Bond University

Research Repository

Do (Australian) jockeys have hot hands?

Wrathall, Robert; Falvey, Rodney; Rajaguru, Gulasekaran

Do (Australian) Jockeys have Hot Hands?

Robert Wrathall

Rod Falvey

Gulasekaran Rajaguru

Bond Business School, Bond University,

Revised July 2019

Abstract

We extend the empirical analysis of hot hands in sports to horse racing, using the winning streaks of a sample of jockeys riding in Australia. Grouping jockeys by strike rate (win percentage), we find evidence of hot hands across almost all strike rates. But considering jockeys individually, only a minority exhibit hot hands. A wagering strategy based on hot hands yields a negative return overall and for most hot hands jockeys, although some do yield a positive return. We conclude that hot hands are present but not ubiquitous and that this is generally recognised in the betting market.

Key Words: Hot Hands, Jockeys, Thoroughbred Horse Racing, bootstrapping

1. Introduction

On the 30th December, 2016, jockey Jeff Lloyd created Australian racing history at Corbould Park Turf Club on the Sunshine Coast. Lloyd rode 7 winners on the 12 race program, bettering the previous record for most metropolitan wins. This feat including winning the last five races on the program. What is also notable about this achievement is that only two of his mounts were favourites. By the end of the afternoon's racing, commentators and spectators observed that Lloyd had 'hot hands' and that today it seemed he could win riding a clothes/saw/hobby/rocking horse. At first glance this does seem to be an extraordinary performance. Lloyd had won 16.5% of his rides in his racing career in Australia prior to that date, and on the basis of this record would have been expected to ride two winners at best. From this perspective it does appear that this was not an ordinary day for Lloyd and that somehow his early wins inspired his subsequent performance – hence the 'hot hands' label.

The hot hands concept has a rather controversial place in sports.[1] In broad terms, hot hands can be defined as 'the belief that the performance of a player during a particular period is significantly better than expected based on the player's overall record' (Gilovich, Vallone & Tversky, 1985; p295). GVT questioned the validity of the hot hands phenomenon and this helped spark an academic debate which has lasted for over three decades. [2] It is not hard to see how such controversy can arise, for while riding five winners in a row seems an improbable outcome for a rider with a 16.5% strike rate (wins/rides), it is not an impossible outcome. It comes down to a question of (beliefs about) likelihoods.

In this paper we extend the empirical analysis of hot hands in sports to horse racing, by looking at the occurrences of winning streaks in the racing records of a sample of jockeys riding in Australia. We compare the actual number of occurrences of each winning streak with the corresponding expected number, given the overall wins and rides of that jockey. We draw two main conclusions. First, if we consider jockeys collectively, grouped by strike rate, the evidence indicates the presence of hot hands

across almost all strike rates. This suggests that an observer who didn't distinguish individual jockeys beyond their strike rates could easily conclude that jockeys had hot hands. Second, if we consider jockeys individually, we find that the majority do not exhibit hot hands. But a significant minority do, although they vary in terms of the number of consecutive wins required to trigger hot hands and the lengths of the streaks for which hot hands persists. Finally, we ask whether these instances of hot hands appear to have been recognised in the betting market. Specifically, we calculate the return on a strategy of placing a $1 to win bet on the next ride of any winning jockey identified as having hot hands. While this wagering strategy yields a negative return overall and for the majority of hot hands jockeys, a small minority do yield a positive return.

2. Hot Hands

We follow the original definition of hot hands (GVT), that an individual's success at a particular sporting task (in our case a winning ride) makes success at the next attempt (i.e. winning on the next ride) more likely than if the first attempt had been a failure.[3] So a basketball player exhibits hot hands if she is more likely to score on a particular shot when her last shot was successful than if it was unsuccessful. Similarly, a golfer sinking a putt, a baseball batter hitting a pitch, a ten-pin bowler bowling a strike, a rugby /American-football player kicking a penalty/field-goal and so on. In principle testing for the presence of hot hands should be straightforward, but in practice such testing often requires information on other situational variables. Not all basketball shots are taken from the same distance and guarded by opponents with identical defensive skills. Likewise, not all putts are of the same length and difficulty; not all baseball pitches and pitchers are alike; not all field-goals are of the same distance etc. However, if one can obtain a large enough sample of tasks these situational variables should wash out, and this is the strategy that we rely upon below. Using an argument analogous to that used by GVT (1985) for basketball, we can say that if each ride is chosen randomly from an ensemble of rides differing in difficulty (due to quality of mount, quality of opposition,

distance of race etc.), this will produce a riding record equivalent to a simple binomial model based on the jockey's strike rate.

A considerable literature on hot hands has arisen in the three decades since GVT (1985) brought the concept from sports folklore into the realm of academic research.[4] The sports covered include basketball (GVT, 1985; Paul and Weinback,2005; Arkes, 2013; Daks et al., 2018), golf (Gilden and Wilson, 1995b; Clark, 2003a; McFall et al, 2009; Livingston, 2012; Stone and Arkes, 2016), volleyball (Raab et al, 2012), baseball (Albright, 1993; Albert, 2008; Green & Zwiebel, 2018), bowling (Dorsey-Palmeteer & Smith, 2004), NFL (Paul et al, 2012), darts (Otting et al, 2018) and even horseshoe pitching (Smith, 2003).[5] The results have been mixed. In a review of twenty years of hot hand research, Bar-Eli, Avugos and Raab (2006) cited 13 studies that did not find hot hands and 11 studies that did. The results from a follow up meta-analysis (Avugos, et al 2012) provided sufficient evidence for the authors to find against the existence of hot hands in sport in general. However, Miller and Sanjurjo (2018) have recently uncovered a bias in the common measure of the conditional dependence of present outcomes on streaks of past outcomes in sequential data used in many studies. Once this bias is corrected for some of the negative conclusions from the earlier studies are reversed. The debate continues.

Taking these results into account, we see our contribution here as threefold. First, we extend the analysis to a new sport – horse racing. The remarks of racing commentators and journalists clearly indicate a belief in the presence of the 'hot hand' in jockey performance, but are these views supported by the data? Second, we look for the presence of hot hands in jockeys individually as well as collectively. The collective evidence suggests that jockeys have hot hands. But does this mean that all jockeys have hot hands all of the time? Finally, is the apparent belief in hot hands by the racing industry reflected in the betting odds. Could we have used hot hands as a successful betting strategy over our sample period?

3. Data and Methodology

We collected data on every jockey who rode in thoroughbred horse races in metropolitan, provincial or country turf clubs in New South Wales, Australia in 2011.[6] This data included information about the jockey's finishing position in all of their races between January 1, 2000 and December 31, 2016. After discarding jockeys with less than 500 rides over this time period, we were left with information from over 1 million rides by the remaining 249 jockeys. The average number of rides per jockey is about 4000, with 18,336 the highest number of rides by a single jockey.

As noted in the introduction, we look for evidence of hot hands by considering winning streaks. Our notion of the hot hand is that a jockey has a higher strike rate in the 'hot' state, which is triggered by a streak of wins, than in the normal (cold) state.[7] It would then seem natural to follow GVT and much of the subsequent literature, in testing for hot hands by comparing the probability of a win following a streak of wins with the probability of a win following an equal length streak of losses. Unfortunately a win is a relatively rare event for a jockey, with the typical jockey having about ten times as many losses as wins and many more losing than winning streaks of any given length.[8] This relative lack of winning streaks reduces the power of tests based on these conditional probabilities.

Instead we rely on the implications of hot hands for the number of winning streaks of any length that will appear in a jockey's racing record. For each jockey we calculate the expected number of occurrences of each length of winning streak based on the total number of rides and wins of that jockey assuming a constant strike rate.[9] We then compare these expected numbers of streaks with the actual numbers for that jockey. If a win triggers the hot state which involves a higher strike rate, then we would expect a jockey having hot hands to have more winning streaks of any given length than if the same jockey did not have hot hands (and hence a constant strike rate).[10]

We illustrate this outcome for the case of 4 rides and 2 wins in Table 1. For a jockey that does not exhibit hot hands the probability of a win is $\frac{1}{2} (= \frac{W}{R})$ on each ride. If the jockey does exhibit hot hands then we consider two states: (i) the cold state (which is the initial state for all jockeys) in which the

probability of a win is also $\frac{1}{2}$; and (ii) the hot state (which is triggered by a win on the previous ride) in which the probability of a win is $q > \frac{1}{2}$. When calculating the conditional probability of an outcome for each individual ride we must keep in mind the outcomes that have gone before and that the total number of wins must be 2 out of the 4 rides.[11] The probability of each sequence is shown in column 2 for a jockey without hot hands (in which case all sequences are equally likely), and in column 3 for a jockey with hot hands (in which case the probability depends on the order of the outcomes). From Table 1 we observe that the hot hand reallocates the probabilities in favour of the sequences with winning streaks, so that if we compare two jockeys, one with and one without hot hands, the jockey with hot hands is more likely to have sequence LWWL or WWLL than the jockey without.

Table 1: Probabilities of each riding sequence with 2 wins in 4 rides, with and without hot hands (HH).

Riding sequence	Probability w/o HH	Probability w HH
LLWW	$\frac{1}{2}\cdot\frac{1}{3}\cdot 1.1 = \frac{1}{6}$	$\frac{1}{2}\cdot\frac{1}{3}\cdot 1.1 = \frac{1}{6}$
LWLW	$\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$	$\frac{1}{2}\cdot\frac{2}{3}\cdot[1-q].1 = \frac{1-q}{3} < \frac{1}{6}$
LWWL	$\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$	$\frac{1}{2}\cdot\frac{2}{3}\cdot q.1 = \frac{q}{3} > \frac{1}{6}$
WLLW	$\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$	$\frac{1}{2}\cdot[1-q].\frac{1}{2}\cdot 1 = \frac{1-q}{4} < \frac{1}{6}$
WLWL	$\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot 1 = \frac{1}{6}$	$\frac{1}{2}\cdot[1-q].\frac{1}{2}\cdot 1 = \frac{1-q}{4} < \frac{1}{6}$
WWLL	$\frac{1}{2}\cdot\frac{1}{3}\cdot 1.1 = \frac{1}{6}$	$\frac{1}{2}\cdot q.1.1 = \frac{q}{2} > \frac{1}{6}$

Notes: (i) q>1/2; (ii) Consider sequence LWLW. For a jockey without hot hands the probability of a loss on the first ride is ½; the probability of a win on the second ride, given that the first ride was a loss and the sequence consists of 2 wins and 2 losses, is 2/3; the probability of a loss on the third ride, given that the first two rides were a win and a loss, is ½; and the probability that the last ride is a win, given that the previous rides have been two losses and a win, is 1. A jockey with hot hands is assumed to start in the cold state and to remain there until a win occurs. The win puts the jockey in the hot state until a loss occurs, in which case the jockey reverts back to the cold state. So, for a jockey with hot hands the probability of loss on the first ride is ½; the probability of a win on the second ride, given a loss on the first, is 2/3; but the probability of a loss on the third ride, given that the win on the second ride put the jockey in the hot state, is 1-q < 1/2; and the probability of a win on the final ride, given the outcomes on the others, is 1. This sequence is less likely for the hot hands jockey.

We consider streak lengths of 2, 3, 4 and 5 wins, which we denote as 2BB, 3BB, 4BB and 5BB (i.e. 2 back to back wins etc.), respectively. Winning streaks longer than this are extremely rare both in fact and expectation as shown in Figure 2 below. A jockey with R rides in our sample period, has R-[k-1]

opportunities to ride kBB wins (e.g. for 2BB these are rides 1 and 2, 2 and 3, ….., R-1 and R). If this jockey had W wins in our sample period her strike rate would be W/R, and the probability of her winning any k consecutive rides is $\prod_{i=0}^{k-1}\left[\frac{W-i}{R-i}\right]$ (e.g. for 2BB, $\frac{W}{R}$ is the probability of winning the first ride and $\frac{W-1}{R-1}$ the probability of winning the second, given that the first was won). Multiplying this by the number of opportunities (R-[k-1]) we obtain the expected number of kBB as

$$E(kBB) = [\ R\text{-}[k\text{-}1]]\left\{\prod_{i=0}^{k-1}\left[\frac{W-i}{R-i}\right]\right\} \tag{1}$$

We can write the actual and expected number of kBB wins for jockey j as Ak_j and Ek_j respectively. Then, under our methodology, a necessary condition for j to exhibit hot hands is that $Ak_j > Ek_j$ for at least one k = 2,..,5.

For the purposes of comparison across jockeys it is useful to normalise both actual and expected streaks by the number of rides. Hence the actual (ak_j) and expected (ek_j) number of kBB streaks *per thousand rides* for jockey j is given by[12]:

$$ak_j = \frac{Ak_j}{R_j}1000$$

and $\quad ek_j = \frac{Ek_j}{R_j}1000 = \frac{[\ R_j-[k-1]]}{R_j}\left\{\prod_{i=0}^{k-1}\left[\frac{W_j-i}{R_j-i}\right]\right\}1000 \approx \left\{\prod_{i=0}^{k-1}\left[\frac{W_j-i}{R_j-i}\right]\right\}1000$

After this normalization and noting that $\frac{[\ R_j-[k-1]]}{R_j} \approx 1$, we see that the differences in expected streaks across jockeys depend (in a rather nonlinear way) primarily on their strike rates.

4. Do Jockeys collectively show hot hands?

We first look at the aggregate evidence to see whether it supports a general impression that jockeys have hot hands. Specifically, is an observer who considers the collective riding records of these jockeys, taking account of differences in strike rates but without paying attention to jockeys individually, likely to conclude that jockeys have hot hands?[13]

To do this we group jockeys by their respective strike rates. Let J(h) be the set of jockeys with strike rate h. We can then define ak_h and ek_h, the actual and expected numbers of streaks of length k of all jockeys with strike rate h, as[14]

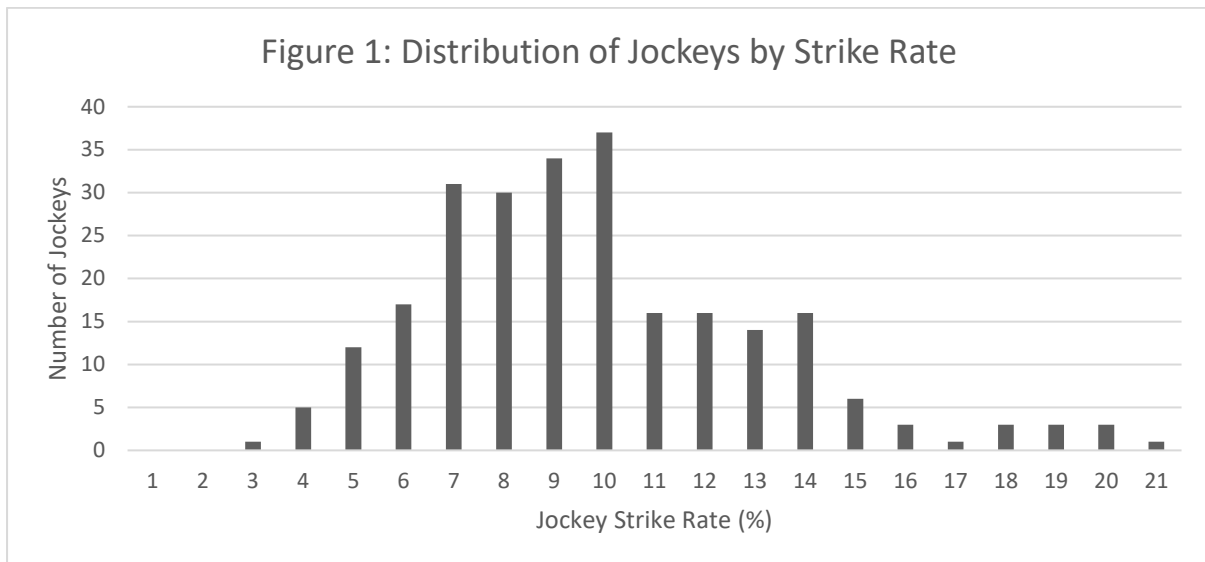$$ak_h = \sum_{j \in J(h)} ak_j \qquad \text{and} \qquad ek_h = \sum_{j \in J(h)} ek_j$$
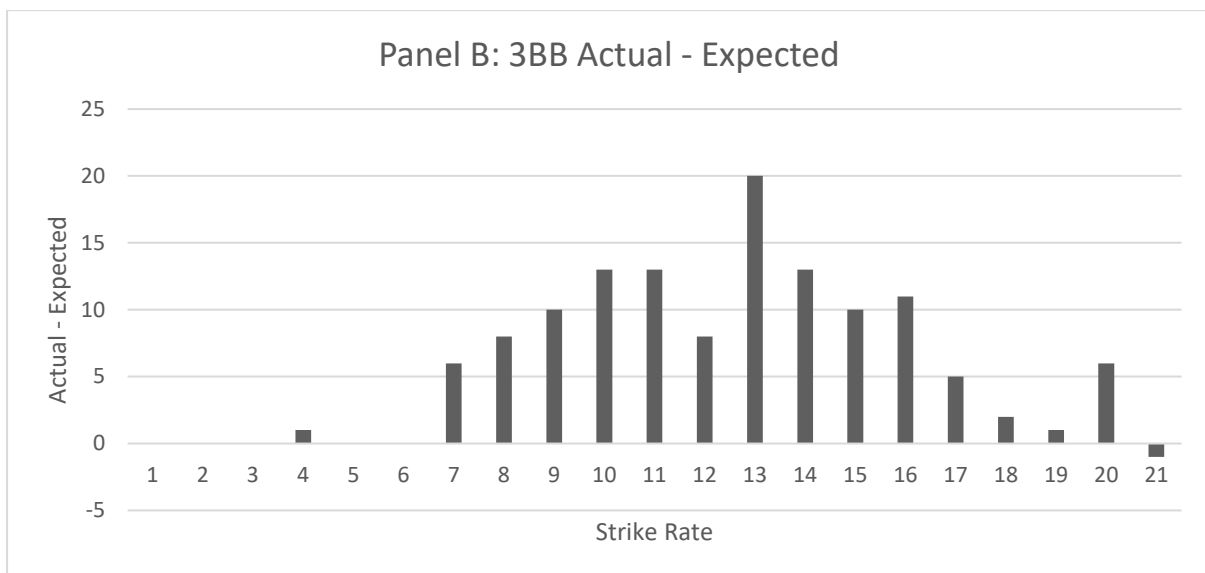

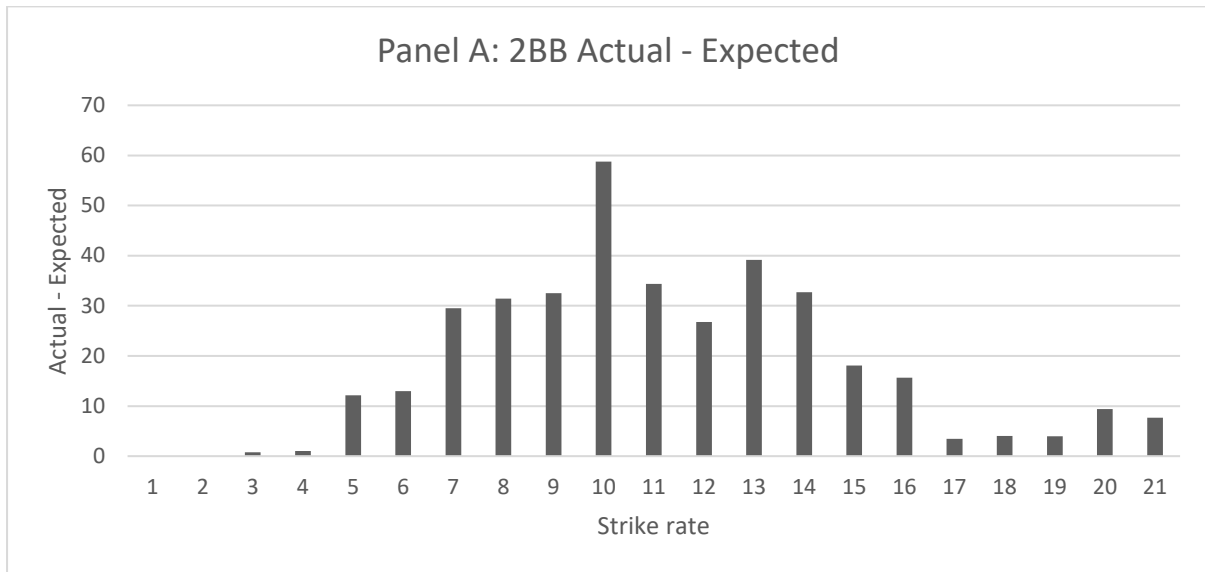Figure 1: Distribution of Jockeys by Strike Rate

Figure 1 plots the distribution of our sample of jockeys across 19 strike rates.[15] This distribution is characterised by a short lower tail and a much longer upper tail. Note that even at this level of aggregation there are three strike rates for which there is only a single jockey (3%, 17% and 21%). Figure 2 then plots the difference between the actual and expected numbers of streaks per thousand rides when Jockeys are collected into groups based on strike rates.[16] In Panel A we see that the actual number of 2BB streaks exceeds the corresponding expected number for all strike rates. In Panels B and C we see that the actual number is at least as great as the expected number for 3BB and 4BB, except for (the single jockey at) the very highest strike rate. In Panel D the actual number is at least as great as the expected number of 5BB for all strike rates. With one exception, jockeys of all strike rates appear to have more winning streaks per thousand rides than we would expect based on their strike rates. This evidence at least is consistent with the notion that a representative jockey exhibits hot hands.

Figure 2: Actual vs Expected Jockey Streak Numbers

Panel A: 2BB Actual - Expected

Panel B: 3BB Actual - Expected

Panel C: 4BB Actual - Expected

Panel D: 5BB Actual - Expected
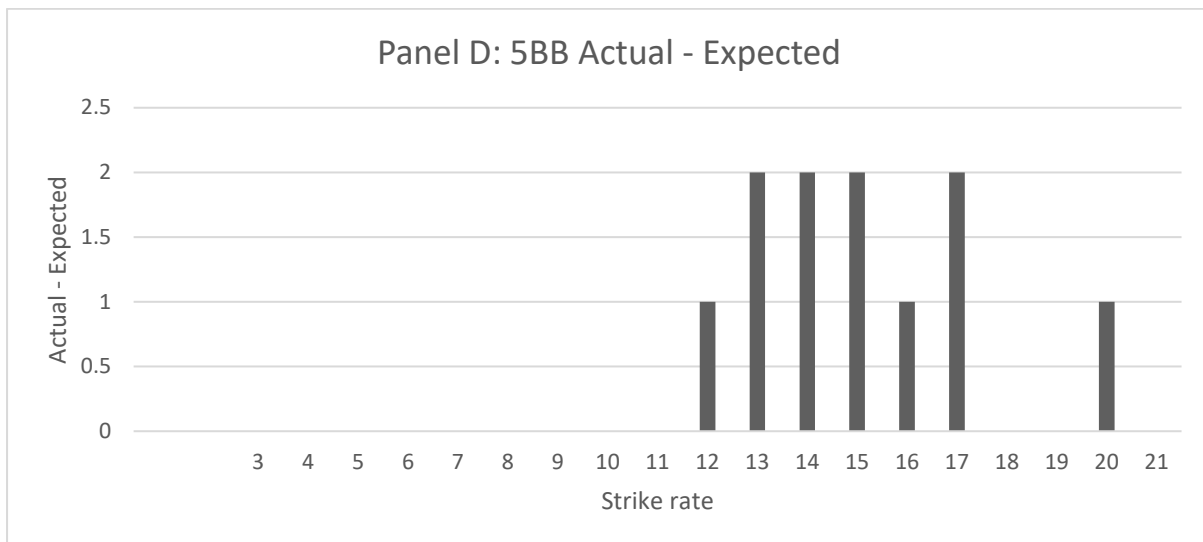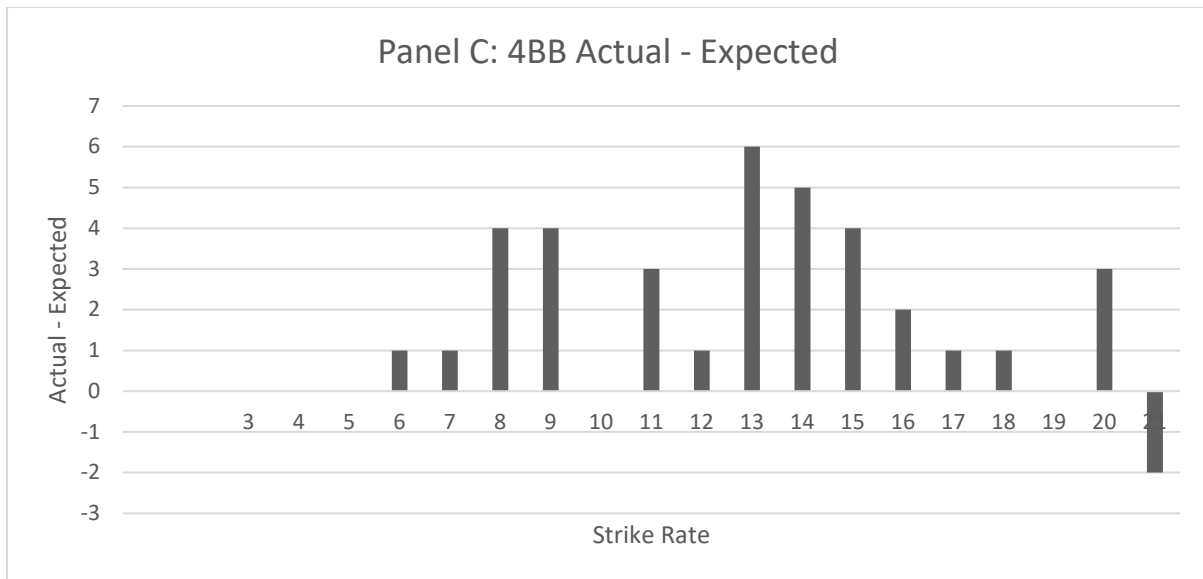
Given that the differences between actual and expected streaks go in the direction of jockeys having hot hands, one can ask whether these differences are significant. To test this we estimate a regression on a cross-section of the 19 strike rates to explore the relationship between actual and expected numbers of each streak. That is for each of the streak lengths we estimate:

$$ak_h = \beta_{0k} + \beta_{1k}\, ek_h + \varepsilon_h \tag{2}$$

Our null hypothesis is that $\beta_{1k} = 1$, for all k = 2,3,4 and 5. The alternative (hot hands) hypothesis is that $\beta_{1k} > 1$, for all k = 2,3,4 and 5. Table 2 reports the results

Table 2: Streak Number Regressions

	Streak Length			
	2BB	3BB	4BB	5BB
Expected (β_{1k}) (Std. Error)	1.136*** (0.014)	1.377*** (0.070)	1.457*** (0.097)	2.068*** (0.188)
Constant (β_{0k}) (Std. Error)	0.773 (2.134)	0.268 (1.069)	1.812 (1.203)	0.138 (0.257)
R^2	0.99	0.95	0.90	0.81
N	19	19	19	19
Test of $\beta_{1k} > 1$	9.06***	5.16***	3.93***	4.47***

Robust Standard errors in parenthesis. *** indicates significance at the 1% level.

For each streak length we can reject the hypothesis that $\beta_{1k} = 1$ in favour of $\beta_{1k} > 1$. The actual number of streaks is significantly larger than the expected number for all four streak lengths, a result consistent with a representative jockey having hot hands.

In Table 3 we explore the relationship between hot hands and the strike rate by regressing the difference between the actual and expected numbers of streaks for the average jockey at each strike rate on a quadratic function of the strike rate. We define n_h as the number of jockeys with strike rate h, so that $Dk_h \equiv \frac{ak_h - ek_h}{n_h}$ is the difference between the actual and expected numbers of streaks of length k for the representative jockey of strike rate h. Our regression is

$$Dk_h = \gamma_{0k} + \gamma_{1k}SR_h + \gamma_{2k}[SR_h]^2 + \varepsilon_h \tag{3}$$

At each streak length we find a positive coefficient on the strike rate and a negative coefficient on its square, indicating that the difference first increases then later declines as the strike rate increases. But only for 2BB and 4BB are both coefficients statistically significant and for these two streak lengths we calculate that the 'peak' strike rate (the strike rate at which the difference between actual and expected streak numbers is largest) is 11 to 12%. Loosely speaking, it is at these strike rates where we might expect to see the 'hottest hands'.

Table 3: Hot Hands and the Strike Rate

		Streak length		
	2BB	3BB	4BB	5BB
SR (std error)	0.65*** (0.16)	0.37*** (0.07)	0.21* (0.11)	0.06 (0.05)
SR^2 (std error)	-0.03*** (0.007)	-0.01 (0.009)	-0.009** (0.004)	-0.001 (0.002)
Constant (std error)	-2.32** (0.78)	-1.56 (0.89)	-0.86*** (0.19)	-0.31 (0.25)
R^2	0.54	0.29	0.18	0.12
N	19	19	19	19
Peak SR	10.83	-	11.67	-
Cross-over SR	17.16	-	15.03	-
Marginal	0.644	0.368	0.208	0.06

Notes: Robust Standard errors in parenthesis. *** and *** denotes rejection at the 10%, 5% and 1% respectively. SR is strike rate, SR^2 is strike rate squared. Peak SR (=$-\gamma_{1k}/2\gamma_{2k}$) is the strike rate at which the estimated difference would be highest. The cross-over SR is the strike rate at which the estimated equation predicts the actual and estimated number of strikes are equal. The Peak and Cross-over SR are not reported for the cases where SR^2 is insignificant. The marginal effect with respect to strike rate is evaluated at the average strike rate.

Finally, given that the weight of this evidence lies in support of these jockeys collectively having hot hands, we might ask whether this information was reflected in the betting odds[17]. To address this, we consider a wagering strategy which involves placing a $1 bet to win on the next ride of any jockey who won on their last ride. We applied this strategy to all the jockeys on their rides within our sample[18] and eagerly tabulated our (hypothetical) winnings. The result was a (fortunately hypothetical) loss of $25,909.29. This outcome is consistent with the notion that the betting market is well aware that the representative jockey may have hot hands.

5. Do individual jockeys have hot hands?

The evidence that we have presented thus far supports the general perception that jockeys, with abilities across the full range of strike rates, exhibit hot hands. Given this it is unsurprising that commentators apply the hot hands label to jockeys having a standout day at the track. However, we should not forget that hot hands are an individual rather than a collective characteristic. As part of his study of hot hands amongst professional golfers, Livingston (2012) briefly reviews models of

momentum from the psychology literature, noting that the key insight is that success need not beget success and failure need not beget failure uniformly for all players. Different personalities react differently to the same event and many factors may impact on how a player reacts to a recent success. Livingston (2012) argues that hot hands may be present in some players but not others and this may explain why it fails to appear in so many aggregated results in the literature. Here we work in the reverse direction, seeking the extent to which the hot hands we observe at the aggregate level translates back to individual behaviour.

Even a casual inspection of the data suggests that *not all* jockeys have hot hands. If we compare the actual and expected numbers of streaks for each jockey individually, without normalising by the number of rides, we do find jockeys for whom the actual is less than the expected[19] for each streak size. Furthermore, a straight comparison of actual with expected streak numbers fails to take into account that a jockey can ride more than her expected number of any streak size in the absence of hot hands (i.e. without any change in her strike rate from one ride to the next). Evidence of hot hands requires not just that a jockey has more streaks than expected given her strike rate, but that she is unlikely to have ridden her actual number of streaks given that strike rate.

To investigate further we construct the relevant statistics for a t-test of the difference between the actual number and expected number of each streak length for each jockey as follows:

$$tk_j = \frac{Ak_j - Ek_j}{\sqrt{var\,(Ek_j)}} \qquad \qquad \textbf{(4)}$$

where var(Ekj) is the sample variance of the expected number of kBB winning streaks for jockey j and is obtained using a bootstrap sample of 100 000 replications from all possible combinations of wins and rides ($C_{W_j}^{R_j}$).

Table A1 in the Appendix records these test statistics and results for all 249 jockeys included in our study. A summary of the aggregated results is presented in Table 4, where we record the numbers of jockeys who do and don't exhibit hot hands at each streak length in columns 2 and 3, respectively. In

the remaining columns we also record how these jockeys progress through the streak lengths. Of the 67 jockeys who exhibit hot hands at 2BB, 43 also exhibit hot hands at 3BB and 24 do not. Of the 182 jockeys who did not exhibit hot hands at 2BB, 24 did at 3BB and so on.

Table 4: Aggregated Hot Hand Results for Individual Jockeys

Streak Length	#with HH	#w/o HH	Progression Patterns															
2BB	67	182	67								182							
3BB	67	182	43				24				24				158			
4BB	45	204	20		23		3		21		12		12		10		148	
5BB	22	227	9	11	2	21	0	3	0	21	4	8	1	11	3	7	3	145

Notes: A shaded cell indicates the number of jockeys with hot hands (HH).

While it is true that 145 jockeys (58% of the total) give no evidence of hot hands at any streak length, a quick glance at Table 3 will dash any hopes that the remaining 104 jockeys form a homogeneous group exhibiting hot hands after any win and for all streak lengths. Only 9 jockeys fit this description. An examination of the progression patterns in Table 4 does allow us to make the following observations, however. First, there is variation across jockeys in the number of consecutive wins needed to trigger hot hands. For 67 one win is enough, but others are late starters - 24 need 2 consecutive wins, 10 need 3 and 3 need 4. Second, there is variation across jockeys in the number of streak lengths for which hot hands remains active once triggered. For 19 jockeys hot hands stays active for all longer streaks (9 jockeys starting at 2BB, 4 at 3BB, 3 at 4BB and 3 at 5BB); for 39 jockeys, hot hands only lasts for one streak length and is never retriggered (21 jockeys at 2BB, 11 at 3BB and 7 at 4BB); for 29 jockeys hot hands lasts for two streak lengths and is never retriggered (21 at 2BB and 8 at 3BB); and lastly for 11 jockeys, hot hands once triggered (at 2BB) lasts for three streak lengths and is never retriggered. Finally, there is a small group of jockeys (6 of 104) that have gaps between their hot hands streaks.

In summary: only 42 percent of individual jockeys exhibit hot hands for at least one streak length, and these jockeys differ both in the number of consecutive wins necessary to trigger hot hands and in the number of consecutive rides for which hot hands persists.

What are the relationships, if any, between hot hands and a jockey's 'experience' (number of rides), 'ability' (strike rate) and gender?[20] Livingston (2012) finds some evidence that golfers with more experience are better able to maintain positive momentum. Is the same true for jockeys? We investigate by constructing the following dummy variables:

$HH_j = 1$, if j exhibits hot hands at some streak length, and = 0 otherwise; and

$Hk_j = 1$, if j exhibits hot hands at streak length k, and = 0 otherwise.

For each of these five dummy variables we run a binary-logit regression with results as reported in Table 5. The results for the HH dummy reported in the second column of Table 5 show that the probability that a jockey has hot hands is positively related to that jockey's experience (rides) but is unrelated to gender. The probability is increasing in the strike rate until 25% after which it declines. The same is true for the dummies showing hot hands at specific streak lengths, except that ability experience seems irrelevant after 3BB. The probability is initially increasing in the strike rate but declines after 14% (2BB) or 17% (3BB or 5BB). The most able jockeys do not necessarily possess the hot hand.

Table 5 Binary Logit Regression Relating Hot Hands to Jockey Characteristics

	Dependent Variable				
	HH	H2	H3	H4	H5
Rides	0.14***	0.15***	0.092*	-0.04	0.065
(000)	(0.05)	(0.005)	(0.05)	(0.063)	(0.102)
SR	25.81*	47.51**	5.14**	52.49**	139.87**
	(14.56)	(22.94)	(2.34)	(23.69)	(62.75)
SR^2	-51.19*	-169.78*	27.06	-158.01*	-407.43*
	(30.01)	(97.10)	(52.28)	(94.76)	(216.80)
Gender	-0.165	-0.636	-0.53	0.079	-0.84
	(0.38)	(0.49)	(0.46)	(0.47)	(1.06)
Cons	-2.88**	-4.46***	-2.14*	-4.89***	-12.91***

	(1.15)		(1.26)		(1.24)		(1.37)		(4.09)	
Peak SR (%)	25		14		(9)		17		17	
Marginal	0.039		0.028		0.021		0.032		0.049	
	Predicting Hot Hands									
	HH=0	HH=1	H2=0	H2=1	H3=0	H3=1	H4=0	H4=1	H5=0	H5=1
Total Obs	145	104	182	67	182	67	204	45	227	22
% Correctly Predicted	63.6	49.25	76.01	34.85	75.27	32.8	82.53	20.8	92.3	20.73
% of jockey's with HH		41.8		26.9		26.9		18.1		8.8

Note: Robust Standard errors in parenthesis. Significance *10%, **5% and ***1%. Rides is number of rides; SR is strike rate (win percentage); Gender is a dummy variable which =1 if a jockey is female. The marginal probability with respect to strike rate is evaluated at the average strike rate and average probability.

The lower part of Table 5 investigates whether these regressions can be used to predict which jockeys exhibit hot hands. From this perspective the results are disappointing. The equation predicts only half of the general HH jockeys correctly and no more than 35% of those at any specific streak length.

Finally, in Table 6 we summarise the results for the individual jockeys of our 'hot hands' wagering strategy discussed earlier (i.e. placing a $1 bet to win on the next ride of any jockey who won on their last ride). Several points can be made here. First, if we concentrate on the 102 jockeys identified as having hot hands, we see that the vast majority (88%) generate a negative return. Again, this is consistent with the hot hands effect being reflected in the market odds, though clearly not all individual cases were identified. Second, we see that there are jockeys without hot hands who also generate a positive return on our hot hands wager. But these are few in number (2) and only comprise just over 1% of these jockeys. Although only 14 of 247 jockeys (5.7%) yield a non-negative return under this strategy, 12 of these have hot hands. Third, since our wagering strategy is specifically intended to cover 2BB hot hands, we separate our hot hands cases into those including 2BB (and possibly longer streak lengths) in column 4 and those excluding 2BB in column 5. Only one jockey provides a positive return in the latter group, a percentage broadly consistent with the No HH category in column 2. In contrast, nearly 17 percent of jockeys in the former group yield a positive return. Finally, we focus on those jockeys with hot hands only at 2BB in column 6. The fact that this group shows the highest percentage (19) of jockeys showing a positive return, may indicate that jockeys whose hot hands have

brief persistence (in this case only one ride) are the hardest for the market to identify.[21] Thus, while the presence of hot hands is broadly reflected in the market odds, there is evidence that the market's identification of hot hands at the individual jockey level is incomplete.

Table 6 Returns from Hot Hands Wagering: Individual Jockeys

Jockeys	No HH	HH	HH Incl.2BB	HH But Not 2BB	HH Only 2BB
# Positive Return	2 (1.4%)	12 (11.8%)	11 (16.7%)	1 (2.8%)	4 (19.0%)
# Negative Return	143 (98.6%)	90 (88.2%)	55 (83.3%)	35 (97.2%)	17 (81.0%)
Total	145	102	66	36	21

6. Further evidence

Our previous tests obtained the expected number of sequences of each length using the assumption of a constant strike rate equal to the jockey's sample average. This strike rate was determined from the jockey's actual number of rides and wins. If this jockey exhibited hot hands, the elevated hot state strike rate triggered by a streak of wins would result in the jockey having a larger number of winning streaks than his (hypothetical) single state equivalent. But even a jockey in a permanent cold state is likely to find her strike rate changing over time - increasing as she gains experience, fluctuating in periods of minor injury and perhaps declining as she ages. Could some of our positive findings for individual jockeys be attributable to these variations?

In order to investigate whether our finding of hot hands for some jockeys can be explained by variations in their strike rates unrelated to hot hands, we consider each jockey's annual racing records over the years 2000 to 2016. In each year we determine the number of rides, the number of wins and the number of 2BB streaks in each year. As noted earlier, winning streaks are relatively rare in horse racing. A typical jockey with a strike rate of 10%, can expect roughly 1 2BB streak for every 100 rides, 1 3BB streak for every 1000 rides, 1 4BB streak for every 10,000 rides etc. Few jockeys in our sample have as many as 1000 rides in any year let alone consistently across years during their racing careers. In considering annual data we therefore restrict attention to 2BB streaks.

We first calculated the expected number of 2BB streaks for each jockey based on their strike rate across all 17 years in this sample and then compared this with the actual number recorded in the data. The number of jockeys for whom the actual 2BB is significantly different from the expected 2BB based on their strike rate over the full period is reported in column 2 of Table 7. [22] We then reran the test aggregating only those years in which the annual strike rate was not significantly different from the 17 year average strike rate. The corresponding results, which are shown in column 3 of Table 7, reveal the extent to which our tests are susceptible to these extraneous strike rate variations. Correcting for changes in the strike rate in this way we find the number of jockeys exhibiting hot hands is reduced by 40% (from 71 to 43). Allowing for variations in the cold state strike rate reduces but does not eliminate the instances of hot hands. [23]

Since the shorter the interval the more stable the strike rate is likely to be, we applied our tests to 2BB streaks at the annual level. This is the shortest interval we can consider while retaining an adequate sample of jockeys with enough rides to generate 2BB winning streaks. We only consider years in which the jockey had at least 200 rides (giving our typical jockey an expected number of at least 2 such streaks in any year considered). [24] The results for this restricted sample are reported in column 4 of Table 7. Over 40% of these jockeys have hot hands for at least one year. The density plot of hot hands years is reported in Panel A of Figure 3. The maximum number of years with hot hands is 4 (the maximum number of consecutive years is 3).

We then add a second constraint that jockeys must meet the minimum ride requirement for at least 8 years during this period. This helps isolate the more experienced jockeys by eliminating the not inconsiderable number who only exceed 200 rides in a few years. The results are shown in column 5 of Table 7 and the density plot of hot hands years is reported in Panel B of Figure 3. Sixty-five percent of these jockeys exhibit hot hands in at least one year. Even though our test has low power on annual data we still find evidence of hot hands for the majority of the most active and experienced jockeys, an outcome consistent with the argument that jockeys with more competitive experience are better able to maintain positive momentum (Livingston, 2012).

Table 7: Incidence of 2BB Hot Hands

	Full Sample		Restricted samples	
			Minimum rides per year	Minimum rides and minimum years
	All Years Combined	Years where ASR=LSR combined	HH in at least one year	HH in at least one year

#Jockeys with HH	71 (29%)	43 (17%)	86 (41%)	49 (65%)
#Jockeys without HH	176 (71%)	204 (83%)	125 (59%)	26 (35%)
Total Jockeys	247	247	211	75

Notes: LSR is lifetime strike rate. ASR is annual strike rate. Minimum rides per year is 200. Minimum years of minimum rides is 8.

Figure 3: Distributions of HH Years
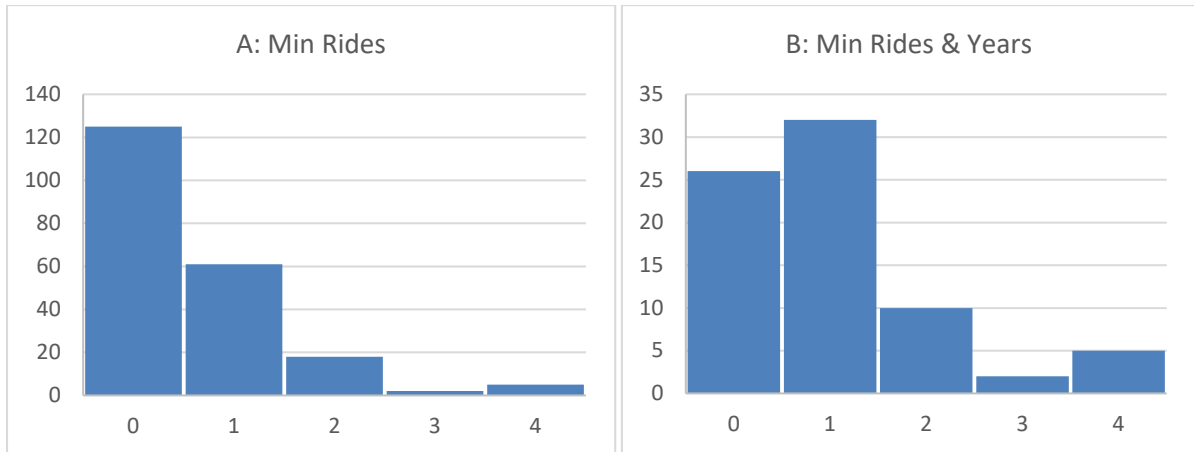
7. Conclusions

So, do Australian Jockeys have hot hands? The evidence that we have presented suggests that enough of them do for an observer who focussed on jockeys' collective performance, taking account of differences in ability only, to answer in the affirmative. For each jockey in our sample we compared their actual numbers of winning streaks of four different lengths with the expected numbers of these streaks based on the jockey's strike rate. When we aggregated these numbers by strike rate we found that the actual almost always exceeded the expected for all streak lengths. We concluded that jockeys collectively have more winning streaks than would be expected based on their strike rates, and this conclusion was confirmed by a regression of actual streaks on expected streaks, at each strike rate. The coefficient on expected streak numbers was significantly greater than unity at all streak lengths.

Do all Australian jockeys have hot hands? Our evidence suggests otherwise. Using a simple t-test of the likelihood of the actual number of streaks being explained by a jockey's strike rate, we found that 58 percent of the jockeys in our sample did not exhibit hot hands at any streak length. The other 42 percent exhibited hot hands for at least one streak length, but only 3.6 percent had hot hands at all

streak lengths. Corrections for variations in the strike rate unrelated to hot hands reduced the instances of hot hands in two-win streaks by about 40 percent. Hot hands appear to be positively related to a jockey's experience (number of rides), but to be unrelated to gender. The probability that a jockey has hot hands is first increasing then decreasing in the strike rate, with a peak that depends on the streak length considered. However, experience and ability alone are not sufficient to predict which jockeys have hot hands. Jockeys with hot hands exhibit a range of behaviour, differing in both the number of consecutive wins required to trigger hot hands and the number of rides for which hot hands persist once triggered.

Is information on hot hands reflected in the betting odds? Broadly, it appears that it is. Our strategy of betting to win on the next ride of any jockey who won on their last ride returned a loss when applied to: (a) all jockeys; (b) all jockeys identified as having hot hands; (c) all jockeys identified as having hot hands at streak lengths of two (and perhaps longer) rides; and even (d) all jockeys identified as only having hot hands at a streak length of two rides. But, that said, our (retrospective) wagering strategy did yield a positive return for twelve percent of the individual jockeys identified as having hot hands. Perhaps there is some scope for a positive return from carefully targeted hot hands wagering.

Finally, we return to the exploits of jockey Jeff Lloyd with which we started this paper. Was Lloyd's achievement of five back-to-back wins on that day an instance of hot hands at play? Unfortunately, Lloyd is not in our sample, but we were able to use information on his 3180 rides and 539 wins in Australia over the period 2011 to 2017 to perform the same hot hands tests as above (see the final row in Appendix Table A1). We found that Lloyd was one of those few jockeys who exhibit hot hands at all streak lengths. In his case the hot hands label seems to have been quite appropriate.

20

References

Albert, J. 1993. Comment. *Journal of the American Statistical Association*, 88: 1184-1188.

Albright, S. C. 1993. A Statistical Analysis of Hitting Streaks in Baseball. *Journal of the American Statistical Association*, 88: 1175-1183.

Arkes, J. 2013. Misses in "Hot Hand" Research. *Journal of Sports Economics*, 14 (4): 401-410.

Avugos, S., Köppen, J., Czienskowski, U., Raab, M., & Bar-Eli, M. 2013. The "hot hand" reconsidered: A meta-analytic approach. *Psychology of Sport and Exercise*, *14(1)*: 21-27

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

Bird, R. & McCrae, M. 2008. The Efficiency of Racetrack Betting Markets: Australian Evidence World Scientific Book Chapters,in: Efficiency of Racetrack Betting Markets. *World Scientific Publishing Co. Pte. Ltd.*, , chapter 57: 575-582.

Brailsford, T. J., Easton, S. A., Gray, P. K. & Gray, S. F. 1995. The Efficiency of Australian Football Betting Markets. *Australian Journal of Management*, 20 (2): 167-195.

Clark, R. D. 2003. Streakiness among professional golfers: Fact or fiction? *International Journal of Sport Psychology*, 34: 63-79.

Daks, A., Desai, N. & Goldberg L. R. 2018 Do the Golden State Warriors have hot hands? *The Mathematical Intelligencer*, 20(4): 1-5.

Dorsey-Palmateer, R. & Smith, G. 2004. Bowlers' Hot Hands. *The American Statistician*: 58**:** 38-45.

Gilden, D. L. & Wilson, S. G. 1995. Streaks in skilled performance. *Psychonomic Bulletin & Review* 2: 260-265.

Gilovich, T., Vallone, R., & Tversky, A. 1985. The Hot Hand in Basketball: On the misperception of random sequences. *The Cognitive Psychology* 17: 295-314.

Green, B. & Zweibel, J. 2018. The Hot Hand fallacy: Cognitive mistakes or equilibrium adjustments? Evidence from Major League Baseball. *Management Science* 64(11): 5315-5348.

Hendricks, D., Patel, J. & Zeckhauser, R. 1993. Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. *The Journal of Finance* 48, 93–130.

Huij, J. & Derwall, J. 2008. "Hot hands" in bond funds. *Journal of Banking & Finance* **32**, 559–572.

Jagannathan, R., Malakhov, A. & Novikov, D. 2010. Do hot hands exist among hedge fund managers? An empirical evaluation. *The Journal of Finance* 65, 217–255.

Liu L, Wang Y, Sinatra R, Giles CL, Song CM, Wang, D. 2018. Hot streaks in artistic, cultural, and scientific careers. *Nature* 559: 396-399.

Livingston, J.A 2012 The hot hand and the cold hand in professional golf. *Journal of Economic Behaviour & Organization* 81: 172-184.

McFall, T. A., Knober, C. R. & Thurman, W. N. 2009. Contests, Grand Prizes, and the Hot Hand. *Journal of Sports Economics*, 10(3): 236-255.

Miller, J.B. & Sanjurjo A. 2018. Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers. *Econometrica* 86: 2019-2047.

Otting, M., Langrock, R., Deutscher, C. & Vianey, L-B. 2018. The hot hand in professional darts. *Proceedings of the 33rd International Workshop on Statistical Modelling*, 1: 231-236.

Parsons, S & Rohde, N. 2015. The hot hand fallacy re-examined: new evidence from the English Premier League. *Applied Economics*, 47(4): 346-357.

Paul, R. J. & Weinbach, A. P. 2005. Better Misperceptions in the NBA The Overbetting of Large Favorites and the "Hot Hand". *Journal of Sports Economics*, 6(4): 636-649.

Paul, R. J., Weinbach, A. P. & Humphreys, B. 2012. Better Belief in the "Hot Hand": Evidence from Detailed Betting Data on the NFL. *Journal of Sports Economics*, 15(6): 390-400.

Raab, M., Gula, B. & Gigerenzer, G. 2012. The hot hand exists in volleyball and is used for allocation decisions. *Journal of Experimental Psychology Applied,* 18**:** 81-94.

Smith, G. 2003. Horseshoe pitchers' hot hands. *Psychonomic Bulletin & Review,* 10**:** 753-758.

Stone, D. F. 2012. Measurement Error and the Hot Hand. *The American Statistician*, 66(1): 61-66

Stone, D.F., Arkes, J. 2016. Reference Points, Prospect Theory, and Momentum on the PGA Tour. *Journal of Sports Economics*, 17(5): 453-482.

Tuckwell, R. H. 1984. Determinants of Betting Turnover, *Australian Journal of Management*, 9(2): 59-66.

Notes

[1] It also has applicability beyond sports. If an individual exhibits hot hands in any endeavour then streaks of success in that endeavour are more likely than for an individual of equal ability without hot hands. If the hot hand phenomenon is real, then managers and investors can potentially profit from knowing which economic agents are currently 'hot', an observation that has prompted investigations of hot hands amongst managers of mutual funds (Hendricks, et al, 1993); bond funds (Huij & Derwall, 2008); and hedge funds (Jagannathan, et al., 2010). Streaks in the careers of artists, movie directors and scientists can also be attributed to temporarily heightened creativity generated by hot hands (Liu et al, 2018).

[2] As Tversky reflected – "I've been in a thousand arguments over the topic, won them all and convinced no-one" (Bar-Eli *et al.*,2006, p535). However, a recent paper by Miller and Sanjurjo (2018) reveals a selection bias in the technique used by GVT and many subsequent researchers, which, once corrected, reverses many of their negative findings. In the light of this Tversky's failure to convince may not be so surprising.

[3] This definition implicitly assumes that players are always in a (latent) 'hot' state which is triggered into action by a 'success'. An alternative view is that players are only sometimes in the 'hot' state and that a success in a 'cold' state would have no effect on the future probability of success (Stone 2012, Arkes 2013). Arkes (2013) notes that tests may fail to detect hot hands even where they are present because the sample size is low and hot hands events are infrequent.

[4] Avugos et al. (2013) found over 250 hot hand papers in a preparatory search for their meta-analysis.

[5] There is also a closely related literature that looks at 'momentum' in team sports. See Parsons and Rhode (2015) for a recent example.

[6] This data was downloaded from RACINGANDSPORTS and RACING.COM on 17/02/2017. It is all publicly available information. In Australia all jockeys are licensed at the State level. Apprentice jockeys work for trainers and at the completion of their contract for training become self-employed, seeking rides from owners and trainers often with the help of a manager/agent. Jockeys are paid a fixed 'mounting fee' for each race ride and are also eligible for 'prize money' which is a share of the purse won by their mount in that race. This share is small - about 6% for a win, less for second or third. While the most successful jockeys can earn large incomes, the average annual income for a jockey in Australia is about $30,000 to $40,000 (Information from various racing websites).

[7] This streak may consist of a single win. Because jockeys' rides are normally required to be named about two days before a race meeting, we can rule out the possibility that hot hands arise because winning jockeys are systematically offered a better ride in their next race(s).

[8] Similarly, Livingston (2012) found that professional golfers rarely had streaks of more than two good holes (a score of birdie or better) or bad holes (a score of bogie or worse).

[9] Our test therefore ignores possible variations in a jockey's strike rate for reasons other than hot hands. This issue is explored in section 6 below.

[10] To be clear, we are comparing the number of winning streaks of each length in a jockey's actual racing record (which might be affected by the presence of hot hands) with the corresponding number that we would expect for a (fictitious) jockey with no hot hands and with the same number of rides *and* wins. Controlling for both wins and rides means that the higher strike rate of the jockey in the hot hand state (if it exists) will be reflected in more winning streaks rather than more wins.

[11] This is the point made by Miller and Sanjurjo (2018).

[12] Note that the ratio of the two variables and the sign of the difference is preserved under this normalisation as $\frac{ak_j}{ek_j} = \frac{Ak_j}{Ek_j}$ and $ak_j - ek_j = \frac{Ak_j - Ek_j}{R_j} 1000$.

[13] An even less discerning observer who viewed all jockeys as interchangeable would conclude that jockeys have hot hands because $\sum ak_j > \sum ek_j$ for all $k = 2,3,4$ and 5. That is 3012>2637 when $k = 2$; 443>316, when $k = 3$; 78>42, when $k = 4$; and 15>6, when $k = 5$.

[14] Note that the jockey data used here is normalised by the number of rides so that the outcomes are not sensitive to the fact that some jockeys had more rides than others.

[15] Jockeys are aggregated into 19 groups based on their strike rates for presentation purposes. All calculations use actual strike rates.

[16] Because the actual streak numbers must be integers, we have rounded the expected streak numbers to the nearest integer, after aggregation by strike rate. This is to avoid rather meaningless differences at low numbers. The (unrounded) expected streak number will always be strictly positive but may be very small at low strike rates and long streak lengths. For example the expected number of 5BB streaks per thousand rides for the jockey with a strike rate of 3% is 0.00002, while the actual number is, not surprisingly, 0. In our view it is more appropriate to record this as actual = expected than as actual < expected.

[17] Brailsford et al. (1995) finds for the Australian Rugby League (ARL) and Australian Football League (AFL) that betting strategies generate significant profits. For further information on racetrack betting in Australia see Tuckwell (1984) and Bird & McCrae (2008).

[18] Our race data also include the date of the meeting, the race number, the name of the horse, the finishing position of the horse and the starting odds of each horse in the race. Two jockeys were excluded because of incomplete data on overseas rides.

[19] Although the number where actual exceeds expected is always larger. Thus $Ak_j > (<)Ek_j$ for 77% of jockeys in 2BB, 65% in 3BB, 38% in 4BB and 14% in 5BB.

[20] These are the jockey characteristics available in our data set.

[21] However confining our wagering strategy to this group (of 21 jockeys) would still yield a negative return overall, as the losses on the other 17 jockeys far outweigh the gains on the 4.

[22] This is the same test as above. Although our data comes from the same sources as before, our selection criteria are different and so we repeat the tests on the new sample. Full results are available on request.

[23] It is perhaps worth reminding ourselves that hot hands itself is a (temporary) increase in the strike rate, so that our attempts to eliminate the effects of unrelated strike rate variations may also erase some evidence of hot hands.

[24] This removed 36 jockeys who did not have at least 200 in any single year from consideration. These jockeys have too few rides in any year to exhibit hot hands even if they have them.