NHH

INSTITUTT FOR FORETAKSØKONOMI

DEPARTMENT OF BUSINESS AND MANAGEMENT SCIENCE

FOR 06 2017

ISSN: 1500-4066

April 2017

Discussion paper

Profitable Robot Strategies in Pari-Mutuel Betting

BY

Petter Bjerksund AND **Gunnar Stensland**

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS

Profitable Robot Strategies in Pari-Mutuel Betting[1]

Petter Bjerksund and Gunnar Stensland[2]

Abstract: We have collected odds and results from 7 474 horse races in Norway and Sweden for a period of approximately 1.5 years. Based on the odds from the win game, we construct a profitable betting strategy for the corresponding triple game. Given a 30% track take, the existence of a profitable strategy is surprising. A robot is typically needed to identify and exploit underrated bets. We argue that the existence of heterogeneous beliefs between players in the market might form a basis for profitable betting strategies. We did expect that bigger pools (more liquidity) would remove this anomaly. That is not the case. More players, and thereby bigger pools, increases the profitability of the system.

1. Introduction

Economists have studied pari-mutuel betting markets for several years. It is a source of insight into the formation of expectation, return for risk, and equilibrium in a market. The idea behind these studies is often to learn something that might also be valid in the more important financial markets for securities.

From previous findings, we know that a *favorite-longshot bias* might exist, that is, the expectation of always betting on the favorite exceeds the expectation of other simple strategies. This phenomenon is referred to as the low odds bias (see Ali (1977), Ash and Quandt (1987), Thaler and Ziemba (1988), Hurley and McDonugh (1995), and Weitzman (1965)).

For an overview of possible explanations of this anomaly, see Ottaviani and Sørensen (2007). Given that the explanation for the *favorite-longshot bias* should ideally explain the anomalies in this paper, the most fruitful explanation investigated is heterogeneous beliefs. The investors (players) do not agree on the win probabilities for the horses. Hence, they do not agree on the probability for a combination of horses. When aggregating heterogeneous beliefs through a pari-mutuel betting pool, underrated combinations might occur.

Based on the odds for each horse winning in a given race (i.e., the win game), we construct a profitable strategy in a second game based on the same race. This second game, called the triple, involves betting on a combination of first, second, and third horse in the same race in the correct order. Given a 30% track take in the triple game, the existence of a profitable strategy is surprising. A similar idea is exploited in the so-called Dr. Z system. For an overview of similar strategies, see Thaler and Ziemba (1988).

The most profitable combinations (high expected value) very often consist of the favorite coming second or third. In this study, we restrict our analysis to the 200 combinations with the lowest odds in each race.

[1] This version: April 3, 2017. We thank Øystein Gjerde for helpful comments and suggestions.
[2] Both: Department of Business and Management Science, Norwegian School of Economics (NHH).
E-mail: Petter.Bjerksund@nhh.no and Gunnar.Stensland@nhh.no

The favorite will be involved in most of them. If we assume heterogeneous beliefs, it might be that all players more or less agree on the ranking of horses, but assign different probabilities to the outcomes. This leads to huge differences in subjective probabilities for the combinations. In this situation, it is easy to construct examples in pari-mutuel betting where the average of subjective probabilities differs from the mathematical probability given by observed odds.

Although it is interesting to develop a "money machine", this is not the main purpose of the paper. We want to show that more bettors and bigger pools do not alter this result. In fact, more investors being involved in the game increases the profitability of the system. This is counter-intuitive since bigger pools make the game interesting for professionals.

In order to make money on this observation you need a robot. Win odds change all the time and so do the odds for the combinations. A computer can easily monitor both and calculate favorable bets in real time. The challenge is to place all bets in the seconds before the race starts. This is a smaller problem for large pools. For large pools, odds are more robust before the race starts.

2. The horse track and data

All the data are collected using the internet (www.rikstoto.no). Every day there are two or more events on some tracks in Scandinavia. Each event consists of from 5 to 12 separate races. More than 90% of the races are harness races, also called trotting. Here a horse that uses a gait that is not allowed more than twice during the race might be disqualified. This influences the probability of the favorite ending second or third. Some horses are viewed as fast but unstable. Most players therefore believe that they will either win or be disqualified.

A typical race day on a track consists of some of following games:

The DD, V4, V5, V65, V75 game: Bet on the first horse in a combination of races (2 (DD) and 4-7 races). The pool in these games might be huge and the winner can take anything from NOK 50 to NOK 20 million. Some of the games also distribute a part of the pool to five out of six correct horses, for example. We will concentrate on the betting in each individual race.

The win ("vinner") game: Bet on a horse to win the race. Let $B = \sum_{i=1}^{n} b_i$ denote the before track take pool in the game, where b_i is total wagers on horse i and n is the number of horses. The total amount available for the winners of the bet is $0.8 \cdot B$. That is, the totalizator takes 20% in this game.

The odds of horse k winning the race are given by $O_k = 0.8 \cdot (B/b_k)$. All these odds are calculated continuously and published on the internet. Typically, new information will be available every 20-30 seconds. It is a problem, however, that a lot of the betting happens in the last minute before the race starts.

The show ("plass") game: Bet on a horse coming first, second or third. This game is utilized by the so-called HZR-system, see Hausch et al. (1981). It is also investigated in Lo et al. (1995). Their idea is to find underrated horses in the show game based on the probabilities in the win game. The empirical part of our paper takes a similar approach. However, our show pool is small and suffers from several biases. Odds are never below 1. Some agents take bets on behalf of the pool. They will earn a percentage of the bets. This in turn leads them (or a friend) to bet heavily on a huge favorite to show. When this favorite shows (first, second or third) the odds for all three horses are 1. This is important since the organizer might lose money in some races.

The exacta box / quinella ("tvilling") game: Bet on the two first horses. The order does not matter.

The triple ("trippel") game: Bet on the three first horses in the correct order. The odds are calculated in the same way as in the win game except that the totalizator take is 30% in this game. The odds are published on the internet and some important rules are set out. The maximum number of horses in a race is 15. With 15 horses, there are 15*14*13 = 2 730 combinations in the triple game. For each race, the odds for the 200 most likely combinations are published on a single web page. It is possible to download odds for other combinations, but in such case you must ask for these combinations separately. For instance, you can ask for every combination where 1 is the first horse. This is a challenge for players but not for a computer program. More important is that the odds that are published before the race starts are capped at 3 000. The realized odds after the race might be all the money placed minus a 30% totalizator take divided by 2 (the minimum bet is 2). The total amount placed on all combinations is also published continuously. The games differ enormously with respect to the total amount. The variation is between NOK 10 000 and NOK 500 000. For a single race, the pool before the track take in the triple game is usually close to that of the win game.

In the empirical analysis in Section 3, we use the odds from the win game to identify underrated combinations in the triple game. The period considered is from December 2015 to March 2017. Every day we have downloaded the odds for the 200 most likely triple combinations in every game, as well as the winner odds, total wagers, and the results.

For practical reasons, not all games in this period are included. We constructed the database such that it allows a maximum of 20 races each day. Whenever there is a choice between two tracks, we include the race track with the biggest pools. Therefore, most of the races are from Sweden. Nonetheless there should be no bias in the sample. Since we have a Norwegian account, some Swedish games are blocked. The idea behind this is that "Norsk Rikstoto" (the Norwegian part of this corporation) tries to lead bets to Norwegian tracks when possible. The total number of races is 7 474.

3. The profitable robot strategy

In financial economics, an arbitrage might involve exploiting two markets for the same outcome. A true arbitrage is where your portfolio gives a positive return whatever happens. This is close to impossible at the horse track, since such a strategy must include shorting of at least one asset (bet). If we were allowed to give odds (sell bets), this might be possible. Another alternative would be a case where we know that a certain horse will fail. However, this case must involve fraud.

As illustrated in the previous section, a given race gives rise to different games. The approach taken in this paper is to exploit the odds in one game (the win) to identify a favorable bet in a second game (the triple) on the same race. The first step is to transform the win odds into probabilities. Let O_k denote the odds for horse k in the win game. The probability p_k for horse k winning is given by

$$p_k = 0.8/O_k \tag{1}$$

where the track take is 20%. Here, we immediately see a problem with low odds. Odds of 1 gives a probability of 0.8. This might be too low. In addition, we have the low odds bias that suggests a different function. In the absence of alternatives, we will use (1) above to translate odds into probabilities.

The next step is to use the probabilities from the win game to compute probabilities in the triple game. Following the Harville-rule, the probability of the combination (k, l, m) in the triple game is

$$p_{k,l,m} = \frac{p_k \cdot p_l \cdot p_m}{(1 - p_k)(1 - p_k - p_l)} \tag{2}$$

(Harville, 1973). The formula is based on independence of outcomes. It is like drawing from a distribution without replacement. There is extensive literature on alternatives to the Harville-rule. See Henery (1991), Stern (1980), and Lo et al. (1995). These studies suggest that the Harville-rule overestimates the probability of the favorite finishing second and third. Against this background, a generalization of the Harville-rule is

$$p_{k,l,m} = p_k \cdot \frac{p_l^{\lambda_2}}{\sum_{s \neq k} p_s^{\lambda_2}} \cdot \frac{p_m^{\lambda_3}}{\sum_{s \neq k,l} p_s^{\lambda_3}} \tag{3}$$

where λ_2 and λ_3 are positive exponents, see Lo et al. (1995). Observe that with $\lambda_2 = \lambda_3 = 1$, Equation (3) translates back to the Harville-rule.

In our empirical analysis, we use Eq. (3) with exponents $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$, which can be motivated empirically by data from Japan (Lo and Bacon-Shone (1992) and Lo (1994).

In the Appendix, we show similar results using the Harville-rule, (2). We also investigate the results of using Eq. (3) with the combination of exponents $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$ that can be motivated by the Henery model and empirically by data from Medowlands and Hong Kong.

The third step is to combine the market odds $O_{k,l,m}$ for each combination (k, l, m) in the triple game with the corresponding probability $p_{k,l,m}$ inferred from the win game by Eqs. (1)-(3). Given this information, we compute the expected pay-off $E_{k,l,m}$ from a one-dollar bet on combination (k, l, m) as

$$E_{k,l,m} = p_{k,l,m} \cdot O_{k,l,m} \quad (4)$$

A high expected pay-off score indicates a profitable combination in the triple game, whereas an expected pay-off score of less than 1 indicates a non-profitable combination.

We investigate all 7 474 games in the following way. For each game, we sort the bets according to expected pay-off score computed from Eq. (4). We pretend to bet in the triple game on all combinations for which the expected pay-off score is within a given interval (we use an interval size of 0.1). The money placed is such that: (i) It should not influence the odds by more than 1%, which we do not adjust for; and (ii) Given success, the money gain is equal for all combinations. So, given that we bet on the combination (k, l, m) in a game, we will put

$$Bet_{k,l,m} = \frac{0.01 \cdot 0.7 \cdot B}{O_{k,l,m}} \quad (5)$$

on this combination, where the total amount invested in the pool is B and the track take in the triple game is 30%.

For each expected pay-off score interval, we add up all bets and all wins. Our win is our bet times the track odds. We understand that track odds will change when we put our money on. On the other hand, so will the total pool. Our bet is such that this effect is less than 1%. Next, we compute the realized win-to-bet ratio. The results for each expected pay-off score interval are reported in Table 1 below. The second and third columns show the total money placed and the realized win, respectively. Column 4 shows the profit and loss, i.e., the difference between money placed and won. Column 5 shows the realized win-to-bet ratio.

Table 1 includes the 200 combinations with the lowest odds in each game. The bottom line in the table shows that betting on all 200 combinations in every triple game gives a return of 71.07%. Given that the track take is 30%, this is as expected.

Table 1 shows that we have some profitable strategies. The return measured as a percentage will increase with the cut-off fraction we choose. Now, suppose that we only include bets with an expected pay-off score of 1.5 or higher (indicated by the dotted line in Table 1). The total money placed is then NOK 172 383 and the total monetary gain is NOK 187 818. Consequently, the net profit is NOK 15 436, which translates into a return of 8.95%.[3]

There are 7 474 races and of these we win 319 times. Hence the probability of winning is 4.27%. We obtain similar results for the two alternative models (see Tables A1 and A2 in the Appendix).

[3] In the first draft of this working paper, we used a sample of 3 000 races. That sample gave us the cut-off rate of 1.5 and an expectation of 11.3%. In order to rule out randomness, we increased the sample in this paper to 7 474.

Table 1: Betting results versus expected pay-off score

The expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$.*

Expected pay-off score	Money placed (NOK)	Realized win (NOK)	Profit and loss (NOK)	Realized win-to-bet ratio
0.0 – 0.1	23 412	6 691	-16 720	0.29
0.1 – 0.2	121 929	67 251	-54 678	0.55
0.2 – 0.3	282 819	156 791	-126 029	0.55
0.3 – 0.4	452 127	282 887	-169 241	0.63
0.4 – 0.5	580 826	424 999	-155 828	0.73
0.5 – 0.6	624 846	396 480	-228 366	0.63
0.6 – 0.7	590 990	415 829	-175 161	0.70
0.7 – 0.8	507 556	372 203	-135 352	0.73
0.8 – 0.9	402 774	298 832	-103 943	0.74
0.9 – 1.0	310 108	230 806	-79 302	0.74
1.0 – 1.1	228 833	184 639	-44 194	0.81
1.1 – 1.2	168 192	120 708	-47 485	0.72
1.2 – 1.3	118 901	99 391	-19 510	0.84
1.3 – 1.4	86 068	70 467	-15 601	0.82
1.4 – 1.5	61 442	47 901	-13 540	0.78
1.5 – 1.6	44 138	54 157	10 019	1.23
1.6 – 1.7	32 003	34 932	2 929	1.09
1.7 – 1.8	23 948	24 567	619	1.03
1.8 – 1.9	18 151	11 933	-6 218	0.66
1.9 – 2.0	12 837	20 357	7 520	1.59
2.0 – 2.1	9 421	9 669	248	1.03
2.1 – 2.2	7 449	5 112	-2 337	0.69
2.2 – 2.3	5 258	4 830	-428	0.92
2.3 – 2.4	4 416	8 474	4 058	1.92
2.4 – 2.5	3 109	2 475	-634	0.80
2.5 – 2.6	2 477	1 249	-1 228	0.50
2.6 – 2.7	1 861	1 598	-262	0.86
2.7 – 2.8	1 536	1 448	-87	0.94
2.8 – 2.9	1 093	1 132	39	1.04
2.9 – 3.0	863	356	-507	0.41
3.3 – 3.1	669	850	181	1.27
3.1 – 3.2	502	432	-70	0.86
3.2 – 3.3	442	0	-442	0.00
3.3 – 3.4	370	1 995	1 625	5.40
3.4 – 3.5	314	1 636	1 321	5.20
3.5 – 3.6	225	0	-225	0.00
3.6 –	1 301	618	-683	0.48
Total	4 733 207	3 363 692	-1 369 514	0.7107

* Sample: The 200 bets with lowest odds in each game.

The total wagers on a race vary enormously (NOK 10 000 - 500 000). For a race with many horses and a small pool, the expectation from playing a low probability combination is low. The minimum bet is NOK 2. Hence, all combinations with a probability lower than $2/(0.7 \cdot B)$ will have a negative expectation, where B denotes the gross pool. Gamblers know this and will therefore probably restrict their betting to the most probable outcomes. If no one has a bet on the winning combination, the pool is carried over to the next race. This boosts the wagers for this race. However, this only happened twice in our sample.

Table 2 shows the result after excluding bets in games where the net pool $(0.7 \cdot B)$ is below some threshold. The second and third columns show the total money placed and the realized win, respectively. Column 4 shows the profit, i.e., the difference between money won and placed. Column 5 shows the realized win-to-bet ratio.

Table 2: Betting results versus minimum net pool

Computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$.*

(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Min. net pool (NOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio	# winning bets	# races included	Winning bets per race	# bets placed	Winning bets per bet
0	172 383	187 818	15 436	1.0895	319	7 474	4.27 %	81 509	0.39 %
20 000	170 088	186 343	16 254	1.0956	310	7 245	4.28 %	78 619	0.39 %
40 000	137 424	154 888	17 464	1.1271	208	5 120	4.06 %	53 281	0.39 %
60 000	85 050	104 338	19 288	1.2268	104	2 643	3.93 %	25 637	0.41 %
80 000	52 457	70 832	18 376	1.3503	55	1 436	3.83 %	12 677	0.43 %
100 000	35 124	52 115	16 991	1.4838	34	882	3.85 %	7 132	0.48 %

* Sample: The 200 bets with the lowest odds in each triple game with an expected pay-off score ≥ 1.5

Column (d) = (c) – (b); Column (e) = (c)/(b) ; Column (h) = (f)/(g) ; Column (j) = (f)/(i).

The first line in Table 2 corresponds to the results discussed above (where all bets with an expected pay-off score of 1.5 or higher are included). It should come as no surprise that the nominal amounts decrease with the minimum pool size, since more bets are excluded. However, it is interesting to observe that the total effect is that the realized win-to-bet ratio increases.

Table 2 indicates that, if we restrict ourselves to bets with an expected pay-off score higher than 1.5, we can increase the realized win-to-bet ratio by excluding games with small gross pool sizes. This is somewhat surprising. In games with low liquidity (a small pool), it is not unlikely that some profitable combinations are forgotten by the investors. We would expect a bigger pool to remove some of these possibilities. On the other hand, a bigger pool will attract more "noise players". Many of the unlikely combinations are interesting because the maximum pay-off increases with the pool size. We obtain similar results for the two alternative models (see Tables A3 and A4 in the Appendix).

Note, from column (h) in Table 2, that the percentage of winning bets is 3.85% if we exclude all games with a net pool of less than NOK 100 000, as compared to 4.27% in the full sample. Since the number of bets per game differs, we have included the total number of bets. On average, there are more than 10 bets per race. However, the win per bet increases when we exclude small pools, see column (j) in Table

2. It increases from 0.39% for the full sample to 0.48% if we include pools larger than 100 000 only. This indicates that large pools give a more accurate win probability.

4. The system

Our findings above suggests a betting strategy that is made conditional on the size of the net pool. Now, suppose that we split the sample in two conditional on the size of the net pool. The results, using NOK 50 000 as the treshold, are reported in Table 3. Observe that the number of races in the two sub-samples are approximately the same.

Table 3: Betting results for small and large net pools

Computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$.*

(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Net pool (KNOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio	# winning bets	# races included	Winning bets per race	# bets placed	Winning bets per bet
0 - 50	62 938	57 520	-5 418	0.9139	167	3 764	4.44 %	45 936	0.36 %
50 -	109 445	130 298	20 853	1.1905	152	3 710	4.10 %	35 573	0.43 %
Total	172 383	187 818	15 436	1.0895	319	7 474	4.27 %	81 509	0.39 %

* Sample: The 200 bets with the lowest odds in each triple game with an expected pay-off score ≥ 1.5

Column (d) = (c) – (b); Column (e) = (c)/(b) ; Column (h) = (f)/(g) ; Column (j) = (f)/(i).

The results for the large net pool bets are as follows: Money placed is NOK 109 445 and total gain is NOK 130 294. The return is 19.05%. In some races, we will bet nothing, while in others, we will bet on several combinations. In total, we will place 35 573 bets. The total number of wins is 152. The win-per-bet ratio is 0.43%. We obtain similar results for the two alternative models (see Tables A5 and A6 in the Appendix).

This brings us to the question of significance. In order to shed some light on this issue, assume that the result is generated by a series of identical and independent draws from a binomial distribution. We want to test the null hypothesis that the win-to-bet ratio is (less or equal to) 1, i.e., that the binomial win probability p reflects that expected win equals money placed.

Suppose that we consider each bet as a binomial trial, i.e., $n = 35\ 573$ trials. The average win from each winning bet is NOK 130 298/152 = NOK 857.22. The average money placed in each bet is NOK 109 445/35 573 = NOK 3.0766. Consequently, the bet win probability is $p = 3.0766/857.22 = 0.3589\%$. Now, let the stochastic variable x represent the number of wins and assume that it is normally distributed with expectation $E(x) = np$ and standard deviation $\sigma(x) = \sqrt{np(1-p)}$. It follows that given the null hypothesis, the probability of observing 152 or more wins is

$$P(x \geq 152) = 1 - N\left(\frac{152 - np}{\sqrt{np(1-p)}}\right) = 1.55\% \tag{6}$$

where $N(\quad)$ is the standard normal cumulative probability function.

Clearly, the outcome of bets that are placed in the same game are not independent. Suppose that we instead consider each race as a binomial trial, i.e., $n = 3\ 710$ trials. The average win is still NOK 857.22, whereas the money placed in each race is NOK 109 445/3 710 = NOK 29.50. Consequently, the race win probability is $p = 29.50/857.22 = 3.4413\%$. By inserting into Eq. (6) above, we now find that the probability of observing 152 or more wins is 1.42%.

We conclude that the null hypothesis that the expectation is (less or equal to) 1 is rejected. Since most of this information is available before the race starts, it is possible to implement this strategy. Computer skills are necessary to automate the process of placing many bets in the final seconds. In addition, more than 10% of the total amount played in a game will often not be available information before the race starts.

5. Characteristics of the profitable bets

We have now identified many profitable strategies. Common to all the profitable cases we have investigated is that our winning bets are on the favorite coming second approximately 50% of the time. Only a small number of winning bets do not involve the favorite. The favorite comes first approximately 20% and third close to 30% of the time. So the question now is: Why are combinations where the favorite comes second underrated?

a) *The bettors do not understand the probabilities for combinations.*
   This explanation is not likely. Some of the players calculate odds in the triple game on the fly. They have no problem with math.

b) *The bettors do not like to put the favorites second.* They watch the race and want to have a horse to cheer for during the race. They focus on finding the winner. If they bet on a combination with another winner, they go for high odds and exclude the favorite from the triple. In addition, there is a certain element of bragging. I want to impress others with my skill. Winning a complicated high odds triple gives status. A big bet on the favorite first also gives status. Betting on the favorite coming second is not skill. No one can predict that outcome.

   This explanation is more likely. It lacks something, however. Today some of the gamblers use the internet. They only care about monetary gains.

c) *The bettors have heterogeneous beliefs.*
   All bettors maximize their expected value given their own subjective probabilities. Sometimes this results in some underrated combinations. Our robot strategy identifies these combinations together with many other combinations with fair or below fair odds. When we restrict our strategies to reasonably high probability bets (among the 200 with the highest track probability)

and combine that with a high "expected return" (1.5), we mostly capture profitable bets. Since we increase the "expected return" factor to far above 1, we exclude the very most probable bets. Therefore, the favorite coming first is underrepresented.

The reason for this is that there is too much work involved in implementing the strategy described above. In addition, the uncertainty is significant. Remember that the winning probability in a single race is less than 5% on average. There are many examples of such anomalies in the world of finance. The explanation given by Pedersen (2015) is "Efficiently Inefficient". The market is efficient to the degree that it takes too much effort to correct prices.

In addition, the totalizator does not like so-called robot play. The organizer of the pool can refuse to pay out if they somehow classify a bettor as a robot player. This has happened in Sweden in connection with the games where bettors bet on the winner in 4 to 7 races, ATG (2015). The idea behind the robot play described in ATG (2015) is the same as we use in the triple here, i.e., finding combinations where the probability of success is greater than the probability implied by the odds.

At first glance, we would expect robot players to increase the pool, remove inefficiencies, and be positive for the pool. In Scandinavia, the monopoly at the track is used to finance the horse owners and everyone involved in the business. No one is allowed to set up their own races.

In order to understand why robot play is banned, we need to investigate what makes a pool. In the last incident of robot play (the pool did not pay the bettor), the organizer used the following argument (AGT, 2015):

> *The given outcome (four horses in four different games) gave relatively low odds. One account (South African internet account) had many winning bets. This player obviously bet more on low odds combinations than on high odds; hence it is a system. The combinations played were such that the probability of winning exceeded the probability indicated by the odds for these combinations.* (Original statement in Swedish)

Note that this is exactly what we are doing in the strategy above. We also put more money on low odds combinations. For a given game, we win the same amount of money regardless of the combination.

6. Heterogeneous beliefs

So why do people play? First, we have the noise trader. He will put money on the game for the excitement and for the opportunity to win a large premium. The rest of the players (the majority) believe that they have skill. They will combine subjective probabilities and odds. This is best illustrated by looking at the expert industry around the races. There are newspapers, TV shows, experts on the internet, tips on the internet in return for money, and so on. If you listen to the experts before the race, they will often say something like: "Number 4 looks attractive compared to the odds." Gamblers are constantly looking for underrated horses, and they form subjective probabilities for each race. Differences in subjective probabilities are enough to feed the pool. Assume that subjective probabilities are robust. That is, final odds do not alter subjective views. In this case, it is easy to construct an example where all players expect a positive return.

Subjective probabilities and pari-mutuel pools are investigated in the literature. Eisenberg and Gale (1959) show that a unique equilibrium exists in a pari-mutuel game with m individuals and n outcomes. Of the many explanations for the "Favorite–Longshot Bias", Ottaviani and Sørensen (2007) mention heterogeneous beliefs. They follow the model given in Ali (1977).

If we accept subjective probabilities for the win game, we should accept the same assumption in the derivative. In this case, the derivative is the triple. In addition, some investors believe in the Harville-rule, while others do not. This adds up to an inefficient market. To remove some of the inefficiency, a robot is required.

The analytical model for heterogeneous beliefs developed by Ali (1977) is a many bettors, two horses model. In Ottaviani and Sørensen (2015), a model with heterogeneous beliefs is used to explain momentum and reversal in the market for securities. Blough (2008) also discusses a betting model with risk-neutral bettors and heterogeneous beliefs.

The insight from these theoretical models that is relevant to this study is the following: Heterogeneous beliefs in a market might result in an equilibrium where the probability of an event indicated by prices differs from the "true" probability of the event. This holds even if all the market participants are risk-neutral. In the following simple example, we try to illustrate this point.

Consider a pari-mutuel win game with three outcomes (horses) and three investors. The investors maximize the expected present value and have subjective probabilities as shown in Table 3. The rightmost column shows the average subjective probability (equal weights).

Table 3: Subjective probabilities in the win game

	Subjective probability			
Winner	Player 1	Player 2	Player 3	Average
Horse 1	0.6	0.5	0.4	1/2
Horse 2	0.25	0.4	0.35	1/3
Horse 3	0.15	0.1	0.25	1/6
SUM	1	1	1	1

Observe from Table 3 that each player ranks horse 1 first. Still, they will play differently. Suppose that they place their bets as shown in Table 4.

Table 4: Bets placed in the win game

	Bet (amount)			
Winner	Player 1	Player 2	Player 3	Sum
Horse 1	90			90
Horse 2		60		60
Horse 3			37.5	37.5
Sum	90	60	37.5	187.5

The total pool is the sum of the bets placed, i.e., 187.5. Let the track take equal 20%. This gives a net pool of 150. First, calculate the odds. Odds for horse 1: 150/90=1.666. Odds for horse 2: 150/60=2.5. Odds for horse 3: 150/37.5=4.

Second, consider the expected pay-off for each player. The expectation for player 1 (subjective prob.) equals 0.6*1.6666*90=90. The expectation for player 2 (subjective prob.) equals 0.4*2.5*60=60. The expectation for player 3 (subjective prob.) equals 0.25*4*37.5= 37.5. The expected pay-off equals the bet for each player. Hence, they all have an expected return equal to 1.

Third, use the odds to create the win probabilities, and compare them with the average subjective probabilities from Table 3 (rightmost column). Horse 1: 0.8/1.66666=0.48 (average 1/2). Horse 2: 0.8/2.5=0.32 (average 1/3). Horse 3: 0.8/4=0.2 (average 1/6). Observe that the win probabilities computed from the odds are different than the average subjective probabilities.

Now, assume that a similar game is repeated over and over again. In exactly 1/3 of the games, player 1 holds the correct probability distribution, while in 1/3 of the games, player 2 holds the correct distribution, and so on. The expectation of NOK 1 placed on horse 1 is $(1/3) \cdot (0.6 + 0.5 + 0.4) = 0.5$. The probability suggested by the odds is 0.48. Hence, we have a low odds bias. In addition, there is a longshot bias.

Why do the players not learn? First of all, it is not the same game that is repeated. It is a similar game. Suppose you are player number 1. Your subjective probability is 0.6 in 1/3 of the games, 0.5 in 1/3 of the games, and 0.4 in 1/3 of the games. As expected, horse number 1 will win 1/2 of the times on average.

Then use the subjective probabilities in the win game (c.f. Table 3) to compute the subjective probabilities for each combination in the triple game, using the Harville-rule. The subjective probabilities are shown in Table 5. The rightmost column shows the average subjective probability for each combination (using equal weights).

Table 5: Subjective probabilities computed from the win game (Harville-rule)

Combination	Subjective probability			Average
	Player 1	Player 2	Player 3	
123	0.375	0.4	0.2333	0.3361
132	0.225	0.1	0.1667	0.1639
213	0.2	0.3333	0.2154	0.2496
231	0.05	0.0667	0.1346	0.0838
312	0.1059	0.0556	0.1333	0.0983
321	0.0441	0.0444	0.1167	0.0684
SUM	1	1	1	1

Next, inspect the following betting behavior as shown in Table 6.

Table 6: Bets placed in the triple game

Combination	Bet (amount)			Sum
	Player 1	Player 2	Player 3	
123	0	74.4602	0	74.4602
132	41.8838	0	0	41.8838
213	0	62.0501	0	62.0501
231	0	0	25.0587	25.0587
312	0	0	24.8201	24.8201
321	0	0	21.7175	21.7175
Sum	41.8838	136.5103	71.5963	249.9905

In the example, we assume a track take of 20%. We can now compute the odds for each combination by dividing the net pool $(249.9905 \cdot 0.8)$ with the bet placed on that combination. Track probabilities are 0.8 divided by the track odds. The results are reported in Table 7 below and compared with the average subjective probability computed above (rightmost column in Table 5).

Table 7: Odds and probabilities in the triple game

Combination	Track odds	Track odds probability	Average subjective probability
123	2.6859	0.2979	0.3361
132	4.7749	0.1675	0.1639
213	3.2231	0.2482	0.2496
231	7.9810	0.1002	0.0838
312	8.0577	0.0993	0.0983
321	9.2088	0.0869	0.0684
Sum		1	1

The expected pay-off score for a given bet as seen from each player's perspective is found by multiplying the track odds with the subjective probability (c.f. Table 5 above). The results are reported in Table 8.

Table 8: Expected pay-off score for each player (using subjective probabilities)

Combination	Track odds	Player 1	Player 2	Player 3
123	2.6859	1.0072	**1.0744**	0.6267
132	4.7749	**1.0744**	0.4775	0.7958
213	3.2231	0.6446	**1.0744**	0.6942
231	7.9810	0.3990	0.5321	**1.0744**
312	8.0577	0.8532	0.4476	**1.0744**
321	9.2088	0.4063	0.4093	**1.0744**

Table 8 shows that all placed bets (numbers in bold face) are profitable given the player's subjective probabilities, i.e., the subjective pay-off score is greater than one. The return is equal to 7.44% for all placed bets in this example. This indicates why the track take might be higher in the triple game than the win game, which is the case on tracks in Scandinavia.

Next, we want to investigate the expected pay-off score given the probabilities that follow from the win game probabilities (see the discussion following Table 5) and the Harville-rule (c.f. Eq. (2)).

Table 9: Expected pay-off score (probabilities from win game and Harville-rule)

Combination	Track odds	Computed prob. from win game	Expected pay-off score
123	2.6859	0.2954	0.7934
132	4.7749	0.1846	0.8815
213	3.2231	0.2259	0.7280
231	7.9810	0.0941	0.7511
312	8.0577	0.12	0.9669
321	9.2088	0.08	0.7367
Sum		1	

We observe from Table 9 that the expected pay-off score differs from combination to combination. This might explain why we predict so many profitable games. Some of the assumed profitable bets predicted simply follow from the fact that we are not able, based on the odds in the win game, to compute correct probabilities for these combinations. If we return to Table 7, we see that the average probability, (average of the subjective probabilities) is different from the probability computed from the mathematical odds in the win game using the Harville-rule.

If, in addition, we introduce more horses and a noise trader in this model, we might find many "profitable" combinations. We assume that the noise trader will bet on high odds with low expected value combinations. This explains the fact that large pool games are more profitable (given our robot) than small pool games in this sample. In a large pool, there is more disagreement or more noise players.

When introducing a robot, we pick some combinations and lower the odds for these. Sometimes these combinations materialize. In these cases, it might happen that all the skilled players will have a lower predicted probability of the outcome than the probability implicitly given by the odds. They complain, and the organizer finds a single internet player with a lot of money on this peculiar combination. The skilled players, who are the backbone of the market, demand action. The action taken by the Swedish pool is to ban so-called robot bets. Inefficiency survives.

7. Conclusion

We have collected odds and results from 7 474 horse races in Norway and Sweden for a period of approximately 1.5 years. Based on the odds from the win game, we construct a profitable betting strategy for the corresponding triple game. Given a 30% track take, the existence of a profitable strategy is surprising. We did expect that bigger pools (more liquidity) would remove this anomaly. This is not the case. More players, which leads to bigger pools, increases the profitability of the system.

In this paper, we have restricted our attention to the 200 combinations in the triple game with the lowest odds. For each race, this information is readily available from a single web page. However, we have conducted a simple test that indicates that it might be profitable to include combinations with higher odds. It is tedious work to obtain information on higher odds combinations from the web. In addition, the reported odds is capped at 3 000. Just before the race starts, the odds for the more unlikely combinations change a lot. It is a question if it is possible to exploit profitable high odds combinations in practice.

References

Ali, Mukhtar M: "Probability and Utility Estimates for Racetrack Bettors." *Journal of Political Economy*. 1977. 85(4). 803–815.

Ash, Peter, and Richard E. Quandt: "Efficiency and Probability in Exotic Bets." *Economica* 54 (August 1887): 289–98.

ATG. http://trav24.no/index.php/Nyhetsarkiv/ATG–med–nye–tiltak–mot–robotspill, 2015.

Blough, Stephen R.: "Differences of Opinion at the Racetrack." In Efficiency of Racetrack Betting Markets" *World Scientific*. 2008. Editors: Hauch, D.B., Lo. V.S. and William T. Ziemba.

Eisenberg, E., and D. Gale: «Consensus of Subjective Probabilities. The Pari–mutuel Method." *Ann. Math. Statist*. 30 1959, 165–168.

Harville, David A.: "Assigning Probabilities to the Outcomes of Multi–Entry Competitions." *J. American Statis. Assoc*. 68 (June 1973): 312–16.

Hausch, Donald B., William T. Ziemba, and Mark Rubinstein: "Efficiency of the Market for Racetrack Betting." *Management Science*. 1981. 27. 1435–52.

Henery, R. J.: “Permutation Probabilities as Models for Horse Races.” *J. Royal Statistical Society B*. 43. 1 1981. 1435–52.

Hurley, William and Lawrence McDonough: “Note on the Hayek Hypothesis and the Favorite–Longshot Bias in Parimutuel Betting.” *American Economic Review*, 1995, 85(4), 949–955.

Lo, V.S.Y.: “Application of Running Time Distribution Models in Japan.” In D.B. Hausch, V.S.Y Lo, and W.T. Ziemba, (Eds.). *Efficiency of Racetrack Betting Markets*, Academic Press, 1994, 237–247.

Lo, V.S.Y., and J. Bacon–Shone: “An Approximation to Ordering Probabilities of Multi–entry Competitions”. Research report 16, Department of Statistics, University of Hong Kong, 1992.

Lo, V. S. Y., John Bacon–Shone, and Kelly Busche: “The Application of Ranking Probability Models to Racetrack Betting.” *Management Science*, Vol. 41, 6, 1995, 1048–59.

Ottaviani, M. and Petter N. Sørensen: «The Favorite–Longshot Bias: An overview of the Main Explanations» *Handbook of Sports and Lottery Markets*. Edited by Donald B. Hausch and William T. Ziemba. North–Holland, Amsterdam, 2008, pp. 83–101.

Ottaviani, M. and Petter N. Sørensen: “Price Reactions to Information with Heterogeneous Beliefs and Wealth Effects: Underreaction, Momentum, and Reversal” *American Economic Review* 2015, 105(1): 1–34.

Pedersen, L. H.: *Efficiently Inefficient*, Princeton University Press, 2015.

Thaler, Richard H. and William T. Ziemba: “Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries.” *Journal of Economic Perspectives*, 1988, 2(2), 161–174.

Weitzman, Martin: “Utility Analysis and Group Behaviour: An Empirical Analysis.” *Journal of Political Economy*, 1965, 73(1), 18–26.

Appendix. Betting results for alternative models

In the empirical analysis in Section 3, the expected pay-off score is based on probabilities computed from the Harville-rule. Table 1 shows the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$. Table A1 shows the betting results following from the Harville-rule.

Table A1: Betting results versus expected pay-off score*
Harville rule (Eq. (2))

Expected pay-off score	Money placed (NOK)	Realized win (NOK)	Profit / loss (NOK)	Realized win-to-bet ratio
0.0 – 0.1	76 863	31 201	-45 661	0.41
0.1 – 0.2	242 981	146 301	-96 680	0.60
0.2 – 0.3	367 633	219 406	-148 226	0.60
0.3 – 0.4	449 251	295 940	-153 311	0.66
0.4 – 0.5	490 709	373 836	-116 873	0.76
0.5 – 0.6	491 699	344 258	-147 441	0.70
0.6 – 0.7	459 749	327 152	-132 597	0.71
0.7 – 0.8	416 699	295 111	-121 588	0.71
0.8 – 0.9	358 054	232 457	-125 597	0.65
0.9 – 1.0	297 129	213 728	-83 402	0.72
1.0 – 1.1	237 589	181 436	-56 153	0.76
1.1 – 1.2	190 116	144 919	-45 197	0.76
1.2 – 1.3	145 813	108 979	-36 834	0.75
1.3 – 1.4	114 938	101 543	-13 395	0.88
1.4 – 1.5	88 512	61 508	-27 004	0.69
1.5 – 1.6	66 159	45 267	-20 892	0.68
1.6 – 1.7	51 156	46 682	-4 475	0.91
1.7 – 1.8	39 454	35 786	-3 668	0.91
1.8 – 1.9	30 648	31 081	434	1.01
1.9 – 2.0	24 134	26 627	2 494	1.10
2.0 – 2.1	17 959	17 461	-498	0.97
2.1 – 2.2	15 048	17 296	2 248	1.15
2.2 – 2.3	11 593	16 001	4 407	1.38
2.3 – 2.4	9 274	10 067	793	1.09
2.4 – 2.5	7 637	6 638	-999	0.87
2.5 – 2.6	5 924	3 869	-2 054	0.65
2.6 – 2.7	4 561	7 434	2 873	1.63
2.7 – 2.8	3 984	2 166	-1 818	0.54
2.8 – 2.9	2 986	3 299	313	1.10
2.9 – 3.0	2 514	2 745	230	1.09
3.3 – 3.1	2 204	2 075	-129	0.94
3.1 – 3.2	1 784	869	-915	0.49
3.2 – 3.3	1 280	466	-814	0.36
3.3 – 3.4	1 162	2 376	1 213	2.04
3.4 – 3.5	1 071	228	-843	0.21
3.5 – 3.6	735	3 014	2 279	4.10
3.6 –	4 214	4 470	256	1.06
Total	4 733 216	3 363 693	-1 369 523	0.7107

* Sample: The 200 bets with lowest odds in each game.

Table A2 shows the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$ (Henery model).

Table A2: Betting results versus expected pay-off score*
Henery model (Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$)

Expected pay-off score	Money placed (NOK)	Realized win (NOK)	Profit / loss (NOK)	Realized win-to-bet ratio
0.0 – 0.1	8 487	840	-7 648	0.10
0.1 – 0.2	52 068	19 770	-32 298	0.38
0.2 – 0.3	202 545	102 193	-100 353	0.50
0.3 – 0.4	465 570	331 696	-133 875	0.71
0.4 – 0.5	706 071	428 734	-277 337	0.61
0.5 – 0.6	778 042	525 230	-252 812	0.68
0.6 – 0.7	694 661	475 739	-218 922	0.68
0.7 – 0.8	548 478	394 679	-153 799	0.72
0.8 – 0.9	401 203	326 232	-74 971	0.81
0.9 – 1.0	282 107	223 809	-58 298	0.79
1.0 – 1.1	189 708	155 618	-34 090	0.82
1.1 – 1.2	127 931	114 978	-12 953	0.90
1.2 – 1.3	86 586	84 456	-2 130	0.98
1.3 – 1.4	58 844	50 768	-8 076	0.86
1.4 – 1.5	39 917	38 184	-1 732	0.96
1.5 – 1.6	26 842	26 143	-699	0.97
1.6 – 1.7	18 416	16 915	-1 501	0.92
1.7 – 1.8	13 052	10 914	-2 139	0.84
1.8 – 1.9	8 954	11 899	2 945	1.33
1.9 – 2.0	6 294	6 859	565	1.09
2.0 – 2.1	4 659	4 473	-186	0.96
2.1 – 2.2	3 284	3 446	162	1.05
2.2 – 2.3	2 336	1 711	-625	0.73
2.3 – 2.4	1 625	2 081	456	1.28
2.4 – 2.5	1 306	850	-456	0.65
2.5 – 2.6	950	1 210	260	1.27
2.6 – 2.7	648	2 024	1 377	3.13
2.7 – 2.8	565	694	129	1.23
2.8 – 2.9	418	0	-418	0.00
2.9 – 3.0	304	326	22	1.07
3.3 – 3.1	237	0	-237	0.00
3.1 – 3.2	250	0	-250	0.00
3.2 – 3.3	143	603	459	4.20
3.3 – 3.4	154	0	-154	0.00
3.4 – 3.5	124	0	-124	0.00
3.5 – 3.6	95	0	-95	0.00
3.6 –	358	618	260	1.73
Total	4 733 233	3 363 693	-1 369 540	0.7101

* Sample: The 200 bets with lowest odds in each game.

Table 2 above shows the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$ for bets with an expected pay-off score greater or equal to 1.5 for various levels of minimum net pool. Tables A3 and A4 show the results for the two alternative models. Table A3 shows the betting results following from the Harville-rule and only including bets with an expected pay-off score greater or equal to 1.8 for various levels of minimum net pool.

Table A3: Betting results versus minimum net pool*
Harville rule (Eq. (2))

(a)	(b)	(c)	(d)	(e)
Min. net pool (NOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio
0	148 711	158 182	9 471	1.06
20 000	146 884	157 528	10 644	1.07
40 000	120 142	131 968	11 826	1.10
60 000	76 071	87 568	11 497	1.15
80 000	47 432	60 724	13 292	1.28
100 000	31 753	44 719	12 965	1.41

* Sample: The 200 bets with lowest odds in each triple game with an expected pay-off score ≥ 1.8.
Column (d) = (c) – (b); Column (e) = (c)/(b).

We observe that the realized win-to-bet ratio increases as bets with small pools are excluded.

Table A4 show the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$ (Henery model) and only including bets with an expected pay-off score greater or equal to 1.8 for various levels of minimum net pool.

Table A4: Betting results versus minimum net pool*
Henery model (Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$)

(a)	(b)	(c)	(d)	(e)
Min. net pool (NOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio
0	32 703	36 795	4 092	1.13
20 000	32 088	36 489	4 401	1.14
40 000	24 454	27 847	3 393	1.14
60 000	13 412	16 403	2 992	1.22
80 000	7 884	8 110	226	1.03
100 000	5 044	6 356	1 312	1.26

* Sample: The 200 bets with lowest odds in each triple game with an expected pay-off score ≥ 1.8.
Column (d) = (c) – (b); Column (e) = (c)/(b).

We observe that the realized win-to-bet ratio increases as bets with small pools are excluded.

Table 3 above shows the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.88$ and $\lambda_3 = 0.81$ for bets with an expected pay-off score greater or equal to 1.5 for small and large net pools (using NOK 50 000 as treshold). Tables A5 and A6 show the results for the two alternative models. Table A5 shows the betting results following from the Harville-rule and only including bets with an expected pay-off score greater or equal to 1.8 for small and large net pools.

Table A5: Betting results for small and large net pools*
Harville rule (Eq. (2))

(a)	(b)	(c)	(d)	(e)
Net pool (KNOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio
0 - 50	51 800	46 560	-5 240	0.8988
50 -	96 914	111 623	14 709	1.1518
Total	148 714	158 183	9 469	1.0637

* Sample: The 200 bets with lowest odds in each triple game with an expected pay-off score ≥ 1.8.
Column (d) = (c) – (b); Column (e) = (c)/(b).

Table A6 show the betting results if the expected pay-off score is based on probabilities computed from Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$ (Henery model) and only including bets with an expected pay-off score greater or equal to 1.8 for small and large net pools.

Table A6: Betting results for small and large net pools *
Henery model (Eq. (3) using $\lambda_2 = 0.76$ and $\lambda_3 = 0.62$)

(a)	(b)	(c)	(d)	(e)
Net pool (KNOK)	Money placed (NOK)	Realized win (NOK)	Profit (NOK)	Win-to-bet ratio
0 - 50	14 321	14 789	469	1.0327
50 -	18 382	22 005	3 623	1.1971
Total	32 703	36 795	4 092	1.1251

* Sample: The 200 bets with lowest odds in each triple game with an expected pay-off score ≥ 1.8.
Column (d) = (c) – (b); Column (e) = (c)/(b).