

Seminar: Experimental Economics WS 2011/2012

Randomization in laboratory basketball

Cornelia Baierl

baierl06@stud.uni-passau

Table of Contents

List of figures:	4
1. Introduction	5
2. Experimental design and course	5
2.1 Course of the basic treatment	6
2.2 Course of the variation treatment “Hot Hand”	7
3. Hypotheses	8
4. Results and analysis	10
5. Limitations	14
6. Conclusion	14
Bibliography	16
Appendix	17

List of figures:

figure 1: payoffmatrix displayed to all subjects	7
figure 2: Hot Hand decision matrix displayed to all subjects	8
figure 3: relative deviation of each strategy from MSE	10
figure 4: predicted frequency played by male and female subjects in comparison to MSE probability	11
figure 5: frequency of selected patterns in a three- strategy experiment	12
figure 6: Plotted frequency of observed 2- tuples (xx) on MSE probability	12
Figure 7: Plotted frequency of observed 4-tuples (xxxx) on MSE frequency	13
Figure 8: Plotted frequency of observed 3-tuples (xxx) an MSE probability	13

1. Introduction

Studies by Ross and Levy (1958) pointed out that in games that require the ability to randomize young children are much more able to do this in comparison to grown-ups. This means that infants hit an identical and independent distribution within small samples quite well - similarly as a computer would randomize strategies. The reason for this phenomenon is that adults are believed to know more about statistical distributions and probability laws and therefore they are liable to the misbelief that also in producing small sequences, the law of large numbers and statistical distribution applies (see Camerer). Moreover, even though the strategies are independently distributed, adults also consider their previous choices and deem the reappearance of a strategy as "improbable". This shows in the fact that certain patterns, e.g. playing the same strategy several times in a row, is believed to be "unlikely" and consequently certain strategies/ patterns are played too less or too often (see Camerer). There is a multitude of studies about human randomization, beginning with the early studies. Then, Atkinson and Suppes (1958) amongst others found out about a perceived misbelief of game theoretic predictions. The studies were continued with more recent studies from the 1980s on (see Camerer).

Randomization in sports is another field related to that issue and was explored amongst others by Palacios-Huerta (2003), who did research about randomization in terms of a zero- sum game in professional football. It is my interest to compare randomization strategies in a laboratory basketball setup to Mixed Strategy Equilibrium (MSE) probabilities. My computer- based experiment, however, took place under laboratory conditions, where subjects should empathize with real players in game situations.

This work is also dedicated to another issue in that it also deals with randomization, namely the Hot Hand phenomenon, originating from basketball. Hot Hand refers to the belief in a lucky streak. There is fair evidence for the existence of the "Hot Hand" in real basketball, (see, e.g., Gilovich et al.), and C.F. Camerer (1989) raises the question if "...the belief in the Hot Hand stems from misunderstanding of random sequences in general...". To me it seems worth examining the question if the Hot Hand phenomenon also exists in laboratory basketball.

This paper is structured the following way: in the adjacent sequence, a description about the experimental design and course is given; part three contains the verbalization of the hypotheses that are examined in the subsequent section. There are certain limitations to the experiment that are dwelled on in part four and last but not least a conclusion about the main results is going to

complete the work.

2. Experimental design and course

The session consisted of two treatments: the Basic Treatment which serves to identify the randomization patterns and the variation treatment called “Hot Hand” that should test for the Hot Hand phenomenon.

2.1 Course of the basic treatment

The experiment took place in November 2011 at the University of Passau. The subjects, female and male students, were introduced to a computer based test using the software z-Tree: Zurich Toolbox for Readymade Economic Experiments Fischbacher (2008). In the beginning of the session the subjects were instructed in German language about the further course of the experiment, which is explained in the following:

At all times, two persons were playing computerized basketball against each other. One player was the offender who had to choose one out of 3 strategies, namely “shoot”, “pass” or “drive”. The other player, who was the defender, could opt for one out of two strategies: “block” or “steal”. At this point it is important to mention that they were playing simultaneously so that they could not guess which decision the opponent took (double blindness). The subjects’ decision was based on a payoff matrix, which is pictured below. It firstly appeared at the last step of instruction and was repeatedly shown in every round. It contains the payoffs (in Nowitzkifranken = NF) according to the decisions both of the offender and the defender. Common randomization games, like matching pennies or stone scissors and paper are so called Zero-sum- games, which means that one player’s gain is the other’s loss and the summarized payoffs are zero, see Camerer (2003:118ff). In my experiment, however, the expected payoff (p^e) for the offender for each strategy is 5. ($p^e(S)=0,5*0+0,5*10=5$; $p^e(P)=0,5*7+0,5*3=5$; $p^e(D)=0,5*10+0,5*0=5$) and for the defender, the expected payoff is approximately 5. ($p^e(B)=0,33*12+0,33*3+0,33*0\approx 5$; $p^e(S)=0,33*0+0,33*5+0,33*10\approx 5$). So considering the constructed payoff the game is similar to a non - constant sum game with one big difference: the strategy “pass” is less risky for the offender than the strategies “shoot” or “drive” since it awards a save payoff of at least 3 NF. Otherwise, in the case that the pass was not stolen and so the offender’s strategy was successful, the payoff is only 7 NF. Therefore, when analyzing the results this just mentioned “riskaverse” decision has to be taken into account.

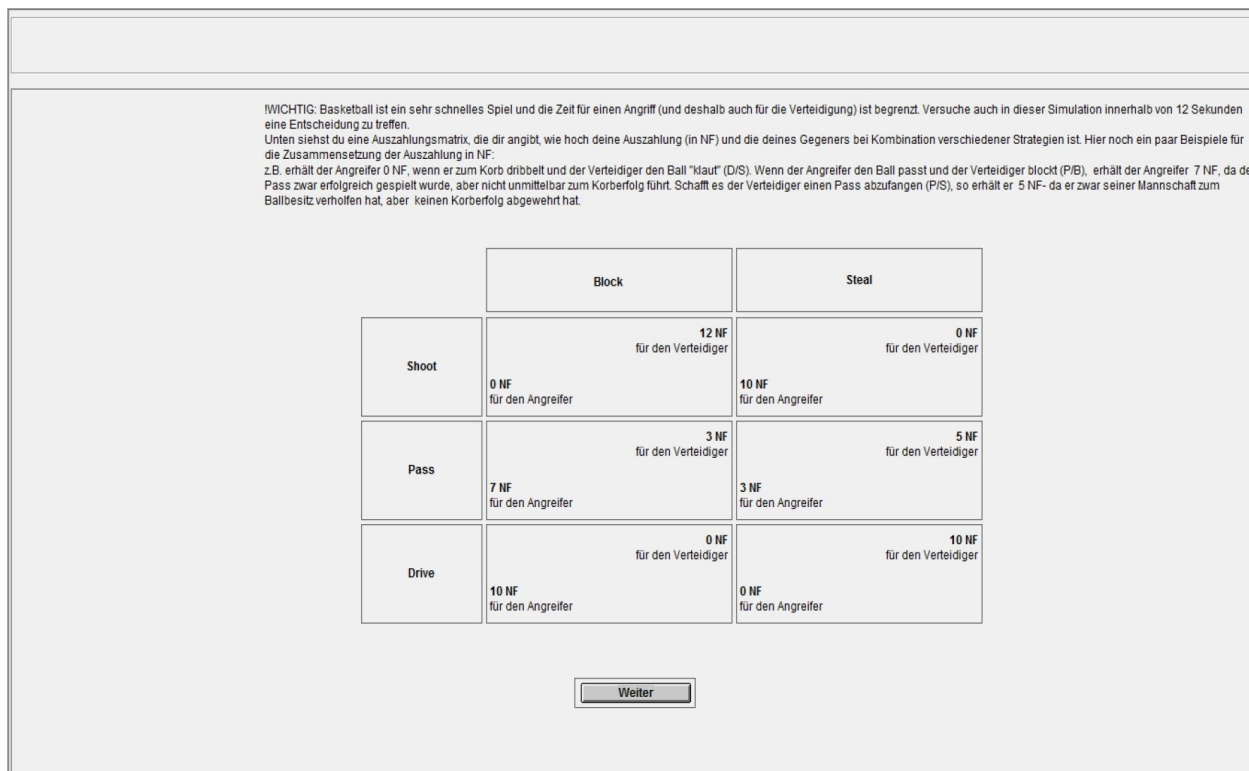


figure 1: Payoffmatrix displayed to all subjects

After each round, the computer matched the results and calculated the payoff for the players, who were neither informed yet about the decision of the opponent nor were they given information about their results after each round. Instead, the subjects played 15 rounds after which a role reversal took place and the offender became defender and vice versa. After having played another 15 rounds, now in their different roles the subjects finished the basic treatment and turned to the next treatment, described in the following.

2.2 Course of the variation treatment “Hot Hand”

Now, all participants were given information about three other players’ performances referring only to the average score of the offense decisions. (One of those players whose information was disclosed was actually the opponent and the other players were another independent couple that was matched together to a group by the computer program.) As one can see in the screenshot below, the announcement of the performance was split into the performance of the first 12 rounds of offense and the last 3 rounds of offense. The subjects were now given the following scenario: they were playing a basketball match which was 3 seconds close to the end and the score was tight- so the last offense action should decide about winning or losing the game. Thus, the participants were asked to

choose one player, according to his or her performance, to pass on the ball so that he or she should take the last offense decision and consequently decide about the outcome of the match. There were further steps taken that don't account for the analysis but completed the game for the students: each person played one further offense and defense round. In the end each player received a payoff that he or she gained during the 30 rounds of offense and defense plus the payoff of the last offense round and additionally the payoff which was gained by the teammate they chose to pass the ball in the last round. All those payoffs were summed up to a final payoff for each player.

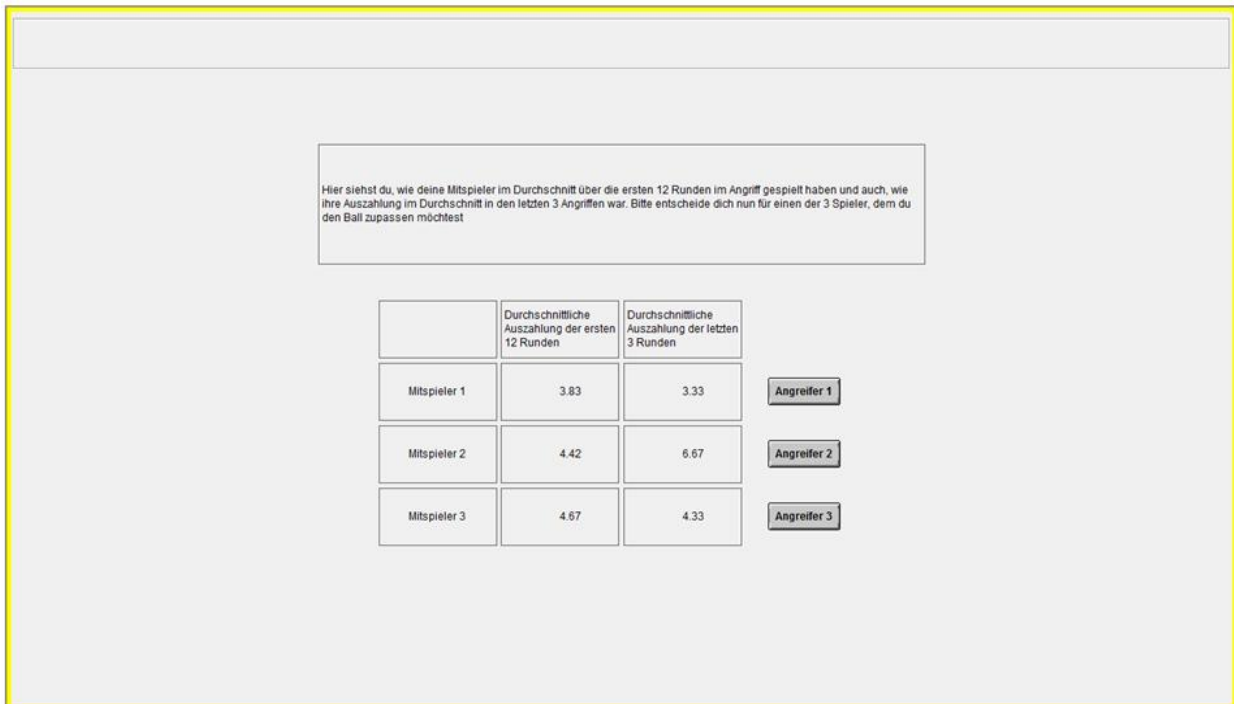


figure 2: Hot Hand decision matrix displayed to all subjects

3. Hypotheses

Basic treatment

According to game theory, when choosing the strategy, both the offender and the defender should play according to the Mixed Strategy Equilibrium (MSE). This means that he/ she chooses the strategies so unpredictably, so that his or her opponent is indifferent about the choice of the strategic answer, see Gibbons (1992: 30)

Making the opponent indifferent in strategies requires good randomization skills. As was already mentioned above, the skill to randomize strategies gets lost with the age. This loss of ability is

attributed to the fact that people learn about probabilities, such as the law of large numbers etc. when growing up. When testing for randomization in sports, Palacios Huerta (2001) observed professional football players over a period of several years. They observed the strategies “right” and “left” which are the corners of the goal the penalty shots entered. They also checked the frequency with which the corners were chosen and the alternation of the chosen corners of the goal. They compared the professionals’ performance to the randomization- performance of non- professionals and found out that this certain ability that gets lost while growing up, can be relearned. Since the subjects in my experiments were all non - professionals in sports they should also show a bias in randomization patterns.

H1: Subjects rather tend to play according to the “law of small numbers” than to play MSE in laboratory sports

Variation treatment: Hot Hand

There are already many studies/ papers even dating back to the 1970s that dealt with the Hot Hand phenomenon, especially applied to basketball. Gilovich, Vallone and Tversky (1985:295) gave the following definition: “the belief in the Hot Hand and the „detection“ of streaks in random sequences is attributed to a general misconception of chance according to which even short random sequences are taught to be.” In order to test this, spectators were asked if players who had scored two or three times successively had a better chance of making a shot. 91% agreed to that, see Gilovich, Vallone and Tversky (1985:297). Fans also believed that it was important to pass the ball to someone who had just made several (two, three or four) shots in a row. To confer this Hot Hand topic to my computer- based experiment, the subjects who played computerized basketball could choose to pass the ball to one of their three teammates in order to complete the offense. Since the subjects were given their teammates’ performance about the first twelve offense actions and the last three offense actions on average the crucial question is according to which calculus the subjects decided? Did they put the emphasis on the long- run average performance (twelve rounds) or even the overall performance or were they only convinced about the performance over the last three offense actions? If the latter one was the case, then the subjects believed somehow in a lucky streak and therefore evidence for the existence of Hot Hand is given.

H2: Subjects in laboratory Basketball tend to believe that other players have a “Hot Hand”

When examining the psychological background of the Hot Hand phenomenon in sports University of

Colonia, Jörn Köppen (2011: 116) found out that male basketball players are more susceptible to the belief in the “Hot Hand”.

H3: In computerized basketball, women tend to play less according to “Hot Hand”

4. Results and analysis

The experiment was accomplished by 25 male and 29 female students of the University of Passau on 21st November 2011. Actually the expected average payoff for every strategy is 5. Therefore, each strategy should have been played with equal frequency. However, if we consider that the strategy “pass” offers less risky outcomes (7 NF with the probability of 50% and 3 NF with the probability of 50%) in comparison to “shoot” (10 NF with a probability of 50% and 0 with a probability of 50%) and “drive” (10 NF with a probability of 50% and 0 with a probability of 50%), the overplaying of the strategy “pass” can be explained through risk aversion. In comparison to “pass”, “shoot” was played 10% less and “drive” was also chosen 6% less than MSE would predict.

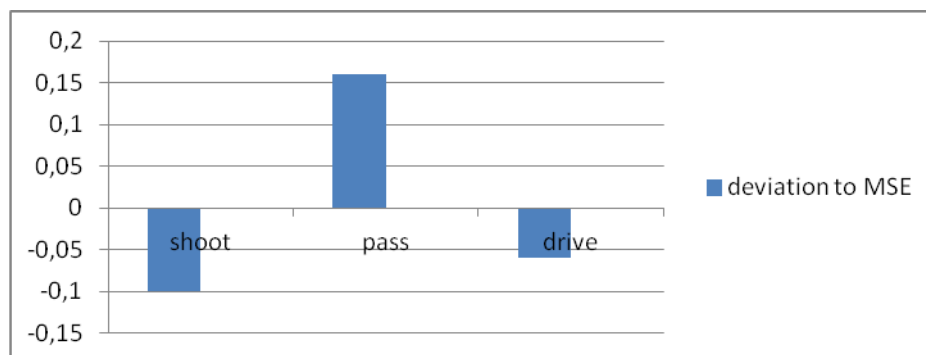


figure 3: relative deviation of each strategy of MSE

Gender differences within the frequencies are outlined in figure 4. The green bars show the MSE probability (actually 33%) for every single strategy, if the “risk aversion” referring to the strategy “pass” is not taken into account. It is yet apparent that male subjects chose “pass” less often than the females did. Instead, male subjects opted more often for “shoot”.

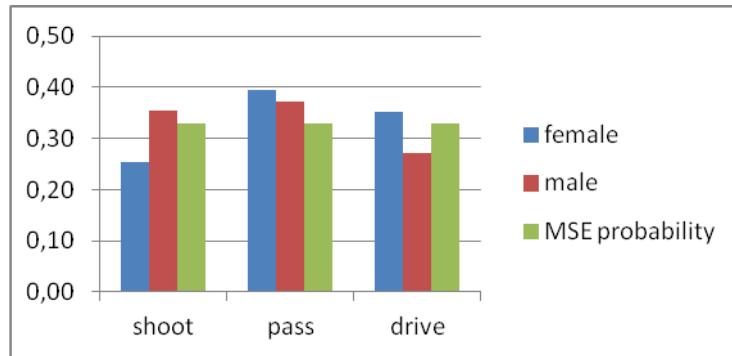


figure 4: Observed frequencies played by male and female subjects in comparison to MSE probabilities

In the following we consider whole patterns of strategies which are shown in figure 5. The letters x y and z stand for the different strategy patterns: xx displays all possible repeating patterns (“shoot” or “pass” or “drive”) the offender played twice in a row, whereby x/yz stands for all possible altering 2-tuples and xyz which indicates all possible altering 3-tuples, containing each strategy “shoot”, “pass” or “drive” only once. The pattern xxx and xxxx denotes any strategy that has been played three or four times in a row respectively. Finally, it becomes clear that the results displayed above in general support the already existing studies cited in Camerer (2003:136): repeated patterns (i.e. xx, xxx and xxxx) are played with too low frequencies and altering strategies tend to be overplayed (which cannot be confirmed with certainty here as the altering 2-tuples are played with lower and the altering 3-tuples are played with higher frequency than the expected MSE frequency).

Pattern type	Predicted frequency relative to MSE	Frequency in treatment	Expected frequency if MSE
xx	lower	0,29	0,33
xy	lower	0,63	0,66
xyz	higher	0,24	0,22
xxx	lower	0,03	0,11

xxxx	lower	0,01	0,04
------	-------	------	------

figure 5: Frequency of selected patterns in a three- strategy experiment

There is also evidence that male and female subjects underplayed and overplayed the strategy patterns in the same way: repeating patterns tended to be played less often than predicted by MSE probability. However, female students, in the aggregate, approximated MSE frequencies better than their male colleagues, especially referring to the altering strategies. (C.f. appendix 1a and 1b).

The before mentioned phenomenon that repeating patterns are played with lower than MSE frequency appears more frequently, the bigger the repeating patterns are. This can be displayed in a plotted diagram where the red line shows the MSE frequency and each point stands for one observation: within the 2-tuples plot, the number of observations over the linear seems to correspond more or less to the number of observations above the line.

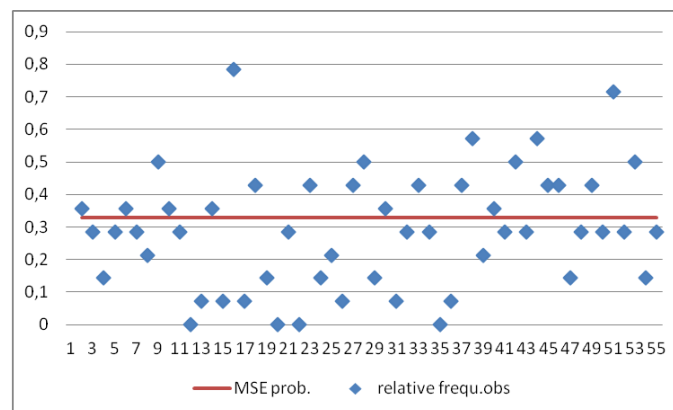


figure 6: Plotted frequency of observed 2-tuples (xx) on MSE probability

However, having a closer look at the 3- tuples plot, one can see that those patterns are played with less than the expected frequency (most of the observations are under the red line).

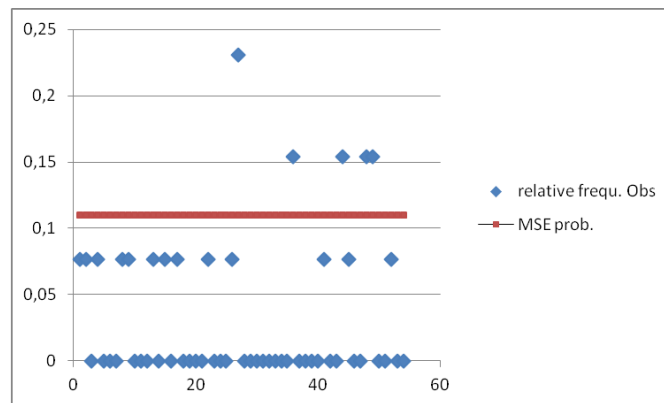


Figure 7: Plotted frequency of observed 3-tuples (xxx) an MSE probability

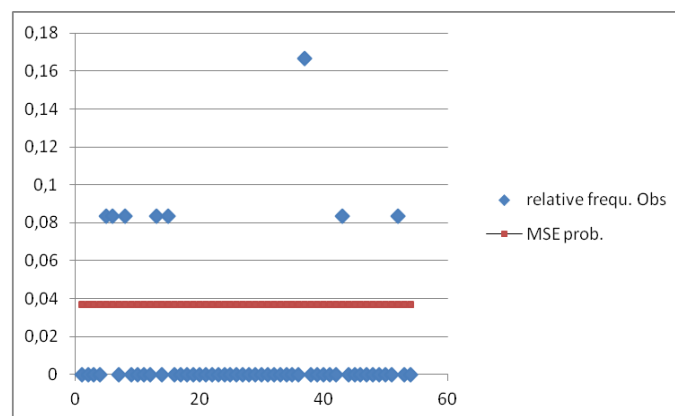


Figure 8: Plotted frequency of observed 4-tuples (xxxx) on MSE frequency

Testing the already described guess for statistical significance, I used a two tailed t-test for only one control sample (t-test type 1) at a significance level of 5% for every single pattern type to test the hypothesis that the arithmetic mean of the observed frequencies equals the arithmetic mean of the MSE frequencies (H_0 : MSE prob. = relative frequency observed). Reckoning the stata© output (c.f. appendix 2b and 2c) one can see that the p value (0,000) for the greater patterns (xxx and xxxx) is smaller than the pre-specified alpha level (1%), yet, for the other patterns this cannot be verified. Consequently, the arithmetic mean of the observed frequencies and the MSE frequencies is significantly different and hence, H_1 can be accepted for at least the pattern types xxx and xxxx.

In order to test whether the Hot Hand also exists in laboratory sports, I ran a Logit- regression with the choice of a player (whom the ball was passed to) as the dependant variable. As explaining

variables, I included “performance of the last 3 rounds” and “total performance”. To investigate whether the players’ gender had an impact on the probability of playing Hot Hand, I included an interaction term of the dummy variable gender (with 0 denoting male, 1 denoting female) with the variable “performance of the last 3 rounds”. Since every subject delivered three observations, the number of observations N increased to 162. If the variable “performance of the last three rounds” had a significantly positive coefficient, the H2: There is a Hot hand phenomenon in laboratory sports, could be accepted.

By contrast, the regression points out that the variable “performance of the last three rounds” has a negative (-0,160)but not significant influence (with a p-value greater than 5%) on the choice of a player, which means that there is no statistic evidence for the Hot Hand Phenomenon and H2 has to be rejected. However, the observed “total performance” of the team mate has a quite strong (0,544) positive impact on the choice of a player which is highly statistically significant ($p < 0,001$). Since the average performance of their team mates was not visible to the subjects, it seems that the students actually calculated the total performance and used that figure as a means of their choice for the teammate.

The interaction term “Hot Hand female” denotes the impact a player’s gender has on the likelihood to be influenced by the performance of the past three rounds, i.e. to play “Hot Hand”. If it was negative and significant, one could say that women tend less to a “Hot Hand decision” and so H3 could be accepted. Yet, the table shows a coefficient that is slightly positive at a value of 0,0247, however, with a low statistical significance, implying that H3 cannot be accepted.

5. Limitations

Several limitations that applied to the experiment could have biased the results. Firstly, the students were aware of the fact that even in the end of the experiment there was no money paid. Possibly they did not strive or concentrate so much as if they had known that a pecuniary remuneration was paid. Furthermore, there was a programming mistake which resulted in the fact that one offender was playing against two defenders and another offender was playing separately. Thirdly, the game was constructed the way that intermediate data after each round was not given so players could not guess out whether their counterpart did play MSE. Having said that, it is important to mention that this approach refers to studies of Rapoport and Budesco (1994) and several other studies that also gave no intermediate data. The reason is that players, given the intermediate data, could have responded to their opponent, rather than playing MSE. Last but not least, while doing the analysis I

recognized that some subjects within the Hot Hand treatment when choosing one of three players chose another subject that did neither have the best performance on total nor on the first 12 rounds, nor on the last 3 rounds. This raises the question whether the subjects did not understand the idea that they should choose the very player they consider to be able to successfully finish the match. They may, however, have been playing deliberately. Unfortunately, a distinction between the two within this experiment is not possible.

6. Conclusion

Finally a short roundup about the main results of my experiment will be given.

The basic treatment whereby the skill to play MSE (i.e. the choice of strategies that makes the opponent indifferent in his or her strategy) should be tested delivered the following main results:

There is evidence for the fact that altering patterns are played with too little frequency and the patterns that include equal strategies are played insufficiently according to MSE probabilities. This means that the larger the patterns of equal type are (xxx), (xxxx), the more the relative deviation to MSE and the patterns of equal strategy type are played too often within the category of both male and female subjects, however, male subjects deviated even more from MSE than the females did.

The variation treatment “Hot Hand” should test if there is also evidence for that belief in a lucky streak. Contrary to the already existing studies that refer to real basketball, in my experiment which was framed within a competitive computer-based basketball experiment, there is no evidence for the belief in the Hot Hand. There is also no statistical evidence that male subjects are more liable to the belief in the Hot Hand.

Bibliography

Atkinson, Richard.C. and Suppes, P. (1958). "An analysis of two- person game situations in terms of statistical learning theory", Technical Report no.8, Stanford University, California.

Budescu, David V. and Rapoport, Ammon. 1994. "Subjective Randomization in One and Two- Person Games" In Journal of Behavioral Decision Making, Vol 7, 261-278 (1994). Chister: John Wiley & Sons. Ltd.

Camerer, Colin F. 1989. "Does the basketball market believe in the 'Hot Hand'?" The American Economic Review, Vol 79, 1257-1261.

Camerer, Colin F. 2003. The roundtable series in behavioral economics, behavioral game theory. Experiments in Strategic Interaction. Princeton, NJ: Princeton University. Press.

Fischbacher, Urs. 2007. z-tree. Zurich Toolbox for Readymade Economic Experiments, Experimental Economics, Vol 10, 171-178 (2007)

Gibbons, Robert. A Primer in Game Theory. Harlow, England: Pearson Education Limited.

Gilovich, Thomas. and Vallone, Robert. and Tversky, Amos. 1985. "The Hot Hand in basketball: On the Misperception of Random Sequences" In Cognitive Psychology, Vol 17, 295-314 (1985). Academic Press Inc.

Köppen, Jörn. 2011. "Das Hot Hand Phänomen im Sport" genehmigte Dissertation, Psychologisches Institut der Deutschen Sporthochschule Köln.

Ross, Bruce.M. and Levy, N. (1958). "Patterned prediction of chance events by children and adults." In Psychological Reports 1958, Vol 4, 87-124 (monogr. supply).

Palacios-Huerta, Ignacio. 2003. "Professionals Play Minimax" In Review of Economic Studies, Vol. (70), 395-415(2003).

Appendix

Pattern type	Predicted frequency relative to MSE	Frequency in treatment	Expected frequency if idd
xx	lower	0,29	0,33
xy	lower	0,61	0,66
xyz	higher	0,26	0,22
xxx	lower	0,03	0,11
xxxx	lower	0,01	0,04

appendix 1a: randomization patterns of male subjects

Pattern type	Predicted frequency relative to MSE	Frequency in treatment	Expected frequency if idd
xx	lower	0,30	0,33
xy	lower	0,65	0,66
xyz	higher	0,23	0,22
xxx	lower	0,04	0,11
xxxx	lower	0,01	0,04

appendix1b: Randomization patterns of female subjects

```

". ttest obs == mse if pattern=="xx", unpaired unequal"

Two-sample t test with unequal variances
-----
Variable |  Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
obs |   54  .2936508  .0244498  .1796687  .2446107  .3426909
mse |   54    .33      0      0      .33      .33
-----+-----
combined |  108  .3118254  .0122938  .1277614  .2874543  .3361965
-----+-----
diff |      -.0363492  .0244498      -.0853893  .0126909
-----+-----
diff = mean(obs) - mean(mse)                t = -1.4867
Ho: diff = 0          Satterthwaite's degrees of freedom = 53

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0715    Pr(|T| > |t|) = 0.1430    Pr(T > t) = 0.9285

```

appendix 2a: T-test for 2-tuple equal strategy

```

". ttest obs == mse if pattern=="xxx", unpaired unequal"

Two-sample t test with unequal variances
-----
Variable |  Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
obs |   54  .034188  .0075174  .0552414  .01911  .049266
mse |   54    .11      0      0      .11      .11
-----+-----
combined |  108  .072094  .0052368  .0544228  .0617126  .0824754
-----+-----
diff |      -.075812  .0075174      -.09089  -.060734
-----+-----
diff = mean(obs) - mean(mse)                t = -10.0849
Ho: diff = 0          Satterthwaite's degrees of freedom = 53

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 1.0000

```

appendix 2b: T-test for 3-tuple equal strategy

```

". ttest obs == mse if pattern=="xxxx", unpaired unequal"

Two-sample t test with unequal variances
-----
Variable |  Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
  obs |   54  .0138889  .0048012  .0352811  .004259  .0235188
  mse |   54   .037    0         0         .037    .037
-----+-----
combined |  108  .0254444  .0026376  .0274106  .0202157  .0306732
-----+-----
  diff |   -0.0231111  .0048012          -0.032741  -0.0134812
-----+-----
  diff = mean(obs) - mean(mse)          t = -4.8137
Ho: diff = 0          Satterthwaite's degrees of freedom = 53

  Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 1.0000

```

appendix 2c: T-test for 4-tuple equal strategy

```

". ttest obs == mse if pattern=="xy", unpaired unequal"

Two-sample t test with unequal variances
-----
Variable |  Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
  obs |   54  .6296296  .0333031  .2447265  .5628322  .6964271
  mse |   54   .66    0         0         .66    .66
-----+-----
combined |  108  .6448148  .0166384  .1729115  .6118311  .6777985
-----+-----
  diff |   -0.0303704  .0333031          -0.0971678  .0364271
-----+-----
  diff = mean(obs) - mean(mse)          t = -0.9119
Ho: diff = 0          Satterthwaite's degrees of freedom = 53

  Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.1830    Pr(|T| > |t|) = 0.3659    Pr(T > t) = 0.8170

```

appendix 2d: T-test for 2-tuple different strategy

```

". ttest obs == mse if pattern=="xyz", unpaired unequal"

Two-sample t test with unequal variances
-----
Variable |  Obs   Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
obs |    54  .2393162  .026409   .1940655   .1863466   .2922859
mse |    54    .22     0         0         .22        .22
-----+-----
combined |   108  .2296581  .0131758  .1369265   .2035387   .2557775
-----+-----
diff |      .0193162  .026409          -.0336534  .0722859
-----+-----
diff = mean(obs) - mean(mse)                t = 0.7314
Ho: diff = 0          Satterthwaite's degrees of freedom = 53

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.7661    Pr(|T| > |t|) = 0.4677    Pr(T > t) = 0.2339

```

appendix 2e: T-test for 3-tuple different strategies

	(1)
	choice
choice	
perf3	-0.160 (0.201)
perftot	0.544** (0.179)
hothandfem	0.0247 (0.067)
ale	
_cons	-5.501*** (1.144)
<i>N</i>	162
pseudo <i>R</i> ²	0.146
<i>AIC</i>	184.1
<i>BIC</i>	196.5

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

appendix 3: Logit regression Hot Hand

