it is a non-reactionary game where players do not interact with each other. This provides a clean test of the hot-hand in the absence of endogenous reactions of other players, and more importantly, we are better able to isolate hot-hand bias by eliminating confounding psychological mechanisms that arise from anticipating opponent strategies. Second, the simplicity of game rules is ideal for studying the hot-hand. That is, while the game is complex enough to induce different strategies (which target to aim), each strategic option has transparent rewards (points) and risks (hit rate), we unambiguously observe every decision a player makes, and we know the objective function of the decision-maker early in the game (hit as many points as possible). I complement this evidence with micro-level data on physician decision-making in childbirth, which has important consequences for public health.

1. Professional Darts
The data is from 2016 World Darts Championship. This tournament includes performances of 32 professional darts players, and all matches are available on Youtube. The sample is restricted

---

1 For a review, see Bar-Eli et al (2006)
3 E.g. Grinblatt et al (1995); Carhart (1997)
4 Croson and Sundali (2005); Guryan and Kearney (2008); Suetens et al (2015)
5 Csapo and Raab (2014); Miller and Sanjurjo (2014); Green and Zwiebel (2015)
to the first nine throws of each leg, where every player is trying to hit as many points as possible. The final sample size is 8,617 throw observations. The figure below shows a heat map of my data sample. The vast majority of throws (81%) landed near Triple-20, which is worth 60 points. These throws are coded as being aimed at Triple-20, 41% of which hit the target successfully. 15% of throws landed near Triple-19 (57 points), 3% were aimed at Triple-18 (54 points), and the remaining 1% of darts were scattered around Triple-17 (51 points) and Bull’s Eye (50 points). For each throw, I code its intended aim as the nearest of the five targets: 60, 57, 54, 51, and 50, and an indicator for whether the throw has hit the intended target.

I find strong evidence of positive autocorrelation in dart performance, suggesting that the hot-hand is not a fallacy in this context. On average, players are 7.9 percentage points more likely to hit the intended target after a prior hit (p<0.001). The results are robust to multiple specifications. I propose one likely explanation for the hot-hand: Following a successful hit, the player is more likely to hit again because one can replicate the physical motion of the previous throw, but with a minimal adjustment to ensure that the second dart does not hit the first dart. When a player has missed, however, the player must make a greater adjustment to the previous motion, which is prone to more errors. Consistent with this story, I find that how badly a player has missed with the previous shot (distance to the intended target) is negatively associated with the hit rate. This mechanism provides a useful way to predict ex-ante whether the hot-hand is likely to exist in a different context. In medicine, for example, a large heterogeneity in patient characteristics makes it less likely that the hot-hand would exist in physician performances.

Next, I test whether players behaviorally respond to the hot-hand. I find that, consistent with the hypothesis, players are 11.4 percentage points more likely to aim at a riskier, higher-reward target after a hit (p<0.001). In an effort to test the optimality of this strategic reaction to the hot-hand, I simulate the decisions of a risk-neutral agent who maximizes the expected number of points, with correct perception of the hot-hand momentum. Across three rational benchmarks, the optimizing agent responds by being 3.0-5.9 percentage points more likely to choose the riskier target after a hit, all significantly less than the observed reaction from professional players in the data. These findings suggest that, even in the presence of hot-hand, players systematically over-respond to streaks, and as a result incur significant costs to their performance. The economic cost of this hot-hand bias is estimated to be about 1.1 points per throw, or $1,305 loss in prize money per match.
2. Childbirth
In childbirth, physicians must decide whether to perform a vaginal delivery or a cesarean section (C-section). A physician prone to the hot-hand fallacy may misinterpret a streak of poor childbirth outcomes that occur by chance, and incorrectly believe that a poor outcome is more likely in subsequent childbirth. As a result, a hot-hand biased physician may be more likely to recommend C-section – a defensive procedure less desirable for mothers not at high risk – to future patients. Mitigating this cognitive mistake can have significant and widespread benefits to public health, by reducing medical costs and improving patient health outcomes.

I use micro-level hospital inpatient data set from Florida in 2014. The data was obtained from the Healthcare Cost and Utilization Project (HCUP). Using procedure codes, I identify 186,279 childbirths, 81,314 of which were C-sections (44%). The C-section rate is high, which may be partly explained by out-of-hospital births that I do not observe in the data. The length of stay in the hospital was used as an indicator for poor patient outcome. The average stay for vaginal delivery and C-section were 2.4 and 3.2 days, respectively. A poor outcome is defined as a stay of 7 days or longer at the hospital (variable mean = 0.02).

The main findings are two-fold. First, I find no evidence of positive autocorrelation in patient outcomes at the physician level, suggesting that physician performances do not exhibit streaks. Second, physicians are 4.0 percentage points more likely to perform a C-section immediately following a poor patient outcome (p=0.025). The regression includes physician fixed effects, and standard errors are clustered at the physician level. The estimated coefficients for outcomes two and three patients ago are 1.1 and 2.1 percentage points, respectively, and are not statistically significant. Estimate suggest that the additional hospital charges incurred by unnecessary C-sections is $19 million per year in the United States.

There are a few alternative mechanisms that can also explain the findings, aside from the hot-hand fallacy. First, if there is a tendency to pass on high-risk mothers to physicians who have recently treated other high-risk mothers, this could generate a positive coefficient estimate. This can be directly tested in the data, using the availability of attending and operating physician IDs. Second, a response to recent experience may reflect a form of (irrational) Bayesian learning about one’s permanent skill as a physician, as opposed to a temporary hot-hand. Finally, a recent bad outcome may make malpractice lawsuits more salient, thus inducing a defensive response from physicians. I am currently waiting for a larger, more comprehensive hospital inpatient data set from New York for the years 2005-2015 to test these alternative mechanisms.

In both the darts and medical contexts, I find consistent evidence of a tendency of professionals to over-respond to their recent experiences. Darts provide a clean and compelling environment to test the hot-hand bias. In medicine, isolating the hot-hand fallacy from other alternative mechanisms is more challenging, but it illustrates a potential area in which the cognitive mistake has far-reaching economic consequences and incur significant costs to public health.