BIASES AND HEURISTICS IN PORTFOLIO MANAGEMENT

DETERMINANTS FOR NON-OPTIMAL PORTFOLIO DIVERSIFICATION

Dissertation

in order to acquire the doctoral degree
from the Faculty of Economic Sciences
at University of Göttingen

submitted by

Ibrahim Filiz
born in Braunschweig

Göttingen, 2018

Thesis Committee

First supervisor:	Prof. Dr. Kilian Bizer
Second supervisor:	Prof. Dr. Markus Spiwoks
Another member of the examination board:	Prof. Dr. Waldemar Toporowski

CONTENTS

CHAPTER 1

INTRODUCTION AND SUMMARY .... - 1 -

CHAPTER 2

PORTFOLIO DIVERSIFICATION: THE INFLUENCE OF HERDING, STATUS-QUO BIAS, AND THE GAMBLER'S FALLACY .... - 13 -

(with Thomas Nahmer, Markus Spiwoks and Kilian Bizer)

Published in:

Financial Markets and Portfolio Management, 32(2), 167-205.

Wolfsburg Working Papers, 17-01, February 2017.

Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 17-2, March 2017.

CHAPTER 3

EMOTIONS AND EXPOSURE TO RISK: THE INFLUENCE OF POSITIVE AND NEGATIVE EMOTIONS ON PORTFOLIO DECISIONS .... - 59 -

Submitted to:

Journal of Economic Behavior and Organization

Published in:

Wolfsburg Working Papers, 18-03, July 2018.

Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-3, June 2018.

CHAPTER 4

MEASUREMENT OF RISK PREFERENCE .... - 97 -

(with Thomas Nahmer, Markus Spiwoks and Zulia Gubaydullina)

Submitted to:

Journal of Risk and Uncertainty

Published in:

Wolfsburg Working Papers, 18-01, June 2018.

Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-1, June 2018.

CHAPTER 5

THE ACCURACY OF INTEREST RATE FORECASTS IN THE ASIA-PACIFIC REGION: OPPORTUNITIES FOR PORTFOLIO MANAGEMENT ....- 127 -

(with Thomas Nahmer, Markus Spiwoks and Kilian Bizer)

Submitted to:
Applied Economics

Published in:
Wolfsburg Working Papers, 18-02, June 2018.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-2, June 2018.

CHAPTER 6

OVERCONFIDENCE: THE INFLUENCE OF POSITIVE AND NEGATIVE AFFECT ....- 168 -

Submitted to:
Journal of Neuroscience, Psychology, and Economics

Published in:
Wolfsburg Working Papers, 17-02, March 2017.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 17-1, January 2017.

Chapter 1

INTRODUCTION AND SUMMARY

"Uncertainty cannot be dismissed so easily in the analysis of optimizing investor behavior. An investor who knew future returns with certainty would invest in only one security, namely the one with the highest future return. If several securities had the same, highest, future return then the investor would be indifferent between any of these, or any combination of these. In no case would the investor actually prefer a diversified portfolio." (Markowitz, 1991)

The question of whether the future value of capital market instruments can be forecasted – in order to be able to solely invest in securities with the best possible performance – has been the subject of a controversial debate for decades (Lakonishok, 1980; Dimson and Marsh, 1984; Fraser and MacDonald, 1993; Krag, 1995; Henze, 2004; Spiwoks, 2004; Benke, 2006; Spiwoks, 2009; Söderlind, 2010). Many of the investigations on forecasting competence carried out until now have led to sobering results.

Markowitz' portfolio theory (1952) is therefore still considered to be definitive for investment decisions. In his groundbreaking work *Portfolio Selection*, Markowitz showed that for risk-averse investors it is usually meaningful to divide their assets up across a range of securities. Nevertheless, in practice investors frequently hold highly under-diversified portfolios (see, for example, Anderson 2013; Hibbert, Lawrence and Prakash 2012; Goetzmann and Kumar, 2008; Meulbroek, 2005; Polkovnichenko, 2005; Huberman and Sengmueller, 2004; Agnew, Balduzzi and Sundén, 2003; Guiso, Haliassos and Japelli, 2002; Benartzi, 2001; Benartzi and Thaler, 2001; Barber and Odean, 2000; Bode, van Echelpoel and Sievi, 1994; Blume and Friend, 1975; Lease, Lewellen and Schlarbaum, 1974).

There are many possible reasons for securities portfolios not being optimally diversified. Experimental economic research in particular has shown that heuristics and cognitive distortions can influence the behavior of investors significantly. This can also make meaningful portfolio diversification significantly more difficult: The correlations between investment alternatives are frequently neglected by investors. Considerable empirical evidence already exists for this phenomenon of correlation neglect (see, for example, Gubaydullina and Spiwoks, 2015; Eyster and Weizsäcker, 2010; Hedesstrom et al., 2006; Benartzi and Thaler, 2001). Many investors distribute their assets evenly across all of the investment alternatives available. This phenomenon, which is known as the 1/N heuristic, is a special form of correlation neglect. In the meantime, there are also significant empirical findings on this issue (see, for example, Fernandes 2013; Morrin et al., 2012; Baltussen and Post, 2011; Hedesstrom et al., 2006, and Benartzi and Thaler, 2001). Many investors allow themselves to be misled by irrelevant information, or attach too much importance to certain items of information (cf. Gubaydullina and Spiwoks, 2015; Kallir and Sonsino, 2009; Goetzmann and Kumar, 2008). Investment decisions are also frequently skewed by an inappropriately strong focus on domestic financial instruments. This phenomenon is known as home bias, and has been proven empirically (cf. Weber et al., 2005; Poterba, 2003; Mitchell and Utkus, 2002). Many investors also tend to see patterns, where in reality there are none. This often leads to random processes being dealt with inappropriately. In this way, the so-called gambler's fallacy can impede optimal decisions on diversification (see, for example, Stöckl et al., 2015; Huber et al., 2010). Emotions can also exert an enormous influence on investment decisions. Grable and Roszkowski (2008), for example, showed in an experimental study that subjects who possessed predominately positive emotions were willing to take greater financial risks. Kuhnen and Knutson (2011) carried out experiments to establish how different moods affected investment decisions. This revealed that subjects with predominantly negative emotions tend to choose low-risk investments, while subjects with predominantly positive emotions tended to favour riskier investments. Subjects whose emotions are positive are more optimistic in relation to their investment decisions. Kaplanski et al. (2015) showed that the mood of investors had an influence on their expectations in terms of returns, and on their perception of risk. The happier the subjects were, the greater were their expectations of their returns, and the lower the presumed risk. Experiments carried out by Lee and Andrade (2014) showed that negative affects promote risk aversion in investment decisions.

Additional research efforts in this field can be found in this dissertation. In particular, the influence of heuristics and cognitive distortions such as the status-quo bias, herd behavior, gambler's fallacy, overconfidence bias, and last but not least the influence of emotions on diversification behavior, are examined here. In addition, a new procedure for determining risk preference is presented, and finally the forecasting skills of analysts in the Asia-Pacific region are studied.

First contribution – Portfolio Diversification: The Influence of Herding, Status-Quo Bias and Gambler's Fallacy

In the last two decades, experimental economic research has identified some heuristics and cognitive distortions (such as the 1/N heuristic, home bias, mental accounting or the illusion of control) which can contribute towards sub-optimal portfolio diversification.[1] However, there is still an enormous need for research.

In the meantime there is a great deal of empirical evidence which shows that the phenomena of herd behavior, status-quo bias and gambler's fallacy can have a significant effect in many financial decision-making situations. In the case of herding behavior, subjects orientate their actions towards those of others, and in this way a herd is formed. In the case of status-quo bias, subjects find it difficult to actively make decisions. Instead, they prefer to leave things as they are. In the case of gambler's fallacy, subjects try to use a sequence of random events to make deductions about future random events, although this is in fact impossible in the case of independent random events.

However, these phenomena have not yet been investigated in connection with diversification decisions. The second chapter (Portfolio Diversification: The Influence of Herding, Status-Quo Bias and Gambler's Fallacy) therefore considers whether herd behavior (orientation towards the majority of one’s fellow players or towards the most successful player (guru)), status-quo bias or gambler’s fallacy provide suitable approaches for explaining why many subjects have sub-optimally diversified portfolios.

[1] For a comprehensive overview see pages 2-3.

The experiment is loosely based on the approach used by Gubaydullina and Spiwoks (2015). The subjects have a choice between two entirely uncorrelated investment alternatives which are identical with regard to their expected returns and the risk involved. The two investment alternatives can both only lead to two events, which each have a probability of occurrence of 50% and are based on a random process. In this situation, the efficient frontier is reduced to a single point (half each of the two investment alternatives), so that the exact extent of the risk aversion of the subjects has no influence on the optimal portfolio decision. A rational and risk-averse subject would therefore always have to choose a combination of half each of the two investment alternatives.

In order to be able to examine possible herd behavior among the subjects, in Treatment 1 the portfolios of all of the players and their investment performance so far are disclosed after each round and presented in a list. In this way, before the subjects might possibly restructure their portfolio, they obtain an overview of the composition of the portfolios of the majority of the subjects in the previous round, and can see the portfolio content selected by the most successful player up to now in the last round. In this way the subjects can align themselves with the opinion of the majority or that of the most successful player up to that point. In Treatment 2 the subjects receive no information about the behavior or the investment performance of the other participants. They are only informed about their own performance, so there is no opportunity for them to align themselves with the majority opinion or with a guru. In order to shed more light on the aspect of status-quo bias, in Treatment 1 the subjects received various portfolio compositions as a starter. In Treatment 3, on the other hand, all of the subjects received the optimal portfolio compositions as a starter. Even before the first round of the game, the subjects can recompose their portfolios free of charge. In order to detect a possible tendency towards gambler’s fallacy, the subjects were asked in all three treatments to explain the basis of their portfolio decisions.

The strategy of always choosing half each of the two investment alternatives only represents the rational strategy in the case of risk-averse investors. Possible divergences from the rational strategy can therefore only be identified if the participants in the experiment are all risk-averse. This is why the Holt and Laury (2002) test was carried out with every single subject, and risk-neutral and risk-loving subjects were thus filtered out of the field of participants.

It could be seen that the majority of portfolio decisions are not based on rational considerations. In addition, we established that herd behavior does not make a significant contribution towards sub-optimal portfolio decisions. The subjects did not follow either the majority or the most successful investor (guru). The behavior of the subjects remained fragmented even in the final round. With regard to status-quo bias, the subjects did not retain their initial portfolios, and they very quickly disposed of their starter packages. Overall, in Treatment 1 the optimal portfolio was not retained any longer than in Treatment 3. The strong divergences from a rational strategy cannot therefore be explained by status-quo bias either.

Ultimately it became clear that gambler's fallacy played an essential role in the explanation of the irrational behavior of the subjects. This phenomenon can be seen clearly if one analyzes the portfolio decisions of the subjects and their justifications. Many subjects seemed to find patterns in a history of random events. However, gambler's fallacy is indisputably also jointly responsible for many sub-optimal portfolio decisions.

Second contribution – Emotions and Exposure to Risk: The Influence of Positive and Negative Emotions on Portfolio Decisions

For a while now it has been clear that emotions also have an effect on decision-making in economic contexts. In recent years, the influence of the weather on market returns at stock exchanges throughout the world has also been thoroughly investigated. While doing so, attempts were also made to create a connection between the weather on the one hand and the mood of capital market protagonists on the other. In some empirical studies, genuine correlations were established. Experimental economic research is therefore increasingly interested in the question of which influence positive and negative emotions have on investment decisions. In the meantime there are a range of findings showing that the mood of investors can influence their investment decisions. However, as yet there have been no studies on whether the mood of investors also has an effect on their diversification behavior and thus on the exposure to risk in their portfolios. This research question is dealt with in the third chapter – Emotions and Exposure to Risk: The Influence of Positive and Negative Emotions on Portfolio Decisions.

The design of the experiment is loosely based on the work of Gubaydullina and Spiwoks (2015). Each subject has to make four investment decisions. In each task the subjects can choose between two different investment alternatives. The subjects have to compile a portfolio which contains four shares. They profit from the dividend payments, which are based on a random process. In each task, four different portfolio compositions are possible which are identical with regard to the expected returns, but whose variance differs. The price movements of the two investment alternatives are ignored in order to create a decision-making situation which is as clear as possible. Mood is influenced by positive film excerpts (in the positive treatment), negative ones (in the negative treatment) and neutral film excerpts (in the neutral treatment). After each film excerpt a manipulation check takes place in order to test whether the intended mood has actually been created among the subjects. In all three treatments it is only the film excerpts which differ. The rest of the experiment is the same in all three treatments. The Holt and Laury (2002) test was used to filter risk-neutral and risk-loving subjects out of the field of participants, because minimum variance portfolios represent the rational strategy only in the case of risk-averse subjects.

It was shown that the deployment of the film excerpts led to the desired results. In each of the three treatments, the desired mood was predominant. To this extent, the approach was very well suited to answering the questions posed in an appropriate way. Overall, the results of the experiment reveal that only a small part of subjects take rational decisions – in other words they always choose the optimal portfolio. This is also reflected by an unnecessarily high-risk exposure. The subjects frequently fail to choose the minimum variance portfolio, although all five portfolio alternatives always exhibit the same expectation value for the payment. A closer look at the results of the experiment shows that the mood of the subjects has an influence on their diversification behavior. This becomes clear when their risk exposure is considered. The average risk exposure in the neutral treatment is significantly lower than in the positive treatment and the negative treatment. In addition, a neutral mood leads to the subjects obtaining higher risk-adjusted payoffs. The average risk-adjusted payment was significantly higher in the neutral treatment than in the positive and the negative treatments.

Third contribution – Measurement of Risk Preference

In chapters two and three, new behavioral anomalies were identified which can be viewed as a reason for sub-optimal portfolio diversification. According to Markowitz's (1952) portfolio theory, a diversified portfolio is only meaningful for risk-averse subjects. In chapters two and three it was therefore necessary to precisely define the risk preferences of the subjects in order to filter risk-neutral and risk-loving subjects out of the field of participants, because decisions which can be absolutely meaningful for a risk-averse subject are frequently completely inconceivable for a risk-loving subject and vice-versa. The well-established Holt and Laury (2002) test was used to determine the risk preferences of the subjects. It can be seen, however, that Holt and Laury's (2002) procedure to determine risk preference clearly also has its weaknesses. The decision-making situation is relatively complex, so that some subjects do not decide according to their actual preference, and make rather spontaneous and ill-considered decisions (for similar observations see Jacobson and Petrie 2009, and Charmes and Viceisza 2011). In addition, clear results are not obtained in every case.[2] The fourth chapter, Measurement of Risk Preference, deals precisely with this issue, and a new procedure for the determination of risk preference is proposed.

In order to exactly define the risk preference of the subjects, the following three features are taken into account in the new procedure. First of all, the new approach is simple and clear. The subjects know precisely which consequences their decisions will have. They have no reason to make spontaneous and ill-considered decisions. Instead they can make conscious decisions which correctly reflect their preferences. Secondly, it makes a clear differentiation between risk-averse, risk-neutral and risk-loving subjects possible, which is by no means the case with the conventional procedures, and thirdly, it takes the influence of loss aversion on risk preference into account.

The new procedure to determine risk preference is a decision between two lotteries. The subjects can choose between taking a card from pile A or one from pile B. Both piles consist of four playing cards each. The subjects are informed that the profit expectation in both piles is identical. In addition, the subjects are made aware of the fact that pile A leads to results which fluctuate slightly around the expected value (low-risk), while pile B leads to results

[2] There are various procedures to determine risk preference which all exhibit certain weaknesses. For an overview see pages 100-107.

which fluctuate considerably around the expected value (high-risk). The subjects have three alternatives (pile A, pile B, or indifference as to whether the card is from A or B). This permits unambiguous conclusions about the three categories of risk preference (risk-averse, risk-neutral and risk-loving). In order to take the third property into account and to investigate the influence of loss aversion on risk preference in more detail, the new approach to establish risk preference is carried out in three variations. In Treatment 1 there is no possibility of loss, in Treatment 2 a small loss is possible, and in Treatment 3 there is the possibility of a large loss.

The results were clear and are largely in line with our expectations. In Treatment 1 (no possibility of loss), less than half of the subjects chose the low-risk variation, and more than half of the subjects chose the high-risk variation. Only very few subjects were indifferent as to whether they chose pile A or pile B. In Treatment 2 (low possibility of loss), just under half of the subjects chose the low-risk variation, and less than half of the subjects chose the high-risk variation. Here again, only a few subjects were indifferent to whether they chose pile A or pile B. In Treatment 3 (significant possibility of loss), however, loss aversion had a marked influence on risk preferences. Here, more than two-thirds of the subjects chose the low-risk variation, and much less than a third of the subjects chose the high-risk variation. Here again, only a few subjects were indifferent to whether they chose pile A or pile B.

The results show in particular how important it is to provide an appropriate possibility of loss when determining risk preference. Only in this way can the risk preferences of the subjects be realistically determined and a clear differentiation made between risk-averse, risk-neutral and risk-loving subjects.

Fourth contribution – The Accuracy of Interest Rate Forecasts in the Asia-Pacific Region: Opportunities for Portfolio Management

Future interest rate trends are of key significance in portfolio management, because bond and share prices are significantly influenced by interest rates. If the general level of interest rates in an economy rises, the prices of most bonds will fall and vice-versa. If one wishes to take a critical look at the current market price of a share, one should establish its future fair value. And in order to establish the future fair value of a share, one has to forecast the predominant future interest rate level, because this simultaneously represents the future discount rate in the determination of the fair value. It is therefore not surprising that the reliability of interest rate forecasts is of great interest to academia and the world of business.

In the last four decades, the quality of interest rate forecasts in the G7 states has already been abundantly researched. However, the results of these studies have largely been very sobering. Only very few studies considered the interest rate forecasts they analyzed to be largely reliable. There have only been a few studies which have dealt with interest rate forecasts for the Asia-Pacific region, and they do not provide a comprehensive impression of the reliability of the forecasts, because either the assessment criteria were insufficient or the period of investigation was too short. In the fifth chapter, The Accuracy of Interest Rate

Forecasts in the Asia-Pacific Region: Opportunities for Portfolio Management, we therefore take an in-depth look at the forecasting competence of analysts in the Asia-Pacific region.

The study analyzes interest rate forecasts from eleven countries in the Asia-Pacific region. The forecast data used comes from the journal Asia Pacific Consensus Forecasts, which is published on a monthly basis. We not only analyzed the consensus forecasts, but also the individual forecasts of the respective institutions. Overall a total of 85,264 forecasts were examined which were published in the period from 1990-2016. The large number of benchmarks deployed in this empirical analysis enables a comprehensive assessment of the quality of the forecasts: (1) Comparison to a naïve forecast with the aid of the Diebold-Mariano test, (2) examination of the forecast direction with the help of the sign accuracy test, (3) examination for systematic forecast errors with the aid of the unbiasedness test, and (4) test for the presence of possible topically-oriented trend adjustments with aid of the TOTA coefficient.

The results are very sobering in part. With a few exceptions, all of the forecast time series are characterized by the phenomenon of topically-oriented trend adjustment. This means that the overwhelming majority of all forecast time series reflect the present rather than the future. This is in line with previous studies (see, for example, Spiwoks et al., 2015; Spiwoks et al., 2010; Spiwoks et al., 2008). These results for topically-oriented trend adjustment are also reflected in the unbiasedness test. Almost all of the forecast time series proved to be biased. A comparison with the naïve forecast shows that only a small part of the forecast time series predicts the future interest rate trend in a significantly more precise way.

However, some of the results of the study are also surprisingly positive. The sign accuracy test reveals that almost half of the forecast time series foresee the future trend (rising or falling interest rates) significantly better than a random walk forecast. This is a notable success which is entirely in contrast to the forecast time series made in other parts of the world (see, for example, Spiwoks et al., 2008). Overall it can be stated that – at least in some of the countries – forecasts of future interest rate trends in the Asia-Pacific region are significantly more successful than those made in other parts of the world. This has consequences for portfolio management: It is conceivable that active portfolio management strategies based on these interest rate forecasts can lead to systematic excess returns in bond investments.

Fifth contribution – Overconfidence: The Influence of Positive and Negative Affect

In economics, overconfidence bias is viewed as one of the reasons for inefficient capital markets. Evidence for this phenomenon has already been found among various groups of individuals such as investors, managers, bankers and other important economic players (see, for example, Ifcher and Zarghamee, 2014; Menkhoff et al., 2013; Barber and Odean, 2001). And the influence of this behavioral anomaly on their portfolio decisions can be tremendous (cf. Deaves et al., 2009; Barber and Odean, 2001; Odean, 1999). However, until now only a few factors are known which might determine the extent of overconfidence. In addition, there are hardly any studies on whether positive or negative emotions have an influence on self-confidence. Whether emotions can impair possible learning effects in relation to self-assessment has not been investigated at all up to now, but the sixth chapter, Overconfidence: The Influence of Positive and Negative Affect, deals with this issue.

The design of the experiment is loosely based on the work of Ifcher and Zarghamee (2014). In a real effort task experiment, the subjects are given tasks to do in five rounds of a game. After every round the subjects assess their own performance. Absolute overconfidence is recorded, and relative overconfidence is also assessed. Subsequently the subjects receive feedback on their actual performance. Every subject finds out how many tasks he or she has solved correctly (absolute overconfidence) and how they have fared in relation to the other subjects (relative overconfidence). Over the course of the five rounds, the subjects thus have the opportunity to learn from their experiences in the preceding rounds and to gradually assess themselves more and more realistically. In other words, learning effects are possible. Mood is influenced by positive film excerpts (in the positive treatment), negative ones (in the negative treatment) and neutral film excerpts (in the neutral treatment). After each film excerpt a manipulation check takes place in order to test whether the intended mood has actually been created among the subjects. In all three treatments it is only the film excerpts which differ. The rest of the experiment is exactly the same in all three treatments.

It can be clearly stated that the deployment of the film excerpts led to the desired results. To this extent, the approach was very well suited to answering the questions posed in an appropriate way. Overall, the results of the study also show that there are no significant differences between the treatments with regard to either absolute overconfidence or relative overconfidence.

In respect of the occurrence of learning effects, however, there is a difference between the moods. When considering absolute overconfidence, learning effects appear in the neutral mood. The participants assess their performance significantly more precisely in the last two rounds of the game than in the first three rounds of the game. By contrast, these significant learning effects do not arise in either a positive or a negative mood. When relative overconfidence is considered, there are no learning effects. The subjects cannot gauge their relative performance over the course of the game any better in either the neutral, positive or negative treatments.

References

Agnew, J. / Balduzzi, P. / Sundén, A. (2003): Portfolio Choice and Trading in a Large 401(k) Plan, American Economic Review, 93(1), 193-215.

Anderson, A. (2013): Trading and Under-Diversification, Review of Finance, 17(5), 1-43.

Andres, P. / Spiwoks, M. (1999): Forecast Quality Matrix: A Methodological Survey of Judging Forecast Quality of Capital Market Forecasts, Journal of Economics and Statistics, 219, 513-542.

Baltussen, G. / Post, G. T. (2011): Irrational Diversification: An Examination of Individual Portfolio Choice, Journal of Financial Quantitative Analysis, 46(5), 1463-1491.

Barber, B. / Odean, T. (2001): Boys will be Boys: Gender, Overconfidence, and common Stock Investments, Quarterly Journal of Economics, 116, 261-292.

Benartzi, S. (2001): Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock, Journal of Finance, 56(5), 1747-1764.

Benartzi, S. / Thaler, R. H. (2001): Naïve Diversification Strategies in Defined Contribution Saving Plans, American Economic Review, 91(1), 79-98.

Benke, H. (2006): Was leisten Kapitalmarktprognosen? Die Sicht eines Stiftungsmanagers, Zeitschrift für das gesamte Kreditwesen, 17, 40-49.

Blume, M. E. / Friend, I. (1975): The Asset Structure of Individual Portfolios and Some Implications for Utility Functions, Journal of Finance, 30(2), 585-603.

Bode, M. / van Echelpoel, A. / Sievi, C. R. (1994): Multinationale Diversifikation: Viel zitiert, kaum befolgt, Die Bank, 94(4), 202-206.

Charness, G. / Viceisza, A. (2011): Comprehension and Risk Elicitation in the field: Evidence From Rural Senegal, IFPRI Discussion Papers, 1135.

Deaves, R. / Lüders, E. / Luo, G. Y. (2009): An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity, Review of Finance, 13, 555-575.

Dimson, E. / Marsh, P. (1984): An Analysis of Brokers' and Analysts' Unpublished Forecasts of UK Stock Returns, Journal of Finance, 39, 1257-1292.

Eyster, E. / Weizsäcker, G. (2011): Correlation Neglect in Financial Decision Making, DIW Discussion Paper, 1104.

Fernandes, D. (2013): The 1/N Rule Revisited: Heterogeneity in the Naïve Diversification Bias, International Journal of Marketing Research, 30(3), 310-313.

Fraser, P. / MacDonald, R. (1993): The Efficiency of CAC Stock Price Forecasts: a Survey Based Perspective, Revue Économique, 44(5), 991-1000.

Goetzmann, W. N. / Kumar, A. (2008): Equity Portfolio Diversification, Review of Finance, 12(3), 433-463.

Grablea, J. E. / Roszkowski, M. J. (2008): The Influence of Mood on the Willingness to Take Financial Risks, Journal of Risk Research, 11(7), 905-923.

Gubaydullina, Z. / Spiwoks, M. (2015): Correlation Neglect, Naïve Diversification, and Irrelevant Information as Stumbling Blocks for Optimal Diversification, Journal of Finance and Investment Analysis, 4(2), 1-19.

Guiso, L. / Haliassos, M. / Japelli, T. (2002): Household Portfolios, MIT Press, Cambridge.

Hedesstrom, T. M. / Svedsater, H. / Garling, T. (2006): Covariation Neglect among Novice Investors, Journal of Experimental Psychology-App., 12(3), 155-165.

Henze, J. (2004): Was leisten Finanzanalysten? Eine empirische Analyse des deutschen Aktienmarktes, Reihe Finanzierung, Kapitalmarkt und Banken, 38, Lohmar.

Hibbert, A. M. / Lawrence, E. R. / Prakash, A. J. (2012): Can Diversification be Learned?, Journal of Behavioral Finance, 13(1), 38-50.

Holt, C. A. / Laury, S. K. (2002): Risk Aversion and Incentive Effects, American Economic Review, 92(5), 1644-1655.

Huber, J. / Kirchler, M. / Stöckl, T. (2010): The Hot Hand Belief and the Gambler's Fallacy in Investment Decisions under Risk, Theory and Decision, 68(4), 445-462.

Huberman, G. / Sengmueller, P. (2004): Performance and Employer Stock in 401(k) Plans, Review of Finance, 8(3), 403-443.

Ifcher, J. / Zarghamee, H. (2014): Affect and Overconfidence: A Laboratory Investigation, Journal of Neuroscience, Psychology, and Economics, 7(3), 125-150.

Jacobsen, S. / Petrie, R. (2009): Learning from Mistakes: What Do Inconsistent Choices over Risk Tell Us?, Journal of Risk and Uncertainty, 38(2), 143-158.

Kallir, I. / Sonsino, D. (2009): The Neglect of Correlation in Allocation Decisions, Southern Economic Journal, 75(4), 1045-1066.

Kaplanski, G. / Levy, H. / Veld, C. / Veld-Merkoulova, Y. (2015): Do Happy People Make Optimistic Investors?, Journal of Financial and Quantitative Analysis, 50(1-2), 145-168.

Krag, J. (1995): Sind Aktienkurse prognostizierbar?, Jenaer Vorträge, 8, Baden-Baden.

Kuhnen, C. M. / Knutson, B. (2011): The Influence of Affect on Beliefs, Preferences, and Financial Decisions, Journal of Financial and Quantitative Analysis, 46(3), 605-626.

Lakonishok, J. (1980): Stock Market Return Expectations: Some General Properties, Journal of Finance, 35, 921-931.

Lease, R. C. / Lewellen, W. G. / Schlarbaum, G. G. (1974): The Individual Investor: Attributes and Attitudes, Journal of Finance, 29(2), 413-433.

Lee, C. J. / Andrade, E. B. (2014): Fear, Excitement, and Financial Risk-Taking, Cognition and Emotion, 29(1), 178-187.

Markowitz, H. M. (1952): Portfolio Selection, Journal of Finance, 7(1), 77-91.

Markowitz, H. M. (1991): Foundations of Portfolio Theory, Journal of Finance, 46(2), 469-477.

Menkhoff, L. / Schmeling, M. / Schmidt, U. (2013): Overconfidence, Experience, and Professionalism: An Experimental Study, Journal of Economic Behavior and Organization, 86, 92-101.

Meulbroek, L. (2005): Company Stock in Pension Plans: How Costly Is It? Journal of Law and Economics, 48(2), 443-474.

Mitchell, O. S. / Utkus, S. P. (2002): The Role of Company Share in Defined Contribution Plans, National Bureau of Economic Research, Working Paper, 9250.

Morrin, M. / Inman, J. J. / Broniarczyk, S. M. / Nenkov, G. Y. / Reuter, J. (2012): Investing for Retirement: The Moderating Effect of Fund Assortment Size on the 1/N Heuristic, Journal of Marketing Research, 49(4), 537-550.

Odean, T. (1999): Do Investors Trade Too Much?, American Economic Review, 89(5), 1279-1298.

Polkovnichenko, V. (2005): Household Portfolio Diversification: A Case for Rank-Dependent Preferences, Review of Financial Studies, 18(4), 1467-1502.

Poterba, J. M. (2003): Employer Share and 401(k) Plans, American Economic Review, 93(2), 398-404.

Söderlind, P. (2010): Predicting Stock Price Movements: Regressions versus Economists, Applied Economics Letters, 17(9), 869-874.

Spiwoks, M. (2004): Die Verwendbarkeit der ZEW-Aktienindex-Prognosen für aktive Portfoliomanagement-Strategien, Jahrbücher für Nationalökonomie und Statistik, 224, 557-578.

Spiwoks, M. (2009): Deceptive Unity: Surprising Characteristics of Pre-Tax Corporate Profit Forecasts, International Research Journal of Finance and Economics, 26, 170-177.

Spiwoks, M. / Bedke, N. / Hein, O. (2008): Forecasting the Past: The Case of US Interest Rate, Financial Markets and Portfolio Management, 22(4), 357-379.

Spiwoks, M. / Bedke, N. / Hein, O. (2010): Topically Orientated Trend Adjustment and Autocorrelation of the Residuals – An Empirical Investigation of the Forecasting Behavior of Bond Market Analysts in Germany, Journal of Money, Investment and Banking, 14, 16-35.

Spiwoks, M. / Gubaydullina, Z. / Hein, O. (2015): Trapped in the Here and Now – New Insights into Financial Market Analyst Behavior, Journal of Applied Finance and Banking, 5(1), 35-56.

Stöckl, T. / Huber, J. / Kirchler, M. / Lindner, F. (2015): Hot Hand and Gambler's Fallacy in Teams: Evidence from Investment Experiments, Journal of Economic Behavior and Organization, 117, 327-339.

Weber, E. U. / Siebenmorgen, N. / Weber, M. (2005): Communication Asset Risk: How Name Recognition and the Format of Historic Volatility Information Affect Risk Perception and Investment Decisions, Risk Analysis, 25(3), 597-609.

Chapter 2

PORTFOLIO DIVERSIFICATION: THE INFLUENCE OF HERDING, STATUS-QUO BIAS, AND THE GAMBLER'S FALLACY

with Thomas Nahmer, Markus Spiwoks and Kilian Bizer
contribution Ibrahim Filiz: 35%

Published in:
Financial Markets and Portfolio Management, 32(2), 167-205.
Wolfsburg Working Papers, 17-01, February 2017.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 17-2, March 2017.

Abstract

This experimental study examines the influence of herding (following the majority of fellow gamblers or the most successful gambler (guru)), status-quo bias, and the gambler’s fallacy on diversification behavior. We find that neither herding nor status-quo bias contributes significantly to non-optimal portfolio choices. The gambler’s fallacy, however, plays an important role in these decisions. Many subjects appear to find patterns in a history of random events and then use these “patterns” to infer the sequence of future events. The gambler’s fallacy is significantly responsible for the fact that the optimal structure of a portfolio is considered in only 37.7% of all choices made by an investor.

Keywords

Behavioral finance, experiments, portfolio choice, non-optimal diversification, herding, guru, status-quo bias, gambler’s fallacy

JEL Classification

G02, G11, D81, D84

1 Introduction

Markowitz (1952) shows that it is useful for risk-averse investors to split capital among different investment instruments. Practice shows, however, that investors often have strongly underdiversified portfolios.[1]

Experimental economic research increasingly addresses the question of why investors seem to find it so difficult to make useful portfolio diversifications. Gubaydullina and Spiwoks (2015) show that many investors have difficulty dealing with the correlations of income return developments.[2] The meaning of the correlations is systematically misjudged. Take the example of 1/n heuristics, where investors distribute their capital equally among all investment alternatives seemingly without noticing or caring how strongly the income returns of these instruments are correlated. Morrin et al. (2012) provide evidence that many subjects tend toward 1/n heuristics (for similar findings, see Fernandes, 2013; Baltussen and Post, 2011). Rieger (2012) reveals that investors systematically miscalculate the probabilities of occurrence. Fellner, Güth, and Maciejovsky (2004) conclude that investors often suffer from an illusion of expertise, and hence overestimate the advantageousness of their own choice of investment. Choi, Laibson, and Madrian (2009) note that diversification decisions are distorted by the phenomenon of mental accounting. Weber, Siebenmorgen, and Weber (2005) detect that investors are subject to a home bias when choosing investment instruments for a portfolio.

We conducted expert discussions with high-ranking bank managers,[3] which revealed other possible reasons for suboptimal diversification decisions. It seems possible that many subjects are distracted from optimal diversification by observing the investment choices of other investors (herding). Being influenced by the successful investment decisions of prominent investors can play a significant role in this process (guru effect). Furthermore, optimal diversification can be hindered by subjects holding on to existing portfolios (status-quo bias). Subjects can also be distracted from meaningful portfolio choices by exclusively following putative patterns of random events (the gambler's fallacy).

The present study therefore addresses the question of whether or not (1) herding, (2) status-quo bias, and (3) the gambler's fallacy do, indeed, sidetrack subjects from making optimum diversification decisions. The literature provides multiple indications of the significant influence that these phenomena can exert on economic decisions.

Looking at the herding literature first, the observation that subjects take their bearings from one another and thereby act as a herd traces as far back as Mackay (1841). Keynes (1936) points out the herding behavior of financial market actors and presents two possible expla-

---

[1] See, e.g., Dimmock *et al.* (2016), Anderson (2013), Hibbert, Lawrence, and Prakash (2012), Goetzmann and Kumar (2008), Meulbroek (2005), Polkovnichenko (2005), Huberman and Sengmueller (2004), Agnew, Balduzzi, and Sundén (2003), Guiso, Haliassos, and Japelli (2002), Benartzi (2001), Benartzi and Thaler (2001), Barber and Odean (2000), Bode, van Echelpoel, and Sievi (1994), Blume and Friend (1975), and Lease, Lewellen, and Schlarbaum (1974).

[2] For similar results, see also Eyster and Weizsäcker (2011), Kallir and Sonsino (2009), and Hedesstrom, Svedsater, and Garling (2006).

[3] We thank Mr. Lothar Henning, Bethmann Bank Frankfurt, and Mr. Frank Weber, Sparkasse Lippstadt, for extensive talks concerning investment behavior of bank customers.

nations for it (reputational herding and investigative herding). Scharfstein and Stein (1990) continue along these lines, adding fuel to a fierce debate that has been raging for the past 25 years. Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) show that herding can even occur when subjects behave rationally and make reasonable decisions (informational cascades). Devenow and Welch (1996) were the first to clearly differentiate between rational herding (reputational herding, investigative herding, and informational cascades) and irrational herding. There are numerous empirical findings that confirm herding behavior among actors on the financial market.[4] Therefore, it seems reasonable to consider herding as a possible origin of non-optimal portfolio diversification. To date, there have not been any experimental studies examining the potential influence of herding on diversification decisions.

A special case of herding is the guru effect. Gurus are highly ranked religious authorities in Hinduism and Buddhism. In Western cultures, the term "guru" also refers to leaders whose followers trust them blindly and uncritically, and the term "guru effect" has been used to describe the situation where private investors rigorously copy the decisions of prominent and very successful investors. The gurus' behavior is closely observed by many actors on the capital market, which is why it can lead to herding. In the research on this phenomenon, capital market simulation with interacting artificial agents (agent-based computational economics) has established itself as a reliable research method. The method has revealed that the network structure of communication among the agents significantly influences events on the capital market. Gurus are so-called super nodes that have numerous direct communication links with other capital market actors and, for this reason, can trigger herding.[5] Furthermore, the guru effect may contribute to distracting investors from optimal diversification decisions. There is as yet no research on how the influence of an investment guru impacts investor's portfolio decisions.

Many people find it difficult to make decisions and thus tend to avoid them, simply leaving things as they are, which is known as the status-quo bias (cf. Samuelson and Zeckhauser, 1988). The psychological processes of this behavior are explained in detail by Anderson (2003). Especially in situations when investors accede to an existing security portfolio (e.g., by inheritance), they often tend to postpone or even completely fail to adjust the portfolio structure. Even if different performances of the stocks in the portfolio lead to an unintended imbalance, many investors, out of dread of adjusting the portfolio, fail to take appropriate action. Aside from dread, this lack of action is often grounded in reluctance to take responsibility for the portfolio's future profit, or lack thereof. Many investors are afraid of regretting

---

[4] Huang, Wu, and Lin (2016), Choi (2016), Galariotis, Rong, and Spyrou (2015), Chang (2013), Kremer and Nautz (2013), Lin, Tsai, and Lung (2013), Belhoula and Naoui (2011), Boyson (2010), Kim and Jegadeesh (2010), Chiang and Zheng (2010), Spiwoks, Bizer, and Hein (2008), Chen, Wang, and Lin (2008), Walter and Weber (2006), Voronkova and Bohl (2005), Spiwoks (2004), Sias (2004), Ennis and Sebastian (2003), Chang, Cheng, and Khorana (2000), Nofsinger and Sias (1999), Wermers (1999), Choe, Kho, and Stulz (1999), Christie and Huang (1995), Lakonishok, Shleifer, and Vishny (1992), Klemkovsky (1977), Kraus and Stoll (1972).

[5] See, e.g., Panchenko, Gerasymchuk, and Pavlov (2013), Hein, Schwind, and Spiwoks (2008, 2012), Tedeschi, Iori, and Gallegati (2009, 2012), and Markose, Alentorn, and Krause (2004). Sumpter, Zabzina, and Nicolis (2012) show that a small number of leaders can heavily influence decisions.

their own actions (cf. Inman and Zeelenberg, 2002; Zeelenberg et al., 2002; Kahnemann and Tversky, 1982). There are now some empirical findings on the status-quo bias as it applies to financial market actors.[6] Numerous experimental studies also provide evidence of the status-quo bias (see, e.g., Geng, 2016; Yen and Chuang, 2008). Hence, it seems reasonable to consider the status-quo bias as a possible reason for non-optimal diversification. There is as yet only one experimental study that directly addresses this topic: Brown and Kagel (2009) yield information on the influence of the status-quo bias on non-optimal portfolio choices.

Looking now at the literature on the gambler's fallacy, we start with the experimental study by Gubaydullina and Spiwoks (2015) that found that irrelevant information can distract subjects from optimal diversification decisions. Considering the history of random events in evaluating random processes seems particularly tempting to many subjects, a phenomenon known as the gambler's fallacy. For example, if a coin toss shows "heads" three times in a row, many people assume that "tails" will show next. The history of unconnected random events, however, does not reveal anything about the future. The possibility for "heads" in the fourth toss is also exactly 50%. The gambler's fallacy has long been a subject of interest,[7] but no work has been done in the context of portfolio diversification.

We conduct an experiment to discover the reasons for insufficient portfolio diversification. In 15 periods, the subjects must make individual decisions about the structure of a portfolio of stocks. By taking a between-subjects approach, we examine the possible influence of herding, the status-quo bias, and the gambler's fallacy. We find that neither herding nor status-quo bias contribute significantly to non-optimal portfolio choices. The gambler's fallacy, however, plays an important role in these decisions.

2 Hypotheses and Experimental Design

2.1. Identification of Optimal Portfolios

Identifying optimal diversification decisions is difficult even in the easiest of cases, when there are only two stocks (A and B) involved that are independent in their income return development. Not only must the efficient frontier of all possible stock combinations be determined, but the investor's indifference curve must be considered (see Figure 1).

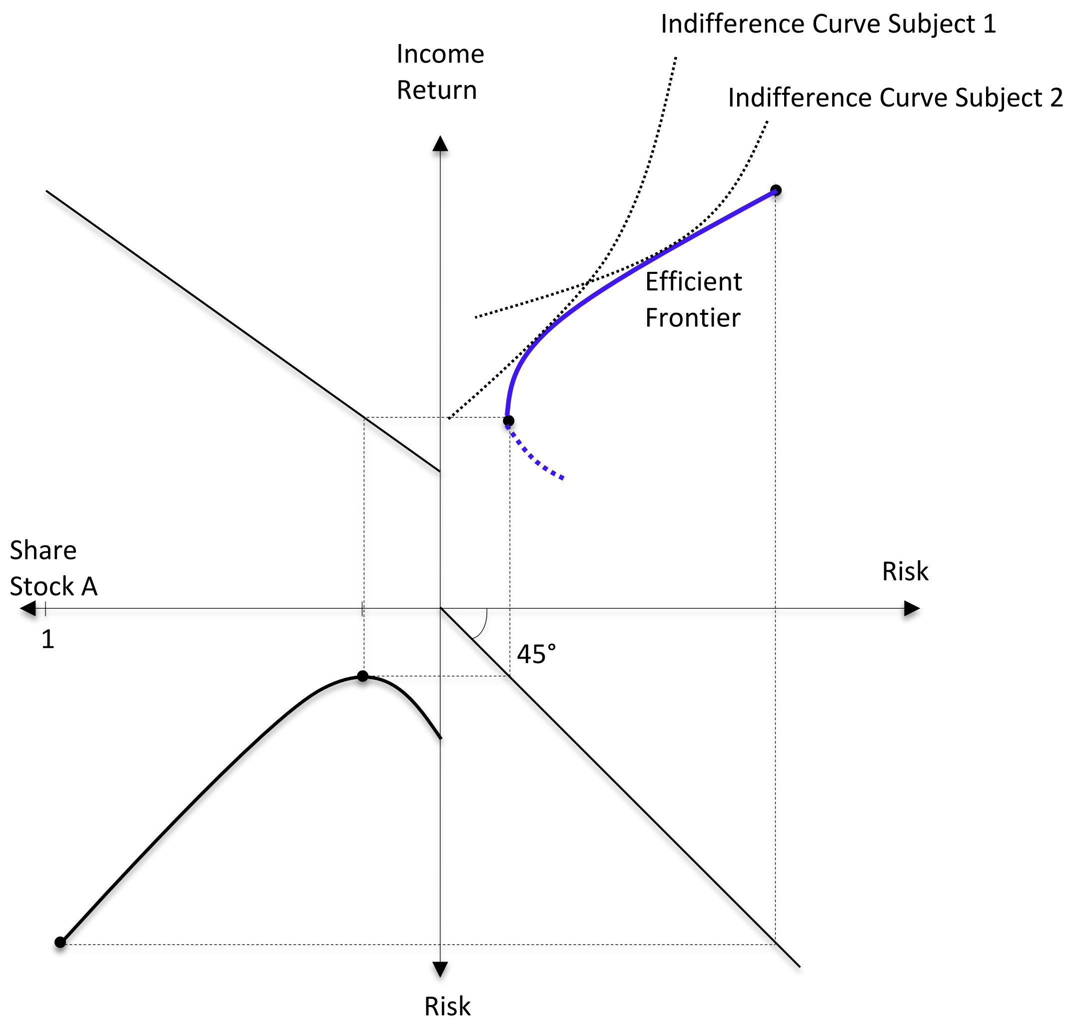
A strongly risk-averse investor (Subject 1) finds his ideal combination of stocks in the lower margin of the efficient frontier. A less risk-averse investor (Subject 2), however, finds his ideal combination of stocks in the upper margin of the efficient frontier. Although there are reliable empirical methods to differentiate between risk-averse, risk-neutral and risk-loving

[6] See, e.g., Freiburg and Grichnik (2013), Bryant, Evans, and Bishara (2012), Gubaydullina, Hein, and Spiwoks (2011), Kempf and Ruenzi (2006), Choi et al. (2004), Agnew, Balduzzi, and Sundén (2003), and Patel, Zeckhauser, and Hendricks (1991).

[7] See e.g. Chen, Moskowitz, and Shue (2016), Suetens, Galbo-Joergensen, and Tyran (2016), Stöckl et al. (2015), Powdthavee and Riyanto (2012), Barron and Leider (2010), Ayton and Fischer (2004), Clotfelter and Cook (1991), Tversky and Kahneman (1971, 1974).

subjects,[8] capturing the exact layout of the indifference curves for a specific subject remains impossible. Some studies try to solve this problem by considering all stock combinations on the efficient frontier as an ideal choice. However, this approach does not consider that only one exact point of the efficient frontier can be deemed the optimal combination of stocks for an individual investor.

Figure 1: Identification of Optimal Stock Combinations in Consideration of the Efficient Frontier and Individual Risk Aversion, or Individual Indifference Curves

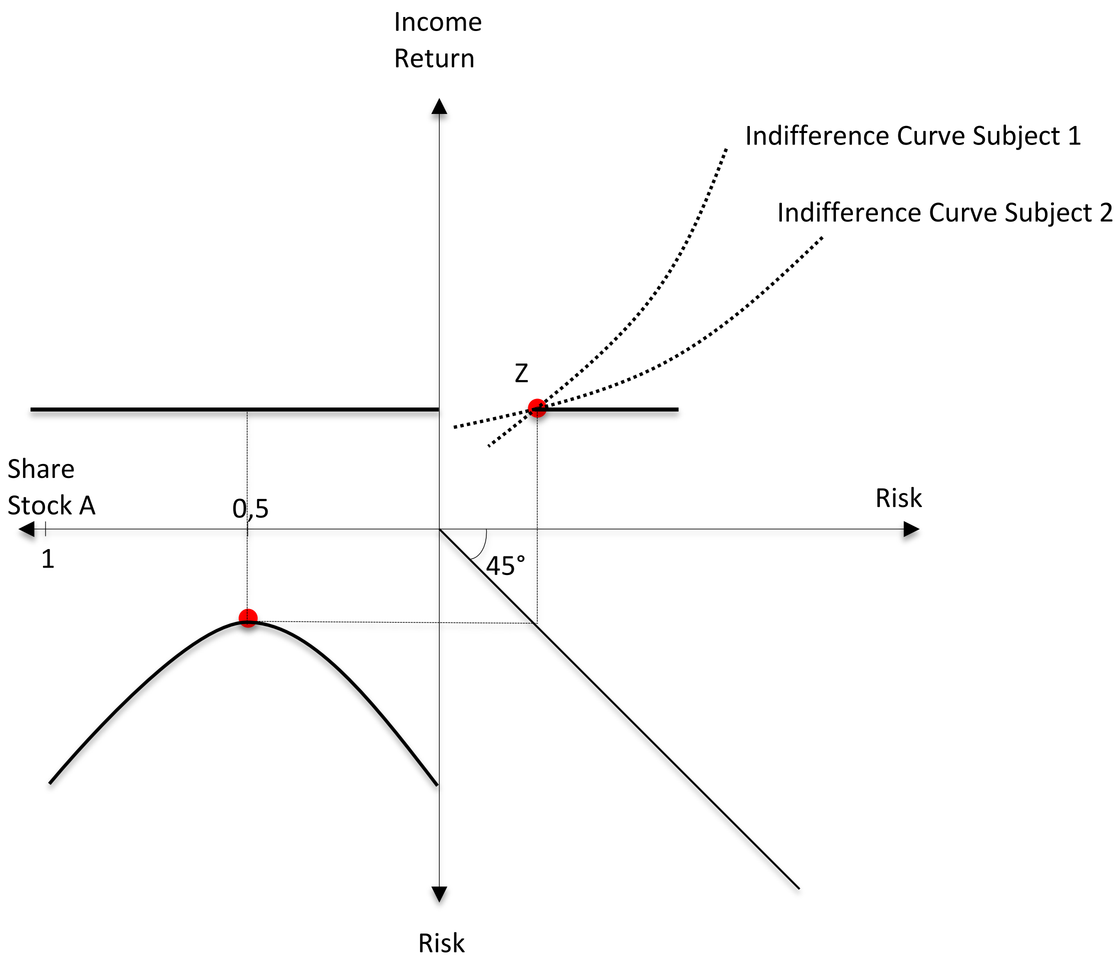


To avoid results so vague that interpretation of them is an exercise in futility, we employ the method of Gubaydullina and Spiwoks (2015): The subjects are offered two entirely uncorrelated alternatives for investment (A and B), which are identical regarding the expected income return and risk. By doing so, the efficient frontier is reduced to a single point (point Z in Figure 2). In this environment, it is of no importance whether a strongly or a less risk-averse subject makes the decision. In both cases, only the exactly equal mix of both invest-

[8] See, e.g., Lönnqvist *et al.* (2015), Charness, Gneezy, and Imas (2013), Crosetto and Filippin (2013), Dohmen *et al.* (2011), Eckel and Grossmann (2002, 2008), Lejuez *et al.* (2003), Holt and Laury (2002), and Gneezy and Potters (1997).

ment alternatives (A and B) can be interpreted as the ideal combination of stocks. Only with this methodology can precise results in an experiment on diversification be obtained.

Figure 2: Precise Identification of an Ideal Combination of Stocks with a Punctiform Efficient Frontier (Point Z).



Ours is an individual decision experiment that follows the approach of Gubaydullina and Spiwoks (2015). The subjects can choose between two different risky securities, A and B. In each period, they assemble a four-stock portfolio. The possible portfolios are AAAA, AAAB, AABB, ABBB, and BBBB. Stock A and stock B both generate an income return of ±0 experimental currency units (ECU) or +7 ECU in each period. Both possible returns occur with a probability of 50% and follow a random process. Stock A's performance is independent of Stock B's performance. Both stocks thus have an expected value of 3.5 ECU. A portfolio consisting of four stocks is expected to generate a return of 14 ECU per period (see Table 1). The ECUs are converted into Euros in the ratio of 10:1, resulting in an expected value of €1.40 per period. Hence, the subjects can profit from their investment behavior if it is successful.

The five possible portfolios may have the same expected income return of 14 ECU per period but the exposure to risk – henceforth expressed by the standard deviation – is different for each portfolio.[9] The standard deviation of the combination AAAA is 14.0, whereas the standard deviation of the combination AABB is only 9.9 (see Table 1).

Table 1: Expected Values and Standard Deviations of the Income Return for the Five Portfolios Considering the Possible Random Events for Stocks A and B in ECU

Random Events	A: +7; B: +7 (p_1 = 0.25)	A: +7; B: ±0 (p_2 = 0.25)	A: ±0; B: +7 (p_3 = 0.25)	A: ±0; B: ±0 (p_4 = 0.25)	*E(r)*	*SD*
AAAA	+28	+28	±0	±0	14	14.0
AAAB	+28	+21	+7	±0	14	11.1
AABB	+28	+14	+14	±0	14	9.9
ABBB	+28	+7	+21	±0	14	11.1
BBBB	+28	±0	+28	±0	14	14.0

p = probability of occurrence; E(r) = expected value of income return; SD = standard deviation.

2.2 Rational Strategy

A rational, risk-averse subject should always choose the combination AABB. Since the expected income returns of the five possible portfolios are identical, it is rational for each risk-averse subject to choose the portfolio with the minimum variance – independent of the degree of the subject's risk aversion.

This choice is intuitive. Regarding the structured components of the given stocks, the subjects can recognize the portfolio with the minimum variance without having to make any mathematical calculations. Using simple plausibility, it can be established that the income return level is most when both stocks A and B are equally represented in the portfolio (see Table 1).

However, considering the numerous empirical findings on the incapability or reluctance of subjects to make reasonable diversification decisions, we expect clear deviations from the rational strategy (always portfolio AABB) to occur in this experiment.

We thus arrive at our first hypothesis.

Hypothesis 1: The subjects are going to behave rationally, which means that they are going to exclusively choose the portfolio with the minimum variance (AABB).

The experiment consists of three treatments. The rational investment strategy (always AABB) is easily realizable in all three treatments. Therefore, the subjects' tendency toward rational diversification decisions is analyzed in all three treatments (see Table 2).

[9] Whether or not the subjects exhibit herding behavior, whether or not they are subject to the status-quo bias, and regardless of whether or not they fall prey to the gambler's fallacy does not influence their payout. If herding, status-quo bias, or the gambler's fallacy lead to frequent deviation from the optimal strategy (always portfolio AABB), only the exposure to risk increases. The expected payout, however, remains unaffected. In fact, the experiment only shows random differences concerning the payouts. Neither consequent herding nor a permanent status-quo bias or a perpetual gambler's fallacy lead to systematically higher payouts than the optimal strategy.

Table 2: Contribution of the Three Treatments to the Objects of Investigation

Treatment	Rational Behavior Hypothesis 1	Herding Hypotheses 2, 3, 4	Status-Quo Bias Hypotheses 5, 6	Gambler's Fallacy Hypotheses 7, 8
Treatment 1 VIEW+0%OPT	X	X	X	X
Treatment 2 NOVIEW+0%OPT	X	X		X
Treatment 3 VIEW+100%OPT	X		X	X

2.3 Herding

As are investigating herding, or the influence of the guru effect, the subjects must be given the opportunity to copy the portfolio decisions of the majority or those of the most successful fellow investor in each period. This results in an experiment that is structured in multiple periods. The portfolios can be rearranged at no cost before the start of each period.

In Treatment 1 (VIEW+0%OPT), the portfolios of each subject as well as their investment success are published in a ranking. Thus, before they decide whether or not to restructure their own portfolio, the subjects are given insight into their fellow investors' portfolio choices in the past period and into the portfolio of the most successful subject. This allows the subjects to follow the majority or the most successful investor (guru). In Treatment 2 (NOVIEW+0%OPT), the subjects do not receive any information about the other subjects' behavior or their investment success. They are informed solely of their own success and therefore do not have the option of following a guru or the majority because neither are detectable (see Figure 3 and Table 2).

Given the numerous empirical findings on the occurrence of herding in the financial market, we expect the portfolios to assimilate during Treatment 1 (VIEW+0%OPT). Thus our second hypothesis reads as follows.

Hypothesis 2: The subjects are not going to converge in the 15 periods of Treatment 1 (VIEW+0%OPT) and will not form a herd.

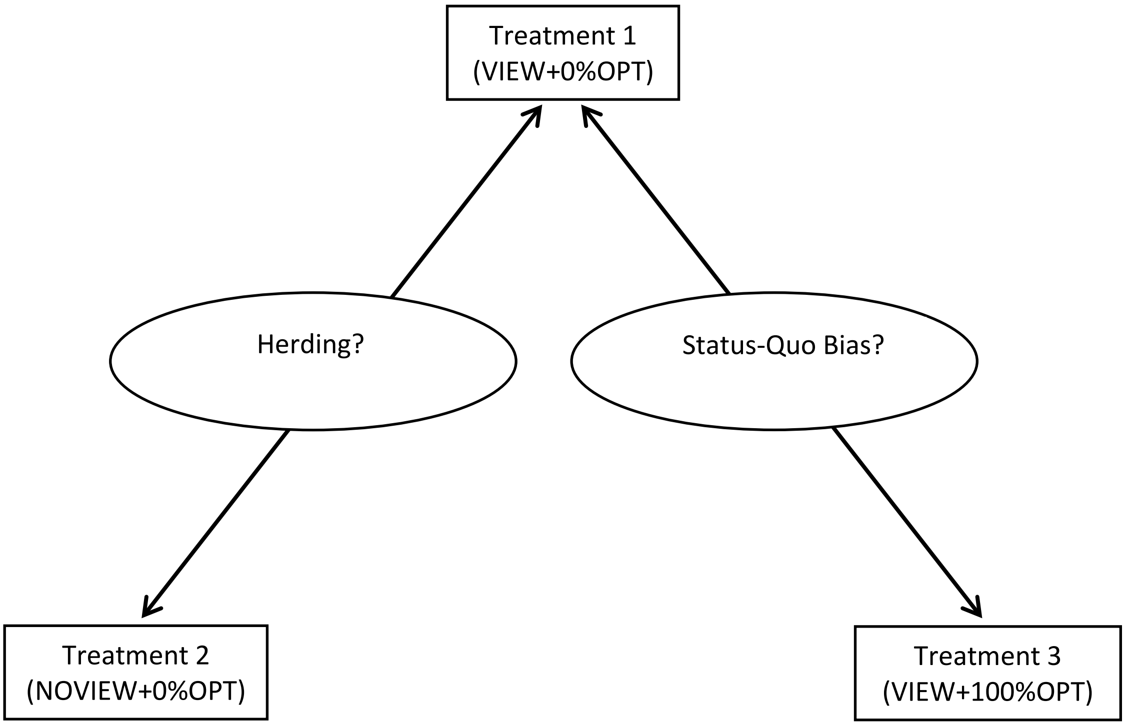
Since the investment behavior and success of the other subjects cannot be observed in Treatment 2, we expect the subjects to be less distracted from the rational strategy (always portfolio AABB). Thus our third hypothesis reads as follows.

Hypothesis 3: The average deviation from the rational strategy (always portfolio AABB) is not going to be stronger in Treatment 1 (VIEW+0%OPT) than in Treatment 2 (NOVIEW+0%OPT).

If deviations from the rational strategy occur more often and are stronger in Treatment 1 (VIEW+0%OPT), this should show in the average exposure to risk.

Hypothesis 4: The average exposure to risk is not significantly higher in Treatment 1 (VIEW+0%OPT) than in Treatment 2 (NOVIEW+0%OPT).

Figure 3: Interrelation of the Three Treatments



2.4 Status-Quo Bias

To reveal, or not, the presence of status-quo bias, we equip the subjects with different stocks in Treatment 1 (VIEW+0%OPT): 25% of the subjects each start with portfolio AAAA, portfolio AAAB, portfolio ABBB, and portfolio BBBB; 0% of the subjects start with the optimal portfolio (AABB). The subjects were allowed to reassemble their portfolios before the first period. In Treatment 3 (VIEW+100%OPT), 100% of the subjects receive the optimal portfolio (AABB) at the beginning of the experiment, which they can again reassemble before it starts (see Figure 3 and Table 2). As the status-quo bias has often been empirically established, we assume that the optimal portfolio (AABB) is more frequently selected during the 15 periods of Treatment 3 (VIEW+100%OPT) than during Treatment 1 (VIEW+0%OPT).

Hypothesis 5: The average deviation from the rational strategy (always portfolio AABB) will not be stronger in Treatment 1 (VIEW+0%OPT) than in Treatment 3 (VIEW+100%OPT).

If the deviations from the rational strategy are stronger and occur more often in Treatment 1 (VIEW+0%OPT), this should show in the average exposure to risk.

Hypothesis 6: The average exposure to risk will not be higher in Treatment 1 (VIEW+0%OPT) than in Treatment 3 (VIEW+100%OPT).

2.5 The Gambler's Fallacy

To detect whether the subjects have fallen prey to the gambler's fallacy, we ask them in all three treatments, as well as between Periods 4 and 5 and Periods 10 and 11, about the reasons for their portfolio choices (in Period 5 and Period 11). The tendency toward the gambler's fallacy can take effect in all three treatments and is therefore investigated in all three (see Table 2).

As the gambler's fallacy is an often observed phenomenon, we expect the subjects to try to detect patterns in the history of random events, which do not exist. For example: "After Stock A has generated a high income return, I will put my faith in Stock B" or "After Stock B has lastly returned no income, I will choose Stock B." Responses like this are evidence of the gambler's fallacy. We expect this type of answer to be given often.

Hypothesis 7: The gambler's fallacy is not going to be one of the main reasons for certain portfolio choices.

Evidence of the gambler's fallacy can also be discovered by looking at the history of the game. If a positive (negative) event for Stock A (Stock B) frequently leads to a reduced (increased) interest in Stock A (Stock B) in the following period, the influence of gambler's fallacy can be inferred.

Hypothesis 8: A positive (negative) income return in the current period does not reduce (increases) the popularity of this stock in the next period.

2.6 Capture of Risk Attitude and Conduction of the Experiment

The exclusive rational strategy for risk-averse investors is to always choose portfolio AABB. Therefore, deviations from the rational strategy will be identified only if risk-averse subjects are admitted to the experiment. To this end, we tested each subject according to Holt and Laury (2002) and cleared the starting field of risk-neutral and risk-loving subjects. To ensure that the task was fully understood by all subjects, we asked them control questions. Only those who answered all control questions correctly were admitted to the experiment. The complete instructions and control questions can be found in Appendix 1.

The experiment was conducted from 19 May 2016 to 27 May 2016 at the Ostfalia Laboratory for Experimental Economic Research (Ostfalia Labor für experimentelle Wirtschaftsforschung *OLEW*) of the Ostfalia University of Applied Sciences in Wolfsburg. One-hundred-eighty-eight subjects took part in the experiment, 38 of whom showed to be risk-neutral or risk-loving. One-hundred-fifty subjects showed risk-averse behavior and thus were admitted to the actual experiment. Fifty-three subjects were exposed to Treatment 1, 46 subjects to Treatment 2, and 51 took part in Treatment 3. The subjects are students of the Ostfalia University of Applied Sciences in Wolfsburg, 84 of whom study at the Faculty of Business (44.7%), 28 at the Faculty for Health Services (14.9%), and 76 at the Faculty of Automotive Engineering (40.4%). Sixteen sessions were conducted in total. Ten to twelve subjects took part in each session.

The experiment was implemented in z-Tree (see Fischbacher, 2007). At the Ostfalia Lab, we used 12 workspaces, each equipped with a monitor, with a wall separating the subjects. The experiments were consistently overseen by a game master to prevent the subjects from communicating with each other or using unauthorized devices (like smartphones). The subjects did not receive a general show-up fee. When assessing their willingness to take risks, an average of €2.18 was paid out to each subject. The actual experiment resulted in a payout of €21.89 on average. In total, the subjects received an average payout of €24.07. The highest payout was €31.85, the lowest was €17.40. The experiment lasted 45 minutes on average. The payout can therefore be deemed highly attractive. All subjects appeared to concentrate and seemed motivated.

3 Results

3.1 Rational Strategy

The results of the experiments partly meet the expectations but also reveal some surprising facts. Hypothesis 1 states that the subjects are going to behave rationally, which means that they are going to exclusively choose the portfolio with the minimum variance (AABB). Tables 3 and 4 clearly show that Hypothesis 1 must be rejected. The optimal portfolio (AABB) is the most frequently chosen alternative in all three treatments, but more than 60% of all portfolio choices can be viewed as less than completely optimal (Table 3). This is also evident when analyzing the exposure to risk (average standard deviation of the portfolios). The t-test (one-sample mean-comparison) shows that, in all three treatments, portfolios with standard deviations significantly higher than the standard deviation of the optimal portfolio are chosen (Table 4). Many subjects thus exhibit non-rational investment behavior. On the other hand, however, extreme portfolios (AAAA or BBBB) were chosen in only 21.4% of all portfolio decisions.

Table 3: Percentage Distribution of the Portfolios in the Three Treatments

Treatment	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios
Rational Strategy for all Three Treatments	0%	0%	100%	0%	0%
Treatment 1 (VIEW+0%OPT)	11.68%	18.73%	39.75%	19.36%	10.44%
Treatment 2 (NOVIEW+0%OPT)	8.68%	22.45%	31.89%	22.46%	14.49%
Treatment 3 (VIEW+100%OPT)	8.36%	21.69%	40.93%	18.29%	10.71%
Total	9.63%	20.88%	37.74%	19.95%	11.77%

Table 4: Exposure to Risk (Average Standard Deviation of the Portfolios)

Treatment	Rational Strategy: Average SD	Actual: Average SD	P-Value
Treatment 1 (VIEW+0%OPT)	9.9	11.37	0.0000***
Treatment 2 (NOVIEW+0%OPT)	9.9	11.49	0.0000***
Treatment 3 (VIEW+100%OPT)	9.9	11.29	0.0000***
Total	9.9	11.40	0.0000***

*** = significant with an error rate of 1%; ** = significant with an error rate of 5%; * = significant with an error rate of 10%; SD = standard deviation.

3.2 Herding

Herding should be reflected in either a quick or gradual assimilation of the subjects' decisions. We therefore examine whether the decisions made by the subjects converge. Table 5 shows that the subjects' behavior continues to be fragmented until the last period, that is, no herding occurs.[10]

Table 5: Percentage Distribution of the Portfolios at the End of the Game in Treatment 1 (VIEW+0%OPT)

Session	Percentage of AAAA Portfolios in Period 15	Percentage of AAAB Portfolios in Period 15	Percentage of AABB Portfolios in Period 15	Percentage of ABBB Portfolios in Period 15	Percentage of BBBB Portfolios in Period 15
Session 1	30.0%	10.0%	50.0%	10.0%	0.0%
Session 2	33.3%	33.3%	22.2%	11.1%	0.0%
Session 3	11.1%	33.3%	44.4%	0.0%	11.1%
Session 4	30.0%	20.0%	40.0%	0.0%	10.0%
Session 5	0.0%	42.9%	28.6%	14.3%	14.3%
Session 6	37.5%	25.0%	0.0%	12.5%	25.0%

In the next step, we compare decisions for the portfolio that was preferred by most subjects to the portfolio that was chosen against the majority opinion. We conclude that the subjects chose the portfolio based on majority opinion in the previous period for a total of 320 times. The portfolio that the majority did not prefer in the previous period was chosen 422 times. Since herding can also be a temporary phenomenon, it makes sense to show the frequencies separately according to periods and to check the significance of the frequencies using a Chi-square goodness-of-fit test (Table 6). In Period 2 we can see, for instance, that 30 subjects follow the majority opinion of the previous period in choosing their portfolio, while only 23 subjects do not do so. However, this difference is insignificant. In those periods with signifi-

[10] The course of the six sessions of Treatment 1 are set out in Tables A-1 to A-6 in Appendix 2.

cant deviations (Periods 8, 9, 10, 11, and 13), those decisions not following the majority opinion always outbalance those that do. Hence, we cannot establish significant herding, evidenced by following the majority opinion, at any point during the experiment.

Table 6: Portfolio Decisions According to and Opposing the Majority Opinion in Treatment 1

Period	Number of Portfolios According to the Majority Opinion in Previous Period	Number of Portfolios Opposing the Majority Opinion in Previous Period	P-Value
2	30	23	0.336
3	25	28	0.680
4	31	22	0.216
5	24	29	0.492
6	24	29	0.492
7	27	26	0.891
8	20	33	0.074*
9	18	35	0.020**
10	18	35	0.020**
11	19	34	0.039**
12	22	31	0.216
13	15	38	0.002***
14	22	31	0.216
15	25	28	0.680
total	320	422	

*** = significant with an error rate of 1%; ** = significant with an error rate of 5%; * = significant with an error rate of 10%.

Neither is there much evidence of a "guru effect" (see Table 7). The guru obviously only impresses the subjects slightly. In 212 cases, their decision follows the previous leader's portfolio; they do not do so in 530 cases. Only in Period 7 are there more subjects following the guru's opinion than those not doing so. The Chi-square goodness-of-fit test shows, however, that the difference is insignificant. In all periods – except for Periods 3 and 7 – the decisions not to choose the portfolio of the most successful subject is observed significantly more often.

Overall, we conclude that there is no herding in Treatment 1 (VIEW+0%OPT). Neither the majority opinion nor the opinion of the leading subject (guru) has much impression on the subjects. Hypothesis 2 states that the subjects are not going to converge in the 15 periods of Treatment 1 and will not form a herd. Hypothesis 2 cannot be rejected.

Contrary to our expectations, the portfolio with the minimum variance is chosen significantly more often in Treatment 1 (VIEW+0%OPT) than in Treatment 2 (NOVIEW+0%OPT). We assumed that, by observing and following their fellow subjects' behavior, the subjects would frequently deviate from choosing the optimal stock combination AABB. In Treatment 2, where the other subjects' behavior cannot be observed, herding is generally impossible. As a matter of fact, the contrary turns out to be the case. While the portfolio with minimum variance (AABB) was chosen in 39.8% of all cases in Treatment 1 (VIEW+0%OPT), it was selected in only 31.9% of all cases in Treatment 2 (NOVIEW+0%OPT). The Wilcoxon-Mann-Whitney test shows that this unexpected difference is even statistically significant. The p-value is 0.021. Hypothesis 3 states that the average deviation from the rational strategy (always portfolio AABB) is not going to be stronger in Treatment 1 than in Treatment 2. Hypothesis 3 cannot be rejected.

Table 7: Portfolio Decisions Following and Not Following the Guru in Treatment 1

Period	Number of Portfolios Corresponding to the Leading Subject of the Previous Period	Number of Portfolios Opposing the Leading Subject of the Previous Period	P-Value
2	6	47	0.000***
3	22	31	0.216
4	12	41	0.000***
5	14	39	0.001***
6	19	34	0.039**
7	28	25	0.680
8	17	36	0.009***
9	7	46	0.000***
10	19	34	0.039**
11	12	41	0.000***
12	9	44	0.000***
13	11	42	0.000***
14	16	37	0.004***
15	20	33	0.074*
total	212	530	

*** = significant with an error rate of 1%; ** = significant with an error rate of 5%; * = significant with an error rate of 10%.

We had expected that the possibility of following other subjects would lead to a significantly higher exposure to risk (average standard deviation of the portfolios) in Treatment 1 (VIEW+0%OPT) than in Treatment 2 (NOVIEW+0%OPT). Since herding did not develop, risk exposure did not increase. The average standard deviation of the portfolios in Treatment 1 (VIEW+0%OPT) was 11.37. The average standard deviation of the portfolios in Treatment 2 (NOVIEW+0%OPT) was 11.49. According to the Wilcoxon-Mann-Whitney test, this difference is not relevant. The p-value is 0.5485. Hypothesis 4 states that the average exposure to risk is not significantly higher in Treatment 1 (VIEW+0%OPT) than in Treatment 2 (NOVIEW+0%OPT). Hence, Hypothesis 4 cannot be discarded.

As an interim result, we conclude that the subjects do not behave rationally in making most of their portfolio choices. The optimal portfolio (AABB) is chosen in only 30 - 40% of all cases (39.8% in Treatment 1, 31.9% in Treatment 2). Herding is clearly not responsible for this. There is neither a lasting orientation toward the portfolio structure chosen by a majority of subjects nor a lasting orientation toward the portfolio structure of the most successful subject.

3.3 Status-Quo Bias

To discover, or not, the presence of any status-quo bias, we now compare Treatment 1 with Treatment 3. In Treatment 1 (VIEW+0%OPT), 25% of the subjects are given portfolio AAAA, 25% portfolio AAAB, 25% portfolio ABBB, and 25% portfolio BBBB before the start of the experiment. No subject receives the optimal portfolio AABB (see Table A-7 in Appendix 3). In Treatment 3 (VIEW+100%OPT), every subject received the portfolio with the minimum variance (AABB) (see Table A-9 in Appendix 3). The subjects are permitted to reassemble their portfolios before the beginning of the first period.

We expected a tendency to follow the status-quo, as has often been found in the literature. That is, in Treatment 3 (VIEW+100%OPT), where all subjects start with the optimal portfolio (AABB), the optimal portfolio structure should have been selected more often than in Treatment 1 (VIEW+0%OPT), in which no subject was provided with an ideally structured portfolio. As a matter of fact, the subjects did not retain their initial portfolios. In Treatment 3 (VIEW+100%OPT), 39.2% of the subjects reassembled their portfolio before the start of the first period, resulting in only 60.8% of the portfolios having the ideal structure (AABB) before the start of the first period (see Figure 5 and Table A-9 in Appendix 3). Over the course of the game, this percentage drops even further.

Figure 4: Percental Distribution of the Portfolios in Treatment 1 (VIEW+0%OPT)

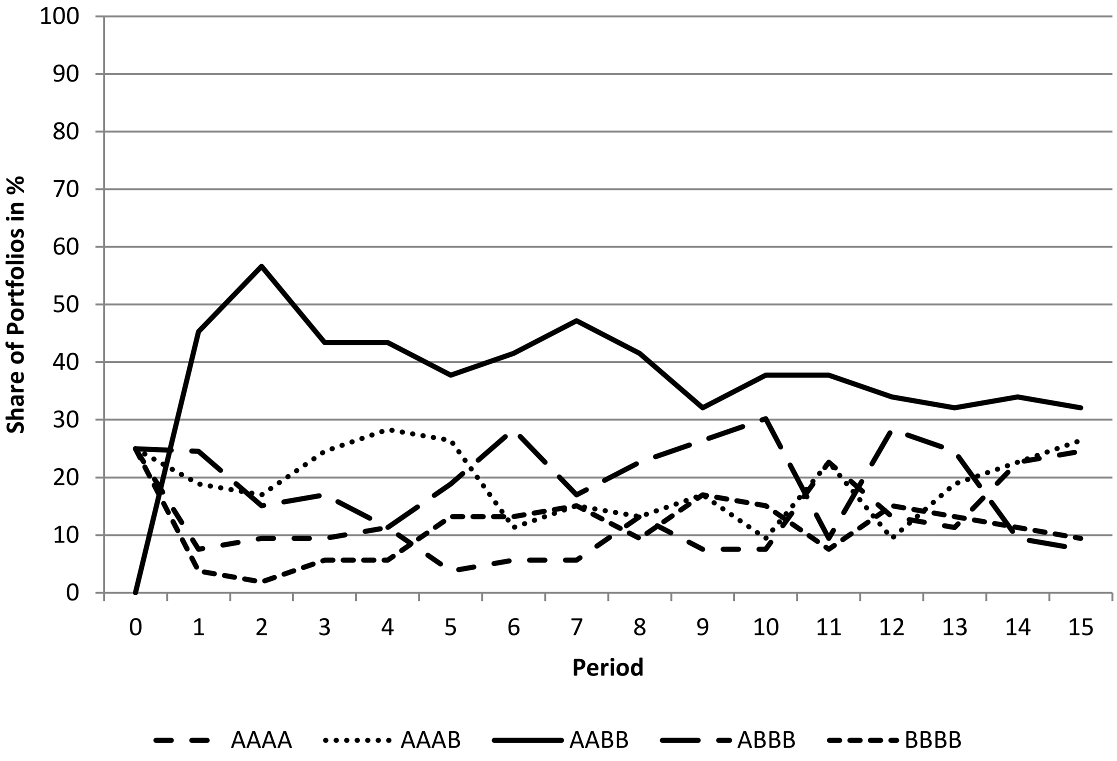
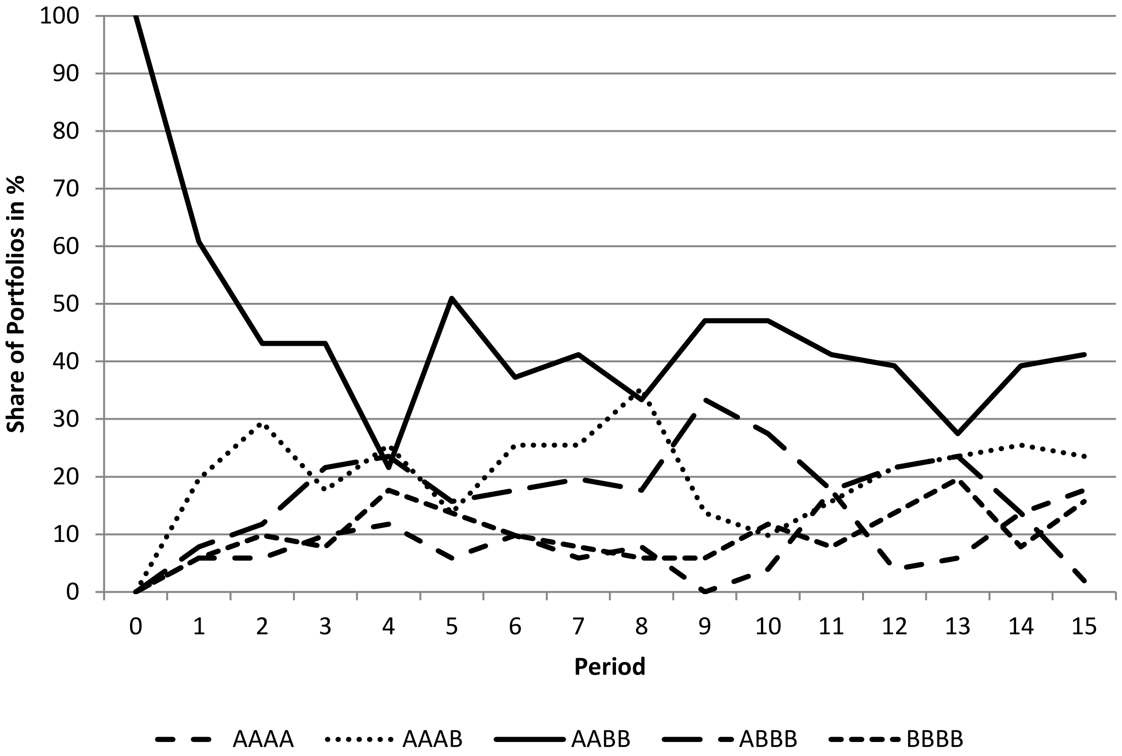


Figure 5: Percental Distribution of the Portfolios in Treatment 3 (VIEW+100%OPT)



In Treatment 1 (VIEW+0%OPT), the share of the optimal portfolio (AABB) increases to 45.3% even before the first period and ranges between 45% and 30% from Period 3 to Period 15 (see Figure 4 and Table A-7 in Appendix 3).

In Treatment 1 (VIEW+0%OPT), 39.8% of all portfolios have the optimal structure. In Treatment 3 (VIEW+100%OPT), only 40.9% of portfolios with this structure remain. This difference is small and non-relevant according to the Wilcoxon-Mann-Whitney test. The p-value is 0.6626. Hypothesis 5 states that the average deviation from the rational strategy (always portfolio AABB) will not be stronger in Treatment 1 (VIEW+0%OPT) than in Treatment 3 (VIEW+100%OPT). Hypothesis 5 cannot be discarded.

We assumed that, given the status-quo bias in Treatment 3 (VIEW+100%OPT), the optimal portfolio (AABB) would be chosen more often than in Treatment 1 (VIEW+0%OPT), which could have resulted in a significantly lower exposure to risk. As a matter of fact, the average exposures to risk (standard deviation) are 11.37 in Treatment 1 and 11.29 in Treatment 3. According to the Wilcoxon-Mann-Whitney test, this difference is not relevant. The p-value is 0.9741. Hypothesis 6 states that the average exposure to risk will not be higher in Treatment 1 (VIEW+0%OPT) than in Treatment 3 (VIEW+100%OPT). Hypothesis 6 cannot be rejected.

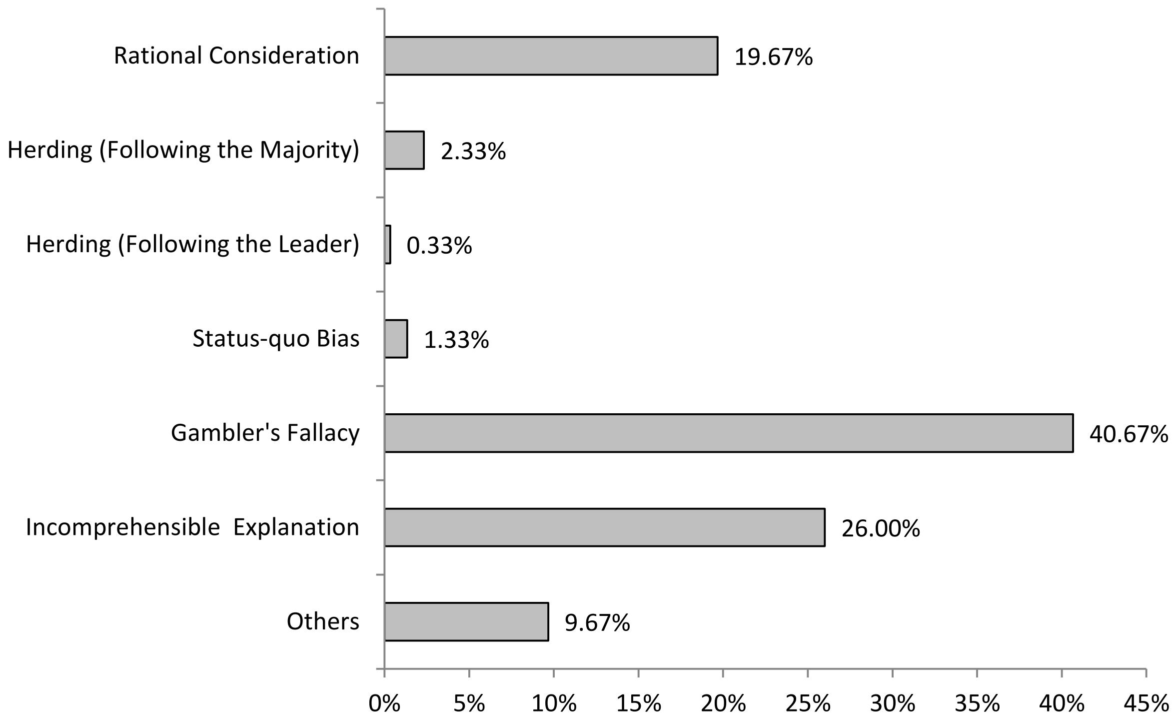
As an intermediate result, we conclude that most subjects do not behave rationally when compiling their portfolios; however, this deviation from the rational strategy is not explained by status-quo bias. Hence, we cannot confirm the result by Brown and Kagel (2009).

3.4 The Gambler's Fallacy

It is plausible to believe that subjects inferred from past random events that a certain sequence of future events would occur, even though this is a nonsensical interpretation of independent random events. To see if this was indeed the case, we assess the reasons the subjects gave for their portfolio choices at the beginning of Period 5 and at the beginning of Period 11. In the experiment, the subjects were asked to provide the following information: "Please give a short explanation for your decision in Period 5 (Period 11)! This explanation does not affect your result! You can openly state your considerations." We expected to receive answers demonstrating that the subjects had indeed fallen prey to the gambler's fallacy.[11]

We differentiate between the following clusters of reasons: (1) Rational consideration (suitable orientation toward the expected value of the income return and the risk); (2) herding (following the majority of investors); (3) herding (following the most successful investor, the guru); (4) status-quo bias (following the present portfolio); (5) the gambler's fallacy (following the history of random events); (6) incomprehensible explanations;[12] and (7) other reasons.[13]

Figure 6: Percentage of the Named Reasons for the Portfolio Choices



[11] Examples of answers demonstrating the gambler's fallacy: "One of the stocks did not show a positive value development which is why I assume that this stock develops positively in the next period." "The value of Stock A has developed more positively than Stock B which is why I hope that Stock B now develops positively." "The performance of Stock B in period 10 was 0 and Stock A showed an added value of 7, which is why I assume that Stock B could gain in the next period." "With Stock A having developed positively, I now invest in Stock B."

[12] Incomprehensible explanations are, for instance, filling the answer box with random letters or only numbers.

[13] Other reasons are, for example: "I have to take a higher risk to achieve a higher ranking."

It can be concluded that subjects often tried to forecast future random events from past random events. This consideration was detected in 40.67% of the answers (see Figure 6). The frequency of this reasoning significantly exceeds a barely accidental distribution of answers. The Chi-square goodness-of-fit test shows a p-value of 0.000. Hypothesis 7 states that the gambler’s fallacy is not going to be one of the main reasons for certain portfolio choices. Hypothesis 7 is rejected. It appears that the gambler’s fallacy contributes significantly to irrational portfolio choices.

The question remains whether a behavioral pattern explaining the gambler’s fallacy can be directly explained by analyzing the portfolio choices. The sequence of random events (income return of Stock A and Stock B) was structured by coin toss in advance and then taken as a basis in all treatments (see Table 8).

Table 8: Sequence of Random Events and Resulting Gambler’s Fallacy

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Random Event of Stock A	±0	+7	±0	+7	+7	±0	±0	+7	±0	±0	+7	+7	±0	±0	+7
Random Event of Stock B	+7	+7	+7	±0	+7	±0	+7	±0	±0	+7	±0	+7	+7	±0	±0
Preference Gambler’s Fallacy		A	A	A	B	B	B	A	B	B	A	B	B	A	A
Actual Overweight of the Stock		A	A	A	B	B	B	A	B	B	A	B	B	A	A
Observations in Field 1 of the 2x2 Matrix		1	2	3				4			5			6	7
Observations in Field 4 of the 2x2 Matrix					1	2	3		4	5		6	7		

The first row of Table 8 lists the periods of the experiment. The second and the third rows list the sequences of random events of Stock A and Stock B (income return of ±0 or +7). The fourth line shows which stock is preferred by the subjects who fall prey to the gambler’s fallacy. In Period 1, for instance, Stock A has an income return of ±0 ECU and Stock B has an income return of +7 ECU. This is the development that the subjects can see before Period 2. This results in a preference for Stock A. Before the start of Period 3, the “errant gambler” again prefers Stock A. This is because, in the previous period, both stocks had a positive income return (+7 ECU) but for Stock A, it is the first in succession, while it is the second in succession for Stock B. Before Period 4, the events of Period 2 are repeated and result in yet another preference for Stock A. Period by period, considerations such as these lead to preferences that are listed in the fourth line of the table.

If all portfolios of all subjects contain more Stocks A than Stocks B, we call this an overweight of Stock A. If all portfolios of all subjects contain more Stocks B than Stocks A, we call this an overweight of Stock B. We can establish that Stock A (Stock B) is being overly weighted if the gambler's fallacy leads to a preference for Stock A (Stock B). And this is exactly what occurs in the fifth row of Table 8 (for more specific detail, see Table A-10 in Appendix 4). In seven periods (Periods 2, 3, 4, 8, 11, 14, and 15), we would expect Stock A to be overly weighted according to the gambler's fallacy, which is indeed the case (see sixth row in Table 8). In seven periods (Periods 5, 6, 7, 9, 10, 12, and 13), we would expect Stock B to be overly weighted according to gambler's fallacy, which is also the case (see seventh row in Table 8). In a 2 x 2 matrix on the predictive accuracy of the gambler's fallacy regarding the overweight of stocks A and B in Periods 2–15, the Chi-square goodness-of-fit test shows that a predictive accuracy of 14:0, which is highly significant (p-value = 0.000).

Hypothesis 8 states that a positive (negative) income return in the current period does not reduce (increases) the popularity of this stock in the next period. Hypothesis 8, therefore, must be rejected. It is obvious that the gambler's fallacy affects the subjects' portfolio choices and thereby contributes to the fact that the rational choice (to always chose the portfolio with the minimum variance) is observed in only 37.7% of all decisions made by the subjects.

This result specifically confirms the study by Stöckl *et al.* (2015) that also demonstrates that investment decisions can be influenced by the gambler's fallacy. We can also conclude that the gambler's fallacy is rather unaffected by general conditions. It occurs regardless of whether information is or is not provided about the other subjects' behavior, and also manifests regardless of whether or not subjects start with an optimal portfolio. Stöckl *et al.* (2015), too, affirm the stability of the gambler's fallacy under various modes of communication or in the event of group decisions.

Barron and Leider (2010) also find indications of the gambler's fallacy. However, in their study, the gambler's fallacy does not prove stable if long historic time series of random events are shown to the subjects. Chen, Moskowitz, and Shue (2016) also find indications of gambler's fallacy, which in their study is stronger in inexperienced subjects than in experienced subjects. Furthermore, Suetens, Galbo-Joergensen, and Tyran (2016) show that the gambler's fallacy can change into a hot-hand fallacy if a certain event reoccurs in unusually long streaks. Our results are in accordance with the studies by Huber, Kirchler, and Stöckl (2010) and Ayton and Fischer (2004). These studies show that the gambler's fallacy manifests in the prognosis of random events. But these studies also show that, when estimating people's success, the hot-hand fallacy is predominant.

4 Conclusion

The present experimental study examines subjects' diversification behavior. It focuses on the research question of whether herding (being guided by most fellow investors or by the most successful investor (guru)), status-quo bias, and / or the gambler's fallacy can explain why many subjects maintain clearly underdiversified portfolios. Although much empirical evidence has been found to explain the influence of these phenomena (herding, status-quo bias, and the gambler's fallacy) on many economic decisions, they have not been experimentally examined in the context of diversification decisions.

This experiment follows the approach by Gubaydullina and Spiwoks (2015): There are only two alternatives for investment (Stock A and Stock B), which can produce only two results. Either they bring an income return of ±0 ECU or a return of +7 ECU per period. Both results occur with a probability of 50%. Stock A's return and Stock B's return are independent events. In this environment, the efficient frontier is reduced to one point (equal mixture of Stock A and Stock B), so that the subjects' degree of risk aversion does not influence optimal portfolio choice.

In Treatment 1 (VIEW+0%OPT), the subjects receive information on their fellow investors in each period. They learn who chose which portfolio and how successful everyone was with their decisions. In Treatment 2 (NOVIEW+0%OPT) this information is not provided. A comparison of Treatment 1 and Treatment 2 is intended to indicate the extent to which subjects are distracted from making optimal portfolio choices due to herding behavior (being guided by most fellow investors or by the most successful investor (guru)). Herding was not observed: The subjects neither followed most of their fellow investors nor did they emulate the most successful investor (guru). We draw the conclusion that herding does not play a significant role in explaining non-optimal portfolio choices.

In Treatment 1 (VIEW+0%OPT), the subjects start with different portfolios. Each of the portfolios AAAA, AAAB, ABBB, and BBBB is given to 25% of the subjects at the start of the experiment. In Treatment 3 (VIEW+100%OPT), 100% of the subjects receive the optimal portfolio (AABB). If status-quo bias is playing a role in diversification decisions, the optimal portfolio should be chosen more often in Treatment 3 (VIEW+100%OPT) and the average exposure to risk should be less than in Treatment 1 (VIEW+0%OPT). However, the experiment did not reveal a significant difference between Treatment 1 and Treatment 3 regarding the choice of optimal portfolios or the average exposure to risk. Hence, status-quo bias does not explain non-optimal portfolio choices. In this respect, our results clearly contradict the study by Brown and Kagel (2009).

Over the course of the experiment, the subjects were asked twice about their reasons for making a certain decision. We conclude from their answers that many subjects are inclined to infer future developments from past random events. If a positive event (+7 ECU) has occurred for Stock A (Stock B), the subjects tend to assume that Stock B (Stock A) will be the optimal choice in the next period. Conversely, if a negative event (±0 ECU) has occurred for Stock A (Stock B), the subjects tend to assume that Stock A (Stock B) will be the optimal choice in the next period. This phenomenon, known as the gambler's fallacy, is clearly visible when analyzing the subjects' portfolio choices. This result particularly confirms the study by Stöckl *et al.* (2015) that shows that investment decisions can be influenced by the gambler's fallacy and that the latter is stable under different general conditions.

Most portfolio choices (62.3%) in this experiment were non-optimal. However, neither herding nor status-quo bias contributed to this irrational behavior. The gambler's fallacy, however, had a substantial influence on the subjects' portfolio choices.

Acknowledgments

We thank the editor and anonymous referees for their helpful and detailed comments, which enormously improved the quality of this paper. We also thank the participants of the 4th International Meeting on Experimental and Behavioral Social Sciences (IMEBESS) at Universitat de Barcelona 2017 for some very constructive comments.

References

Agnew, J. / Balduzzi, P. / Sundén, A. (2003): Portfolio Choice and Trading in a Large 401(k) Plan, American Economic Review, 93(1), 193-215.

Anderson, A. (2013): Trading and Under-Diversification, Review of Finance, 17(5), 1-43.

Anderson, C. J. (2003): The Psychology of Doing Nothing: Forms of Decision Avoidance Result from Reason and Emotion, Psychological Bulletin, 129(1), 139-167.

Ayton, P. / Fischer, I. (2004): The Hot-Hand Fallacy and the Gambler's Fallacy: Two Faces of Subjective Randomness? Memory and Cognition, 32(8), 1369-1378.

Baltussen, G. / Post, G. T. (2011): Irrational Diversification: An Examination of Individual Portfolio Choice, Journal of Financial Quantitative Analysis, 46(5), 1463-1491.

Banerjee, A. V. (1992): A Simple Model of Herd Behavior, Quarterly Journal of Economics, 107(3), 797-817.

Barber, B. M. / Odean, T. (2000): Trading is Hazardous to your Wealth: The Common Stock Investment Performance of Individual Investors, Journal of Finance, 55(2), 773-806.

Barron, G. / Leider, S. (2010): The Role of Experience in the Gambler's Fallacy, Journal of Behavioral Decision Making, 23(1), 117-129.

Belhoula, M. / Naoui, K. (2011): Herding and Positive Feedback Trading in American Stock Market: A Two-Directional Behavior of Investors, International Journal of Business Management, 6(9), 244-252.

Benartzi, S. (2001): Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock, Journal of Finance, 56(5), 1747-1764.

Benartzi, S. / Thaler, R. H. (2001): Naïve Diversification Strategies in Defined Contribution Saving Plans, American Economic Review, 91(1), 79-98.

Bikhchandani, S. / Hirshleifer, D. / Welch, I. (1998): Learning from the Behavior of Others: Conformity, Fads and Informational Cascades, Journal of Economic Perspectives, 12(3), 151-170.

Blume, M. E. / Friend, I. (1975): The Asset Structure of Individual Portfolios and Some Implications for Utility Functions, Journal of Finance, 30(2), 585-603.

Bode, M. / van Echelpoel, A. / Sievi, C. R. (1994): Multinationale Diversifikation: Viel zitiert, kaum befolgt, Die Bank, 94(4), 202-206.

Boyson, N. M. (2010): Implicit Incentives and Reputational Herding by Hedge Fund Managers, Journal of Empirical Finance, 17(3), 283-299.

Brown, A. L. / Kagel, J. H. (2009): Behavior in a Simplified Stock Market: The Status-Quo Bias, the Disposition Effect and the Ostrich Effect, Annals of Finance, 5(1), 1-14.

Bryant, L. L. / Evans, J. / Bishara, J. (2012): Status Quo Bias in the Mutual Fund Market: Consumer Inertia Amidst Benefitless Fee Increases, Banking and Finance Review, 4(1), 89-107.

Chang, C. H. (2013): The Relationship between the Disposition Effect and Herding Behavior: Evidence from Taiwan's Information Technology Stocks, International Journal of Economics and Management Engineering, 7(7), 1971-1975.

Chang, E. C. / Cheng, J. W. / Khorana, A. (2000): An Examination of Herd Behavior in Equity Markets: An International Perspective, Journal of Banking and Finance, 24(10), 1651-1679.

Charness, G. / Gneezy, U. / Imas, A. (2013): Experimental Methods: Eliciting Risk Preferences, Jorunal of Economic Behavior and Organization, 87, 43-51.

Chen, D. / Moskowitz, T. J. / Shue, K. (2016): Decision Making under the Gambler's Fallacy: Evidence from Asylum Judges, Loan Officers and Baseball Umpires, Quarterly Journal of Economics, 131(3), 1181-1242.

Chen, Y. / Wang, C. / Lin, F. (2008): Do Qualified Foreign Institutional Investors Herd in Taiwan's Securities Market? Emerging Markets Finance and Trade, 44(4), 62-74.

Chiang, T. C. / Zheng, D. (2010): An Empirical Analysis of Herd Behavior in Global Stock Markets, Journal of Banking and Finance, 34(8), 1911-1921.

Choe, H. / Kho, B. C. / Stulz, R. M. (1999): Do Foreign Investors Destabilize Stock Markets? The Korean Experience in 1997, Journal of Financial Economics, 54(2), 227-264.

Choi, J. J. / Laibson, D. / Madrian, B. C. / Metrick, A. (2004): For Better or For Worse: Default Effects and 401(k) Savings Behavior, in: Wise, D. A.: Perspectives in the Economics of Aging, University of Chicago Press, 81-121.

Choi, J. J. / Laibson, D. / Madrian, B. C. / Metrick, A. (2009): Reinforcement Learning and Savings Behavior, Journal of Finance, 64(6), 2515-2534.

Choi, S. (2016): Herding Among Local Individual Investors: Evidence from Online and Offline Trading, Economics Letters, 144, 4-6.

Christie, W. G. / Huang, R. D. (1995): Following the Pied Piper: Do Individual Returns Herd Around the Market? Financial Analysts Journal, 51(4), 31-37.

Clotfelter, C. T. / Cook, P. J. (1991): The "Gambler's Fallacy" in Lottery Play, NBER Working Paper, 3769.

Crosetto, P. / Filippin, A. (2013): The "Bomb" Risk Elicitation Task, Journal of Risk Uncertainty, 47(1), 31-65.

Devenow, A. / Welch, I. (1996): Rational Herding in Financial Economics, European Economic Review, 40, 603-615.

Dimmock, S. G. / Kouwenberg, R. / Mitchell, O. S. / Peijnenburg, K. (2016): Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence, Journal of Finance and Economics, 119(3), 559-577.

Dohmen, T. / Falk, A. / Huffman, D. / Sunde, U. / Schupp, J. / Wagner, G. G. (2011): Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences, Journal of the European Economic Association, 9(3), 522-550.

Eckel, C. C. / Grossman, P. J. (2002): Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk, Evolution and Human Behavior, 23(4), 281-295.

Eckel, C. C. / Grossman, P. J. (2008): Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices, Journal of Economic Behavior and Organization, 68(1), 1-17.

Ennis, R. M. / Sebastian, M. D. (2003): A Critical Look at the Case for Hedge Funds, Journal of Portfolio Management, 29(4), 103-112.

Eyster, E. / Weizsäcker, G. (2011): Correlation Neglect in Financial Decision Making, DIW Discussion Paper, 1104.

Fellner, G. / Güth, W. / Maciejovsky, B. (2004): Illusion of Expertise in Portfolio Decisions: An Experimental Approach, Journal of Economic Behavior and Organization, 55(3), 355-376.

Fernandes, D. (2013): The 1/N Rule Revisited: Heterogeneity in the Naïve Diversification Bias, International Journal of Marketing Research, 30(3), 310-313.

Fischbacher, U. (2007): Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, Experimental Economics, 10(2), 171-178.

Freiburg, M. / Grichnik, D. (2013): Institutional Reinvestments in Private Equity Funds as a Double-Edged Sword: The Role of the Status-Quo Bias, Journal of Behavioral Finance, 14(2), 134-148.

Galariotis, E. C. / Rong, W. / Spyrou, S. I. (2015): Herding on Fundamental Information: A Comparative Study, Journal of Banking and Finance 50, 589-598.

Geng, S. (2016): Decision Time, Consideration Time and Status-Quo Bias, Economic Inquiry, 54(1), 433-449.

Gneezy, U. / Potters, J. (1997): An Experiment on Risk Taking and Evaluation Periods, Quarterly Journal of Economics, 112(2), 631-645.

Goetzmann, W. N. / Kumar, A. (2008): Equity Portfolio Diversification, Review of Finance, 12(3), 433-463.

Gubaydullina, Z. / Hein, O. / Spiwoks, M. (2011): The Status-Quo Bias of Bond Market Analysts, Journal of Applied Finance and Banking, 1(1), 31-51.

Gubaydullina, Z. / Spiwoks, M. (2015): Correlation Neglect, Naïve Diversification, and Irrelevant Information as Stumbling Blocks for Optimal Diversification, Journal of Finance and Investment Analysis, 4(2), 1-19.

Guiso, L., / Haliassos, M. / Japelli, T. (2002): Household Portfolios, MIT Press, Cambridge.

Hedesstrom, T. M. / Svedsater, H. / Garling, T. (2006): Covariation Neglect among Novice Investors, Journal of Experimental Psychology-App., 12(3), 155-165.

Hein, O. / Schwind, M. / Spiwoks, M. (2008): Frankfurt Artificial Stock Market – A Microscopic Stock Market Model with Heterogeneous Interacting Agents in Small-World Communication Networks, Journal of Economic Interaction and Coordination, 3(1), 59-71.

Hein, O. / Schwind, M. / Spiwoks, M. (2012): Network Centrality and Stock Market Volatility: The Impact of Communication Topologies on Prices, Journal of Finance and Investment Analysis, 1(1), 199-232.

Hibbert, A. M. / Lawrence, E. R. / Prakash, A. J. (2012): Can Diversification Be Learned?, Journal of Behavioral Finance, 13(1), 38-50.

Holt, C. A. / Laury, S. K. (2002): Risk Aversion and Incentive Effects, American Economic Review, 92(5), 1644-1655.

Huang, T. C. / Wu, C. C. / Lin, B. H. (2016): Institutional Herding and Risk-Return Relationship, Journal of Business Research, 69(6), 2073-2080.

Huber, J. / Kirchler, M. / Stöckl, T. (2010): The Hot Hand Belief and the Gambler's Fallacy in Investment Decisions under Risk, Theory and Decision, 68(4), 445-462.

Huberman, G. / Sengmueller, P. (2004): Performance and Employer Stock in 401(k) Plans, Review of Finance, 8(3), 403-443.

Inman, J. J. / Zeelenberg, M. (2002): Regret in Repeat Purchase Versus Switching Decisions: The Attenuating Role of Decision Justifiabilit, Journal of Consumer Research, 29(1), 116-128.

Kahneman, D. / Tversky, A. (1982): The Psychology of Preferences, Scientific American, 246(1), 160-173.

Kallir, I. / Sonsino, D. (2009): The Neglect of Correlation in Allocation Decisions, Southern Economic Journal, 75(4), 1045-1066.

Kempf, A. / Ruenzi, S. (2006): Status Quo Bias and the Number of Alternatives: An Empirical Illustration from the Mutual Fund Industry, Journal of Behavioral Finance, 7(4), 204-213.

Keynes, J. M. (1936): The General Theory of Employment, Interest and Money, Macmillan, London.

Kim, W. / Jegadeesh, N. (2010): Do Analysts Herd? An Analysis of Recommendations and Market Reactions, Review of Financial Studies, 23(2), 901-937.

Klemkosky, R. C. (1977): The Impact and Efficiency of Institutional Net Trading Imbalances, Journal of Finance, 32(1), 79-86.

Kraus, A. / Stoll, H. R. (1972): Parallel Trading by Institutional Investors, Journal of Financial Quantitative Analysis, 7(5), 2107-2138.

Kremer, S. / Nautz, D. (2013): Causes and Consequences of Short-Term Institutional Herding, Journal of Banking and Finance, 37(5), 1676-1686.

Lakonishok, J. / Shleifer, A. / Vishny, R. W. (1992): The Impact of Institutional Trading on Stock Prices, Journal of Finance and Economics, 32(1), 23-43.

Lease, R. C. / Lewellen, W. G. / Schlarbaum, G. G. (1974): The Individual Investor: Attributes and Attitudes, Journal of Finance, 29(2), 413-433.

Lejuez, C. W. / Aklina, W. M. / Zvolenskyb, M. J. / Pedullac, C. M. (2003): Evaluation of the Balloon Analogue Risk Task (BART) as a Predictor of Adolescent Real-World Risk-Taking Behaviours, Journal of Adolescence, 26(4), 475-479.

Lin, W. T. / Tsai, S.-C. / Lung, P.-Y. (2013): Investors' Herd Behavior: Rational or Irrational? Asia-Pacific Journal of Financial Studies, 42(5), 755-776.

Lönnqvist, J.-E. / Verkasalo, M. / Walkowitz, G. / Wichardt, P. C. (2015): Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison, Journal of Economical Behavioral Organization, 119, 254-266.

Mackay, C. (1841): Extraordinary Popular Delusions and the Madness of Crowds, Richard Bentley, London.

Markose, S. / Alentorn, A. / Krause, A. (2004): Dynamic Learning, Herding and Guru Effects in Networks, Working Paper, Economics Department, University of Essex.

Markowitz, H. (1952): Portfolio Selection, Journal of Finance, 7(1), 77-91.

Meulbroek, L. (2005): Company Stock in Pension Plans: How Costly Is It? Journal of Law and Economics, 48(2), 443-474.

Morrin, M. / Inman, J. J. / Broniarczyk, S. M. / Nenkov, G. Y. / Reuter, J. (2012): Investing for Retirement: The Moderating Effect of Fund Assortment Size on the 1/N Heuristic, Journal of Marketing Research, 49(4), 537-550.

Nofsinger, J. / Sias, R. (1999): Herding and Feedback Trading by Institutional and Individual Investors, Journal of Finance, 54(6), 2263-2295.

Panchenko, V. / Gerasymchuk, S. / Pavlov, O. V. (2013): Asset Price Dynamics with Local Interactions under Heterogeneous Beliefs, Journal of Economic Dynamics and Control, 37(12), 2623-2642.

Patel, J. / Zeckhauser, R. / Hendricks, D. (1991): The Rationality Struggle: Illustrations from Financial Markets, American Economic Review, 81(2), 232-236.

Polkovnichenko, V. (2005): Household Portfolio Diversification: A Case for Rank-Dependent Preferences, Review of Financial Studies, 18(4), 1467-1502.

Powdthavee, N. / Riyanto, Y. E. (2012): Why Do People Pay for Useless Advice? Implications of Gambler's and Hot-Hand Fallacies in False-Expert Setting, IZA Discussion Paper, 6557.

Rieger, M. O. (2012): Why Do Investors Buy Bad Financial Products? Probability Misestimation and Preferences in Financial Investment Decision, Journal of Behavioral Finance, 13(2), 108-118.

Samuelson, W. / Zeckhauser, R. (1998): Status-Quo Bias in Decision Making, Journal of Risk Uncertainty, 1(1), 7-59.

Scharfstein, D. S. / Stein, J. C. (1990): Herd Behavior and Investment, American Economic Review, 80(3), 465-479.

Sias, R. W. (2004): Institutional Herding, Review of Financial Studies, 17(1), 165-206.

Spiwoks, M. (2004): External Triggered Herding bei Rentenmarkt-Analysten, Finanzmarkt und Portfolio Management, 18(1), 58-83.

Spiwoks, M. / Bizer, K. / Hein, O. (2008): Anchoring Near the Lighthouse: Bond Market Analysts' Behavior Co-ordination by External Signal, European Journal of Economics, Finance and Administrative Sciences, 13, 169-191.

Stöckl, T. / Huber, J. / Kirchler, M. / Lindner, F. (2015): Hot Hand and Gambler's Fallacy in Teams: Evidence from Investment Experiments, Journal of Economic Behavior and Organization, 117, 327-339.

Suetens, S. / Galbo-Joergensen, C. / Tyran, J.-R. (2016): Predicting Lotto Numbers: Experiment on the Gambler's Fallacy and the Hot-Hand Fallacy, Journal of the European Economic Association, 14(3), 584-607.

Sumpter, D. J. T. / Zabzina, N. / Nicolis, S. C. (2012): Six Predictions About the Decision Making of Animal and Human Groups, Managerial and Decision Economics, 33(5-6), 295-309.

Tedeschi, G. / Iori, G. / Gallegati, M. (2009): The Role of Communication and Imitation in Limit Order Markets, European Physical Journal B, 71, 489-497.

Tedeschi, G. / Iori, G. / Gallegati, M. (2012): Herding Effects in Order Driven Markets: The Rise and Fall of Gurus, Journal of Economic Behavior and Organization, 81(1), 82-96.

Tversky, A. / Kahneman, D. (1971): Belief in the Law of Small Numbers, Psychological Bulletin, 76(2), 105-110.

Tversky, A. / Kahneman, D. (1974): Judgment under Uncertainty: Heuristics and Biases, Science, 185, 1124-1131.

Voronkova, S. / Bohl, M. T. (2005): Institutional Traders' Behavior in an Emerging Stock Market: Empirical Evidence on Polish Pension Fund Investors, Journal of Business Finance and Accounting, 32(7-8), 1537-1560.

Walter, A. / Weber, F. M. (2006): Herding in the German Mutual Fund Industry, European Financial Management, 12(3), 375-406.

Weber, E. U. / Siebenmorgen, N. / Weber, M. (2005): Communication Asset Risk: How Name Recognition and the Format of Historic Volatility Information Affect Risk Perception and Investment Decisions, Risk Analysis, 25(3), 597-609.

Wermers, R. (1999): Mutual Fund Herding and the Impact on Stock Prices, Journal of Finance, 54, 581-622.

Yen, H. R. / Chuang, S. C. (2008): The Effect of Incidental Affect on Preference for the Status Quo, Journal of the Academy of Marketing Science, 36(4), 522-537.

Zeelenberg, M. / van den Bos, K. / van Dijk, E. / Pieters, R. (2002): The Inaction Effect in the Psychology of Regret, Journal of Personality and Social Psychology, 82(3), 314-327.

Appendix 1: Instructions, Control Questions, Screenshot, Assessment of Risk Attitude

Instructions (Treatment 1 and Treatment 3)

The Game

By investing in stocks, you can benefit from the income return. There are two securities (Stock A and Stock B) to choose from. In each period of the game, the value of Stock A is changed by +0 experimental currency units (ECU) or +7 experimental currency units (ECU); both events occur with a probability of 50%. The same is true for Stock B. Both securities have an expected value of +3.50 ECU per period. The performances of both stocks are independent random events. 15 periods will be played in total and in each period, you have 4 units at your disposal that must be invested. The 4 units can be invested in the following portfolio combinations:

Portfolio 1: AAAA

Portfolio 2: AAAB

Portfolio 3: AABB

Portfolio 4: ABBB

Portfolio 5: BBBB

The game starts at period 0. At period 0, you have the possibility to restructure the portfolio that was randomly assigned to you. The performance of the stocks for the single periods were previously determined by 15 x two coin tosses (one toss for Stock A and one toss for Stock B). “Heads” signifies a good period (+7 ECU) and “tails” signifies a weak period (+0 ECU). You can earn up to 420 ECU in the 15 periods.

For maximum transparency, you will be shown your results and all subjects will be ranked after each period. The ranking is established according to the total earnings. You will therefore be able to compare the performance of your portfolio to the performance of the other subjects’ portfolios. You can also earn up to €3.85 in a lottery, detailed information on which you will receive in due course.

The Payout

In the 15 periods, you can earn up to 420 ECU with the securities. 1 ECU equals €0.10. The maximum payout is €42.00 (420 x 0.10). Up to €3.85 from the lottery are added to this sum. In total, you can earn up to €45.85. You will receive your money at the end of the experiment.

Please note:

- Please keep quiet during the experiment!
- Please do not look at your seatmate's monitor!
- Auxiliary devices (calculators, smartphones etc.) are not allowed. All electronic devices must be switched off!

Instructions (Treatment 2)

The Game

By investing in stocks, you can benefit from the income return. There are two securities (Stock A and Stock B) to choose from. In each period of the game, the value of Stock A is changed by +0 experimental currency units (ECU) or +7 experimental currency units (ECU); both events occur with a probability of 50%. The same is true for Stock B. Both securities have an expected value of +3.50 ECU per period. The performances of both stocks are independent random events. 15 periods will be played in total and in each period, you have 4 units at your disposal that must be invested. The 4 units can be invested in the following portfolio combinations:

Portfolio 1: AAAA

Portfolio 2: AAAB

Portfolio 3: AABB

Portfolio 4: ABBB

Portfolio 5: BBBB

The game starts at period 0. At period 0, you have the possibility to restructure the portfolio that was randomly assigned to you. The performance of the stocks for the single periods were previously determined by 15 x two coin tosses (one toss for Stock A and one toss for Stock B). "Heads" signifies a good period (+7 ECU) and "tails" signifies a weak period (+0 ECU). You can earn up to 420 ECU in the 15 periods. For maximum transparency, you will be shown your results after each period. You can also earn up to €3.85 in a lottery, detailed information on which you will receive in due course.

The Payout

In the 15 periods, you can earn up to 420 ECU with the securities. 1 ECU equals €0.10. The maximum payout is €42.00 (420 x 0.10). Up to €3.85 from the lottery are added to this sum. In total, you can earn up to €45.85. You will receive your money at the end of the experiment.

Please Note:

- Please keep quiet during the experiment!
- Please do not look at your seatmate's monitor!
- Auxiliary devices (calculators, smartphones etc.) are not allowed. All electronic devices must be switched off!

Control Questions

Control questions (tick the box):

1. What is your task in this game?

   ʘ Solving mathematical problems.

   ʘ Investing in stocks and taking part in a lottery. (correct)

   ʘ Giving economic forecasts.

2. How many different securities are there to choose from and how many free stocks do you receive?

   ʘ There are 4 different securities to choose from and I receive 2 free stocks.

   ʘ There are 2 different securities to choose from and I receive 2 free stocks.

   ʘ There are 2 different securities to choose from and I receive 4 free stocks. (correct)

3. On what does the payout depend in the 15 periods?

   ʘ On the dividend payouts.

   ʘ On the performance of the stocks. (correct)

   ʘ On the DAX market trend.

4. How many different combinations of the portfolio are possible?

   ʘ 2

   ʘ 4

   ʘ 5 (correct)

Screenshot Treatment 1 (recreated to increase readability)

You are subject 10. Your result in period 1:

The performance of Stock A in Period 1 is	The performance of Stock B in Period 1 is	Your payout for Period 1 is	Your total payout so far amounts to
±0 ECU	+7 ECU	14 ECU	14 ECU

Previous results of all subjects:

Ranking according to Previous Success	Subjects	Pay-out so far	Last Portfolio Composition
1	3	28 ECU	BBBB
1	5	28 ECU	BBBB
3	4	21 ECU	ABBB
4	7	14 ECU	AABB
4	9	14 ECU	AABB
4	10	14 ECU	AABB
7	1	7 ECU	AAAB
7	6	7 ECU	AAAB
9	2	0 ECU	AAAA
9	8	0 ECU	AAAA

There are still two stocks to choose between (Stock A and Stock B). The value developments for both stocks are independent random processes with the possible results +7 ECU (good period development) and ±0 ECU (unprofitable period development).

You receive 4 stocks. You can choose whether to receive 4 Stocks A, 3 Stocks A and one Stock B, 2 Stocks A and 2 Stocks B, one Stock A and three Stocks B or 4 Stocks B. The value developments of the stocks in period 2 were previously determined by two coin tosses (one toss for Stock A and one toss for Stock B). "Heads" means a good development in the period whereas tails stands for a negative period development (±0 ECU).

Do you want to reallocate your portfolio? If not, please click on "O.K."! If so, please choose your preferred portfolio and then click on "O.K."!

I choose:
ʘ 4 Stocks A
ʘ 3 Stocks A + 1 Stock B
ʘ 2 Stocks A + 2 Stocks B
ʘ 1 Stock A + 3 Stocks B
ʘ 4 Stocks B

O.K.

Instructions to Determine the Risk Preference

Each decision is a choice between "version A" and "version B". Each version is comparable to a lottery with different payouts and different probabilities of occurrence.

You have 10 decisions to make and enter. One of these decisions will be considered to determine your payout from the lottery as follows: After you entered all your decisions, a ten-sided dice is thrown to select one of the 10 decisions. Each of the decisions therefore has a 10% probability of being chosen. The selected lottery (A or B) is then played. The probability of occurrence is simulated by an urn containing table tennis balls: In an urn with 10 table tennis balls the number of orange balls determines the probability of the higher payout.

Example for decision no. 8: In an urn with 10 table tennis balls are 8 orange-colored and 2 white balls. The probability that a randomly picked ball is orange is therefore 80%. If the picked ball is orange, you will receive €2.00 in version A and €3.85 in version B. If the picked ball is white, you will receive €1.60 in version A and €0.10 in version B. So: You make 10 decisions (either A or B), one of them is randomly selected (with a dice) and played (with an urn and 10 table tennis balls) – the result will determine your payout from the lottery. Please answer the following control questions about the lottery before making any decisions.

Risk level	Version A:				Version B:				Your Decision A or B
	p(€2.00)	Payout	p(€1.60)	Payout	p(€3.85)	Payout	p(€0.10)	Payout	
1	10%	€2.00	90%	€1.60	10%	€3.85	90%	€0.10	
2	20%	€2.00	80%	€1.60	20%	€3.85	80%	€0.10	
3	30%	€2.00	70%	€1.60	30%	€3.85	70%	€0.10	
4	40%	€2.00	60%	€1.60	40%	€3.85	60%	€0.10	
5	50%	€2.00	50%	€1.60	50%	€3.85	50%	€0.10	
6	60%	€2.00	40%	€1.60	60%	€3.85	40%	€0.10	
7	70%	€2.00	30%	€1.60	70%	€3.85	30%	€0.10	
8	80%	€2.00	20%	€1.60	80%	€3.85	20%	€0.10	
9	90%	€2.00	10%	€1.60	90%	€3.85	10%	€0.10	
10	100%	€2.00	0%	€1.60	100%	€3.85	0%	€0.10	

Control Questions to Determine the Risk Preference

Control questions (tick the box):

1. What is the minimum and the maximum payout in the lottery?

   ʘ The minimum payout is €0.00 and the maximum payout is €1.60.

   ʘ The minimum payout is €0.10 and the maximum payout is €3.85. (correct)

   ʘ The minimum payout is €0.10 and the maximum payout is €1.60.

2. If the dice selects the 7th decision and you choose version A and have drawn a white table tennis ball from the urn, what is your payout?

ʘ €0.00

ʘ €2.00

ʘ €1.60 (correct)

3. How many white table tennis balls are in the urn if the dice chooses the tenth decision?

ʘ 10

ʘ 0 (correct)

ʘ 5

4. How many orange table tennis balls are in the urn if the dice chooses the fourth decision?

ʘ 6

ʘ 0

ʘ 4 (correct)

Appendix 2: Results of Treatment 1 per Session

Table A-1: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 1)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distrib.	25%	25%	0%	25%	25%	-	-
Period 1	20.0%	0.0%	60.0%	20.0%	0.0%	-	-
Period 2	10.0%	20.0%	70.0%	0.0%	0.0%	AABB	BBBA
Period 3	20.0%	10.0%	30.0%	10.0%	30.0%	AABB	AABB
Period 4	20.0%	10.0%	40.0%	20.0%	10.0%	AABB, BBBB	BBBB
Period 5	0.0%	10.0%	50.0%	10.0%	30.0%	AABB	AAAA
Period 6	0.0%	10.0%	60.0%	20.0%	10.0%	AABB	AAAB
Period 7	0.0%	10.0%	80.0%	0.0%	10.0%	AABB	AABB
Period 8	10.0%	20.0%	40.0%	10.0%	20.0%	AABB	AABB
Period 9	10.0%	0.0%	40.0%	10.0%	40.0%	AABB	BBBA
Period 10	10.0%	10.0%	30.0%	20.0%	30.0%	AABB, BBBB	BBBB, AABB
Period 11	20.0%	10.0%	50.0%	10.0%	10.0%	AABB, BBBB	BBBB
Period 12	20.0%	0.0%	50.0%	20.0%	10.0%	AABB	AAAA
Period 13	20.0%	10.0%	40.0%	20.0%	10.0%	AABB	BBBB
Period 14	30.0%	0.0%	40.0%	30.0%	0.0%	AABB	AAAA
Period 15	30.0%	10.0%	50.0%	10.0%	0.0%	AABB	AAAA

In Session 1 of Treatment 1, 60% of the subjects have already settled on portfolio AABB. This number increases to 70% in Period 2. Subject to some variations, it increases to 80% in Period 7. In Period 10, however, it recedes to 30%.

Table A-2: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 2)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distribution	25%	25%	0%	25%	25%	-	-
Period 1	0.0%	22.2%	44.4%	33.3%	0.0%	-	-
Period 2	0.0%	0.0%	66.7%	33.3%	0.0%	AABB	ABBB
Period 3	0.0%	11.1%	55.6%	33.3%	0.0%	AABB	AABB, ABBB
Period 4	11.1%	11.1%	55.6%	22.2%	0.0%	AABB	ABBB
Period 5	0.0%	11.1%	55.6%	22.2%	11.1%	AABB	ABBB
Period 6	11.1%	0.0%	44.4%	44.4%	0.0%	AABB	AABB, ABBB
Period 7	11.1%	0.0%	55.6%	33.3%	0.0%	AABB, ABBB	AABB, ABBB
Period 8	11.1%	22.2%	44.4%	11.1%	11.1%	AABB	ABBB
Period 9	11.1%	22.2%	33.3%	22.2%	11.1%	AABB	AAAB
Period 10	11.1%	11.1%	44.4%	22.2%	11.1%	AABB	AABB
Period 11	22.2%	22.2%	44.4%	11.1%	0.0%	AABB	BBBB, ABBB
Period 12	0.0%	11.1%	44.4%	33.3%	11.1%	AABB	AAAA
Period 13	11.1%	11.1%	11.1%	55.6%	11.1%	AABB	BBBB
Period 14	22.2%	44.4%	22.2%	0.0%	11.1%	ABBB	BBBB, ABBB
Period 15	33.3%	33.3%	22.2%	11.1%	0.0%	AAAB	BBBB, AAAB

In Session 2 of Treatment 1 we again observe a trend toward the stock combination AABB. In the first period, 44% of the subjects chose this portfolio structure, and 67% do so in the second period. Subsequently, however, these decisions recede to 11% in Period 13, only to end at 22% in the last period.

Table A-3: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 3)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distribution	25%	25%	0%	25%	25%	-	-
Period 1	11.1%	33.3%	33.3%	11.1%	11.1%	-	-
Period 2	11.1%	0.0%	77.8%	11.1%	0.0%	AAAB, AABB	BBBB
Period 3	11.1%	44.4%	44.4%	0.0%	0.0%	AABB	AAAA
Period 4	0.0%	22.2%	66.7%	0.0%	11.1%	AAAB, AABB	AAAA, AAAB, AABB
Period 5	0.0%	11.1%	44.4%	22.2%	22.2%	AABB	AAAB
Period 6	0.0%	0.0%	33.3%	22.2%	44.4%	AABB	BBBB
Period 7	0.0%	22.2%	55.6%	0.0%	22.2%	BBBB	BBBB
Period 8	11.1%	0.0%	55.6%	11.1%	22.2%	AABB	AAAB, AABB, BBBB
Period 9	11.1%	11.1%	22.2%	44.4%	11.1%	AABB	AAAA
Period 10	0.0%	0.0%	33.3%	44.4%	22.2%	ABBB	BBBB
Period 11	33.3%	0.0%	33.3%	22.2%	11.1%	ABBB	BBBB
Period 12	11.1%	11.1%	66.7%	0.0%	11.1%	AABB	AAAA, BBBB
Period 13	0.0%	0.0%	55.6%	22.2%	22.2%	AABB	AAAA, BBBB
Period 14	11.1%	22.2%	55.6%	0.0%	11.1%	ABBB	BBBB
Period 15	11.1%	33.3%	44.4%	0.0%	11.1%	AAAB	AAAA, AAAB

In Session 3 of Treatment 1, an increasing number of subjects choose portfolio BBBB. None of the subjects in Period 3 decides for this combination. In Period 4, 11% make this decision, and as much as 22% do so in Period 5, and even 44% in Period 6. This development, however, is not continued but collapses rather quickly.

Table A-4: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 4)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distribution	25%	25%	0%	25%	25%	-	-
Period 1	0.0%	30.0%	40.0%	30.0%	0.0%	-	-
Period 2	10.0%	30.0%	50.0%	10.0%	0.0%	AABB	ABBB
Period 3	0.0%	20.0%	50.0%	30.0%	0.0%	AABB	AAAA, AABB, ABBB
Period 4	10.0%	40.0%	40.0%	10.0%	0.0%	AABB	ABBB
Period 5	0.0%	40.0%	30.0%	20.0%	10.0%	AAAB, AABB	AAAB
Period 6	0.0%	30.0%	40.0%	20.0%	10.0%	AAAB	AAAB
Period 7	10.0%	50.0%	20.0%	10.0%	10.0%	AABB	AAAB
Period 8	10.0%	20.0%	55.0%	20.0%	0.0%	AAAB	BBBB
Period 9	0.0%	30.0%	30.0%	30.0%	10.0%	AABB	AAAA
Period 10	0.0%	10.0%	40.0%	40.0%	10.0%	AAAB, AABB, ABBB	AABB
Period 11	10.0%	40.0%	40.0%	10.0%	0.0%	AABB, ABBB	AABB
Period 12	10.0%	10.0%	20.0%	40.0%	20.0%	AAAB, AABB	AAAA
Period 13	0.0%	50.0%	40.0%	0.0%	10.0%	ABBB	BBBB
Period 14	30.0%	20.0%	40.0%	0.0%	10.0%	AAAB	AABB
Period 15	30.0%	20.0%	40.0%	0.0%	10.0%	AABB	AAAA

In Session 4 of Treatment 1, similar results are achieved. The portfolio structure ABBB is increasingly considered by the subjects. In Period 7, this portfolio is selected by 10% of the subjects, and by 20%, 30%, and 40% percent in the subsequent periods. This development stops abruptly after that. In Periods 13 to 15, this combination is no longer chosen.

Table A-5: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 5)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distribution	25%	25%	0%	25%	25%	-	-
Period 1	0.0%	14.3%	71.4%	14.3%	0.0%	-	-
Period 2	0.0%	28.6%	57.1%	14.3%	0.0%	AABB	ABBB
Period 3	0.0%	28.6%	42.9%	28.6%	0.0%	AABB	AAAB
Period 4	0.0%	28.6%	42.9%	14.3%	14.3%	AABB	ABBB
Period 5	0.0%	71.4%	14.3%	14.3%	0.0%	AABB	AAAB
Period 6	0.0%	14.3%	57.1%	28.6%	0.0%	AAAB	AAAB
Period 7	0.0%	0.0%	57.1%	28.6%	14.3%	AABB	ABBB
Period 8	0.0%	14.3%	28.6%	57.1%	0.0%	AABB	ABBB
Period 9	0.0%	14.3%	42.9%	42.9%	0.0%	ABBB	AAAB
Period 10	0.0%	14.3%	28.6%	57.1%	0.0%	AABB, ABBB	AABB
Period 11	14.3%	42.9%	28.6%	0.0%	14.3%	ABBB	AAAB, ABBB
Period 12	14.3%	28.6%	0.0%	57.1%	0.0%	AAAB	AAAB
Period 13	0.0%	14.3%	42.9%	42.9%	0.0%	ABBB	AAAB, ABBB
Period 14	0.0%	14.3%	28.6%	28.6%	28.6%	AABB, ABBB	ABBB
Period 15	0.0%	42.9%	28.6%	14.3%	14.3%	AABB, ABBB, BBBB	AAAB, ABBB

In Session 5 of Treatment 1, the combination ABBB is again popular for some time. The number of subjects deciding for this portfolio increases from Period 5 (14%), over Periods 6 and 7 (29%) to Period 8 with 57%. Subject to some variations, this choice recedes to 14% in Period 15.

Table A-6: Percentage Distribution of the Portfolios in the Game (Treatment 1, Session 6)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios	Most Frequently Chosen Portfolio(s) in the Previous Period	Portfolio(s) of the Leader(s) in the Previous Period
Initial Distribution	25%	25%	0%	25%	25%	-	-
Period 1	12.5%	12.5%	25.0%	37.5%	12.5%	-	-
Period 2	25.0%	25.0%	12.5%	25.0%	12.5%	AABB	BBBB
Period 3	25.0%	37.5%	37.5%	0.0%	0.0%	AAAA, AAAB, ABBB	ABBB
Period 4	25.0%	62.5%	12.5%	0.0%	0.0%	AAAB, AABB	AABB
Period 5	25.0%	25.0%	25.0%	25.0%	0.0%	AAAB	AAAB
Period 6	25.0%	12.5%	12.5%	37.5%	12.5%	AAAA, AAAB, AABB, ABBB	AAAA
Period 7	12.5%	0.0%	12.5%	37.5%	37.5%	ABBB	ABBB
Period 8	37.5%	0.0%	25.0%	37.5%	0.0%	ABBB, BBBB	BBBB
Period 9	12.5%	25.0%	25.0%	12.5%	25.0%	AAAA, ABBB	AABB
Period 10	25.0%	12.5%	50.0%	0.0%	12.5%	AAAB, AABB, BBBB	BBBB
Period 11	37.5%	25.0%	25.0%	0.0%	12.5%	AABB	BBBB
Period 12	25.0%	0.0%	12.5%	25.0%	37.5%	AAAA	AAAA
Period 13	37.5%	25.0%	0.0%	12.5%	25.0%	BBBB	BBBB
Period 14	37.5%	37.5%	12.5%	0.0%	12.5%	AAAA	BBBB
Period 15	37.5%	25.0%	0.0%	12.5%	25.0%	AAAA, AAAB	AAAA

In Session 6 of Treatment 1, no herding can be established for portfolio AAAB. In Period 1, this combination is chosen by 13% of the subjects, by 25% in Period 2, by 38% in Period 3, and by as much as 63% in Period 4. In the following period, this development suddenly stops. As early as Period 7, this portfolio is no longer chosen by any of the subjects.

Appendix 3: Percentage Distribution of the Portfolios in the Game

Table A-7: Percentage Distribution of the Portfolios in the Game (Treatment 1)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios
Initial Distribution	25%	25%	0%	25%	25%
Period 1	7.5%	18.9%	45.3%	24.5%	3.8%
Period 2	9.4%	17.0%	56.6%	15.1%	1.9%
Period 3	9.4%	24.5%	43.4%	17.0%	5.7%
Period 4	11.3%	28.3%	43.4%	11.3%	5.7%
Period 5	3.8%	26.4%	37.7%	18.9%	13.2%
Period 6	5.7%	11.3%	41.5%	28.3%	13.2%
Period 7	5.7%	15.1%	47.2%	17.0%	15.1%
Period 8	13.2%	13.2%	41.5%	22.6%	9.4%
Period 9	7.5%	17.0%	32.1%	26.4%	17.0%
Period 10	7.5%	9.4%	37.7%	30.2%	15.1%
Period 11	22.6%	22.6%	37.7%	9.4%	7.5%
Period 12	13.2%	9.4%	34.0%	28.3%	15.1%
Period 13	11.3%	18.9%	32.1%	24.5%	13.2%
Period 14	22.6%	22.6%	34.0%	9.4%	11.3%
Period 15	24.5%	26.4%	32.1%	7.5%	9.4%

Table A-8: Percentage Distribution of the Portfolios in the Game (Treatment 2)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios
Initial Distri-bution	25%	25%	0%	25%	25%
Period 1	2.2%	15.2%	52.2%	21.7%	8.7%
Period 2	8.7%	30.4%	17.4%	34.8%	8.7%
Period 3	4.3%	28.3%	32.6%	21.7%	13.0%
Period 4	10.9%	34.8%	15.2%	19.6%	19.6%
Period 5	8.7%	13.0%	50.0%	13.0%	15.2%
Period 6	15.2%	13.0%	39.1%	17.4%	15.2%
Period 7	13.0%	19.6%	30.4%	28.3%	8.7%
Period 8	4.3%	39.1%	28.3%	10.9%	17.4%
Period 9	8.7%	21.7%	21.7%	21.7%	26.1%
Period 10	4.3%	13.0%	37.0%	26.1%	19.6%
Period 11	15.2%	26.1%	26.1%	17.4%	15.2%
Period 12	0.0%	10.9%	39.1%	32.6%	17.4%
Period 13	4.3%	19.6%	28.3%	32.6%	15.2%
Period 14	15.2%	30.4%	26.1%	21.7%	6.5%
Period 15	15.2%	21.7%	34.8%	17.4%	10.9%

Table A-9: Percentage Distribution of the Portfolios in the Game (Treatment 3)

Period	Percentage of AAAA Portfolios	Percentage of AAAB Portfolios	Percentage of AABB Portfolios	Percentage of ABBB Portfolios	Percentage of BBBB Portfolios
Initial Distribution	0%	0%	100%	0%	0%
Period 1	5.9%	19.6%	60.8%	7.8%	5.9%
Period 2	5.9%	29.4%	43.1%	11.8%	9.8%
Period 3	9.8%	17.6%	43.1%	21.6%	7.8%
Period 4	11.8%	25.5%	21.6%	23.5%	17.6%
Period 5	5.9%	13.7%	51.0%	15.7%	13.7%
Period 6	9.8%	25.5%	37.3%	17.6%	9.8%
Period 7	5.9%	25.5%	41.2%	19.6%	7.8%
Period 8	7.8%	35.3%	33.3%	17.6%	5.9%
Period 9	0.0%	13.7%	47.1%	33.3%	5.9%
Period 10	3.9%	9.8%	47.1%	27.5%	11.8%
Period 11	17.6%	15.7%	41.2%	17.6%	7.8%
Period 12	3.9%	21.6%	39.2%	21.6%	13.7%
Period 13	5.9%	23.5%	27.5%	23.5%	19.6%
Period 14	13.7%	25.5%	39.2%	13.7%	7.8%
Period 15	17.6%	23.5%	41.2%	2.0%	15.7%

Appendix 4: Further Results

Table A-10: Manifestation of the Gambler's Fallacy in the Portfolio Choices of All Treatments

1	2	3	4	5	6	7	8	9	10	11
Period	Random Event of the Previous Period for Stock A	Random Event of the Previous Period for Stock B	Expected Preference with Gambler's Fallacy	Percentage of AAAA Portfolios (W)	Percentage of AAAB Portfolios (X)	Percentage of AABB Portfolios	Percentage of ABBB Portfolios (Y)	Percentage of BBBB Portfolios (Z)	Surplus (W+X) – (Y+Z)	Answers the Expectation
2	±0	+7	A	8.00%	25.33%	40.00%	20.00%	6.67%	6.67%	Yes
3	+7	+7	A	8.00%	23.33%	40.00%	20.00%	8.67%	2.67%	Yes
4	±0	+7	A	11.33%	29.33%	27.33%	18.00%	14.00%	8.67%	Yes
5	+7	±0	B	6.00%	18.00%	46.00%	16.00%	14.00%	-6.00%	Yes
6	+7	+7	B	10.00%	16.67%	39.33%	21.33%	12.67%	-7.33%	Yes
7	±0	±0	B	8.00%	20.00%	40.00%	21.33%	10.67%	-4.00%	Yes
8	±0	+7	A	8.67%	28.67%	34.67%	17.33%	10.67%	9.33%	Yes
9	+7	±0	B	5.33%	17.33%	34.00%	27.33%	16.00%	-20.67%	Yes
10	±0	±0	B	5.33%	10.67%	40.67%	28.00%	15.33%	-27.33%	Yes
11	±0	+7	A	18.67%	21.33%	35.33%	14.67%	10.00%	15.33%	Yes
12	+7	±0	B	6.00%	14.00%	37.33%	27.33%	15.33%	-22.67%	Yes
13	+7	+7	B	7.33%	20.67%	29.33%	26.67%	16.00%	-14.67%	Yes
14	±0	+7	A	17.33%	26.00%	33.33%	14.67%	8.67%	20.00%	Yes
15	±0	±0	A	19.33%	24.00%	36.00%	8.67%	12.00%	22.67%	Yes

Table A-10 shows that the gambler's fallacy did indeed consistently influence the subjects' portfolio choices. The first column lists those periods of the game in which the gambler's fallacy could take effect. No random events existed before Period 1, which is why the gambler's fallacy could not take effect before the portfolio choice in Period 2. The second and third columns of Table A-10 list the random events of the previous periods. The fourth column shows which stock is preferred by the subjects who fall prey to the gambler's fallacy. In Period 1, for instance, Stock A has an income return of ±0 ECU and Stock B has an income return of +7 ECU. This is the development that the subjects can see before Period 2. This results in a preference for Stock A. Before the start of Period 3, the "errant gambler" again prefers Stock A because, in the previous period, both stocks had a positive income return (+7 ECU) but for Stock A, it is the first in succession, while it is the second in succession for Stock B. Before Period 4, the events of Period 2 are repeated and result in yet another preference for Stock A. Period by period, considerations such as these lead to preferences that are listed in the fourth column of Table A-10.

Columns 5, 6, 7, 8, and 9 of Table A-10 show the percentages of the five possible portfolio compilations (AAAA, AAAB, AABB, ABBB and BBBB) for all three treatments (Treatment 1, Treatment 2, and Treatment 3; for detailed results, see Tables A-7, A-8, and A-9 in Appendix

3). Column 10 of Table A-10 displays an unweighted spread, calculated by subtracting the portfolio decisions that chose Stock B (Columns 8 (Y) and 9 (Z)) from those that prefer Stock A (Columns 5 (W) and 6 (X)). We calculate: (W + X) – (Y + Z). In Period 2 this means: (8.00 + 25.33) – (20.00 + 6.67) = 6.67.

If the subtraction results in a positive value, it can be concluded that the subjects preferred Stock A to Stock B when compiling their portfolios. If the subtraction results in a negative value, it can be concluded that the subjects preferred Stock B to Stock A when assembling their portfolios. Consequently, we can expect that a preference for Stock A results in a positive balance and that a preference for Stock B results in a negative balance. In Column 11 of Table A-10, we analyze which periods match our expectations, and we can see that this occurs without exception. When in Column 4 the Stock A (Stock B) is listed, then the balance in Column 10 is positive (negative).

Considering that the subjects fall prey to the gambler's fallacy and accordingly develop preferences for Stock A or Stock B (Column 4 of Table A-10), then we can find a suitable explanation for the surplus of Stock A (positive indication in Column 10) or, respectively, for the surplus of Stock B (negative indication in Column 10) in all periods.

Table A-11: Variances and Standard Deviations of the Five Portfolios Considering the Actual Events for Stock A and Stock B in Experimental Currency Units (ECU)

Period	Events		Performance of Portfolios				
	Stock A	Stock B	AAAA	AAAB	AABB	ABBB	BBBB
1	±0	+7	±0	+7	+14	+21	+28
2	+7	+7	+28	+28	+28	+28	+28
3	±0	+7	±0	+7	+14	+21	+28
4	+7	±0	+28	+21	+14	+7	±0
5	+7	+7	+28	+28	+28	+28	+28
6	±0	±0	±0	±0	±0	±0	±0
7	±0	+7	±0	+7	+14	+21	+28
8	+7	±0	+28	+21	+14	+7	±0
9	±0	±0	±0	±0	±0	±0	±0
10	±0	+7	±0	+7	+14	+21	+28
11	+7	±0	+28	+21	+14	+7	±0
12	+7	+7	+28	+28	+28	+28	+28
13	±0	+7	±0	+7	+14	+21	+28
14	±0	±0	±0	±0	±0	±0	±0
15	+7	±0	+28	+21	+14	+7	±0
		Variance	209.07	115.27	84.00	115.27	209.07
		Stand. Dev.	14.46	10.74	9.17	10.74	14.46

Chapter 3

EMOTIONS AND EXPOSURE TO RISK: THE INFLUENCE OF POSITIVE AND NEGATIVE EMOTIONS ON PORTFOLIO DECISIONS

Submitted to:
Journal of Economic Behavior and Organization

Published in:
Wolfsburg Working Papers, 18-03, July 2018.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-3, June 2018.

Abstract

This experimental study addresses the question of whether positive and negative emotions have an influence on diversification behavior, and it reveals that only a small part of subjects take rational decisions and always choose the optimal portfolio. In addition, the study shows that the mood of subjects has an influence on their portfolio decisions and thus also on their exposure to risk. The average risk of the portfolio – measured against the standard deviation of the returns – is lower in the treatment entitled 'neutral' than in the treatments entitled 'positive' and 'negative'.

Keywords

positive affect; negative affect; mood; emotions; risk exposure; laboratory experiment; portfolio choice; investment decisions; correlation neglect; information processing; investor rationality

JEL classification

C91, D81, G11, G41

1 Introduction

The danger of having a portfolio of securities which is not optimally diversified was shown in September 2015, for example: Due to the Volkswagen Group's Dieselgate scandal, German car shares suffered considerable losses within a period of only a few days. Those who had a high proportion of automobile industry shares in their portfolio rapidly lost up to a third of the value of their portfolio. Markowitz (1952) showed that for risk-averse subjects it makes sense to hold diversified securities portfolios. Nevertheless, in reality many subjects hold insufficiently diversified securities portfolios (see, for example, Ackert et al., 2015; Anderson, 2013; Hibbert et al., 2012; Ackert et al., 2011; Goetzmann and Kumar, 2008; Meulbroek, 2005; Polkovnichenko, 2005; Huberman and Sengmueller, 2004; Agnew et al., 2003; Poterba 2003; Mitchell and Utkus 2002; Guiso et al., 2002; Benartzi, 2001; Benartzi and Thaler, 2001; Barber and Odean, 2000; De Bondt, 1998; Kelly, 1995; Bode et al., 1994; French and Poterba, 1991; Blume and Friend, 1975; Lease et al., 1974).

There are many possible reasons for securities portfolios not being optimally diversified. Experimental economic research has already presented findings on this subject: The correlations between investment alternatives are frequently neglected by investors. Considerable empirical evidence already exists for this phenomenon of correlation neglect (see, for example, Gubaydullina and Spiwoks, 2015; Eyster and Weizsäcker, 2010; Hedesstrom et al., 2006; Benartzi and Thaler, 2001). Many investors distribute their assets evenly across all of the investment alternatives available. This phenomenon, which is known as the 1/N heuristic, is a special form of correlation neglect. In the meantime, there are also significant empirical findings on this issue (see, for example, Fernandes 2013; Morrin et al., 2012; Baltussen and Post, 2011; Hedesstrom et al., 2006, and Benartzi and Thaler, 2001). Many investors allow themselves to be misled by irrelevant information, or attach too much importance to certain information (cf. Gubaydullina and Spiwoks, 2015; Kallir and Sonsino, 2009; Goetzmann and Kumar, 2008). Investment decisions are frequently skewed by an inappropriately strong focus on domestic financial instruments. This phenomenon is known as home bias, and has also been proven empirically (cf. Weber et al., 2005; Poterba, 2003; Mitchell and Utkus, 2002). Many investors also tend to see patterns where in reality there are none. This often leads to random processes being dealt with inappropriately. In this way, the so-called gambler's error can impede optimal decisions on diversification (see, for example, Filiz et al., 2018; Stöckl et al., 2015; Huber et al., 2010).

The influence of emotions on decision-making is now well-established in the literature (for an overview see, for example, George and Dane, 2016; Lerner et al., 2015; Vohs et al., 2007; Baker and Wurgler, 2007; Baumeister et al., 2007; Pham, 2007; Shiv et al., 2005; Nofsinger, 2005; Lucey and Dowling, 2005; Daniel et al., 2002; Hirshleifer, 2001; Loewenstein et al., 2001; Isen, 2000; Loewenstein, 2000; Schwarz, 2000; Elster, 1998; Bless et al., 1996; Elster, 1996; Johnson and Tversky, 1983).

In recent decades, the effects of sunshine, rain, cloud cover, wind strength, storms and other meteorological factors on market returns at share exchanges worldwide have been thoroughly investigated (Kim, 2017; Kaustia and Rantapuska, 2016; Apergis et al., 2016; Bassi et al., 2013; Lu and Chou, 2012; Mirza et al., 2012; Floros, 2011; Symeonidis et al., 2010; Kang et al., 2010; Shu and Hung, 2009; Chang et al., 2008; Keef and Roush, 2007; Chang et al., 2006; Dowling and Lucey, 2005; Cao and Wei, 2005; Tufan and Hamarat, 2004; Krivelyova and Robotti, 2003; Hirshleifer and Shumway, 2003; Kamstra et al., 2003; Pardo and Enric, 2002; Krämer and Runde, 1997; Saunders, 1993). While doing so, attempts were also made to create a connection between the weather and the mood of capital market protagonists. Hirshleifer and Shumway (2003) showed that share market returns on days when the sun shined in the morning were higher on average than on days with bad weather. This result was explained by sunshine favoring a positive atmosphere among investors. Kamstra et al. (2003) established that share market returns varied according to the length of the day, which has been interpreted in a similar way to the results Hirshleifer and Shumway (2003). Kaustia and Rantapuska (2016) carried out a similar study – however, they only observed a weak connection between the effect of the length of a respective day and investment decisions.

Experimental economic research is increasingly interested in the question of which influence positive and negative emotions have on investment decisions. Grable and Roszkowski (2008), for example, showed in an experimental study that subjects whose positive emotions predominate were willing to take greater financial risks. Kuhnen and Knutson (2011) carried out experiments to establish how different moods affected investment decisions. This revealed that subjects with predominantly negative emotions tend to choose low-risk investments. Subjects with predominantly positive emotions, on the other hand, tended to favour riskier investments. Subjects whose emotions are positive are more optimistic in relation to their investment decisions. Kaplanski et al. (2015) showed that the mood of investors had an influence on their expectations in terms of returns, and on their perception of risk. The happier the subjects were, the greater their expectations were of their returns, and the lower their presumed risk is of stock market investments. Experiments carried out by Lee and Andrade (2014) showed that negative affects promote risk aversion in investment decisions. Lahav and Meer (2012) as well as Andrade et al. (2016) used experiments to examine the effect of emotions on speculative bubbles, whereby they established that speculative bubbles were larger in the case of positive affects than with negative affects. Breaban and Noussair (2018) followed a similar approach, though their findings were not as clear-cut as those of Lahav and Meer (2012) or of Andrade et al. (2016).

As one can see, there are a range of findings showing that the mood of investors can influence their investment decisions. However, as yet there have been no studies on whether the mood of investors also has an effect on their diversification behavior and thus on the exposure to risk in differently composed portfolios. This research topic has, however, now been addressed by this study.

Chapter 2 deals with the design of the experiment. In Chapter 3, hypotheses are elaborated, and in Chapter 4 the results are presented and analyzed. In Chapter 5 the most important results of the investigation are summarized.

2 Experimental Design

2.1 Diversification Decisions

Markowitz (1952) proceeds from a very simple starting point: A choice between two risky securities. The first security (x_1) has a comparatively low expected return (e_1) and a comparatively low-risk exposure (s_1). The second security (x_2) has a higher expected return (e_2) and a higher risk exposure (s_2). As long as the two securities x_1 and x_2 are not fully positively correlated in terms of the level of their returns, the returns of the portfolios develop in a proportional way while their risk exposure is disproportionately low. When viewing expected returns and risk exposure simultaneously, the so-called efficient frontier emerges. The efficient frontier shows all the possible combinations of expected returns and risk which can be considered efficient. However, for a specific investor, only one point on this efficient frontier represents the optimal combination of securities. Which point that is depends on the shape of the field of indifference curves of the investor in question. However, it has not been possible until now to determine the exact characteristic of the indifference curve field of a specific subject. In order to nevertheless be able to differentiate between diversification decisions which are suboptimal and optimal, the approach by Gubaydullina and Spiwoks (2015) is useful: In this method there are two securities x_1 and x_2 which both offer the same return (e1 = e2). In this way, the efficient frontier is reduced to a single point (cf. Gubaydullina and Spiwoks 2015, Figure 2). In a decision-making situation of this kind, the exact characteristic of the indifference curve field of an investor is no longer significant. It suffices to know whether the investor should be categorized as risk-averse in order to be able to differentiate between optimal and suboptimal diversification. There are various well-established procedures for discovering whether a subject is risk-loving, risk-neutral or risk-averse. In this study, the approach used by Holt and Laury (2002) is followed.

2.2 Tasks

Each subject has to make four investment decisions (Tasks 1-4), from whose success he or she is directly affected.

In Task 1 there are two different securities to choose from (share A and share B). The subjects have to compile a portfolio which contains four shares. The possible portfolios are thus AAAA, AAAB, AABB, ABBB and BBBB. The subjects profit from the dividend payments. The price trends of the two shares are ignored in order to create a decision-making situation which is as clear as possible. The expectation value of the returns is thus solely based on the dividend payments. The dividend payments (= expectation value of the returns) of the two shares A and B are identical ($e_A = e_B$ = €1.50). However, they exhibit different risk profiles. Whereas share A generates €3 or €0, share B yields either €1 or €2 ($s_A > s_B$). Both events have a probability of occurrence of 50%. Whether a favorable or an unfavorable event occurs depends – in both companies – on the economic situation. The yield of the two shares are accordingly not independent of each other, they are entirely positively correlated (correlation coefficient = +1). The subjects are informed of these circumstances. Test questions are used to ensure that the subjects have understood this point of departure.

In Task 1, the subjects are informed about the movements of returns in the past ten years. The intention is that in this way they will obtain a specific impression of the possible events – of the completely positive correlation and of the different risk profiles of the two securities A and B.

Table 1: Dividend Payments of the Past Ten Years for Share A and Share B

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Share A	€3	€0	€3	€0	€0	€0	€3	€3	€0	€3	€?
Share B	€2	€1	€2	€1	€1	€1	€2	€2	€1	€2	€?

By mixing the two securities A and B, no reduction in risk can be achieved in view of the entirely positive correlation of the dividend payments. The ideal portfolio for risk-averse investors is thus BBBB (Table 2).

Table 2: Expectation Value for the Dividend Payments and Variance of the Possible Portfolios for Task 1

Portfolio Composition	AAAA	AAAB	AABB	ABBB	BBBB
Expectation Value of Dividend Payment	€6	€6	€6	€6	€6
Variance	36.0	25.0	16.0	9.0	4.0

At the end of the experiment, and in the presence of the subjects, the actual dividends of the shares A and B for the year 2016 are determined randomly by tossing a coin. The subjects then receive a payoff in the amount of the dividend payment of their portfolio. If there is a favorable random event (dividend of share A = €3 and dividend of share B = €2), the person who has chosen the portfolio AABB receives a payment of €10 (2 x €3 + 2 x €2). The person who has chosen the portfolio AAAA receives a payment of €12 (4 x €3). If an unfavorable random event occurs (dividend of share A = €0 and dividend of share B = €1), the person who has selected the portfolio AABB receives €2 (2 x €0 + 2 x €1). The person who has chosen the portfolio AAAA receives €0 (4 x €0).

In Task 2, on the other hand, a choice can be made between two different investment alternatives (share X and share Q). The subjects are asked to compile a portfolio consisting of four shares. The possible portfolios are thus XXXX, XXXQ, XXQQ, XQQQ and QQQQ. The subjects profit from the dividend payments. The price trends of the two shares are ignored in order to create a decision-making situation which is as clear as possible. The expectation value of the returns is thus solely based on the dividend payments. The dividend payments (= expectation value of the returns) of the two shares X and Q are identical ($e_X = e_Q$ = €1.00). The risk exposure of the two shares X and Q is also identical ($s_X = s_Q$). The two shares pay a dividend of either €0 or €2. With both shares the probability of the occurrence of these two events is 50%. The dividend payments of the shares X and Q are based on independent random processes (correlation coefficient = 0). The subjects are informed about these circumstances. Test questions are used to ensure that the subjects have understood this point of departure.

In Task 2, the subjects are informed about the course of the returns in the past ten years (Table 2). The intention is that in this way they will obtain a specific impression of the possible events. In addition, the intention is to make them realize that the dividend payments of the shares are entirely uncorrelated.

Table 3: Dividend Payments of the Past Ten Years for Share X and Share Q

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Share X	€0	€0	€2	€0	€0	€0	€2	€2	€2	€2	€?
Share Q	€0	€2	€2	€2	€0	€2	€0	€0	€2	€0	€?

By mixing the two shares X and Q, a significant reduction of risk exposure can be achieved in view of the uncorrelated movements in the dividend payments. The ideal portfolio for risk-averse investors is thus XXQQ (Table 4).

Table 4: Expectation Value for the Dividend Payments and Variance of the Possible Portfolios for Task 2

Portfolio Composition	QQQQ	QQQX	QQXX	QXXX	XXXX
Expectation Value of Dividend Payment	€4	€4	€4	€4	€4
Variance	16.0	8.8	6.4	8.8	16.0

At the end of the experiment, and in the presence of the subjects, the actual dividends of the shares X and Q for the year 2016 are determined randomly (by tossing a coin). At the end of the game, the subjects then receive a payoff in the amount of the dividend payment of their portfolio. If there is an unfavorable random event for share X (dividend of share X = €2) and an unfavorable random event for share Q (dividend of share Q = €0), the person who has chosen portfolio XXQQ receives a payment of €4 (2 x €2 + 2 x €0). The person who has chosen the portfolio XXXQ receives a payment of €6 (3 X €2 + 1 X €0).

<u>Task 3</u> is similar to Task 1. There is a choice between two different securities (share G and share H). The subjects have to compile a portfolio which contains four shares. The possible portfolios are thus GGGG, GGGH, GGHH, GHHH and HHHH. The subjects profit from the dividend payments. The price trends of the two shares are ignored in order to create a decision-making situation which is as clear as possible. The expectation value of the returns is thus solely based on the dividend payments. The dividend payments (= expectation value of the returns) of the two shares G and H are identical ($e_G = e_H =$ €1.50). However, they exhibit dif-

ferent risk profiles. Whereas share G generates either €3 or €0, share H yields either €1 or €2 (sG > sH). Both events have a probability of occurrence of 50%. Whether a favorable or an unfavorable event occurs depends – in both companies – on the economic situation. The yield of the two shares are accordingly not independent of each other; they are entirely positively correlated (correlation coefficient = +1). The subjects are informed about these circumstances. Test questions are used to ensure that the subjects have understood this point of departure.

No reduction in risk can be achieved by mixing the two securities G and H given the entirely positive correlation of the dividend payments. The ideal portfolio for risk-averse investors is thus HHHH (Table 5).

Table 5: Expectation Value for the Dividend Payments and Variance of the Possible Portfolios for Task 3

Portfolio Composition	GGGG	GGGH	GGHH	GHHH	HHHH
Expectation Value of the Dividend Payment	€6	€6	€6	€6	€6
Variance	36	25	16	9	4

Task 4 is similar to Task 2. There is a once again a choice between two different investment alternatives (share E and share F). The subjects are asked to compile a portfolio consisting of four shares. The possible portfolios are thus EEEE, EEEF, EEFF, EFFF and FFFF. The subjects profit from the dividend payments. The price trends of the two shares are ignored in order to create a decision-making situation which is as clear as possible. The expectation value of the returns is thus solely based on the dividend payments. The dividend payments (= expectation value of the returns) of the two shares E and F are identical ($e_E = e_F$ = €1.00). The risk exposure of the two shares E and F is also identical ($s_E = s_F$). The two shares have a dividend of either €0 or €2. With both shares the probability of the occurrence of these two events is 50%. The dividend payments of the shares E and F are based on independent random processes (correlation coefficient = 0).

By mixing the two shares E and F, a significant reduction of risk exposure can be achieved given the uncorrelated movement of the dividend payments. The ideal portfolio for risk-averse investors is thus EEFF (Table 6).

Table 6: Expectation Value for the Dividend Payments and Variance of the Possible Portfolios for Task 4

Portfolio Composition	EEEE	EEEF	EEFF	EFFF	FFFF
Expectation Value of the Dividend Payment	€4	€4	€4	€4	€4
Variance	16	8.8	6.4	8.8	16

2.3 Treatments

As this study investigates whether mood has effects on the diversification decisions of subjects and thus on the risk exposure of their portfolios, the mood of the subjects is influenced with brief film excerpts. Emotional film excerpts are a common and effective method to create emotions.[1] Film excerpts are also being increasingly used to influence moods in economic experiments (see, for example, Andrade et al., 2016; Oswald et al., 2015; Ifcher and Zarghamee, 2014; Lahav and Meer, 2012; Ifcher and Zarghamee, 2011; Schaefer et al., 2010; Rottenberg et al., 2007; Kirchsteiger et al., 2006; Gross and Levenson, 1995).

The selection of short film excerpts deployed here are taken from the study by Schaefer et al. (2010), in which more than 70 film excerpts were assessed with regard to their ability to create emotions. The film clips are clearly assigned to the moods which are desired. There are film excerpts which evoke negative emotions, and there are film excerpts which evoke positive emotions. There are also film excerpts which do not influence the mood of the viewers. These film sequences are described as neutral.

In order to create positive emotions, the following film excerpts were used in our experiment: (1) Benny and Joon (122 seconds): Benny (Johnny Depp) plays the fool in a café. (2) Life is Beautiful (266 seconds): A mother and son are re-united after the Second World War. (3) Dead Poets Society (163 seconds): All of the students in a class stand on their desks to show their solidarity with Mr. Keating (Robin Williams), who has just been fired. (4) Forrest Gump (121 seconds): Father and son are reunited. (5) Dinner for Schmucks (101 seconds): Complex humorous scenes.

In order to create negative emotions, the following film excerpts were used in our experiment: (1) Schindler's List (101 seconds): The SS storm a house and shoot everyone in it. (2) The Piano (42 seconds): A person's finger is chopped off deliberately with an axe. (3) The Blair Witch Project (232 seconds): Final scene in which the protagonists are seemingly killed. (1) Schindler's List (76 seconds): Bodies are burned in a concentration camp. (5) Saving Private Ryan (327 seconds): A war scene at Omaha Beach in the Second World War.

The film excerpts used here which do not affect the emotions of the subjects (neutral) are as follows: (1) The Lover (43 seconds): Marguerite (Jane March) gets into a car. She drives to a house in a busy street and knocks on a door. A Chinese man opens and she goes in. (2) Blue (40 seconds): A man is clearing up the drawers of his desk. A woman is walking along a street and says hello to another woman. (3) Train ride (58 seconds): A train travels through a green landscape.[2] (2) Blue (25 seconds): A woman goes up an escalator carrying a crate. (5) Blue (16 seconds): A person holds a piece of aluminum foil out of the window of a moving car.

In the experiment, three treatments are compared. In the negative treatment, the subjects watch a film excerpt which evokes negative emotions before making their portfolio decisions. In the positive treatment, the subjects watch a film excerpt which evokes positive emotions before making their portfolio decisions. In the neutral treatment, the subjects

[1] There are various ways of influencing mood. This also includes real situations, memories and imagination, noises and music, presents, film clips and the so-called Velten technique (cf. Westermann et al., 1996).

[2] The film clip train ride is from the study by Gendolla and Krüsken, 2002.

watch a film excerpt which does not have any effect on them before making their portfolio decisions.

In all three treatments it is only the film excerpts which differ. The rest of the experiment is the same, so in all three treatments the subjects have to carry out Tasks 1-4.

2.4 Sequence of the Tasks and Procedure of the Experiment

After the subjects have read the thorough instructions, their mood is measured before the experiment with the following question:

How are you feeling now? Please mark the adequate number!

1—2—3—4—5—6—7—8—9—10

very bad very good

Then the test questions about the decision-making situations (Tasks 1-4) are posed. Only those who answer the test questions correctly are allowed to participate in the experiment. This ensures that the subjects understand which decision can lead to which consequences for their payoff.

Subsequently – before the first diversification task – their mood is influenced with a corresponding film clip. Depending on the treatment, a film clip is shown which evokes positive emotions (positive treatment), negative emotions (negative treatment), or no emotions (neutral treatment). After the film excerpt a manipulation check takes place to test whether the intended mood has been created among the subjects.

For the manipulation check the following question is posed:[3]

Which emotions did you experience while watching the movie clip?
Please mark one number accordingly!

1—2—3—4—5—6—7—8—9—10

very negative very positive

Subsequently the subjects have to make their first diversification decision in Task 1. After this decision has been made, the subjects are shown the second film clip. Once again, depending on the treatment, a film clip is shown which evokes positive emotions (positive treatment), negative emotions (negative treatment), or no emotions (neutral treatment).

[3] Similar manipulation checks were also carried out in the studies by Andrade et al., 2016; Lahav and Meer, 2012; Rottenberg et al., 2007; Kirchsteiger et al., 2006.

After the film excerpt a manipulation check again takes place to test whether the intended mood has really been created among the subjects. Subsequently the subjects have to make their second diversification decision in Task 2. The same procedure is repeated until all four diversification decisions have been made.

Then the experiment examines – using the approach of Holt and Laury (2002) – whether the subjects should be categorized as risk-averse, risk-neutral or risk-loving. This is absolutely necessary, because the portfolios with the lowest possible risk exposure (in Task 1: BBBB; in Task 2: XXQQ; in Task 3: HHHH, and in Task 4: EEFF) are only optimal for risk-averse investors. Risk-loving subjects, on the other hand, would normally always choose the portfolio with the highest variance. Risk-neutral subject are indifferent with regard to all of the portfolios, as their expectation value is always the same for all five possible portfolio alternatives.

Before the subjects make the ten decisions on their preference for lottery A or lottery B, depending on the treatment a film excerpt is shown which evokes positive emotions (positive treatment), negative emotions (negative treatment), or no emotions at all (neutral treatment). After the film excerpt a manipulation check takes place again to test whether the intended mood has really been created among the subjects. Then the subjects make their ten decisions regarding lottery A or B.

After the subjects have made their decisions, the drawing of the random events for the securities of Tasks 1-4 and the draw for the lottery according to Holt and Laury (2002) are made. Following this the payoff is made – dependent on their choice of portfolio in the Tasks 1-4, and depending on their success in the lottery according to Holt and Laury (2002). In the diversification decision tasks, up to €40 can be earned, and in the lottery up to €3.85. In addition, every participant receives a show-up fee of €1.50, so overall the subjects can earn up to €45.35. On average, the participants earn €24.44. The maximum was €38.35 and the minimum was €11.60. The survey takes an average of 40 minutes, so the payment can be described as attractive. Without exception, the subjects gave the impression of being concentrated and committed.

The experiment was carried out at the Ostfalia University of Applied Sciences in Germany with students from the Faculties of Automotive Engineering, Public Health Services and Business in the period from 12 August 2015 to 24 September 2015 in the Ostfalia Laboratory for Experimental Economic Research (OLEW). Overall, 123 students took part in the experiment in 37 sessions. From the Faculty of Business, 60 students took part in the experiment (48.78%), while from the Faculty of Automotive Engineering 40 students (32.52%), and from the Faculty of Public Health Services 23 students (18.70%) participated. Of these, 45 were women (36.59%) and 78 were men (63.41%). The number of participants per treatment was as follows: 44 participants in the negative treatment, 39 in the neutral treatment and 40 participants in the positive treatment. The participants were 24.3 years old on average.

In order to test the hypotheses, the subjects who proved to be risk-neutral or risk-loving in the procedure used by Holt and Laury (2002) were eliminated from the experiment. In addition, those subjects who chose variant A in the tenth decision in the test according to Holt and Laury (2002) were also eliminated, because it must be assumed that subjects who do not choose variant B in the tenth decision of the lottery have not really understood the decision-making situation. This is because in the tenth decision, variant B is clearly superior to variant A – regardless of the risk preferences of the subject. The number of participants after the necessary thinning out was as follows: 26 participants in the negative treatment, 24 in the neutral treatment and 25 participants in the positive treatment.

The experiment was programmed with z-Tree (cf. Fischbacher, 2007). The instructions, test questions and screenshots from the experiment can be found in the appendices.

3 Hypotheses

As the expectation values of the five portfolio alternatives are the same, the optimal diversification for risk-averse subjects is to choose the portfolio with the minimum variance. However, there are many empirical findings which show that in practice a large number of portfolios are not optimally diversified (see the introduction). It cannot therefore be expected that risk-averse subjects will always choose the portfolio composition which minimizes risk exposure. Hypothesis 1 is therefore: Risk-averse subjects will not always choose the minimum variance portfolio in Tasks 1-4. Null hypothesis 1 is thus: All (risk-averse) subjects will only choose the minimum variance portfolio in Tasks 1-4.

There are numerous empirical findings showing that positive emotions can reduce the perception of risk in investment decisions (see Conte et al., 2018; Kaplanski et al., 2014; Stanton et al., 2014; Kuhnen and Knutson, 2011; Shu, 2010; Grable and Roszkowski, 2008; Yuen and Lee, 2003; Forgas, 1998). It can thus be expected that the minimum variance portfolio will be selected less frequently in the positive treatment than in the neutral treatment. Hypothesis 2 is therefore: The average variance of the selected portfolios is higher in the positive treatment than in the neutral treatment. Null hypothesis 2 is thus: The average variance of the selected portfolios is not higher in the positive treatment than in the neutral treatment.

There are some empirical findings which show that negative emotions also have an inhibitive effect on subjects who based their own decisions on rational considerations (see, for example, Conte et al., 2018; Gambetti and Giusbert, 2012; Lee and Andrade, 2011; Pham, 2007; Kliger and Levy, 2003; Tiedens and Linton, 2002; Lerner and Keltner, 2001; Leith and Baumeister, 1996). It can thus be expected that in the negative treatment the minimum variance portfolio is chosen less frequently than in the neutral treatment. Hypothesis 3 is therefore: The average variance of the selected portfolios in the negative treatment is higher than in the neutral treatment. Null hypothesis 3 is thus: The average variance of the selected portfolios in the negative treatment is not higher than in the neutral treatment.

If it is correct that positive and negative emotions contribute equally to a weakening of rational behavior, then it can be presumed that there are no significant differences with regard to portfolio decisions in the positive and negative treatments. Hypothesis 4 is therefore: The average variance of the selected portfolios is the same in the positive and negative treatments. Null hypothesis 4 is thus: There is a significant difference in the average variance between the positive and negative treatments.

If it is correct that a neutral mood tends to contribute more towards making meaningful investment decisions than a positive or negative mood, this must also be reflected in the risk-adjusted payments, i.e. in the performance of the subjects. Hypothesis 5 is therefore: Subjects who make their investment decisions in the neutral treatment will obtain higher risk-adjusted payments than subjects in the positive and negative treatments. Null hypothesis 5 is thus: Subjects who take part in the neutral treatment will not obtain significantly higher risk-adjusted payments than subjects in the positive and negative treatments.

4 Results of the Experiment

4.1 The Effectiveness of Influencing Mood

First, we consider whether it has been possible to create the respective desired mood in the three treatments. The first measurement of mood takes place before the presentation of the first film excerpt (round 0). In Figure 1 it can clearly be seen that the mood in round 0 – i.e. before the targeted creation of a mood – was rather good in all three treatments. The median of round 0 in all three treatments was 8.

The average mood in the five rounds in which a film excerpt was presented to create a certain mood was as follows: In the negative treatment it was 3.09 (SD 1.42), in the neutral treatment it was 5.67 (SD 1.14) and in the positive treatment it was 7.32 (SD 1.36). In Figure 1 it can be clearly seen in the box plots that the creation of a specific mood in the individual treatments was successful.

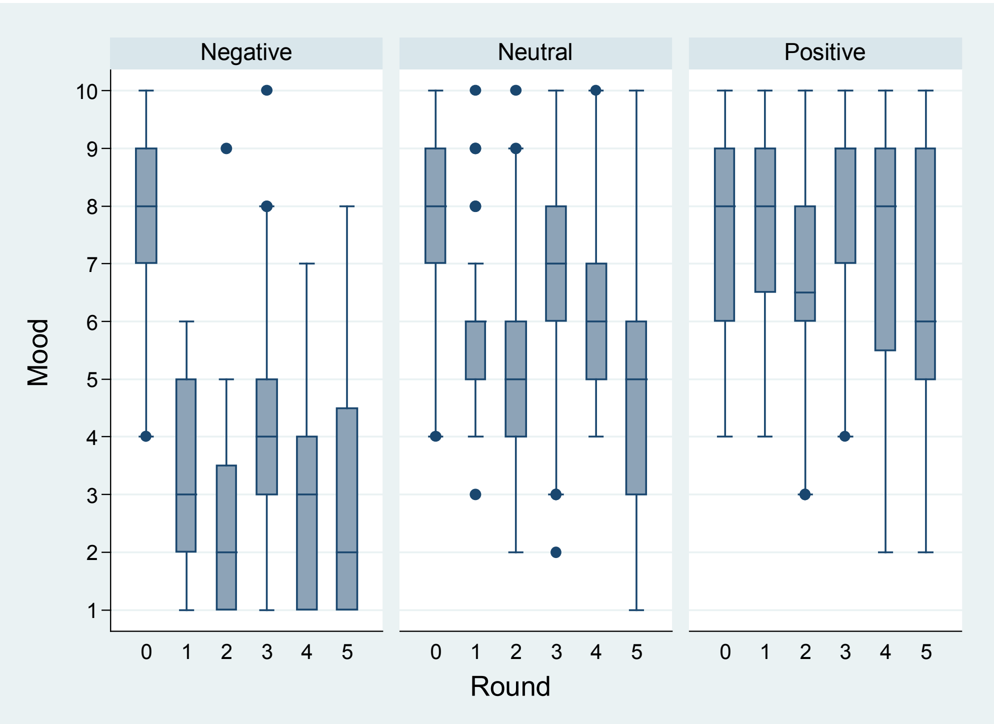


Figure 1: Box Plots on the Mood of the Subjects in the Respective Rounds of the Game According to Treatments

Figure 2 also shows box plots on the mood of the subjects in the three treatments. It summarizes the five rounds in which film excerpts were presented to manipulate moods.

It is clearly recognizable that the mood of the subjects varies considerably between the three treatments. The box for the negative treatment extends from 2.1 to 3.8. The box for the neutral treatment varies from 4.8 to 6.0. The box for the positive treatment varies from 6.2 to 8.2.

The fact that mood manipulation with the aid of the film clips worked is also shown by Table 7. The negative treatment shows significantly lower mood values than the positive treatment (z= -7.466, p= 0.0000; Mann-Whitney U test). The negative treatment shows significantly lower mood values than the neutral treatment (z= -6.417, p= 0.0000; Mann-Whitney U test). The positive treatment shows significantly higher average mood values than the neutral treatment (z= 5.089, p= 0.0000; Mann-Whitney U test).

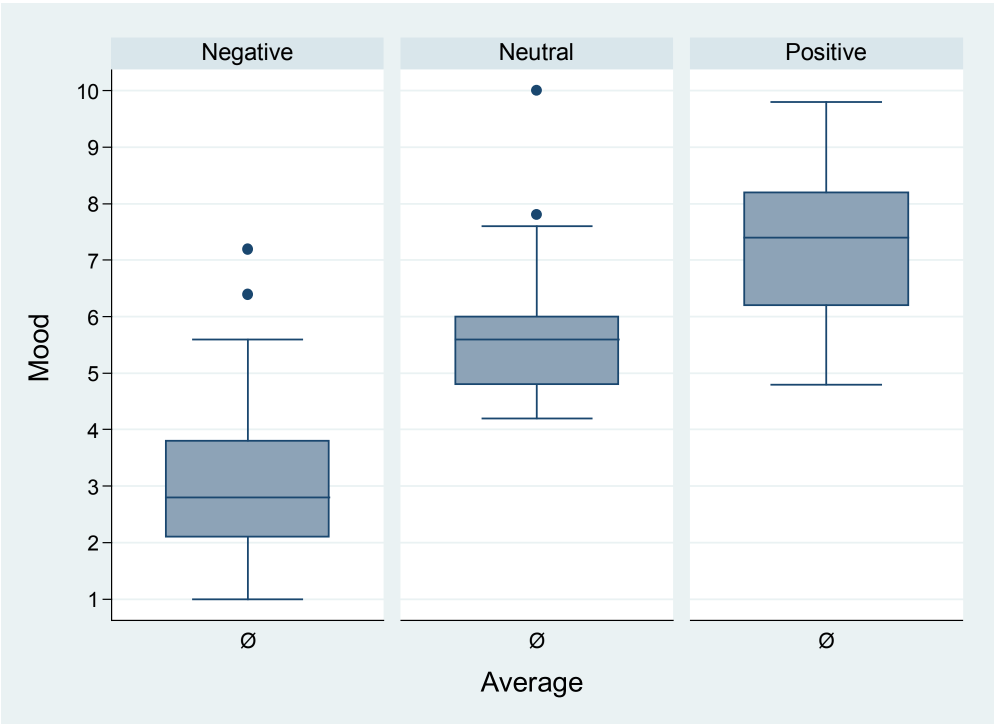


Figure 2: Box Plots of Mood after the Treatments (Summary of Rounds 1-5)

Overall it can be stated that the deployment of the film excerpts led to the desired results. In each of the three treatments, the desired mood was predominant. In the positive treatment positive emotions prevailed. In the neutral treatment, a generally average mood was present, and in the negative treatment, negative emotions predominated. To this extent, the approach was very well suited to answering the questions posed.

Table 7: Average Mood of the Subjects in the Respective Rounds

		Average Mood per Round (Standard Deviation)						
Treatment	#	Before the Experiment	Round 1	Round 2	Round 3	Round 4	Round 5	Ø
Negative Treatment	44	7.84 (1.57)	3.36 (1.62)	2.39 (1.78)	4.00 (1.97)	2.82 (1.59)	2.87 (2.06)	3.09*** (1.42)
Neutral Treatment	39	7.92 (1.36)	5,79 (1.49)	5.05 (1.73)	6.62 (1.99)	6.00 (1.49)	4.87 (2.04)	5.67*** (1.14)
Positive Treatment	40	7,43 (1.60)	7,55 (1.66)	6.85 (2.03)	8.18 (1.74)	7.33 (2.35)	6.68 (2.31)	7.32*** (1.36)

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).
For the calculation of the average values (Ø), the figures of rounds 1-5 were used.

4.2 Rational Strategy

Now we will take a look at the percentage distribution of the portfolios in the three treatments (Table 8). In the upper part of Table 8, Tasks 1 and 3 are listed. For risk-averse investors, the alternatives BBBB and HHHH respectively represent a rational strategy in the Tasks 1 and 3, because for all of the possible portfolio structures, the expectation value of the payment is identical. However, the risk exposure (variance) in the portfolios BBBB and HHHH is significantly lower than in the other portfolio alternatives. In all three treatments there are clear deviations from the rational strategy. In the negative treatment, only 13.46% of the participants chose the optimal portfolio. In the neutral treatment, this even falls to 8.33%. And in the positive treatment, only 12% of the participants chose the optimal portfolio.

In the upper part of the table, Tasks 2 and 4 are illustrated. The rational strategy here would be to choose a mix of two portfolios. The portfolios QQXX and EEFF have the same expectation value for the payment as the four other portfolio alternatives, but the risk exposure (variance) is considerably lower here than in the other four portfolio alternatives. The subjects were obviously able to deal significantly better with this starting position. In the negative treatment, 69.23% of the risk-averse participants chose the optimal portfolio. In the neutral treatment, this even rose to 75%. In the positive treatment, 70% of the risk-averse subjects chose the optimal portfolio.

Table 8: Percentage Distribution of the Portfolios in the Three Treatments

	Percentage Distribution of the Portfolios				
Portfolios Task 1	AAAA	AAAB	AABB	ABBB	BBBB
Portfolios Task 3	GGGG	GGGH	GGHH	GHHH	HHHH
Variance	36	25	16	9	4
Rational Strategy	0.00%	0.00%	0.00%	0.00%	100.00%
Negative Treatment	7.69%	19.24%	26.92%	32.69%	13.46%
Neutral Treatment	0.00%	10.42%	31.25%	50.00%	8.33%
Positive Treatment	4%	26%	26%	32%	12%
Portfolios Task 2	QQQQ	QQQX	QQXX	QXXX	XXXX
Portfolios Task 4	EEEE	EEEF	EEFF	EFFF	FFFF
Variance	16	8.8	6.4	8.8	16.0
Rational Strategy	0%	0%	100%	0%	0%
Negative Treatment	5.77%	13.46%	69.23%	5,77%	5.77%
Neutral Treatment	6.25%	6.25%	75.00%	10,42%	2.08%
Positive Treatment	10%	6.00%	70.00%	6,00%	8.00%

Null hypothesis 1 is: All risk-averse subjects will choose the minimum variance portfolios in Tasks 1-4. This null hypothesis clearly has to be rejected. Hypothesis 1 can thus be viewed as confirmed for the time being. The subjects do not always choose the optimal portfolio. In other words, they do not always take a rational approach. These results are in line with those of Ackert et al., 2015; Gubaydullina and Spiwoks, 2015; Ackert et al. 2011; Eyster and Weizsäcker, 2010; Goetzmann and Kumar, 2008, and Hedesstrom et al., 2006.

This is also reflected by an unnecessarily high-risk exposure. In Table 9, the average variance for the minimum variance portfolio and the average variance of the portfolios chosen by the subjects are compared.

Table 9: Average Variance of the Rational Strategy and the Average Variance of the Portfolios Chosen by the Subjects in the Three Treatments

Treatment	Rational Strategy Average Variance	Actual: Average Variance	T-Test P-Value
Negative treatment	5.2	11.67	0.0000***
Neutral treatment	5.2	10.02	0.0000***
Positive treatment	5.2	11.94	0.0000***

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

Here it is also revealed that the subjects frequently fail to choose the minimum variance portfolio, although all five portfolio alternatives exhibit the same expectation value for the payment. This way of evaluating the data also leads to the rejection of null hypothesis 1.

4.3 The Influence of Mood on Portfolio Decisions

In order to compare the average risk exposure (variance) of the three treatments, first, the average risk exposure of each individual subject is established and entered into a histogram of the average risk exposure according to treatments (Figure 3).

It is noticeable that in the neutral treatment the distribution takes the form of a peak, which is clearly skewed to the right. By contrast, distribution in the negative and positive treatments shows a broader spread and is only slightly skewed to the right. This means that the average risk exposure of the subjects in the neutral treatment was lower than the average risk exposure of subjects in the positive and negative treatments. This also becomes clear when viewing the distribution of the average variances below ten. In the neutral treatment, significantly more than half of the subjects are below this limit. In the positive and negative treatments, on the other hand, significantly less than half of the subjects are below this limit.

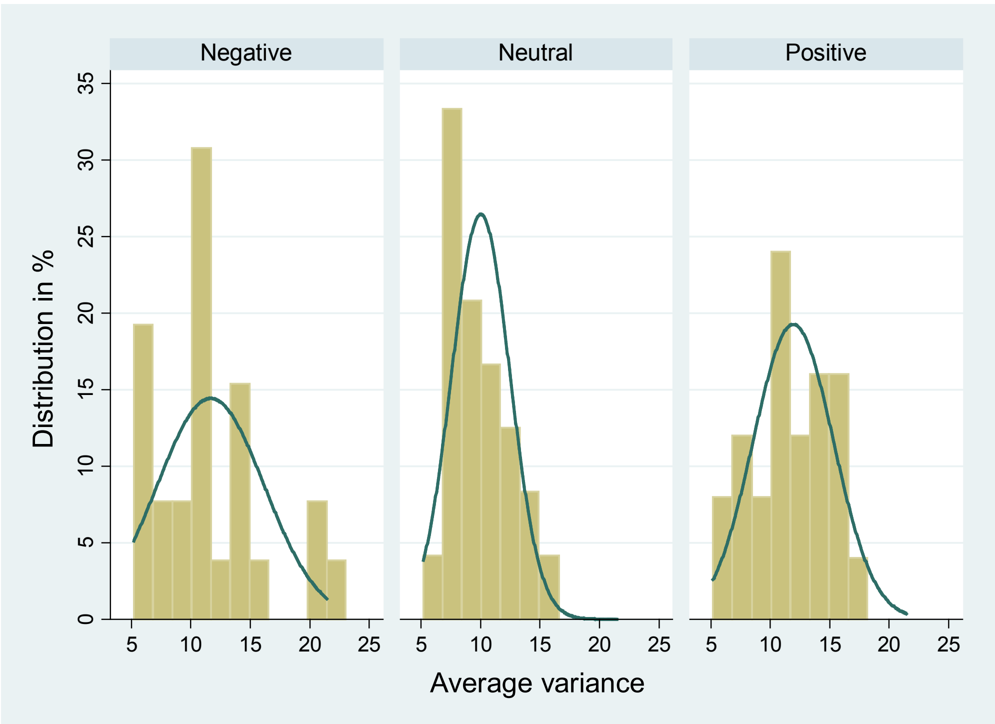


Figure 3: Percentage Distribution of the Average Variance in the Three Treatments

The small but easily recognizable differences in average risk exposure in the three treatments are also shown when viewing the box plots (Figure 4). In the neutral treatment, the median – at 9.45 – is clearly below the medians of the positive (11.7) and negative (11.2) treatments. When viewing figures 3 and 4, differences between the neutral treatment on the one hand and the positive and negative treatments on the other are recognizable.

Whether these differences are significant was examined with the Wilcoxon rank sum test. First, the positive treatment was compared to the negative treatment.

The average risk exposure (variance) was significantly higher in the positive treatment at 11.94 than in the neutral treatment at 10.02. In the Wilcoxon rank sum test, the difference – with a p-value of 0.0294 – proved significant (Table 10). Null hypothesis 2 thus has to be rejected. Hypothesis 2 states that the average variance of the selected portfolios in the positive treatment is higher than in the neutral treatment. This hypothesis can be viewed as confirmed for the meantime.

Table 10: Risk Exposure (Average Variance of the Portfolios) in the Positive and Neutral Treatments

Average Variance Positive Treatment	Average Variance Neutral Treatment	Wilcoxon Rank Sum Test P-Value
11.94	10.02	0.0294**

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

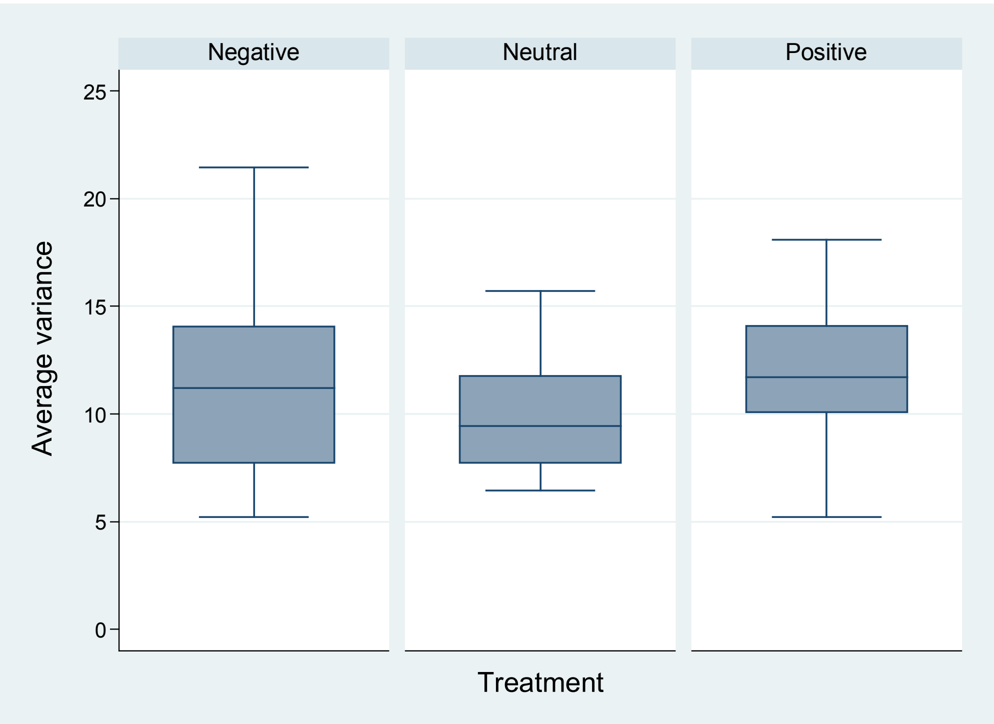


Figure 4: Average Variance in the Three Treatments

The findings of Grable and Roszkoswki (2008) as well as those of Kuhnen and Knutson (2011), who discovered that there is a greater willingness to take risks when positive emotions are present, are thus confirmed.

Next, we examined whether the risk exposure in the negative treatment is also recognizably higher in the negative treatment than in the neutral treatment. The average variance in the negative treatment (11.67) was significantly higher than in the neutral treatment (10.02), but in the Wilcoxon rank sum test this difference proved to be non-significant (Table 11). Null hypothesis 3 therefore cannot be rejected. Hypothesis 3 states that the average variance in the negative treatment is higher than in the neutral treatment. This expectation was not confirmed in a statistically significant way.

Table 11: Risk Exposure (Average Variance of the Portfolios) in the Negative and Neutral Treatments

Average Variance Negative Treatment	Average Variance Neutral Treatment	Wilcoxon Rank Sum Test P-Value
11.67	10.02	0.2420

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

This leaves the question of whether the average risk exposure of the positive treatment and the negative treatment deviate significantly from each other. The average variance of the negative treatment was 11.67, and that of the positive treatment was 11.94 (Table 12). This relatively small difference revealed itself to be insignificant in the Wilcoxon rank sum test.

Table 12: Risk Exposure (Average Variance of the Portfolios) in the Negative and Positive Treatments

Average Variance Negative Treatment	Average Variance Positive Treatment	Wilcoxon Rank Sum Test P-Value
11.67	11.94	0.3956

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

Null hypothesis 4 thus has to be rejected. Hypothesis 4 states that the average variance of the selected portfolios is not significantly different in the positive and negative treatments. This hypothesis can thus be viewed as confirmed for the meantime.

4.4 The Influence of Mood on Risk-Adjusted Returns

Finally we examined whether the mood of the subjects was also reflected in their risk-adjusted payoffs (performance). To do so, we deployed a simplified performance benchmark. The risk-adjusted payoff (*RA*) is determined as follows.

(1)

$$\text{RA} = \frac{\text{Payoff}}{1 + (\frac{Variance}{100})}$$

In order to be able to compare the performance (risk-adjusted payoff) of the subjects in the three treatments, first, the performance of each individual subject was determined and then entered into a histogram according to treatments (Figure 5). It is noticeable that in the neutral treatment the distribution takes the form of a slight peak which is skewed to the left. By contrast, distribution in the negative and positive treatments is slightly skewed to the right. In addition, in the neutral treatment the risk-adjusted payoffs of the subjects were over 20. In the positive and negative treatments on the other hand, over 40% and over 30% of the subjects respectively were below 20.

When viewing the box plots, the differences between the three treatments become clear (Figure 6). In the neutral treatment, the median – at 26.55 – is clearly above the medians of the positive (21.26) and negative (23.18) treatments. This means that the performance of the subjects in the neutral treatment was better than the performance of subjects in the positive and negative treatments. Whether these differences are significant was then examined with the Wilcoxon rank sum test. First of all the positive treatment was compared to the neutral treatment, and then the negative and the neutral were compared.

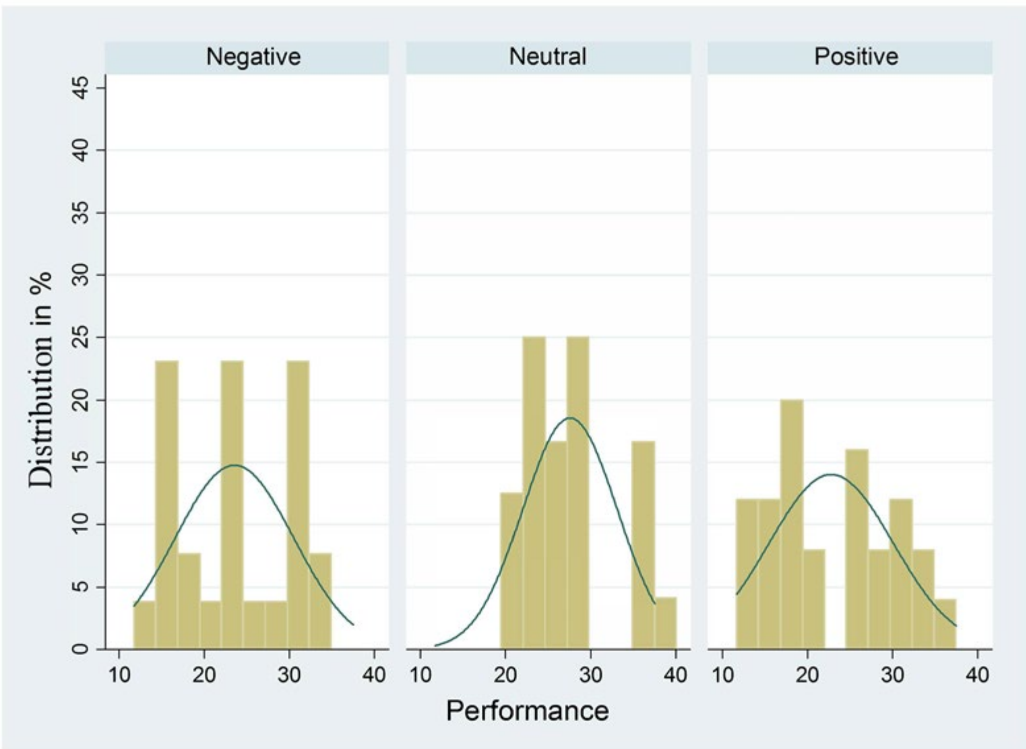


Figure 5: Percentage Distribution of Performance (Risk-Adjusted Payoff) in the Three Treatments

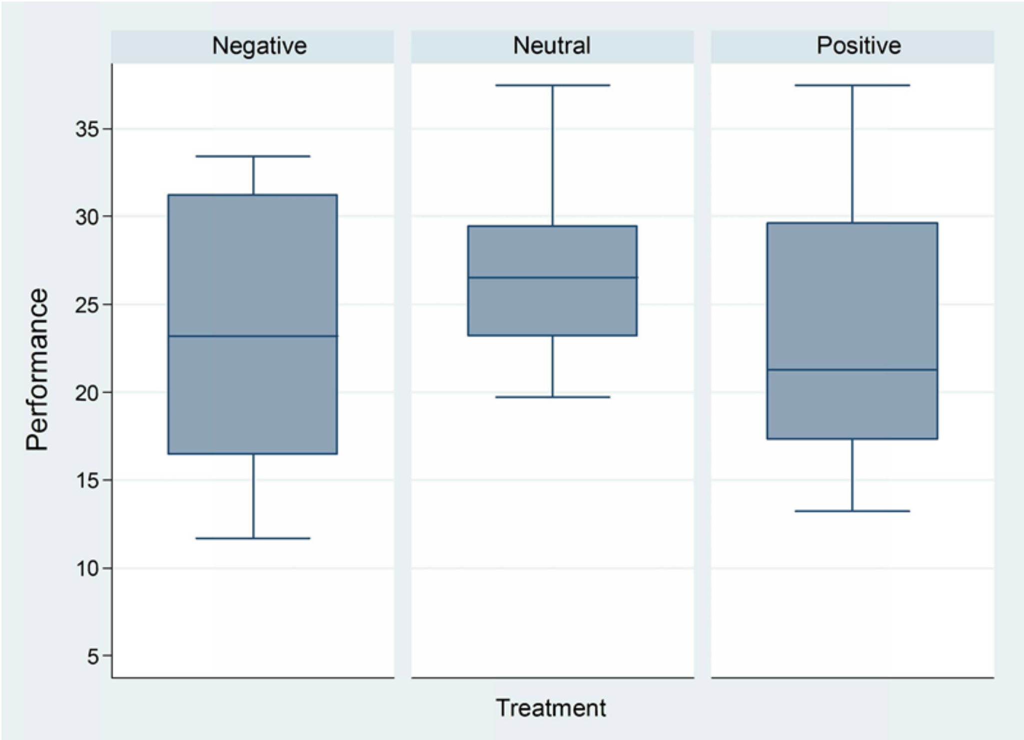


Figure 6: Risk-Adjusted Payoffs in the Three Treatments

The average performance was significantly higher at 27.52 in the neutral treatment than in the positive treatment at 22.81. In the Wilcoxon rank sum test, the difference – with a p-value of 0.0285 – proved to be significant (Table 13). The average performance was also significantly higher at 27.52 in the neutral treatment than in the positive treatment at 23.53. In the Wilcoxon rank sum test, the difference – with a p-value of 0.0545 – again proved to be significant (Table 14).

Table 13: Performance (Risk-Adjusted Payoff) in the Positive and Neutral Treatments

Average Performance Positive Treatment	Average Performance Neutral Treatment	Wilcoxon Rank Sum Test P-Value
22.81	27.52	0.0285**

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

Table 14: Performance (Risk-Adjusted Payoff) in the Negative and Neutral Treatments

Average Performance Negative Treatment	Average Performance Neutral Treatment	Wilcoxon Rank Sum Test P-Value
23.53	27.52	0.0545*

The significant values are highlighted (*** p<0.01; ** p<0.05; * p<0.1).

Thus, null hypothesis 5 clearly has to be rejected. Hypothesis 5 states that the subjects who make their investment decisions in the neutral treatment will obtain higher risk-adjusted payments than the subjects in the positive and negative treatments. This hypothesis can be viewed as confirmed for the meantime.

5 Summary and Conclusion

For risk-averse subjects it is usually meaningful to diversify their portfolios (Markowitz, 1952). However, practice shows that many subjects have poorly diversified securities portfolios. Furthermore, some studies show that personal mood can have an influence on investment decisions. However, until now the question of whether positive and negative emotions have an influence on diversification behavior and thus on risk exposure has not been investigated.

The design of this experiment is loosely based on the work of Gubaydullina and Spiwoks (2015). Each subject has to take four investment decisions. In each task the subjects can choose between two different securities. The subjects have to compile a portfolio which contains four shares, and they profit from the dividend payments. The dividend payments of the shares are based on a random process. The price movements of the two shares are ignored in order to create a decision-making situation which is as clear as possible. Mood is influenced by positive (in the positive treatment), negative (in the negative treatment) and neutral film excerpts (in the neutral treatment). Manipulation checks show whether the influencing of mood by the film clips has been successful.

The results of the experiment show that the mood of the subjects has an influence on their diversification behavior. This becomes clear as soon as risk exposure is considered: The average risk exposure in the neutral treatment (10.02) is lower than in the positive treatment (11.94) and the negative treatment (11.67).

In addition, a neutral mood leads to the subjects obtaining higher risk-adjusted payoffs. The average risk-adjusted performance was significantly higher at 27.52 in the neutral treatment than in the positive treatment at 22.81 and in the negative treatment at 23.53.

Positive and negative moods lead to higher risk exposure and to lower risk-adjusted payoffs than a neutral mood. From this one can derive the recommendation that investment processes should as far as possible not be affected by emotions. It would seem wise to develop systematic investment rules and to strictly orientate the investment process towards them, or to only let groups of persons (such as an investment committee) take investment decisions rather than individuals.

Acknowledgments

I thank the participants of the 2016 German Association for Experimental Economic Research (GfeW) annual conference for very constructive comments. I also thank Markus Spiwoks, Julia Brüggemann, John Jermyn, Varun Vruddhula and Sebastian Klein for their very helpful comments. This work was supported by the German Association for Experimental Economic Research (GfeW).

References

Ackert, L. S. / Mazzotta, S. / Qi, L. (2011): An Experimental Investigation of Asset Pricing in Segmented Markets, Southern Economic Journal, 77(3), 585-598.

Ackert, L. S. / Church, B. K. / Qi, L. (2015): An Experimental Examination of Portfolio Choice, Review of Finance, 19(5), 1-21.

Agnew, J. / Balduzzi, P. / Sundén, A. (2003): Portfolio Choice and Trading in a Large 401(k) Plan, American Economic Review, 93(1), 193-215.

Anderson, A. (2013): Trading and Under-Diversification, Review of Finance, 17(5), 1699-1741.

Andrade, E. B. / Odean, T. / Lin, S. (2016): Bubbling with Excitement: An Experiment, Review of Finance, 20(2), 1-20.

Apergis, N. / Gabrielsen, A. / Smales, L. A. (2016): (Unusual) Weather and Share Returns – I am Not in the Mood for Mood: Further Evidence from International Markets, Financial Markets and Portfolio Management, 30(1), 63-94.

Baker, M. / Wurgler, J. (2007): Investor Sentiment in the Share Market, Journal of Economic Perspectives, 21(2), 129-151.

Baltussen, G. / Post, G. T. (2011): Irrational Diversification: An Examination of Individual Portfolio Choice, Journal of Financial and Quantitative Analysis, 46(5), 1463-1491.

Barber, B. M. / Odean, T. (2000): Trading is Hazardous to your Wealth: The Common Share Investment Performance of Individual Investors, Journal of Finance, 55(2), 773-806.

Bassi, A. / Colacito, R. / Fulghieri, P. (2013): O Sole Mio: An Experimental Analysis of Weather and Risk Attitudes in Financial Decisions, Review of Financial Studies, 26(7), 1824-1852.

Baumeister, R. F. / Vohs, K. D. / DeWall, C. N. / Zhang, L. (2007): How Emotion Shapes Behavior: Feedback, Anticipation, and Reflection, rather than Direct Causation, Personality and Social Psychology Review, 11(2), 167-203.

Benartzi, S. (2001): Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Share, Journal of Finance, 56(5), 1747-1764.

Benartzi, S. / Thaler, H. (2001): Naive Diversification Strategies in Defined Contribution Saving Plans, American Economic Review, 91(1), 79-98.

Bless, H. / Schwarz, N. / Kemmelmeier, M. (1996): Mood and Stereotyping: The Impact of Moods on the use of General Knowledge Structures, European Review of Social Psychology 7(1), 63-93.

Blume, M. E. / Friend, I. (1975): The Asset Structure of Individual Portfolios and some Implications for Utility Functions, Journal of Finance, 30(2), 585-603.

Bode, M. / van Echelpoel, A. / Sievi, C. R. (1994): Multinationale Diversifikation: Viel zitiert, kaum befolgt, Die Bank, 94(4), 202-206.

Breaban, A. / Noussair, C. N. (2018): Emotional State and Market Behavior, Review of Finance, 22(1), 279-309.

Cao, M. / Wei, J. (2005): Share Market Returns: A Note on Temperature Anomaly, Journal of Banking and Finance, 29(6), 1559-1573.

Chang, S. C. / Chen, S. S. / Chou, R. K. / Lin, Y. H. (2008): Weather and Intraday Patterns in Share Returns and Trading Activity, Journal of Banking and Finance, 32(9), 1754-1766.

Chang, T. / Nieh, C. C. / Yang, M. J. / Yang, T. Y. (2006): Are Share Market Returns Related to the Weather Effects? Empirical evidence from Taiwan, Physica A: Statistical Mechanics and its Applications, 364, 343-354.

Conte, A. / Levati, M. V. / Nardi, C. (2018): Risk Preferences and the Role of Emotions, Economica, 85(338), 305-328.

Daniel, K. / Hirshleifer, D. / Teoh, S. H. (2002): Investor Psychology in Capital Markets: Evidence and Policy Implications, Journal of Monetary Economics, 49(1), 139-209.

De Bondt, W. F. M. (1998): A Portrait of the Individual Investor, European Economic Review, 42(3-5), 831-844.

Dowling, M. / Lucey, B. M. (2005): Weather, Biorhythms, Beliefs and Share Returns – some Preliminary Irish Evidence, International Review of Financial Analysis, 14(3), 337-355.

Elster, J. (1998): Emotions and Economic Theory, Journal of Economic Literature, 36(1), 47-74.

Elster, J. (1996): Rationality and the Emotions, Economic Journal, 106(438), 1386-1397.

Eyster, E. / Weizsäcker, G. (2010): Correlation Neglect in Financial Decision-Making, DIW Discussion Paper, 1104.

Fernandes, D. (2013): The 1/N Rule Revisited: Heterogeneity in the Naïve Diversification Bias, International Journal of Research in Marketing, 30(3), 310-313.

Filiz, I. / Nahmer, T. / Spiwoks, M. / Bizer, K. (2018): Portfolio Diversification: The Influence of Herding, Status-quo Bias, and the Gambler's Fallacy, Financial Markets and Portfolio Management, 32(2), 167-205.

Fischbacher, U. (2007): Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, Experimental Economics, 10(2), 171-178.

Floros, C. (2011): On the Relationship between Weather and Share Market Returns, Studies in Economics and Finance, 28(1), 5-13.

Forgas, J. P. (1998): On being Happy and Mistaken: Mood Effects on the Fundamental Attribution Error, Journal of Personality and Social Psychology, 75(2), 318-331.

French, K. R. / Poterba, J. M. (1991): Investor Diversification and International Equity Markets, American Economic Review, 81(2), 222-226.

Gambetti, E. / Giusberti, F. (2012): The Effect of Anger and Anxiety Traits on Investment Decisions, Journal of Economic Psychology, 33(6), 1059-1069.

George, J. M. / Dane, E. (2016): Affect, Emotion, and Decision Making, Organizational Behavior and Human Decision Processes, 136, 47-55.

Goetzmann, W. N. / Kumar, A. (2008): Equity Portfolio Diversification, Review of Finance, 12(3), 433-463.

Grablea, J. E. / Roszkowski, M. J. (2008): The Influence of Mood on the Willingness to Take Financial Risks, Journal of Risk Research, 11(7), 905-923.

Gross, J. J. / Levenson, R. W. (1995): Emotion Elicitation Using Films, Cognition and Emotion, 9(1), 87-108.

Gubaydullina, Z. / Spiwoks, M. (2015): Correlation Neglect, Naïve Diversification, and Irrelevant Information as Stumbling Blocks for Optimal Diversification, Journal of Finance and Investment Analysis, 4(2), 1-19.

Guiso, L. / Haliassos, M., / Jappelli, T. (2002): Household Portfolios, MIT Press, Cambridge.

Hedesstrom, T. M. / Svedsäter, H. / Gärling, T. (2006): Covariation Neglect among Novice Investors, Journal of Experimental Psychology: Applied, 12(3), 155-165.

Hibbert, A. M. / Lawrence, E. R. / Prakash, A. J. (2012): Can Diversification be Learned?, Journal of Behavioral Finance, 13(1), 38-50.

Hirshleifer, D. (2001): Investor Psychology and Asset Pricing, 56(4), 1533-1597.

Hirshleifer, D. / Shumway, T. (2003): Good Day Sunshine: Share Returns and the Weather, Journal of Finance, 58(3), 1009-1032.

Holt, C. A. / Laury, S. K. (2002): Risk Aversion and Incentive Effects, American Economic Review 92(5), 1644-1655.

Huber, J. / Kirchler, M. / Stöckl, T. (2010): The Hot Hand Belief and the Gambler's Fallacy in Investment Decisions under Risk, Theory and Decision, 68(4), 445-462.

Huberman, G. / Sengmueller, P. (2004): Performance and Employer Share in 401(k) Plans, Review of Finance, 8(3), 403-443.

Ifcher, J. / Zarghamee, H. (2011): Happiness and Time Preference: The Effect of Positive Affect in a Random-Assignment Experiment, American Economic Review, 101(7), 3109-3129.

Ifcher, J / Zarghamee, H. (2014): Affect and Overconfidence: A Laboratory Investigation, Journal of Neuroscience, Psychology, and Economics, 7(3), 125-150.

Isen, A. (2000): Positive Affect and Decision Making, Lewis, M. / Haviland-Jones, J., Handbook of Emotion, New York.

Johnson, E. J. / Tversky, A. (1983): Affect, Generalization, and the Perception of Risk, Journal of Personality and Social Psychology, 45(1), 20-31.

Kallir, I / Sonsino, D. (2009): The Neglect of Correlation in Allocation Decisions, Southern Economic Journal, 75(4), 1045-1066.

Kamstra, M. J. / Kramer, L. A. / Levi, M. D. (2003): Winter Blues: A SAD Share Market Cycle, American Economic Review, 93(1), 324-343.

Kang, S. H. / Jiang, Z. / Lee, Y. / Yoon, S.-M. (2010): Weather Effects on the Returns and Volatility of the Shanghai Share Market, Physica A: Statistical Mechanics and its Applications, 389(1), 91-99.

Kaplanski, G. / Levy, H. / Veld, C. / Veld-Merkoulova, Y. (2015): Do Happy People Make Optimistic Investors?, Journal of Financial and Quantitative Analysis, 50(1-2), 145-168.

Kaustia, M. / Rantapuska, E. (2016): Does Mood Affect Trading Behavior?, Journal of Financial Markets, 29, 1-26.

Keef, S. P. / Roush, M. L. (2007): Daily Weather Effects on the Returns of Australian Share Indices, Applied Financial Economics, 17(3), 173-184.

Kelly, M. (1995): All their Eggs in one Basket: Portfolio Diversification in U.S. Households, Journal of Economic Behavior and Organization, 27, 87-96.

Kim, J. H. (2017): Share Returns and Investors' Mood: Good Day Sunshine or Spurious Correlation?, International Review of Financial Analysis, 52, 94-103.

Kirchsteiger, G. / Rigottii, L. / Rustichini, A. (2006): Your Morals Might be your Moods, Journal of Economic Behavior and Organization, 59, 155-172.

Kliger, D. / Levy, O. (2003): Mood-Induced Variation in Risk Preferences, Journal of Economic Behavior and Organization, 52, 573-584.

Krämer, W. / Runde, R. (1997): Shares and the Weather: An Exercise in Data Mining or Yet Another Capital Market Anomaly?, Empirical Economics, 22(4), 637-641.

Krivelyova, A. / Robotti, C. (2003): Playing the Field: Geomagnetic Storms and International Share Markets, Working Paper Federal Reserve Bank of Atlanta, 2003-5.

Kuhnen, C. M. / Knutson, B. (2011): The Influence of Affect on Beliefs, Preferences, and Financial Decisions, Journal of Financial and Quantitative Analysis, 46(3), 605-626.

Lahav, Y. / Meer, S. (2012): The Effect of Induced Mood on Prices in Asset Markets – Experimental Evidence, Working Paper, Ben-Gurion University of the Negev.

Lease, R. C. / Lewellen, W. G. / Schlarbaum, G. G. (1974): The Individual Investor: Attributes and Attitudes, Journal of Finance, 29(2), 413-433.

Lee, C. J. / Andrade, E. B. (2014): Fear, Excitement, and Financial Risk-Taking, Cognition and Emotion, 29(1), 178-187.

Lee, C. J. / Andrade, E. B. (2011): Fear, Social Projection, and Financial Decision Making, Journal of Marketing Research, 48(SPL), 121-129.

Leith, K. P. / Baumeister, R. F. (1996): Why do Bad Moods Increase Self-Defeating Behavior? Emotion, Risk Tasking, and Self-Regulation, Journal of Personality and Social Psychology, 71(6), 1250-1267.

Lerner J. S. / Keltner D. (2001): Fear, Anger and Risk, Journal of Personality and Social Psychology, 81(1), 146-159.

Lerner, J. S. / Li, Y. / Valdesolo, P. / Kassam, K. S. (2015): Emotion and Decision Making, Annual Review of Psychology, 66(1), 799-823.

Loewenstein, G. (2000): Emotions in Economic Theory and Economic Behavior, American Economic Review, 90(2), 426-432.

Loewenstein, G. / Weber, E. / Hsee, C. / Welch, N. (2001): Risk as Feelings, Psychological Bulletin, 127(2), 267-286.

Lu, J. / Chou, R. K. (2012): Does the Weather have Impacts on Returns and Trading Activities in Order-Driven Share Markets? Evidence from China, Journal of Empirical Finance, 19(1), 79-93.

Lucey, B. M. / Dowling, M. (2005): The Role of Feelings in Investor Decision-Making, Journal of Economic Surveys, 19(2), 211-237.

Markowitz, H. (1952): Portfolio Selection, Journal of Finance, 7(1), 77-91.

Meulbroek, L. (2005): Company Share in Pension Plans: How Costly Is It?, Journal of Law and Economics, 48(2), 443-474.

Mitchell, O. S. / Utkus, S. P. (2002): The Role of Company Share in Defined Contribution Plans, National Bureau of Economic Research, Working Paper 9250.

Mirza, H. H. / Asghar, A. / Jam E Kausar, M. / Mushtaq, N (2012): Share Market Returns and Weather Anomaly: Evidence from an Emerging Economy, Journal of Economics and Behavioral Studies, 4(5), 239-244.

Morrin, M. / Inman, J. J. / Broniarczyk, S. M. / Nenkov, G. Y. / Reuter, J. (2012): Investing for Retirement: The Moderating Effect of Fund Assortment Size on the 1/N Heuristic, Journal of Marketing Research, 49(4), 537-550.

Nofsinger, J. R. (2005): Social Mood and Financial Economics, Journal of Behavioral Finance 6(3), 144-160.

Oswald, A. / Proto, E. / Sgroi D. (2015): Happiness and Productivity, Journal of Labor Economics, 33(4), 789-822.

Pardo, A. / Enric, V. (2002): Spanish Share Returns: Where is the Weather Effect?, European Financial Management, 9(1), 117-126.

Pham, M. T. (2007): Emotion and Rationality: A Critical Review and Interpretation of Empirical Evidence, Review of General Psychology, 11(2), 155-178.

Polkovnichenko, V. (2005): Household Portfolio Diversification: A Case for Rank-Dependent Preferences, Review of Financial Studies, 18(4), 1467-1502.

Poterba, J. M. (2003): Employer Share and 401(k) Plans, American Economic Review, 93(2), 398-404.

Rottenberg, J. / Ray, R. D. / Gross, J. J. (2007): Emotion Elicitation Using Films, Coan, J. A. / Allen, J. J. B., Handbook of Emotion Elicitation and Assessment, London, 9-28.

Schaefer, A. / Nils, F. / Sanchez, X. / Philippot, P. (2010): Assessing the Effectiveness of a Large Database of Emotion-Eliciting Films: A new Tool for Emotion Researchers, Cognition and Emotion, 24(7), 1153-1172.

Saunders, E. M. (1993): Share Prices and Wall Street Weather, American Economic Review, 83(5), 1337-1345.

Schwarz, N. (2000): Emotion, Cognition, and Decision Making, Cognition and Emotion, 14(4), 433-440.

Shiv, B. / Loewenstein, G. / Bechara, A. / Damasio, H. / Damasio, A. R. (2005): Investment Behavior and the Negative Side of Emotion, Psychological Science, 16(6), 435-439.

Shu, H.-C. (2010): Investor Mood and Financial Markets, Journal of Economic Behavior and Organization, 76(2), 267-282.

Shu, H.-C. / Hung, M. W. (2009): Effect of Wind on Share Market Returns: Evidence from European Markets, Applied Financial Economics, 19(11), 893-904.

Stanton, S. J. / Reeck, C. / Huettel, S. A. / Labar, K. S. (2014): Effects of Induced Mood on Economics Choices, Judgmend and Decision Making, 9(2), 167-175.

Symeonidis, L. / Daskalakis, G. / Markellos, R. N. (2010): Does the Weather Affect Share Market Volatility?, Finance Research Letters, 7(4), 214-223.

Stöckl, T. / Huber, J. / Kirchler, M. / Lindner, F. (2015): Hot Hand and Gambler's Fallacy in Teams: Evidence from Investment Experiments, Journal of Economic Behavior and Organization, 117, 327-339.

Tiedens, L. Z. / Linton, S. (2002): Judgment under Emotional Certainty and Uncertainty: The Effects of Specific Emotions on Information Processing, Journal of Personality and Social Psychology, 81(6), 973-988.

Tufan, E. / Hamarat, B. (2004): Do Cloudy Days Affect Share Exchange Returns: Evidence from the Istanbul Share Exchange, Journal of Naval Science and Engineering, 2(1), 117-126.

Vohs, K. / Baumeister, R. / Loewenstein, G. (2007): Do Emotions Help or Hurt Decision Making? A Hedgefoxian Perspective, New York.

Weber, E. U. / Siebenmorgen, N. / Weber, M. (2005): Communication Asset Risk: How Name Recognition and the Format of Historic Volatility Information Affect Risk Perception and Investment Decisions, Risk Analysis, 25(3), 597-609.

Westermann, R. / Spies, K. / Stahl, G. / Hesse, A. (1996): Relative Effectiveness and Validity of Mood Induction Procedures: A Meta-Analysis, European Journal of Social Psychology, 26(4), 557-580.

Yuen, K. S .L. / Lee, T. M. C. (2003): Could Mood State Affect Risk-Taking Decisions? Journal of Affective Disorders, 75(1), 11-18.

Appendix 1: Instructions for the Experiment

The Game

In the first section you will take diversification decisions for a portfolio in four different tasks. In each task you receive four free shares. Two companies are available to choose from (e.g. company K and company L). You can then choose whether you want to have 4 K shares, 4 L shares, 3 K shares and 1 L share, 3 L shares and 1 K share, or 2 K shares and 2 L shares. The dividend payments which your four shares yield in 2016 are paid out to you. Movements in the prices of the shares are of no significance to you. In the second section you make ten decisions for a lottery draw.

Before each task you are shown a short film excerpt lasting no longer than a minute. In each task you have five minutes time to enter your decisions. You will be given detailed information on the tasks in the respective sections.

Payment

You will receive a basic payment of €1.50. In addition, you can receive dividend payments up to €40 in the first section, and up to €3.85 in the lottery in the second section. In total you can thus earn up to €45.35. Payment is made at the end of the experiment.

Information

Please remain quiet during the experiment.

Please do not look at your neighbor's screen.

<u>No</u> aids are permitted (calculators, smartphones etc.). All electronic devices must be switched off.

Please note the respective time limits given on the upper right of the screen. If you do not enter anything during this time you will not receive any payment for the respective task.

Appendix 2: Test Questions for the Game

Multiple choice test questions:

Test question 1: What is your task in this game?

ʘ Solving mathematical problems.

ʘ Making diversification decisions and participating in a lottery. (correct)

ʘ Making economic forecasts.

Test question 2: How many companies are represented in each task and how many free shares do you receive?

ʘ Four companies are represented in each task, and there are 2 free shares to choose from.

ʘ Two companies are represented in each task, and there are 2 free shares to choose from.

ʘ Two companies are represented in each task, and there are 4 free shares to choose from. (correct)

Test question 3: What does the payment in the first section depend on?

ʘ On the movement of the share prices.

ʘ On the dividend payments. (correct)

ʘ On the level of the DAX.

Test question 4: How many possibilities are there for the diversification of your portfolio in each task?

ʘ 2

ʘ 4

ʘ 5 (correct)

Appendix 3: Test Questions on the Lottery

Multiple-choice test questions:

Test question 1: How high are the minimum and maximum payoffs in the lottery?

ʘ The minimum payment is €0.00 and the maximum payment is €1.60.

ʘ The minimum payment is €0.10 and the maximum payment is €3.85. (correct)

ʘ The minimum payment is €0.10 and the maximum payment is €1.60.

Test question 2: If the roll of the dice selects the 7th decision, you have chosen variant A in the 7th decision, and you have drawn a white table tennis ball from the pot, how much is your payoff?

ʘ €0

ʘ €2

ʘ €1.60 (correct)

Test question 3: If the roll of the dice selects the 10th decision, how many white table tennis balls are in the pot?

ʘ 10

ʘ 0 (correct)

ʘ 5

Test question 4: If the roll of the dice selects the fourth decision, how many yellow table tennis balls are in the pot?

ʘ 6

ʘ 0

ʘ 4 (correct)

Appendix 4: Measuring Risk Preferences Based on Holt and Laury (2002)

The lottery test used by Holt and Laury (2002) was slightly modified here. In the lottery there are 10 risk stages and two variants. Variant A (risk-averse) with the possibility of payments of €1.60 or €2, and variant B (risk-loving) with the payment possibilities €0.10 or €3.85. The lower payment is made when a white table tennis ball is drawn, and the higher payment if a yellow table tennis ball is drawn from the pot (and replaced). The risk levels indicate how many white and yellow balls are in the pot (Table 4). For every risk stage the subjects have to decide whether they want to choose variant A or B.[4] The risk level is determined with a ten-sided die. The higher the risk level, the higher the probability of receiving a larger payment. Two examples: At risk level 1, 9 white table tennis balls and one yellow table tennis ball are in the pot. At risk level 8, 9 two white table and eight yellow table tennis balls are in the pot.

Table A-1: Lottery

Risk Level	Variant A:				Variant B:				Decision A or B
	p(€2) yellow		p(€1.60) white		p(€3.85) yellow		p(€0.10) white		
1	10%	€2	90%	€1.60	10%	€3.85	90%	€0.10	
2	20%	€2	80%	€1.60	20%	€3.85	80%	€0.10	
3	30%	€2	70%	€1.60	30%	€3.85	70%	€0.10	
4	40%	€2	60%	€1.60	40%	€3.85	60%	€0.10	
5	50%	€2	50%	€1.60	50%	€3.85	50%	€0.10	
6	60%	€2	40%	€1.60	60%	€3.85	40%	€0.10	
7	70%	€2	30%	€1.60	70%	€3.85	30%	€0.10	
8	80%	€2	20%	€1.60	80%	€3.85	20%	€0.10	
9	90%	€2	10%	€1.60	90%	€3.85	10%	€0.10	
10	100%	€2	0%	€1.60	100%	€3.85	0%	€0.10	

After the experiment the dividend payments are determined by tossing a coin. In order to determine the winnings of the lottery, first, a ten-sided die is thrown to decide the risk level. Then the table tennis balls are placed in the pot in accordance with the risk level and every subject draws a table tennis ball from the pot. The ball is then returned to the pot.

[4] Before the selection of the lottery, the subjects have answered four test questions. The test questions check whether the subjects have understood the lottery.

Appendix 5: Screenshot of the Experiment with Z-Tree (Reconstructed in Order to Improve Readability)

Screenshot 1: Measurement of Mood before the Experiment (Reconstructed in Order to Improve Readability)

How are you feeling now? Please mark the adequate number!

very bad Θ Θ Θ Θ Θ Θ Θ Θ Θ Θ very good

1–2–3–4–5–6–7–8–9–10

O.K.

Figure A-1: Measurement of Mood before the Experiment

Screenshot 2: Test Questions (Reconstructed in Order to Improve Readability)

Please answer the following test questions about the game:

Test question 1: What is your task in this game?

Θ Solving mathematical problems.

Θ Making diversification decisions and participating in a lottery.

Θ Making economic forecasts.

Test question 2: How many companies are represented in each task and how many free shares do you receive?

Θ 4 companies are represented in each task, and there are 2 free shares to choose from.

Θ 2 companies are represented in each task, and there are 2 free shares to choose from.

Θ 2 companies are represented in each task, and there are 4 free shares to choose from.

Test question 3: What does the payment in the first section depend on?

ʘ On the movement of the share prices.

ʘ On the dividend payments.

ʘ On the level of the DAX.

Test question 4: How many possibilities are there for the diversification of your portfolio in each task?

ʘ 2

ʘ 4

ʘ 5

O.K.

Figure A-2: Test Questions

Screenshot 3: Manipulation Check after the Attempt to Influence Mood (Reconstructed in Order to Improve Readability)

Which emotions did you experience while watching the movie clip?

Please mark one number accordingly!

very negative ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ very positive

1–2–3–4–5–6–7–8–9–10

O.K.

Figure A-3: Manipulation Check after the Attempt to Influence Mood

Screenshot 4: Diversification Decision (Reconstructed in Order to Improve Readability)

You can choose between two shares (share A and share B) of a specific sector of industry. You can read in the table how high the dividend payments for both shares were during the past 10 years. If the economic situation in the sector is good, the dividend of share A is €3 and that of share B is €2. If the economic situation in the sector is poor, the dividend of share A is €0 and that of share B is €1. The economic trend in this sector can vary from year to year and has to be viewed as a random process: The probability of a good or poor economic situation is 50% respectively.

Year	**Share A**	**Share B**
2006	**€3**	**€2**
2007	**€0**	**€1**
2008	**€3**	**€2**
2009	**€0**	**€1**
2010	**€0**	**€1**
2011	**€3**	**€2**
2012	**€3**	**€2**
2013	**€3**	**€2**
2014	**€0**	**€1**
2015	**€3**	**€2**
2016	**€?**	**€?**

You receive four free shares. You can choose whether you want to have 4 A shares, 4 B shares, 3 A shares + 1 B share, 3 B shares + 1 A share or 2 A shares and two B shares. The dividend payments which your four shares yield in 2016 are paid out to you. The dividend payments for 2016 are determined by tossing a coin. If it is heads, this means a good economic situation, while tails means a poor economic situation or a weak year. Movements in the prices of the shares are of no significance to you.

Make your selection now. **I select:**

ʘ 4 A shares.
ʘ 4 B shares.
ʘ 3 A shares + 1 B share.
ʘ 3 B shares + 1 A share.
ʘ 2 A shares + 2 B shares.

Please give brief reasons for your selection. These reasons have no effect on the payment, so you can write down your thoughts openly and honestly.

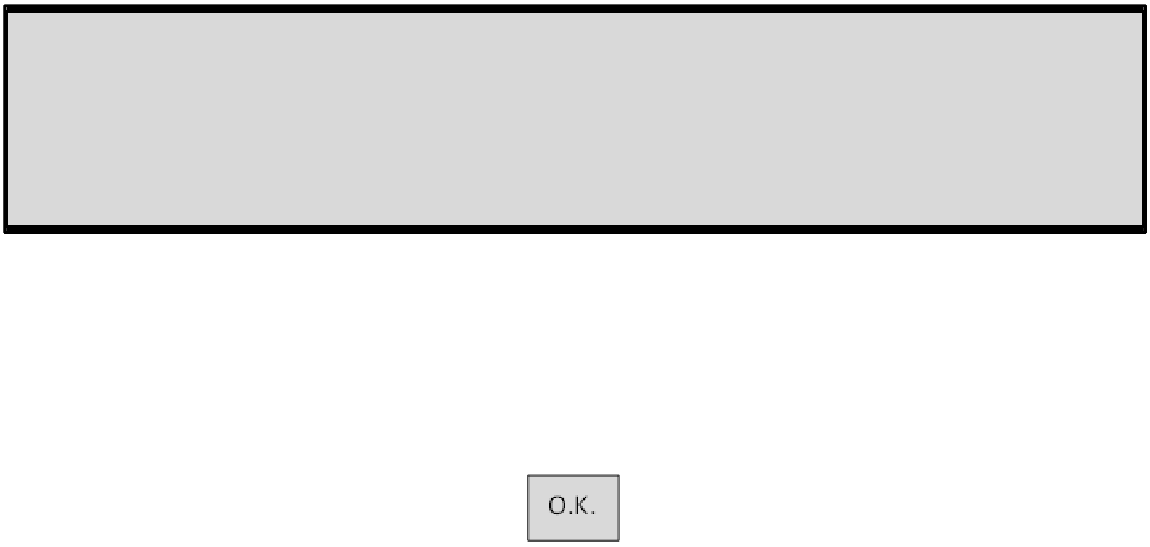


Figure A-4: Diversification Decision

Screenshot 5: Lottery (Reconstructed in Order to Improve Readability)

You make your decisions on the next page. Each decision is a choice between variant A and variant B. Each variant is a type of lottery with different payoff sums and probabilities of occurrence. You make ten decisions. Enter your respective decision in the right-hand column of the table. One of these decisions will be used to determine your payoff in the lottery. This is done as follows: After you have made all ten decisions, a ten-sided dice is thrown to determine which of the ten decisions will be used. Each of the decisions thus has the same 10% probability of being used. Then the lottery you have chosen (A or B) is played.

The probability of occurrence is simulated with the help of a pot with table tennis balls: In a pot with 10 table tennis balls, the number of yellow balls indicates the probability with which the higher payoff sum will occur. Example for decision no. 8: In a pot with 10 table tennis balls, 8 are yellow and 2 are white. The probability that a randomly drawn table tennis ball is yellow is thus 80%. If the table tennis ball drawn card is yellow, you receive €2 in variant A and €3.85 in variant B. If, however, the table tennis ball drawn is white, you receive €1.60 in variant A and €0.10 in variant B. You thus make ten decisions (either for lottery A or B). One of these is randomly chosen (with a die) and played (with a pot and ten table tennis balls) – the result determines your payoff in the lottery. Please answer the following test questions about the lottery before you make your decisions.

Risk Level	Variant A:				Variant B:			
	p(€2) yellow		p(€1.60) white		p(€3.85) yellow		p(€0.10) white	
1	10%	€2	90%	€1.60	10%	€3.85	90%	€0.10
2	20%	€2	80%	€1.60	20%	€3.85	80%	€0.10
3	30%	€2	70%	€1.60	30%	€3.85	70%	€0.10
4	40%	€2	60%	€1.60	40%	€3.85	60%	€0.10
5	50%	€2	50%	€1.60	50%	€3.85	50%	€0.10
6	60%	€2	40%	€1.60	60%	€3.85	40%	€0.10
7	70%	€2	30%	€1.60	70%	€3.85	30%	€0.10
8	80%	€2	20%	€1.60	80%	€3.85	20%	€0.10
9	90%	€2	10%	€1.60	90%	€3.85	10%	€0.10
10	100%	€2	0%	€1.60	100%	€3.85	0%	€0.10

Test question 1: How high are the minimum and maximum payoffs in the lottery?

ʘ The minimum payoff is €0.00 and the maximum payoff is €1.60.

ʘ The minimum payoff is €0.10 and the maximum payoff is €3.85.

ʘ The minimum payoff is €0.10 and the maximum payoff is €1.60.

Test question 2: If the roll of the dice selects the 7th decision, you have chosen variant A in the 7th decision, and you have drawn a white table tennis ball from the pot, how high is your payoff?

ʘ €0.00

ʘ €2.00

ʘ €1.60

Test question 3: If the roll of the dice selects the 10th decision, how many white table tennis balls are in the pot?

ʘ 10

ʘ 0

ʘ 5

Test question 4: If the roll of the dice selects the 4th decision, how many yellow table tennis balls are in the pot?

ʘ 6

ʘ 0

ʘ 4

O.K.

Figure A-5: Lottery

Screenshot 6: Field for the Entry of the Lottery Decisions (Reconstructed in Order to Improve Readability)

Risk Level	Variant A:				Variant B:			
	p(€2) yellow		p(€1.60) white		p(€3.85) yellow		p(€0.10) white	
1	10%	€2	90%	€1.60	10%	€3.85	90%	€0.10
2	20%	€2	80%	€1.60	20%	€3.85	80%	€0.10
3	30%	€2	70%	€1.60	30%	€3.85	70%	€0.10
4	40%	€2	60%	€1.60	40%	€3.85	60%	€0.10
5	50%	€2	50%	€1.60	50%	€3.85	50%	€0.10
6	60%	€2	40%	€1.60	60%	€3.85	40%	€0.10
7	70%	€2	30%	€1.60	70%	€3.85	30%	€0.10
8	80%	€2	20%	€1.60	80%	€3.85	20%	€0.10
9	90%	€2	10%	€1.60	90%	€3.85	10%	€0.10
10	100%	€2	0%	€1.60	100%	€3.85	0%	€0.10

Now please make the ten decisions:

Which variant would you rather play – A or B?

No. 1: ʘ A ʘ B

No. 2: ʘ A ʘ B

No. 3: ʘ A ʘ B

No. 4: ʘ A ʘ B

No. 5: ʘ A ʘ B

No. 6: ʘ A ʘ B

No. 7: ʘ A ʘ B

No. 8: ʘ A ʘ B

No. 9: ʘ A ʘ B

No.10: ʘ A ʘ B

O.K.

Figure A-6: Field for the Entry of the Lottery Decisions

Chapter 4

MEASUREMENT OF RISK PREFERENCE

with Thomas Nahmer, Markus Spiwoks and Zulia Gubaydullina
contribution Ibrahim Filiz: 35%

Submitted to:
Journal of Risk and Uncertainty

Published in:
Wolfsburg Working Papers, 18-01, June 2018.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-1, June 2018.

Abstract

The procedures previously used to determine risk preference (risk-averse, risk-neutral or risk-loving) exhibit a number of weaknesses. In part, they are so complex and sophisticated that the subjects frequently give spontaneous, ill-considered answers. In this way, their actual risk preference can often not be correctly determined. In addition, in this process there are situations and circumstances in which it is not possible to clearly assign subjects to one of the three categories of risk preference. In addition, with the previous approaches, loss aversion – which has an important influence on risk preference – is not taken into consideration, or only insufficiently. We propose here a new procedure to determine risk preference which is (1) extremely simple and clear, which (2) enables unambiguous differentiation between risk-averse, risk-neutral and risk-loving subjects, and which (3) takes the influence of loss aversion on risk preference into account in an appropriate way.

Keywords

Risk preference, loss aversion, portfolio choice, diversification behavior, behavioral finance, experimental research

JEL Classification

B49, C91, G11, G40

1 Introduction

Markowitz (1952) shows that for risk-averse subjects it usually makes sense to hold diversified securities portfolios. However, there are many empirical findings which reveal that under-diversified portfolios are very frequently held.[1] Experimental economic research examines this contradiction and finds many reasons why suboptimal decisions are frequently made with regard to diversification.[2]

Meaningful experimental results on diversification behavior can normally only be obtained if clarity about the risk preference of the subjects can be achieved: Because that which is meaningful for a risk-averse subject can be complete nonsense for a risk-loving subject, and vice-versa. In the meantime there are a whole range of procedures available to determine risk preference.[3]

In our view, a good procedure for determining risk preference must above all comply with three criteria:

1. It must be a simple and clear procedure.
2. It must be possible to clearly and unambiguously differentiate between risk-averse, risk-neutral and risk-loving subjects.
3. The influence of loss aversion on risk preference should not be neglected.

We consider these three criteria to be key. (1) The procedure has to be simple and clear so that we can really record the risk preferences of the subjects. In the case of complex and confusing decision-making situations, subjects frequently lose patience and then give spontaneous, ill-considered answers. This can sometimes lead to a blurring of their risk preference rather than it being revealed. (2) In the previous approaches used, there are certain combinations of circumstances in which risk-neutral, risk-averse and risk-loving subjects make – with good reason – the same decisions. In that case, it is not possible to differentiate between the three forms of risk preference. (3) As we will show later on, risk preference is significantly determined by the possibility of suffering losses. Procedures to measure risk preference which do not contain the possibility of losses systematically underestimate the proportion of risk-averse subjects.

Our study is divided up into four sections. First, we evaluate the previous approaches against the background of the three criteria we have postulated. In the following chapter we present our new procedure to measure risk preference. Using an experimental investigation we sub-

---

[1] See, for example, Dimmock et al. (2016), Anderson (2013), Hibbert, Lawrence and Prakash (2012), Goetzmann and Kumar (2008), Meulbroek (2005), Polkovnichenko (2005), Huberman and Sengmueller (2004), Agnew, Balduzzi and Sundén (2003), Guiso, Haliassos and Japelli (2002), Benartzi (2001), Benartzi and Thaler (2001), Barber and Odean (2000), Bode, van Echelpoel and Sievi (1994), Blume and Friend (1975), and Lease, Lewellen and Schlarbaum (1974).

[2] See, for example, Filiz et al. (2018), Gubaydullina and Spiwoks (2015), Fernandes (2013), Morrin et al. (2012), Rieger (2012), Eyster and Weizsäcker (2011), Baltussen and Post (2011), Kallir and Sonsino (2009), Hedesstrom, Svedsater and Garling (2006), Fellner, Güth and Maciejovsky (2004), Choi, Laibson and Madrian (2009), Weber, Siebenmorgen and Weber (2005).

[3] See, for example, Lönnqvist et al. (2015), Charness, Gneezy and Imas (2013), Crosetto and Filippin (2013), Eckel and Grossmann (2002, 2008), Lejuez et al. (2002), Holt and Laury (2002), and Gneezy and Potters (1997).

sequently show that loss aversion should not be neglected when measuring risk preference. In the final chapter we summarize the most important results of the investigation.

2 The Previous Approaches and their Weaknesses

In the following section we discuss the approaches of Holt and Laury (2002), Eckel and Grossman (2008), and Crosetto and Filippin (2013). In addition, we briefly consider the approaches used by Lejuez et al. (2002), Gneezy and Potters (1997), the DOSPERT questionnaire created by Weber, Blais and Betz (2002), and the socio-economic panel (Schupp and Wagner, 2002; Wagner, Burkhauser and Behringer, 1993).

2.1 The Multiple Price List Method of Holt and Laury (2002)

In the multiple price list method of Holt and Laury (2002), subjects are asked to make ten decisions choosing between two lotteries in each case (Table 1). As the first decision, lottery A ($2.00 with a probability of 10% or $1.60 with a probability of 90%) is set against lottery B ($3.85 with a probability of 10% and $0.10 with a probability of 90%). The subject has to decide whether they would play lottery A or lottery B. This is followed by the other nine comparisons between lottery A and lottery B. From the sequence of the ten decisions, conclusions about the risk preferences of the subject are then drawn.

Table 1: The Lottery Alternatives of Holt and Laury (2002)

	Lottery A				Lottery B			
No.	Prob.	Event	Prob.	Event	Prob.	Event	Prob.	Event
1	10%	$2.00	90%	$1.60	10%	$3.85	90%	$0.10
2	20%	$2.00	80%	$1.60	20%	$3.85	80%	$0.10
3	30%	$2.00	70%	$1.60	30%	$3.85	70%	$0.10
4	40%	$2.00	60%	$1.60	40%	$3.85	60%	$0.10
5	50%	$2.00	50%	$1.60	50%	$3.85	50%	$0.10
6	60%	$2.00	40%	$1.60	60%	$3.85	40%	$0.10
7	70%	$2.00	30%	$1.60	70%	$3.85	30%	$0.10
8	80%	$2.00	20%	$1.60	80%	$3.85	20%	$0.10
9	90%	$2.00	10%	$1.60	90%	$3.85	10%	$0.10
10	100%	$2.00	0%	$1.60	100%	$3.85	0%	$0.10

Prob. = probability of occurrence; event = random event.

The main problem of this approach is the complexity of the decision-making situation. Neither the expected returns nor the extent of the risk exposure of the alternatives A and B are clearly recognizable for the subjects. Accordingly, many subjects decide randomly or based on a gut feeling. In this situation, it frequently occurs that ten decisions are made where the decision-making process cannot be clearly interpreted. Jacobson and Petrie (2009) as well as Charnes and Viceisza (2011) show that between 55% and 75% of the decision-making processes cannot be clearly interpreted. Charness et al. (2018) and Dave et al. (2010) also point out additional uncertainties in the interpretation of results.

The approach used by Holt and Laury (2002) becomes somewhat clearer if one considers the expected returns and the risk (standard deviation) of the ten lottery alternatives (Table 2). In the first lottery alternative, lottery A has an expected return of $1.64 and a standard deviation of 0.12. Lottery B has an expected return of $0.48 and a standard deviation of 1.13. Risk-neutral subjects orientate themselves solely towards expected returns and therefore decide in favor of alternative A. A risk-averse subject will also decide in favor of alternative A, because here the expected returns is higher and at the same time the risk is lower than that of alternative B. But how would a risk-loving subject decide? Expected returns would speak for alternative A, but the risk speaks for alternative B. How a risk-loving subject decides therefore depends on the extent of their appetite for risk. Subjects with a great appetite for risk will choose alternative B because the higher risk more than compensates for the lower expected return. Subjects with a mild appetite for risk will choose alternative A because the higher expected return more than compensates for the lower risk.

Table 2: The Expected Returns and Risk (Standard Deviation) of the Lottery Alternatives used by Holt and Laury (2002) and the Preferences of Risk-Neutral, Risk-Averse and Risk-Loving Subjects

	Lottery A		**Lottery B**		Preference	Preference	Preference
No.	E(A)	SD	E(B)	SD	Risk-Neutral	Risk-Averse	Risk-Loving
1	1.64	0.12	0.48	1.13	A	A	A or B
2	1.68	0.16	0.85	1.50	A	A	A or B
3	1.72	0.18	1.23	1.72	A	A	A or B
4	1.76	0.20	1.60	1.84	A	A	A or B
5	1.80	0.20	1.98	1.88	B	A or B	B
6	1.84	0.20	2.35	1.84	B	A or B	B
7	1.88	0.18	2.73	1.72	B	A or B	B
8	1.92	0.16	3.10	1.50	B	A or B	B
9	1.96	0.12	3.48	1.13	B	A or B	B
10	2.00	0.00	3.85	0.00	B	B	B

E(A) = expected returns of lottery A; E(B) = expected returns of lottery B; SD = standard deviation.

In the fifth decision, risk-neutral subjects choose alternative B, because the expected return of $1.98 is higher than that of alternative A ($1.80). Risk-loving subjects also choose alternative B because here both expected return and risk are higher than in alternative A. But how will risk-averse subjects react? The expected return would speak for alternative B, but the risk speaks for alternative A. How the subject decides now depends on the extent of their risk aversion. If they are highly risk-averse, they will choose alternative A because the lower risk offsets the lower expected return. If, however, they are only slightly risk-averse, they will decide in favor of alternative B, because the higher expected return more than compensates for the higher risk.

Now the following question arises: How should subjects be classified who always prefer alternative A in the first four decisions and then prefer alternative B in the last six decisions? These can be either risk-neutral, risk-averse or risk-loving subjects (see Table 2). It cannot therefore be guaranteed that they will be unambiguously assigned to one of the three possible categories of risk preference (risk-averse, risk-neutral or risk-loving).

The approach used by Holt and Laury (2002) therefore does not satisfy any of the three requirements which we formulated at the beginning for reliable determining of risk preference: (1) It is complex and unclear. (2) It does not lead to a clear differentiation between risk-neutral, risk-averse and risk-loving subjects. (3) It does not take the possibility of losses into account.

2.2 The Approach used by Eckel and Grossman (2008)

The approach used by Eckel and Grossman (2008) has the advantage that the decision-making situation is significantly clearer than in the case of Holt and Laury (2002). The subjects decide in favor of one of five possible lotteries. In each lottery there are two possible events which each have a probability of occurrence of 50%. From lottery 1 to lottery 5, the expected values rise, as do the risks (Table 3, Figure 2).

In the loss treatment the participants receive $6 for filling in a questionnaire[4] in the run-up to the lottery. They can lose part of this $6 in lottery 4 and all of it in lottery 5. In order to remunerate all subjects uniformly, the expected values are $6 higher in the no-loss treatment. The approach used by Eckel and Grossman (2008) thus also takes the possibility of losses into consideration.

Table 3: Lottery Alternatives in Eckel and Grossman (2008)

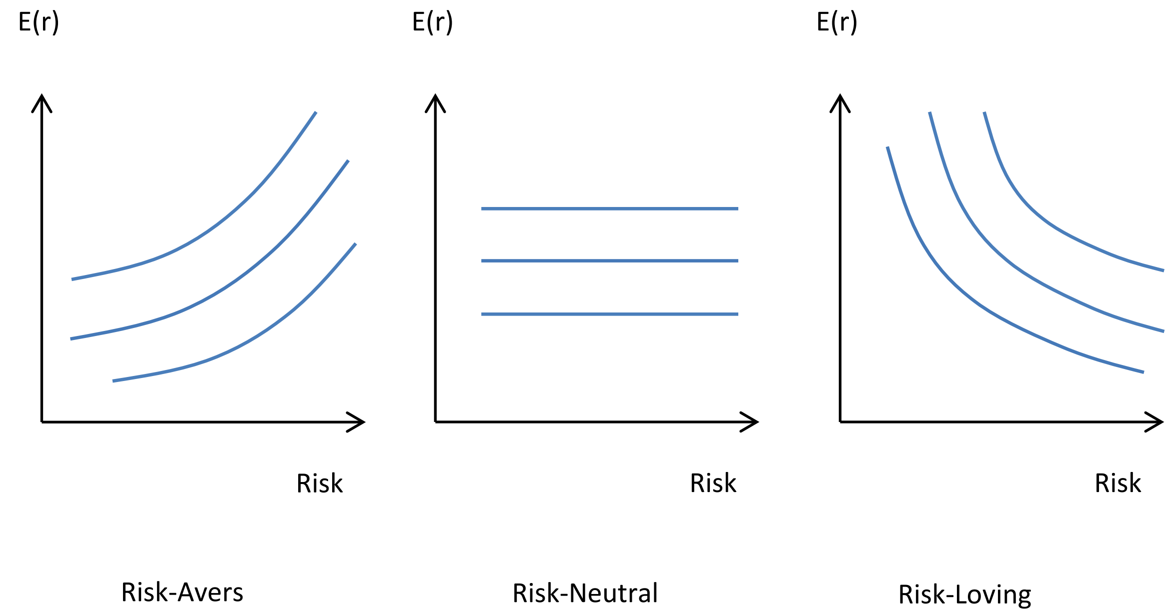
No.	Event	Prob	Return Loss	Return No-Loss	E(r) Loss	E(r) No-Loss	Risk SD
1	A	50%	$10	$16	$10	$16	0
	B	50%	$10	$16			
2	A	50%	$18	$24	$12	$18	6
	B	50%	$6	$12			
3	A	50%	$26	$32	$14	$20	12
	B	50%	$2	$8			
4	A	50%	$34	$40	$16	$22	18
	B	50%	$-2	$4			
5	A	50%	$42	$48	$18	$24	24
	B	50%	$-6	$0			

Event = possible random event; Prob = probability of occurrence; Return Loss = payoff of the coincidental events in the loss treatment; Return No-loss = payoff of the coincidental events in the no-loss treatment; E(r) loss = expected value of the payoff in the loss treatment; E(r) no loss = expected value of the payoff in the no-loss treatment; SD = standard deviation.

[4] Zuckerman's sensation-seeking scale. See Zuckerman (1979, 1994).

The approach deployed by Eckel and Grossman (2008) is problematic in that the assignment of the subjects to the three categories of risk preference (risk-averse, risk-neutral and risk-loving) is by no means clear. This becomes apparent when one considers that risk-averse, risk-neutral and risk-loving subjects exhibit fundamentally diverging indifference curves. Risk-averse subjects have rising indifference curves, whereas risk-neutral subjects have absolutely horizontal indifference curves and risk-loving subjects have falling indifference curves (Figure 1).

Figure 1: The Form of the Indifference Curves for Risk-Averse, Risk-Neutral and Risk-Loving Subjects

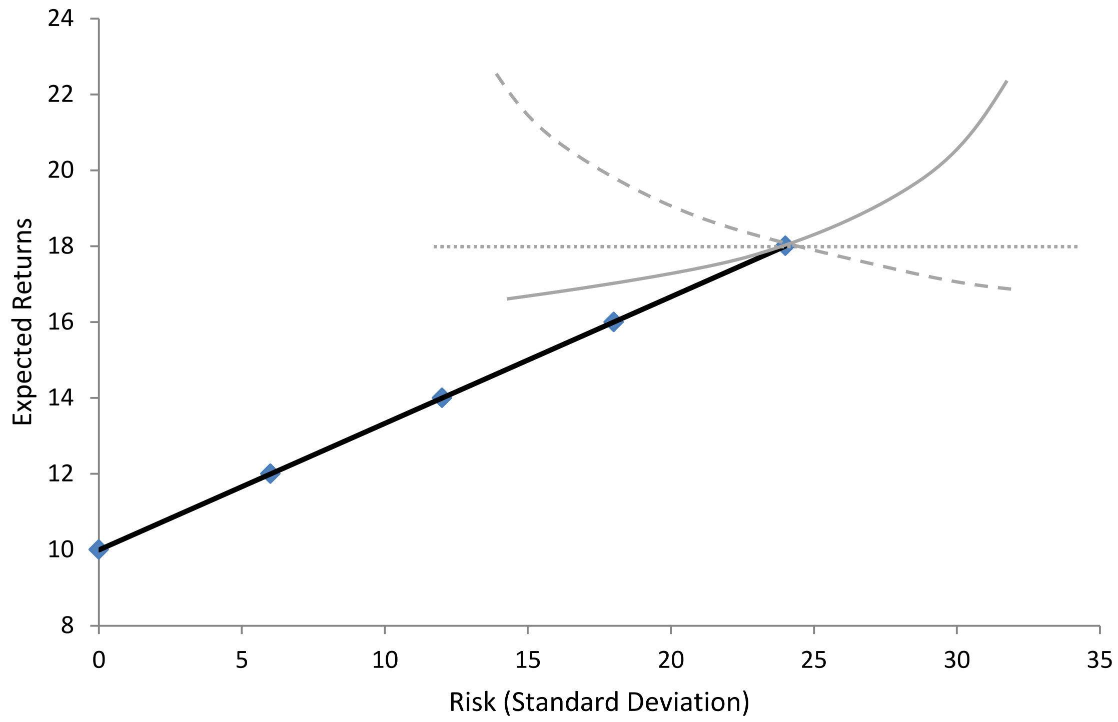


E(r) = expected value of return; risk (standard deviation).

If the space of possibilities which results from the five lotteries is considered, the following becomes recognizable: All of the subjects who choose lottery 5 can be risk-averse as well as risk-neutral or risk-loving (Figure 2).

The approach used by Eckel and Grossman (2008) thus manages to fulfill two of the three criteria we have put forward: It is a simple and clear decision-making situation and the possibility of making losses is also taken into account. However, the unambiguous identification of risk-neutral, risk-averse and risk-loving subjects is not possible.

Figure 2: Space of Possibilities in Eckel and Grossman (2008) as well as the Indifference Curves of a Risk-Averse (Unbroken Grey Line), a Risk-Neutral (Dotted Grey Line) and a Risk-Loving Subject (Dashed Grey Line)

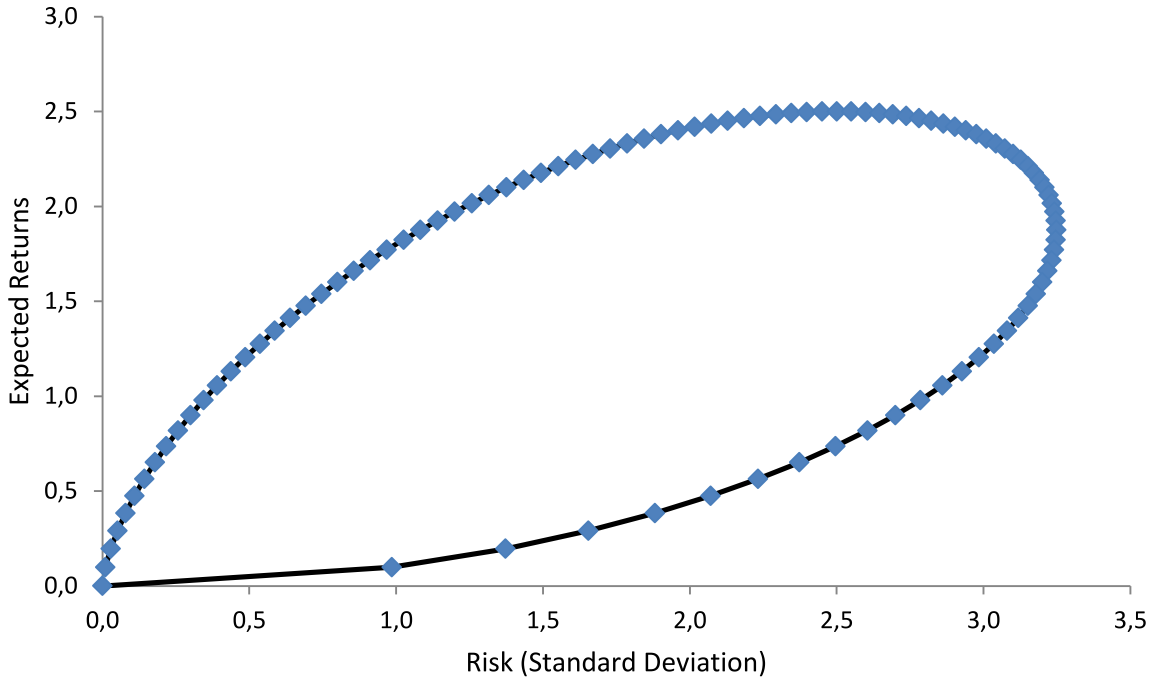


2.3 The Approach used by Crosetto and Filippin (2013)

Crosetto and Filippin (2013) have proposed the most interesting approach yet to determine risk preferences. In this approach, the participants are faced with the following decision-making situation: They have to decide how many of a total of 100 boxes they want to collect. One of the boxes contains a 'bomb'. The participants receive a payoff of €0.10 per box. After they have decided on a number of boxes (static version) or have ended the game by pressing the 'stop button' (dynamic version),[5] a number between 1 and 100 is drawn from an urn. If the number drawn is ≤ the number of collected boxes, the 'bomb' has exploded and the money is gone. If the number drawn is > the number of collected boxes, the subject receives a payment based on the multiplication of the number of boxes collected by €0.10. It can be expected that the subjects want to win as much money as possible. The more boxes they collect, the higher the payoff. At the same time, the risk of encountering the 'bomb' (number drawn ≤ the number of collected boxes) rises. The subjects thus have to weigh up how much risk is meaningful to them. The space of possibilities of this decision-making situation is shown in Figure 3.

[5] Crosetto and Filippin deployed a static basic variation and a dynamic variant. In the static variant the subjects only see a picture of 100 boxes and have to decide how many they want to collect. In the dynamic PC version the 100 boxes are shown on the screen. By pressing a start button the participants trigger the collection of one box per second until they press the stop button.

Figure 3: Space of Possibilities in Crosetto and Filippin (2013)



From the first to the 50th box, expected returns rises gradually. At the same time the risk also increases steadily. From the 50th to the 75th boxes, the risk continues to rise, whereas expected returns falls. From the 75th to the 100th boxes, the risk as well as expected returns both decrease. The highest expected return is achieved if one collects exactly 50 boxes. Risk-averse subjects – depending on their risk aversion – will choose between one and 50 boxes. Risk-loving subjects will choose between 50 and 75 boxes. Risk-neutral subjects will always collect exactly 50 boxes, because expected return reaches its maximum level there. The efficient frontier of the space of possibility thus extends from one to 75 boxes. The section from 76 to 100 boxes, however, is the non-efficient part of the space of possibility.

The great advantage of this approach is the enormous clarity of the decision-making situation. In addition, loss opportunities can also be implemented easily, which Crosetto and Filippin (2013) in fact do in one of the treatments.

Nevertheless, some criticism can be made: (1) If a subject collects exactly 50 boxes it is not possible to recognize whether they are risk-averse, risk-neutral or risk-loving. While it is true that all risk-neutral subjects will collect exactly 50 boxes, one cannot conclude that all subjects who collect 50 boxes are risk-neutral. In view of the maximum expected return, slightly risk-averse or slightly risk-loving subjects could also consider 50 boxes to be the most attractive option.[6] (2) The decision-making situation is indeed very clear, but it is not simple. How many subjects recognize that the maximum expected return can be found at exactly 50 boxes? And how many subjects realize what the risk (standard deviation) is for the 100 different

[6] Around 14% of the subjects decide to collect exactly 50 boxes. This means that a notable proportion of the subjects cannot be assigned unambiguously to one of the three categories (risk-averse, risk-neutral and risk-loving).

possibilities? A considerable amount of calculating is required to work that out. (3) How should subjects who collect more than 75 boxes be characterized? Those persons who move in the non-efficient part of the space of possibilities are also either risk-averse, risk-neutral or risk-loving. There is no other possibility. However, which of these three alternatives they fit into cannot be said, because each subject who collects more than 75 boxes is obviously not aware of the shape of the space of possibilities.

The three requirements we have put forward for a good process to determine risk preferences are not completely fulfilled here. The decision-making situation is clear, but it is not exactly simple. It is not possible in every case to unambiguously assign subjects to one of the three categories of risk preference (risk-averse, risk-neutral and risk-loving). On the positive side, introducing a risk of loss is simple, which Crosetto and Filippin (2013) in fact do in one of the treatments.

2.4 Further Approaches

The method used by Lejuez et al. (2002) aims to create a relative comparison of risk preference between two or more subjects. However, his aim is not to assign them to one of the three categories of risk preference (risk-averse, risk-neutral and risk-loving). The decision-making situation is designed as follows: A balloon and a pump are shown on a computer screen. With every click of a mouse, the balloon is pumped up a bit more and the participant receives €0.05. Their credit is shown on a temporary account. The subject can stop pumping at any time. If the balloon bursts, the credit accumulated is lost. A total of 90 rounds of the game are played, in which there are three different colored balloons (blue, yellow and orange). The three colors represent different probabilities of bursting. The subjects are only informed that the three different-colored balloons have a different bursting point, and that the balloon can even burst on the first pump. The average number of pumps made is used as an indicator for risk preference. As no advance information is provided about expected returns and risk, this method is not suitable for assigning subjects to one of the three categories of risk preference: Only a relative comparison between subjects can take place. For example, it can be established that subject A acts more cautiously than subject B. However, whether subject A is risk-averse and subject B is risk-loving remains unclear. Subject A could be strongly risk-averse and subject B could be slightly risk-averse. Or subject A is slightly risk-loving and subject B is highly risk-loving. This remains unclear.

The method used by Gneezy and Potters (1997) examines which proportion of their portfolio subjects invest in a risky asset. To do so, they are asked which proportion of 200 cents they want to bet on in a lottery which there is a probability of two thirds that they will lose the amount and a probability of one third that they will win two and a half times the amount. So if they win they retain the amount they wager plus two and a half times the amount as winnings. The lottery thus has a positive expected value. A total of nine rounds are played. In treatment H, the participants decide separately for each round which proportion of the 200 cents they want to bet. In treatment L, decisions are made in advance for each of three rounds of the game. The amount which is wagered thus remains constant for three rounds. Depending on the treatment, the participants are informed about the (aggregated) results after one or three lotteries and then they bet again. It is shown that the average amount placed as a bet in treatment L (decision in advance) is greater than in treatment H (separate decision for each round). The results reveal that an investment period spread over several periods leads to a larger proportion of the investor's assets being invested in a risky asset. In an adapted form, Charness and Gneezy (2010) established that the participants of the experiment would pay in order to have more frequent opportunities to change the composition of their portfolio. However, the structure of the experiment is not suited to assigning the subjects to one of the three categories of risk preference (risk-averse, risk-neutral and risk-loving). Once again, the approach can only be used to establish that subject A acts more cautiously than subject B. The same issue arises as in the case of Lejuez et al. (2002).

Another way of determining individual risk preference is to interview the subjects. A good example of this is the domain-specific risk taking questionnaire (DOSPERT) developed by Weber, Blais and Betz (2002). The questionnaire relates to a large number of high-risk activities or behaviors from five fields: (1) Sports and leisure, (2) health, (3) social issues, (4) ethics, and (5) finances. The questionnaire records the probability of the respondents taking risks, their perception of these risks and of the benefit which might result from the risks taken. A total of 40 topics are evenly distributed over five fields, whereby only the field of finance is subdivided into (a) gambling and (b) investment risks. The participants estimate their own risk preference on a scale from 1 (low-risk) to 5 (high-risk). Assignation to one of the three categories of risk preference (risk-averse, risk-neutral and risk-loving) is not possible on the basis of this questionnaire. Once again, this approach can only be used to establish that subject A acts more cautiously than subject B. The same issues arise as in the case of Lejuez et al. (2002).

Another example of surveying risk preference within the framework of a questionnaire is the socio-economic panel (SOEP). Schupp and Wagner (2002) as well as Wagner, Burkhauser and Behringer (1993) describe the approach used in the questionnaire. The idea is that the interviewees provide information about their general risk preferences. Assignation to one of the three categories of risk preference (risk-averse, risk-neutral and risk-loving) is not possible on the basis of this questionnaire. The same applies to the differentiated versions of the SOEP approach (Schupp and Wagner, 2002; Wagner, Burkhauser and Behringer, 1993). These approaches can only be used to establish that subject A acts more cautiously than subject B. The same issues arise as in the case of Lejuez et al. (2002).

Lönnqvist et al. (2015) examine the time stability of various procedures for the measurement of risk preferences, while Charness, Gneezy and Imas (2013) compare different procedures for the measurement of risk preferences. However, they do not provide a different approach for the identification of risk-neutral, risk-averse and risk-loving subjects.

3 The New Approach

We propose a procedure for differentiating between risk-averse, risk-neutral and risk-loving subjects which is very clear and simple and which makes it possible to assign subjects unambiguously to the three categories of risk preference.

It deals with a decision to choose between two lotteries.[7] The subjects take a card – they can choose between taking a card from pile A or one from pile B. Both piles consist of four playing cards each. In pile A there are two cards which lead to a profit of €+4, and two cards which lead to a profit of €+6 (Figure 4). In pile B there are two cards which lead to no profit €±0, and two cards which lead to a profit of €+10 (Figure 5).

Figure 4: The Four Cards in Pile A

Figure 5: The Four Cards in Pile B

The subjects are informed that the expected return in both piles is identical at €+5. In addition, the subjects are made aware of the fact that pile A leads to results which fluctuate slightly around the expected value (low-risk), while pile B leads to results which fluctuate considerably around the expected value (high-risk). The two piles of cards containing four cards each are not only shown on the screen, but can also be seen as real playing cards on the table of the game leader. The subjects are informed that the pile of cards which they decide for (A or B) will be shuffled and that they then have to take a card. The entire survey is programmed in z-Tree (Fischbacher, 2007). However, we have decided not to program random events in z-Tree, but to carry them out analogously. In this way we want to counteract the possible suspicion that it could be a manipulated random event. The subjects see the

[7] We were inspired by Bechara et al. (1994) here.

playing cards and can be sure that there is a probability of exactly 50% that the favorable event (€+6 in pile A and €+10 in pile B) will occur. In addition, they can also be sure that there is a probability of exactly 50% that the unfavorable event (€+4 in pile A and €±0 in pile B) will occur (Table 4).

Table 4: Lottery Alternatives in the New Approach

Pile	Prob	Return	E(r)	Risk (SD)
A	50%	€+4	€+5	1.0
	50%	€+6		
B	50%	€±0	€+5	5.0
	50%	€+10		

Prob = probability; E(r) = expected value of the return; SD = standard deviation.

The shuffling of the cards is left to a machine in order to avoid the suspicion that the game leader has an influence on random events. After the cards have been shuffled, the subjects have to take one of the four cards from the pile they have chosen. They then receive the payment which is noted on this card. Test questions are used to ensure that the subjects understand the circumstances. The instructions for the game, the test questions and selected screenshots can be viewed in the appendix.

The two lotteries lead to a clear space of possibilities (Figure 6) which also permits an unambiguous assignment of the subjects to the three categories of risk preference (risk-averse, risk-neutral and risk-loving).

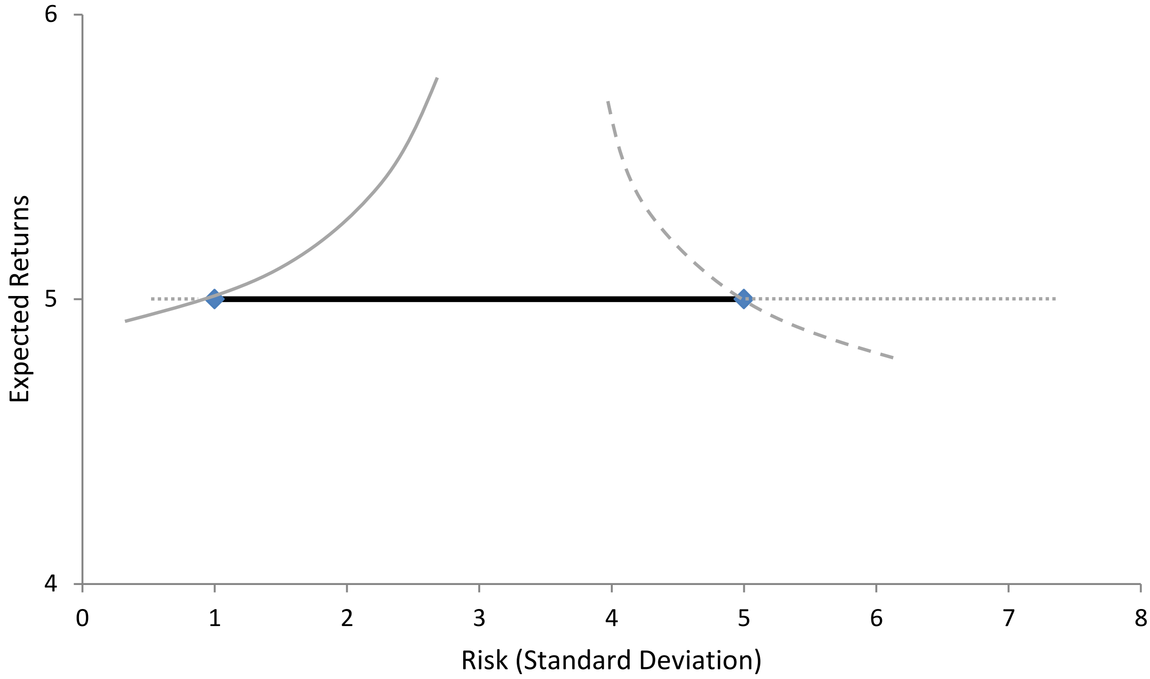
The space of opportunities consists of only two points. The left point shows the expected return and risk profile of pile A (low-risk). The right point shows the expected return and risk profile of pile B (high-risk).

The decision options for the subjects are:

- I would like to take a card from pile A.
- I would like to take a card from pile B.
- I would like to take a card. I don't mind which pile I take one from.

If one takes into account the form of the indifference curves for risk-averse, risk-neutral and risk-loving subjects (Figure 1), the decision made by the subjects leads to an unambiguous assignment to one of the three categories of risk preference: Risk-averse subjects prefer pile A. Risk-loving subjects prefer pile B. Risk-neutral subjects are indifferent as to whether they choose pile A or B (Figure 6).

Figure 6: Space of Possibilities of the New Approach as well as the Indifference Curves of a Risk-Averse (Unbroken Grey Line), a Risk-Neutral (Dotted Grey Line) and a Risk-Loving Subject (Dashed Grey Line)



In this way, two out of the three criteria for a suitable procedure to record the risk preference of the test persons are fulfilled: (1) The decision-making situation is very clear and simple. The subjects know precisely which consequences their decision will have. They do not have to decide on the basis of a gut feeling, but can make well thought-out, conscious decisions corresponding to their preferences. (2) The three alternatives (pile A, pile B or indifference as to whether the card is from A or B) permit unambiguous conclusions about the three categories of risk preference (risk-averse, risk-neutral and risk-loving). In the following chapter we will also take the influence of loss aversion on risk preference into account.

4 Taking Loss Aversion into Account in the New Approach

There is hardly another phenomenon in behavioral economics which has been the subject of as much research as loss aversion (for a comprehensive overview see, for example, Kahneman 2011, Chapter 29; see also Rabin 2000; Fehr and Goette 2007; Tom et al., 2007). Frequently, subjects are strongly influenced in their actions by the effort to avoid losses. One would expect that risk preference would also be influenced by the possibility of the threat of losses. However, this presumption has not yet been confirmed. Eckel and Grossman (2008) and Crosetto and Filippin (2013) have both included treatments with the possibility of losses. In spite of this, notable effects on the risk preferences of the subjects could not be observed in either study.

Mukherjee et al. (2017) showed that in the case of small amounts up to $4, a profit had a greater positive influence on the well-being of the participants than an equally high loss had in negative terms. Here, an evaluation scale ranging from 0 (= no effect) to 5 (= very strong effect) was used. In the case of an amount of $25, however, the negative perception of a loss was more intense than the positive feeling of an equally large profit. These results indicate that loss aversion might only have an influence on risk preference when larger amounts are involved. In Eckel and Grossman (2008), losses between $-2 and $-6 can occur. And in Crosetto and Filippin (2013), losses of €-2.50 can occur.

We will now carry out the new approach to establish risk preference in three variations in order to investigate the influence of loss aversion on risk preference in more detail. In Treatment 1, there is no possibility of making a loss. In Treatment 2, a loss of €-2.50 can be made. In Treatment 3, a loss of €-25 can be made (Table 5).

Table 5: Random Events, Expected Values and Standard Deviations in Treatments 1-3

Treatment	Pile	Prob.	Return	E(r)	SD
1	A	50%	€+4	€+5	1.0
		50%	€+6		
	B	50%	€±0	€+5	5.0
		50%	€+10		
2	A	50%	€+4	€+5	1.0
		50%	€+6		
	B	50%	-2.5	€+5	7.5
		50%	€+12.5		
3	A	50%	€+4	€+5	1.0
		50%	€+6		
	B	50%	€-25	€+5	30
		50%	€+35		

Prob. = probability; E(r) = expected value of the return; SD = standard deviation.

The results obtained by Mukherjee et al. (2017), Eckel and Grossman (2008) and Crosetto and Filippin (2013) lead us to expect that there will be no significant differences between Treatment 1 and Treatment 2. The possible losses of €-2.50 are presumably too small to have an influence on the risk preferences of the subjects. The first hypothesis is therefore as follows:

<u>Hypothesis 1:</u> In Treatment 2, not more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 1.

The first null hypothesis which will have to be examined is therefore:

<u>Null hypothesis 1:</u> In Treatment 2, significantly more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 1.

The results of Mukherjee et al. (2017), however, give reason to presume that the danger of losses of €-25 can have an influence on the risk preferences of the subjects. The second hypothesis is therefore as follows:

Hypothesis 2: In Treatment 3, more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 1.

The second null hypothesis to be examined is therefore:

Null hypothesis 2: In Treatment 3, not more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 1.

If the presumption is correct that the possibility of a small loss does not really impress subjects, whereas that of a larger loss has a significant influence on risk preferences, it must also be possible to establish a difference between Treatment 2 and Treatment 3. Our third hypothesis is therefore as follows:

Hypothesis 3: In Treatment 3, more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 2.

The third null hypothesis which will have to be examined is therefore:

Null hypothesis 3: In Treatment 3, not more (less) subjects will prove to be risk-averse (risk-loving) than in Treatment 2.

In our experiment we conduct a between-subjects comparison. A total of 157 students of the Ostfalia University of Applied Sciences in Wolfsburg took part in the experiment. 53 subjects played Treatment 1, 52 subjects played Treatment 2, and 52 subjects played Treatment 3. 53 women (33.76%) and 104 men (66.24%) took part. 72 of the subjects study business management (45.86%), 69 subjects study vehicle construction (43.95%) and 16 students study health care (10.19%). The experiment was carried out from 4-10 April 2018 in the Ostfalia Laboratory for Experimental Economic Research (OLEW) in Wolfsburg in Germany. The experiment is programmed in z-Tree. Only the playout of random events is carried out in an analogue way by taking a card from the respective selected pile.[8]

The actual experiment is preceded by a real effort task. We give the subjects a task which is not enjoyable and which requires a considerable amount of time. The subjects are supposed to view the task as work which is paid for with an appropriate amount (€25). The subjects have to encode a total of 175 three-letter words in sequences of numbers. When they have encoded a word correctly, the next word appears. This real effort task is based on Erkal, Gangadharan and Nikiforakis (2011). In order to make it more demanding, Benndorf, Rau and Solch (2014) change the assignment of numbers to letters for every word. We used this approach.

In addition, we consider it to be important that payment for the real effort task made is in cash and is carried out directly afterwards and before the actual experiment (the selection of one of the two lotteries). Willingness to spend is noticeably reduced if payment is made in cash in comparison to credit or debit cards (see, for example, Prelic and Semester, 2001; Runnemark et al., 2015). It has also been shown that impulsive purchase behavior is restricted when a person is handling cash (see, for example, Thomas, Kaushik and Seenivasan, 2011). From this we conclude that immediate cash payment after the real effort task leads

[8] We also chose this path in order to obtain maximum credibility with regard to an uninfluenced random process (see also Chapter 3).

to the subjects perceiving the amount as their own hard-earned money. In this way, the so-called house money effect[9] is probably avoided or at least considerably reduced.

We pay the subjects a show-up fee of €2. For the coding work (real effort task) the subjects earn €25, for which they require between 35 and 60 minutes. In the actual experiment the subjects earn an average of €5.56. Overall the subjects thus earn an average of €32.56. Reading the instructions for the game, answering the test questions, carrying out the coding work, deciding between piles A and B and taking a card take up between 60 and 90 Minutes. The payment they receive is therefore at an appropriate, average level. The subjects gave the impression of being very attentive and motivated.

The results were clear and are largely in line with our expectations (Table 6). In Treatment 1 (no possibility of loss), only 21 out of 53 subjects (39.62%) chose the low-risk variation (pile A). 28 subjects (52.83%) chose the risky variant (pile B). Four subjects (7.55%) were indifferent to whether they chose a card from pile A or B. In Treatment 2 (possibility of a small loss), 25 out of 52 subjects (48.08%) chose the low-risk variation (pile A). 24 subjects (46.15%) chose the risky variant (pile B). Three subjects (5.77%) were indifferent to whether they chose a card from pile A or B. In Treatment 2, the low-risk pile A was chosen more often and the risky pile B was chosen less frequently than in Treatment 1. However, Pearson's chi squared test showed this difference to be insignificant, with a p-value of 0.418 (Table 7). The null hypothesis 1 therefore has to be rejected. This confirmed our presumption (hypothesis 1) that a possibility of a small loss of €-2.50 does not have any significant influence on the risk preferences of the subjects. This result is in line with the findings of Crosetto and Filippin (2013) and those of Eckel and Grossman (2008).

Table 6: Results of the Selection Decision According to Treatments

Treatment	Pile A	Pile B	Decision for Pile A Number	Decision for Pile B Number	Indifferent Number	Decision for Pile A in %	Decision for Pile B in %	Indifferent in %
1	€+4/€+6	€±0/€+10	21	28	4	39.62%	52.83%	7.55%
2	€+4/€+6	€-2.5/€+12.5	25	24	3	48.08%	46.15%	5.77%
3	€+4/€+6	€-25/€+35	36	11	5	69.23%	21.15%	9.62%

Table 7: Results of Pearson's Chi Squared Test

Comparison	P-Value
Treatment 1 (no possibility of loss) versus Treatment 2 (possibility of a small loss)	0.418
Treatment 1 (no possibility of loss) versus Treatment 3 (possibility of significant loss)	0.001
Treatment 2 (low possibility of loss) versus Treatment 3 (possibility of significant loss)	0.009

[9] Thaler and Johnson (1990) showed that subjects take more risks when they have previously made a profit or if start-up capital is made available to them. This applies as long as their earlier profit or start-up capital have not been used up, and they are playing with 'house money', as it were.

In Treatment 3, however (possibility of a significant loss) a marked influence on risk preferences can be noted. Here, 36 out of 52 subjects (69.23%) chose the low-risk variation (pile A). Only eleven subjects (21.15%) chose the risky variant (pile B). Five subjects (9.62%) were indifferent to whether they chose a card from pile A or B. This is a marked difference in comparison to Treatment 1 (no possibility of loss). Pearson's chi squared test also shows this difference with a p-value of 0.001 (Table 7). Null hypothesis 2 clearly has to be rejected. Our presumption that the possibility of a higher loss (€-25) leads more often to risk-averse behavior (hypothesis 2) is thus confirmed.

The comparison between Treatment 2 (risk of a small loss) and Treatment 3 (risk of high loss) also reveals considerable differences. Pearson's chi squared test shows this difference with a p-value of 0.009 (Table 7). It is therefore clear that null hypothesis 3 also has to be rejected. Our presumption that a risk of a high loss influences the risk preference of subject considerably more than a risk of a low loss (hypothesis 3) is thus confirmed. Crosetto and Filippin (2013) had already expressed the presumption that a probability of a high loss would have an effect on the measurement of risk preferences. Our results entirely confirm this presumption.

Overall, it can be stated that taking a probability of a substantial loss into account leads to a more realistic recording of the three categories of risk preference (risk-averse, risk-neutral and risk-loving).

5 Conclusion

Experimental research on diversification behavior requires a clear differentiation between risk-averse, risk-neutral and risk-loving subjects, because decisions which can be absolutely meaningful for a risk-loving subject are completely inconceivable for a risk-averse subject and vice-versa. Robust findings in experimental research on diversification can only be obtained if it is known how to categorize the risk preferences of the subject. Differentiating between risk-neutral, risk-averse and risk-loving subjects is, however, a demanding task. The approach used by Holt and Laury (2002) has undoubtedly received the most attention. We have also used this procedure on several occasions (see, for example, Filiz et al., 2018; Gubaydullina and Spiwoks 2015). However, we also had the impression that not all subjects dedicate themselves to the task with the necessary concentration, and in view of its complexity ultimately make spontaneous decisions which are not well-thought out (for similar observations see Jacobson and Petrie, 2009; Charmes and Viceisza, 2011).

The approach used by Eckel and Grossman (2008) is significantly simpler and clearer, and that deployed by Crosetto and Filippin (2013) even more so. However, all three procedures exhibit the weakness that in certain situations it is not possible to differentiate in an unambiguous and reliable way between risk-averse, risk-neutral and risk-loving subjects. In addition, in these three approaches the influence of loss aversion on risk preference is not taken into consideration, or not sufficiently.

In the form of our Treatment 3 (probability of a substantial loss) we are proposing a new approach to discriminate between risk-averse, risk-neutral and risk-loving subjects which is (1) extremely simple and clear, and which (2) permits the clear assignment of subjects to the three categories of risk preference, and (3) takes the influence of loss aversion on risk preference into account in an appropriate way.

Acknowledgments

We thank the participants of the 2018 CARF Lucerne annual conference for their useful comments and suggestions. We also thank the participants of the doctoral research seminar hosted by Kilian Bizer at the University of Göttingen in 2017. Furthermore, we are grateful to the participants of the 2018 German Association for Experimental Economic Research (GfeW) annual meeting at the University of Paderborn. For the design and creation of the playing cards, we would like to thank Cihan Celik.

References

Agnew, J. / Balduzzi, P. / Sunden, A. (2003): Portfolio Choice and Trading in a Large 401(k) Plan, American Economic Review, 93(1), 193-215.

Anderson, A. (2013): Trading and Under-Diversification, Review of Finance, 17(5), 1-43.

Baltussen, G. / Post, G. T. (2011): Irrational Diversification: An Examination of Individual Portfolio Choice, Journal of Financial and Quantitative Analysis, 46(5), 1463-1491.

Barber, B. M. / Odean, T. (2000): Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, Journal of Finance, 55(2), 773-806.

Bechara, A. / Damasio, A. R. / Damasio, H. / Anderson, S. W. (1994): Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex, Cognition, 50, 7-15.

Benartzi, S. (2001): Excessive Extrapolation and the Allocation of 401(K), Accounts to Company Stock, Journal of Finance, 56(5), 1747-1764.

Benartzi, S. / Thaler, R. H. (2001): Naïve Diversification Strategies in Defined Contribution Saving Plans, American Economic Review, 91(1), 79-98.

Benndorf, V. / Rau, H. A. / Sölch, C. (2014): Minimizing Learning Behavior in Experiments with Repeated Real-Effort Tasks, SSRN, 2503029.

Blume, M. E. / Friend, I. (1975): The Asset Structure of Individual Portfolios and some Implications for Utility Functions, Journal of Finance, 30(2), 585-603.

Bode, M. / van Echelpoel, A. / Sievi, C. R. (1994): Multinationale Diversifikation: Viel Zitiert, Kaum Befolgt, Die Bank, 94(4), 202-206.

Charness, G. / Eckel, C. / Gneezy, U. (2018): Complexity in Risk Elicitation may the Conclusions: A Demonstration Using Gender Differences, Journal of Risk and Uncertainty, 56, 1-17.

Charness, G. / Gneezy, U. / Imas, A. (2013): Experimental Methods: Eliciting Risk Preferences, Journal of Economic Behavior and Organization, 87, 43-51.

Charness, G. / Viceisza, A. (2011): Comprehension and Risk Elicitation in the Field: Evidence from Rural Senegal, IFPRI Discussion Papers, 1135.

Choi, J. J. / Laibson, D. / Madrian, B. C. / Metrick, A. (2004): For Better or for Worse: Default Effects and 401(K) Savings Behavior, Wise, D. A. (Ed.), Perspectives in the Economics of Aging, University of Chicago Press, 81-121.

Crosetto, P. / Filippin, A. (2013): The 'Bomb' Risk Elicitation Task, Journal of Risk and Uncertainty, 47(1), 31-65.

Dave, C. / Eckel, C. C. / Johnson, C. A. / Rojas, C. (2010): Eliciting Risk Preferences: When is Simple Better? Journal of Risk and Uncertainty, 41(3), 219-243.

Dimmock, S. G. / Kouwenberg, R. / Mitchell, O. S. / Peijnenburg, K. (2016): Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence, Journal of Financial Economics, 119, 59-577.

Eckel, C. C. / Grossman, P. J. (2002): Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk, Evolution and Human Behavior, 23(4), 281-295.

Eckel, C. C. / Grossman, P. J. (2008): Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices, Journal of Economic Behavior and Organization, 68(1), 1-17.

Erkal, N. / Gangadharan, L. / Nikiforakis, N. (2011): Relative Earnings and Giving in a Real-Effort Experiment, American Economic Review, 101(7), 3330-3348.

Eyster, E. / Weizsäcker, G. (2011): Correlation Neglect in Financial Decision Making, DIW Discussion Paper, 1104.

Fehr, E. / Goette, L. (2007): Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment, American Economic Review, 97(1), 298-317.

Fernandes, D. (2013): The 1/N Rule Revisited: Heterogeneity in the Naïve Diversification Bias, International Journal of Research in Marketing, 30(3), 310-313.

Filiz, I. / Nahmer, T. / Spiwoks, M. / Bizer, K. (2018): Portfolio Diversification: The Influence of Herding, Status-Quo Bias and the Gambler's Fallacy, Financial Markets and Portfolio Management, 32(2), 167-205.

Fischbacher, U. (2007): Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, Experimental Economics, 10(2), 171-178.

Gneezy, U. / Potters, J. (1997): An Experiment on Risk Taking and Evaluation Periods, Quarterly Journal of Economics, 112(2), 631-645.

Goetzmann, W. N. / Kumar, A. (2008): Equity Portfolio Diversification, Review of Finance, 12(3), 433-463.

Gubaydullina, Z. / Spiwoks, M. (2015): Correlation Neglect, Naïve Diversification, and Irrelevant Information as Stumbling Blocks for Optimal Diversification, Journal of Finance and Investment Analysis, 4(2), 1-19.

Guiso, L. / Haliassos, M. / Japelli, T. (2002): Household Portfolios, MIT Press, Cambridge.

Hedesstrom, T. M. / Svedsater, H. / Garling, T. (2006): Covariation Neglect among Novice Investors, Journal of Experimental Psychology: Applied, 12(3), 155-165.

Hibbert, A. M. / Lawrence, E. R. / Prakash, A. J. (2012): Can Diversification be Learned? Journal of Behavioral Finance, 13(1), 38-50.

Holt, C. A. / Laury, S. K. (2002): Risk Aversion and Incentive Effects, American Economic Review, 92(5), 1644-1655.

Huberman, G. / Sengmueller, P. (2004): Performance and Employer Stock in 401(K) Plans, Review of Finance, 8(3), 403-443.

Filiz, I. (2018): Overconfidence: Der Einfluss Positiver und Negativer Effekte, Sofia Diskussionsbeiträge zur Institutionenanalyse, 17(1).

Jacobsen, S. / Petrie, R. (2009): Learning from Mistakes: What Do Inconsistent Choices over Risk Tell Us? Journal of Risk and Uncertainty, 38(2), 143-158.

Kahneman, D. (2009): Thinking, Fast and Slow, London.

Kallir, I. / Sonsino, D. (2009): The Neglect of Correlation in Allocation Decisions, Southern Economic Journal 75(4), 1045-1066.

Lease, R. C. / Lewellen, W. G. / Schlarbaum, G. G. (1974): The Individual Investor: Attributes and Attitudes, Journal of Finance, 29(2), 413-433.

Lejuez, C. W. / Read, J. P. / Kahler, C. W. / Richards, J. B. / Ramsey, S. E. / Stuart, G. L. / Strong, D. R. (2002): Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (Bart), Journal of Experimental Psychology: Applied, 8(2), 75-84.

Lönnqvist, J. E. / Verkasalo, M. / Walkowitz, G. / Wichardt, P. C. (2015): Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison, Journal of Economic Behavior and Organization, 119, 254-266.

Markowitz, H. (1952): Portfolio Selection, Journal of Finance, 7(1), 77-91.

Meulbroek, L. (2005): Company Stock in Pension Plans: How Costly is it? Journal of Law and Economics, 48(2), 443-474.

Morrin, M. / Inman, J. J. / Broniarczyk, S. M. / Nenkov, G. Y. / Reuter, J. (2012): Investing for Retirement: The Moderating Effect of Fund Assortment Size on the 1/N Heuristic, Journal of Marketing Research, 49(4), 537-550.

Mukherjee, S. / Pammi, V. S. C. / Sahay, A. / Srinivasan, N. (2017): Is Loss-Aversion Magnitude-Dependent? Measuring Prospective Affective Judgments Regarding Gains and Losses, Judgment and Decision Making, 12(1), 81-89.

Polkovnichenko, V. (2005): Household Portfolio Diversification: A Case for Rank-Dependent Preferences, Review of Financial Studies, 18(4), 1467-1502.

Prelec, D. / Simester, D. (2001): Always Leave Home Without it: A Further Investigation of the Credit-Card Effect on Willingness to Pay, Marketing Letters, 12(1), 5-12.

Rabin, M. (2000): Risk Aversion and Expected-Utility Theory: A Calibration Theorem, Econometrica, 68(5), 1281-1293.

Rieger, M. O. (2012): Why Do Investors Buy Bad Financial Products? Probability Misestimation and Preferences in Financial Investment Decision, Journal of Behavioral Finance, 13(2), 108-118.

Runnemark, E. / Hedman, J. / Xiao, X. (2015): Do Consumers Pay More Using Debit Cards Than Cash?, Electronic Commerce Research and Applications, 14(5), 285-291.

Schupp, J. / Wagner, G. G. (2002): Maintenance of and Innovation in Long-Term Panel Studies, the Case of the German Socio-Economic Panel (Gsoep), DIW Discussion Paper, 276.

Thaler, R. H. / Johnson, E. J. (1990): Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice, Management Science, 36(6), 643-660.

Thomas, M. / Kaushik Desai, K. / Seenivasan, S. (2011): How Credit Card Payments Increase Unhealthy Food Purchases: Visceral Regulation of Vices, Journal of Consumer Research, 38(1), 126-139.

Tom, S. M. / Fox, C. R. / Trepel, C. / Poldrack, A. R. (2007): The Neural Basis of Loss Aversion in Decision-Making under Risk, Science, 315(5811), 515-518.

Wagner, G. G. / Burkhauser, R. V. / Behringer, F. (1993): The English Language Public Use File of the German Socio-Economic Panel, Journal of Human Resources, 28(2), 429-433.

Weber, E. U. / Blais, A. R. / Betz, N. E. (2002): A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors, Journal of Behavioral Decision Making, 15(4), 263-290.

Weber, E. U. / Siebenmorgen, N. / Weber, M. (2005): Communication Asset Risk: How Name Recognition and the Format of Historic Volatility Information Affect Risk Perception and Investment Decisions, Risk Analysis, 25(3), 597-609.

Zuckermann, M. (1979): Sensation Seeking: Beyond the Optimal Level of Arousal, John Wiley and Sons Inc, New York.

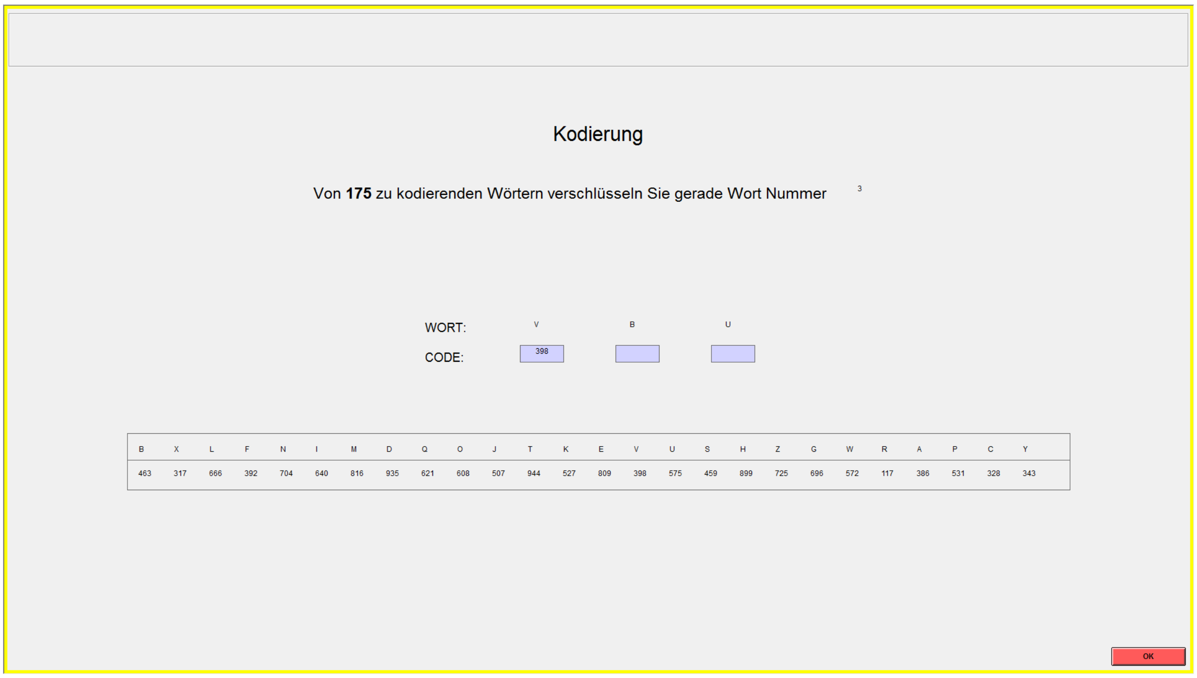
Zuckermann, M. (1994): Behavioral Expressions and Biosocial Bases of Sensation Seeking, Cambridge University Press, Cambridge.

Appendix: Instructions for the Game, Test Questions and Selected Screen-shots

The main instructions are from the study by Benndorf, Rau and Solch (2014).

The Task

In this game, you can earn money by completing a task. The task consists of **coding 175 words** into numbers. For every correctly coded word you receive credit of **14 cents.** In total you can earn up to **€25**. In the task, **three upper case letters** correspond to one word. Every upper-case letter must be assigned to a number. The coding for this can be found in the table below. Please take a look at the photo on the screen:



In this example, the participant has already coded two words correctly. Now the three upper case letters have to be coded: **V, B and U**. The solution is provided in the table:

- V is 398. (see the number entered above by the participant)
- B is 463.
- U is 575.

To enter please click on the blue box below the first upper-case letter.

When all three figures have been entered please click on OK with the mouse.

The computer then checks whether **ALL** upper-case letters have been correctly coded into figures, i.e. whether all three figures were entered correctly. Only then is the word considered to be correct. If the wrong number is entered, the computer points this out (in red letters) after pressing the OK button. The current word remains on the screen until the correct number is entered. However, your previous entries (in the three fields under the upper-case letters) are all deleted. The table remains the same, i.e. the numbers assigned to the letters remain identical. In the same way, the position of the upper-case letters in the table does not change.

When the correct number is entered you receive the next randomly drawn word (again consisting of three upper-case letters). The table continues to be randomly 're-shuffled': New three-digit figures are randomly selected and entered into the table as new mappings for the upper-case letters. The position of the upper-case letters in the table is randomly rearranged. Please note that all 26 upper-case letters of the German alphabet are used.

Following on from this task you will take part in a lottery. You will be given detailed information about the lottery when the time comes.

Payment

- Basic payment of **€2.**
- For every correctly coded word you receive credit of **14 cents**. In total you can earn up to **€25** (175 x 14.2857 cents).
- The money you might win in the lottery is added to this.

Important Information

- Please remain quiet during the game!
- Do not look at your neighbor’s screen!
- No aids are permitted (calculators, smartphones etc.) All electronic devices must be switched off!

Test Questions (Treatment 1)

Test question 1: How much is the minimum and maximum payment when you choose pile A?

a. The minimum payment is €+6 and the maximum payment is €+35.
b. The minimum payment is €+4 and the maximum payment is €+6. (correct)
c. The minimum payment is €+4 and the maximum payment is €+35.

Test question 2: How much is the minimum and maximum payment when you choose pile B?

a. The minimum payment is €±0 and the maximum payment is €+35.
b. The minimum payment is €+4 and the maximum payment is €+10.
c. The minimum payment is €±0 and the maximum payment is €+10. (correct)

Test question 3: How many different piles of cards are there?

a. 1

b. 2 (correct)

c. 3

Test question 4: How high is the probability of occurrence of the best and worst possible results in the lottery?

a. 100%

b. 0%

c. 50% (correct)

Screenshot 1: Real Effort Task (Reconstructed in Order to Improve Readability)

Coding

Out of **175** words to be coded you are currently encoding word number **1**

Word:	**Y**	**E**	**S**
Code:			

T	C	Y	R	K	Q	V	Z	N	X	D	A	M	U	S	P	W	G	J	B	F	I	Q	H	L	E
735	105	494	343	737	691	825	865	771	587	204	606	362	769	155	345	287	288	165	977	208	200	177	291	791	660

O.K.

Figure A-1: Real Effort Task

Screenshot 2: Instructions and Test Questions on the Lottery in Treatment 1 (Reconstructed in Order to Improve Readability)

<u>Lottery</u>

You can now take a card as part of a lottery. There are two piles of cards to choose from (pile A and pile B). The cards are drawn by hand.

There are four cards in pile A. Two cards lead to a payment of **€+4** and two cards lead to a payment of **€+6.**

There are four cards in pile B. Two cards lead to a payment of **€±0** and two cards lead to a payment of **€+10.**

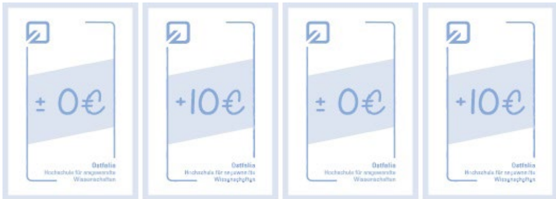
Pile A

(consists of four cards)

Best possible event: €+6 (Probability: 50%)
Worst possible event: €+4 (Probability: 50%)
Expected value: €+5
Low-risk (results fluctuate slightly around the expected value)

Pile B

(consists of four cards)



Best possible event: €+10 (Probability: 50%)
Worst possible event: €±0 (Probability: 50%)
Expected value: €+5
High-risk (results fluctuate considerably around the expected value)

Please answer the following test questions about the lottery:

Test question 1: How much is the minimum and maximum payment when you choose pile A?

ʘ The minimum payment is €+6 and the maximum payment is €+35.

ʘ The minimum payment is €+4 and the maximum payment is €+6.

ʘ The minimum payment is €+4 and the maximum payment is €+35.

Test question 2: How much is the minimum and maximum payment when you choose pile B?

ʘ The minimum payment is €+/-0 and the maximum payment is €+35.

ʘ The minimum payment is €+4 and the maximum payment is €+10.

ʘ The minimum payment is €+/-0 and the maximum payment is €+10.

Test question 3: How many different piles of cards are there?

ʘ 1

ʘ 2

ʘ 3

Test question 4: How high is the probability of occurrence of the best and worst possible results in the lottery?

ʘ 100%

ʘ 0%

ʘ 50%

O.K.

Figure A-2: Instructions and Test Questions on the Lottery in Treatment 1

Screenshot 3: Decision-Making Options on the Lottery in Treatment 1 (Reconstructed in Order to Improve Readability)

<u>Lottery</u>

You can now take a card as part of a lottery. There are two piles of cards to choose from (pile A and pile B). The cards are drawn by hand.

There are four cards in pile A. Two cards lead to a payment of **€+4** and two cards lead to a payment of **€+6.**

There are four cards in pile B. Two cards lead to a payment of **€±0** and two cards lead to a payment of **€+10**.

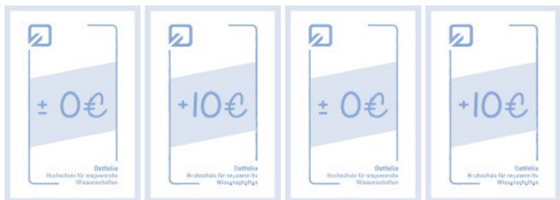
Pile A

(consists of four cards)

Best possible event: €+6 (Probability: 50%) Worst possible event: €+4 (Probability: 50%) Expected value: €+5 Low-risk (results fluctuate slightly around the expected value)

Pile B

(consists of four cards)



Best possible event: €+10 (Probability: 50%) Worst possible event: €±0 (Probability: 50%) Expected value: €+5 High-risk (results fluctuate considerably around the expected value)

You now have to decide which pile you want to take a card from.

Make your selection now. Click on one of the three alternatives.

ʘ I would like to take a card from pile A.

ʘ I would like to take a card from pile B.

ʘ I would like to take a card, but I don't mind which pile I take one from.

O.K.

Figure A-3: Decision-Making Options on the Lottery in Treatment 1

Chapter 5

THE ACCURACY OF INTEREST RATE FORECASTS IN THE ASIA-PACIFIC REGION: OPPORTUNITIES FOR PORTFOLIO MANAGEMENT

with Thomas Nahmer, Markus Spiwoks and Kilian Bizer
contribution Ibrahim Filiz: 35%

Submitted to:
Applied Economics

Published in:
Wolfsburg Working Papers, 18-02, June 2018.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 18-2, June 2018.

Abstract

We analyzed interest rate forecasts from Australia, China, Hong Kong, India, Indonesia, Malaysia, New Zealand, Singapore, South Korea, Taiwan and Thailand. We assessed 532 forecast time series with a total of 85,264 individual interest rate forecasts. To do so, we carried out a comparison to naïve forecasts and investigated the forecast time series for topically-orientated trend adjustments. In addition, we deployed the sign accuracy test and the unbiasedness test. The results are very sobering in part: 95.9% of all forecast time series are characterized by the phenomenon of topically-orientated trend adjustments, and 99.4% of all forecast time series proved to be biased. Only a small proportion of the forecast time series (3.6%) reflected the future interest rate trend significantly more precisely than a naïve forecast. However, at the same time some of the results of the study are surprisingly positive. The sign accuracy test revealed that 48.3% of all forecast time series predict the interest rate trend significantly better than a random walk forecast.

Keywords

Interest rate forecasts, survey forecasts, forecast accuracy, portfolio management, topically-orientated trend adjustment behavior.

JEL Classification

E44, E47, G11, G12, G15, G17, G21, F37

1 Introduction

Future interest rate trends are of key significance for almost all investment decisions on the capital markets. If the general level of interest rates in an economy rises, the prices of most bonds will fall. Only securities with a short residual term to maturity and floating-rate bonds remain largely unaffected by such developments. The longer the term to maturity and the lower the coupon, the greater the fall in bond prices. If the general level of interest rates falls, the opposite effect occurs, and the prices of most bonds rise.

Interest rate trends are, however, also of great importance for investments in the stock market. The fair value of a share is the total of all discounted future profits that this share provides. If one wishes to take a critical look at the current market price of a stock, it is wise not to compare its current price with the current fair value, but rather to determine the *future* fair value of the stock. If the current value is significantly below the fair value which it will have at the end of the investment horizon, it is an attractive investment. However, in order to establish the future fair value of a stock, one has to forecast the predominant future interest rate level, because this simultaneously represents the future discount rate in the determination of fair value.

In the case of international portfolios, exchange rate movements also have to be taken into account. The interest rate parity theory shows that interest rate trends at home and abroad have great significance for exchange rates.

As a rule, financial market analyzes normally begin with a forecast of interest rate trends, because bond and share prices – and ultimately also exchange rates – are significantly influenced by interest rates. It is therefore not surprising that the accuracy of interest rate forecasts have been of great interests to academics and business professionals for a long time now. In the past 40 years, around 50 studies have already been published on the accuracy of survey-based interest rate forecasts (see the comprehensive synoptic overview in Table 20 in the appendix). Some trends have emerged in these studies:

Only a few studies considered the interest rate forecasts they analyzed to be largely reliable.[1] These were largely forecasts on the base rates of central banks or forecasts of short-term market rates such as the three-month money market rate. However, for portfolio management, interest rate forecasts for bonds with maturities of at least a year are primarily of interest, because active portfolio management strategies can be realized much more easily in this segment. There are also some studies with mixed findings.[2] Here again, it is mostly forecasts for short-term interest rates which come off well, while more than half of the stud-

[1] See, for example, Throop (1981), Tabak and Feitosa (2008), Baghestani and Marchon (2012), Knüppel and Schultefrankenfeld (2013) and Pierdzioch (2015).

[2] See, for example, Dua (1988), Zarnowitz and Braun (1992), Cho (1996), Gosnell and Kolb (1997), Greer (2003), Scheier and Spiwoks (2006), Goodhart and Lim (2008), Spiwoks, Bedke and Hein (2008), Chun (2009), Spiwoks, Bedke and Hein (2010), Jongen, Verschoor and Wolff (2011), Kunze, Kramer and Rudschuk (2013), Kunze and Gruppe (2014), Baghestani and Danila (2014), Beechay and Österholm (2014), Oliver and Pasaogullari (2015), and Miah, Khalifa and Hammoudeh (2016).

ies take a very critical view of the quality of the interest rate forecasts which they examined.[3]

Until now, US interest rate forecasts have been the main focus of research, although European interest rate forecasts – particularly British and German ones – have also been frequently examined. In the Asia-Pacific region there have been several studies focusing on Japan,[4] but otherwise there has only been little published research dealing with interest rate forecasts for the Asia-Pacific region. Goodhart and Lim (2008) looked at interest rate forecasts in New Zealand, while Baghestani, Arzaghi and Kaya (2015) analyzed their Australian counterparts. Jongen, Verschoor and Wolff (2011) investigated forecasts of interest rate trends in Hong Kong, Indonesia, Malaysia, New Zealand, Singapore and Taiwan for the period 1995-2009. However, this study limited itself to assessment criteria which do not provide a comprehensive impression of the accuracy of the forecasts. Instead it compared survey forecasts with random walk forecasts or implicit forward rate forecasts. Miah, Khalifa and Hammoudeh (2016) analyzed interest rate forecasts from China, Hong Kong, India, Indonesia, the Philippines, Singapore, South Korea, Taiwan and Thailand, among others. They examined the period 2001-2012 and applied the efficiency test and the unbiasedness test. As a data basis they used the survey forecasts of Fx4casts.com. For us, this is definitely the most interesting study among those which have previously been carried out. However, we refer to a different data basis (Asia Pacific Consensus Forecasts), and in addition to interest rate forecasts from China, Hong Kong, India, Indonesia, Singapore, South Korea, Taiwan and Thailand, we also look at forecasts from Malaysia, Australia and New Zealand. We also take a longer period of time into account in our analysis (1990-2016).[5] In addition, we use a far more comprehensive set of tools for the evaluation of the quality of forecasts: (1) Comparison to a naïve forecast with the aid of the Diebold-Mariano test, (2) examination of the forecast direction with the help of the sign accuracy test, (3) examination for systematic forecast errors with the aid of the unbiasedness test, and (4) test for the presence of possible topically-oriented trend adjustments with aid of the TOTA coefficient.

Unlike many of the previous studies we not only examine the time series of the consensus forecasts, but also the forecast time series of the banks, investment companies, associations, consulting firms and industrial companies which participated in the survey. It cannot be ruled out that individual institutions might succeed in making forecasts which are more reliable than the consensus forecasts. Limiting the analysis to consensus forecasts might therefore mislead us. We evaluated a total of 532 time series with 85,264 interest rate forecasts.

[3] See, for example, Friedman (1980), Belongia (1987), Simon (1989), Hafer and Hein (1989), Francis (1991), Hafer, Hein and MacDonald (1992), Domian (1992), Ilmanen (1996), Kolb and Stekler (1996), Baghestani, Jung and Zuchegno (2000), Albrecht (2000), Spiwoks (2003), Brooks and Gray (2004), Benke (2004), Mose (2005), Baghestani (2005), Benke (2006), Spiwoks and Hein (2007), Mitchell and Pearce (2007), Spiwoks, Bedke and Hein (2009), Gubaydullina, Hein and Spiwoks (2011), Schwarzbach, Kunze, Rudschuck and Windels (2012), Chortareas, Jitmaneeroj and Wood (2012), Butter and Jansen (2013), Spiwoks, Gubaydullina and Hein (2015), and Kunze, Wegener, Bizer and Spiwoks (2017).

[4] See, for example, Gosnell and Kolb (1997), Spiwoks and Hein (2007), Gubaydullina, Hein and Spiwoks (2011), Jongen, Verschoor and Wolff (2011), Butter and Jansen (2013), Spiwoks, Gubaydullina and Hein (2015), and Baghestani, Arzaghi and Kaya (2015).

[5] The Australian interest rate forecasts start in 1990. The other time series only begin later.

In this respect, the study can be viewed as the most comprehensive analysis by far of interest rate forecasts in the Asia-Pacific region.

Some surprising results were revealed in the process, which certainly opens up opportunities for active portfolio management strategies. For example, 61.5% of the forecast time series on the interest rates of Indian state bonds with ten years remaining to maturity (forecast horizon: 13 months) predict the future interest rate trend (rising or falling) significantly better than a random walk forecast. With forecasting results of this kind, it should be possible to systematically obtain excess returns.

The study is divided into five sections. In chapter 2 the data basis is described in detail. In chapter 3 the methods used are presented. The results of the study are shown in chapter 4. In the final chapter a summary of the study is provided.

2 Data Basis

Bates and Granger (1969) were the first to raise the question of whether better results could be achieved via a suitable combination of several forecasts than by means of choosing the (presumably) best forecast. The idea behind this is that many forecasts contain useful elements of information which are not found in other forecasts and which can be brought together in a consensus forecast (see, for example, Thiele 1993). This idea triggered a lively debate on the possibilities and limitations of suitable combinations of forecasts, which culminated in 1989 in a special edition of the Journal of Forecasting and the International Journal of Forecasting. As a result of this debate, the company Consensus Economics created the specialist journal Consensus Forecasts. Since October 1989 it has been published on a monthly basis. Every month, Consensus Economics interviews more than 700 leading academics from the fields of economics and business for their forecasts in relation to various economic indicators for over 85 countries. Alongside the forecasts of these experts, Consensus Economics also publishes a consensus mean, which is the arithmetical average of the experts' forecasts.

The interest rate forecasts for Australia, China, Hong Kong, India, Indonesia, Malaysia, New Zealand, Singapore, South Korea, Taiwan and Thailand which are analyzed here come from the regularly published journal Asia Pacific Consensus Forecasts. We examined the forecasts which were published there in the period from January 1990 to December 2015. The forecasts relate to the period from April 1990 to the end of December 2016. The data for Australia in the initial years is from the journal Consensus Forecasts. For the period after the establishment of the journal Asia Pacific Consensus Forecasts in 1995, the Australian interest rate forecasts are taken from that periodical. We evaluated a total of 532 time series with 85,264 interest rate forecasts. There is a detailed overview in Table 1.

Asia Pacific Consensus Forecasts differentiates between two forecast horizons: In the journal, the forecasts are occasionally described as three-month forecasts and twelve-month forecasts. In reality, however the forecast horizons are of four and thirteen months. This can be seen in the following example: In the edition of January 2015, which was available in around mid-January, the forecasts for the end of April 2015 and the end of January 2016 were published. The forecasts themselves are handed in by the participating institutions at the beginning of January. The actual period of time from the beginning of January 2015 to the end of April 2015 is four months, while the period of time from the beginning of January 2015 to the end of January 2016 is thirteen months.

Table 1: Data Used from the Journal Asia Pacific Consensus Forecasts

Country	Subject of Forecast	Number of Time Series Analyzed	Number of Forecasts Analyzed	Results in the Table
Australia	10 Year Government Bond Yield	42	7,871	3
	Three Month Interest Rates	42	8,115	4
China	One-Year Base Lending Rate	30	3,507	5
Hong Kong	Prime Lending Rate	30	5,159	6
	Three Month Interest Rates	38	6,077	7
India	10 Year Government Bond Yield	26	3,809	8
	Three Month Interest Rates	24	4,196	9
Indonesia	10 Year Government Bond Yield	24	3,595	10
Malaysia	Base Lending Rate	30	4,374	11
	Three Month Interest Rates	36	5,842	12
New Zealand	10 Year Government Bond Yield	36	6,566	13
	Three Month Interest Rates	36	6,552	14
Singapore	Prime Lending Rate	30	3,876	15
	Three Month Interest Rates	38	5,906	16
South Korea	Three Year Interest Rates	28	3,194	17
Taiwan	10 Year Government Bond Yield	16	2,103	18
Thailand	Three Month Interest Rates	26	4,522	19
∑		532	85,264	

We analyzed all of the forecast time series which have at least 80 items of data. We did not take time series with less than 80 observations into consideration. Under certain circumstances, time series which are too short or contain too large gaps can lead to inconclusive results in the procedures used to measure the quality of forecasts.

3 Methods

The following statistical tools were used to measure the quality of forecasts: Comparison to a naïve forecast with the aid of the Diebold-Mariano test (3.1), examination of the forecast direction with the help of the sign accuracy test (3.2), the test for the unbiasedness of the forecasts (3.3) and the test for the presence of topically-orientated trend adjustments with the help of the TOTA coefficient (3.4).

3.1 Comparison to a Naïve Forecast with the Aid of the Diebold-Mariano Test

The French mathematician Pierre Simon Laplace (1814) introduced the principle of indifference (also known as the principle of insufficient reason) into the literature: A black box emits a figure x, and then the subject is requested to forecast which figure the black box will emit next. In view of the subject's complete lack of knowledge regarding the processes going on in the black box, it is not possible to give a single reason why the next figure should be larger than x. They can also not give a single reason why the next figure should be smaller than x. The only thing which an unknowing but sensible person can do is to forecast the figure x again for the future. In this way, a naïve forecast (everything remains the same) is understandable as long as one has no insight into the processes which lead to the figures which need to be forecast. Ever since it was identified, the naïve forecast has been considered to be the rock bottom in terms of forecast accuracy. Even a very rudimentary understanding of the processes at play should lead to better accuracy than that offered by a naïve forecast.

Simple measurements of forecast quality (such as mean absolute error or mean squared error) enable us to make a comparison with a naïve forecast. However, these simple approaches do not permit an assessment of statistical significance. This deficit is avoided by using the Diebold-Mariano test (Diebold and Mariano, 1995). To do so, we calculate the mean squared error (MSE) for the time series of the expert prognoses and for the time series of the naïve forecasts. The test statistics of the Diebold-Mariano test are defined as follows:

$$DM = \frac{\frac{1}{T}\sum(V(P_{t1}) - V(P_{t2}))}{\sqrt{\hat{\gamma}\, d/T}}$$

T = number of observations

V = loss function

P_1 = naïve forecast

P_2 = expert forecast

$\sqrt{\hat{\gamma}\, d/T}$ = joint spread of the two loss functions

The null hypothesis tested in this way is that the naïve forecast (P_1) and the expert forecast (P_2) have the same accuracy. Neither one of the two alternatives thus provides a clearly better result. The numerator is the mean deviation between the loss functions V of the two forecast approaches to be compared. Normally a squared loss function is assumed; in other words, the squared errors of the two forecast approaches are compared (P_1 and P_2). The denominator is the joint spread of the two loss functions. This is estimated on the basis of the long-term autocovariances of the loss functions. In the case of large samples, this test value is asymptotically normally distributed.

3.2 Sign Accuracy Test

The sign accuracy test (Merton, 1981; Henriksson and Merton, 1981) is another widespread tool for evaluating forecasts. In this procedure, the extent of a forecasted change is not the issue. It only examines whether the general direction of the forecasts (rising or falling) is correct. The forecasts are then entered into a 2x2 matrix (Table 2).

Table 2: 2x2 Contingency Table

	Actual Event: Interest Rates Rise	Actual Event: Interest Rates Fall	∑
Forecast: Interest Rates Rise	N_{11}	N_{12}	N_1
Forecast: Interest Rates Fall	N_{21}	N_{22}	N_2
∑	N_1	N_2	N

On the one hand, a differentiation is made between whether an interest rate increase or an interest rate fall was forecast; on the other hand, a differentiation is also made between whether an interest rate rise or an interest rate fall has actually occurred. The principal diagonal in the 2x2 matrix (N_{11} and N_{22}) indicates the forecasts which are correct regarding the trend direction. The secondary diagonal (N_{12} and N_{21}) indicates the forecasts which are incorrect regarding the trend direction. A chi squared test is now applied to examine whether the distribution frequency of the four fields is significantly different from a random walk forecast (cf. Diebold and Lopez, 1996; Joutz and Stekler, 2000). If this is the case, a comparison between the number of observations in the principal diagonals and the secondary diagonals must be carried out to establish whether the forecasts are significantly better or significantly worse than a random walk forecast.

3.3 Unbiasedness Test

The unbiasedness test using the Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969) can check whether the forecast errors are systematic. According to the theory of rational expectations, this should not be the case. The Mincer-Zarnowitz regression takes the following form:

$$A_t = \alpha + \beta P_t + u_t$$

A_t = event which has actually occurred (dependent variable)

α = constant

P_t = forecast of the actual event at the moment in time *t*

β = coefficient of the respective forecasts

u_t = error term at the moment in time *t*

Based on this equation, forecasts are considered unbiased if α is not significantly different to 0, and β is not significantly different to 1. In addition, the error term u_t may not be autocorrelated.

Forecasts are considered unbiased when, with a low probability of error, the joint hypothesis of α = 0 and β = 1 does not have to be rejected. This is checked by using the Wald test. A further condition is the absence of autocorrelations in the value of the error term u_t, which is examined with the Durbin-Watson test. If, according to these criteria, a forecast time series is based on rational expectations, Granger and Newbold (1973) argue that this by no means signifies that the forecasts are perfect. They merely do not exhibit systematic errors.

3.4 Topically-Orientated Trend Adjustment

In order to answer the question of whether forecasters have oriented themselves towards current levels when drawing up interest rate forecasts, the TOTA coefficient is used as a statistical benchmark (Andres and Spiwoks, 1999). Topically-orientated trend adjustment (TOTA) is present when forecasts reflect the present more strongly than the future. In the most unfavorable case, the future-oriented character of such forecasts may be lost entirely.

The TOTA coefficient is the quotient of two coefficients of determination (R^2_A and R^2_B). The R^2_A measures the correlation between the forecasts at the time of their validity and the actual events. The R^2_B measures the correlation between the forecasts at the time of their appearance and the actual events. The TOTA coefficient takes the following form:

$$TOTA\ coefficient = \frac{R^2_{\text{forecasts (validity date); actual events}}}{R^2_{\text{forecasts (issue date); actual events}}} = \frac{R^2_A}{R^2_B}$$

If the TOTA coefficient has a value of < 1, topically-orientated trend adjustment is given, and forecasts reflect the present more strongly than the future.

The TOTA coefficient and the unbiasedness test are closely related. If a forecast time series is characterized by the phenomenon of topically-orientated trend adjustment, the forecast error u_t is normally not randomly distributed (cf. Spiwoks, Bedke and Hein, 2010). Forecast time series which have a TOTA coefficient of < 1 are therefore normally biased.

4 Results

510 of the 532 forecasts analyzed have a TOTA coefficient of < 1 (see Tables 3-19). 95.9% of all the forecast time series analyzed are therefore characterized by the phenomenon of topically-orientated trend adjustments. If interest rates rise, expectations regarding future interest rates will therefore normally be revised upwards. If interest rates fall, expectations regarding future interest rates will therefore usually be revised downwards. In this way, the

forecast time series ultimately reflect current interest rates more strongly than future ones. Expressed more pointedly, it could be said that the experts are forecasting the present rather than the future. This is consistent with the results of earlier studies. In an analysis of 1,182 forecast time series of the G7 countries and five other European countries, a total of 98.5% of all forecast time series studied exhibited a topically-orientated trend adjustment (see Spiwoks, Gubaydullina and Hein, 2015).

These sobering findings are also reflected in the unbiasedness test. 529 of the 532 forecasts analyzed exhibit bias (see Tables 3-19). In 99.4% of all forecast time series studied, either α differs significantly from 0, or β differs significantly from 1, or the error term u_t proves to be autocorrelated.

Even unbiased forecasts can exhibit dramatic forecasting errors. The term unbiased merely states that forecasting errors are not of a systematic nature. A systematic forecasting error is, for example, a continuous over – or underestimation of the subject of the forecast ($\alpha \neq 0$). A different kind of systematic forecasting error is present when small actual events are constantly overestimated (or underestimated), and major actual events are constantly underestimated (or overestimated) ($\beta \neq 1$). Systematic forecasting errors are also present when the error term u_t reveals a pattern. This is usually the case when topically-orientated trend adjustment is present (cf. Spiwoks, Bedke and Hein 2010). However, other systematic forecasting errors can also lead to the error term u_t proving to be autocorrelated. Biased forecast time series are thus a reflection of systematic errors in drawing up the forecasts. This is true for 99.4% of all the forecasts we considered.

An expert's forecast can be viewed as largely worthless if it cannot bear comparison with the respective naïve forecast. A naïve forecast requires no specialist knowledge and is available free of charge to everyone at any time. One should, however, expect that forecasts made by highly-paid financial market experts are more exact than naïve forecasts. In many of the forecast subjects and forecast horizons examined here, the experts' forecasts – compared to the mean squared forecast error – are indeed more precise than naïve forecasts.[6] A total of 175 out of 532 forecast time series (32.9%) exhibit lower mean squared forecast errors than the respective naïve forecasts. However, the Diebold-Mariano test shows that only 19 out of 532 forecast time series (3.6%) contain significantly better forecasting results than naïve forecasts. The experts who forecast the prime lending rate in Hong Kong are particularly successful. 14 out of 30 forecast time series (46.7%) predict the interest rate trend significantly

[6] In the forecasts of the prime lending rate in Hong Kong, it can be seen that 26 out of 30 forecast time series (86.7%) were superior to the naïve forecast. In the case of the forecasts of the 3-month rate in Hong Kong, 24 out of 38 forecast time series (63.2%) were superior to a naïve forecast, while in the forecasts of the 3-month rate in India, at least the forecasts with a 13 month forecast horizon were highly successful: 9 out of 12 forecast time series (75%) were more precise than the corresponding naïve forecast. Among the forecasts of 10-year interest rates in Indonesia, the forecasts with a horizon of 13 months were once again very successful. 9 out of 12 forecast time series (75%) are more exact than a naïve forecast. Forecasts of the base lending rate in Malaysia were more successful than a naïve forecast in 18 out of 30 cases (60%), which is also the case for forecasts of 3-month interest rates in Malaysia in 19 out of 36 cases (52.8%). The forecasts of the 3-month rate in New Zealand were more precise than a naïve forecast in 25 out of 36 cases (69.4%). Among the forecasts of the 3-month rate in Thailand with a forecast horizon of 13 months, 9 out of 13 forecast time series (69.2%) were superior to the naïve forecast.

better than a naïve forecast (Table 6). Apart from this there are only five individual cases in which the time series of expert forecasts are significantly more precise than the time series of the respective naïve forecasts.

The sign accuracy test merely reveals whether forecasts were in the right direction (rising or falling). For the sign accuracy test, however, it is completely irrelevant whether forecasts predict the extent of future trends. The findings here are surprisingly favorable. In 248 out of 513 forecast time series (48.3%), the future trend (rising or falling interest rates) has been grasped significantly better than by a random walk forecast (see Tables 3-19). This is also a remarkable success in comparison to the findings of many previous studies. For example, Spiwoks, Bedke and Hein (2008) established a success rate of only 19.9% among US interest rate forecasts.

In the case of Australian 3-month interest rates with a forecast horizon of four months, 13 out of 21 forecast time series (61.9%) were significantly better in predicting the future trend direction (rising or falling) than a random walk forecast (Table 4). The forecasts for the base lending rate in China are very conspicuous: 29 out of 30 forecast time series (96.7%) predict the future interest rate trend significantly better than a random walk forecast (Table 5). This result is even surpassed by forecasts for the prime lending rate in Hong Kong. All 30 forecast time series (100%) reflect the future interest rate trend significantly more precisely than a random walk forecast (Table 6). The forecasts for three-month interest rates in Hong Kong, with a 13 month forecast horizon, are also very successful. 14 out of 19 forecast time series (73.7%) predict the interest rate trend significantly better than a random walk forecast (Table 7). Forecasts for the three-month rate in India are equally successful. In 17 out of 24 forecast time series (70.8%), the future trend (rising or falling interest rates) is reflected significantly better than by a random walk forecast (Table 9). The base lending rate in Malaysia is also forecasted successfully: 23 out of 28 forecast time series (82.1%) predict the future interest rate trend significantly better than a random walk forecast (Table 11). The forecasts for the three-month rate in New Zealand similarly predict the future interest rate trend significantly better (rising or falling) in 25 out of 36 cases (69.4%) than a random walk forecast (Table 14). Among the forecasts for three-month interest rates in Thailand, it is particularly those with a forecast horizon of four months that are successful. 10 out of 13 forecast time series (76.9%) predict the future trend significantly more precisely than a random walk forecast (Table 19).

In the case of 19 out of 532 forecast time series, the sign accuracy test could not be carried out, because frequencies of < 1 occur in one or several fields of the 2x2 contingency table. In these cases, however, the chi squared distribution is no longer a suitable test statistic (see, for example, Spiwoks, Bedke and Hein, 2009).

Overall, it can be stated that forecasting three-month interest rates is considerably easier than ten-year interest rates. Only 15.3% of the forecast time series on 10-year rates (Australia, India, Indonesia, New Zealand, Taiwan) predict the future trend (rising or falling interest rates) significantly more precisely than a random walk forecast, whereas in the case of three-month interest rates (Australia, Hong Kong, India, Malaysia, New Zealand, Singapore, Thailand) the figure is 57.1%. This coincides with the findings which have been obtained in other parts of the world. For example, in the case of US interest rate forecasts, Spiwoks, Bedke and Hein (2008) showed that only 8.8% of all forecast time series on 10-year interest rates were significantly more successful than a random walk forecast, while in the case of three-month interest rates the figure was 30.9%.

The interest rates for short maturities are influenced considerably more by the actions of central banks than the interest rates for long maturities. In addition, central banks frequently provide an outlook on their future base rate policies. It can be that careful observation of central bank policy benefits forecasts of three-month interest rates, but not those for ten-year interest rates (cf. Spiwoks, Bedke and Hein 2008, p. 376). That would explain the variations in the success of forecasts.

Table 3: Results of the Measurement of Forecast Quality for Australia (10-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
ANZ	440	0.813	-	0.004	o	0.163	0.000	0.000	0.406	-	0.034	o	0.388	0.000	0.000
BIS Shrapnel	442	0.837	-	0.000	o	0.821	0.000	0.000	0.559	-	0.000	-	0.048	0.000	0.000
BT Financial Group	450	0.808	-	0.025	o	0.803	0.000	0.000	0.431	-	0.024	o	0.108	0.000	0.000
Centre of Policy St.	205	0.461	-	0.011	o	0.994	0.000	0.000	0.022	-	0.049	o	0.743	0.000	0.000
Citigroup	245	0.796	-	0.002	o	0.686	0.000	0.000	0.296	-	0.079	o	0.334	0.000	0.000
Commonwealth B.	444	0.785	-	0.000	-	0.029	0.000	0.000	0.401	-	0.015	o	0.642	0.000	0.000
Deloitte Acc. Econ.	428	0.806	-	0.017	o	0.290	0.000	0.000	0.424	o	0.111	-	0.011	0.000	0.000
Deutsche Bank	124	0.349	o	0.148	o	0.856	0.000	0.000	0.256	-	0.075	o	0.107	0.000	0.000
Goldman Sachs	228	0.743	-	0.009	o	0.542	0.000	0.000	0.415	o	0.183	o	0.564	0.000	0.000
HSBC	286	0.814	-	0.096	o	0.277	0.000	0.000	0.623	-	0.031	o	0.178	0.000	0.000
JPMorgan Chase	403	0.760	-	0.000	-	0.094	0.000	0.000	0.361	-	0.000	o	0.198	0.000	0.000
Macquarie	386	0.752	-	0.035	-	0.064	0.000	0.000	0.363	-	0.001	-	0.025	0.000	0.000
Merrill Lynch	300	0.847	-	0.018	o	0.889	0.000	0.000	0.570	o	0.242	o	0.266	0.000	0.000
Moody’s Analytics	206	0.755	-	0.006	o	0.952	0.000	0.000	0.341	o	0.106	o	0.924	0.000	0.000
Nation. Australia B.	411	0.825	-	0.011	o	0.713	0.000	0.000	0.499	-	0.018	o	0.943	0.000	0.000
Nomura	328	0.587	-	0.023	o	0.675	0.000	0.000	0.137	-	0.024	-	0.011	0.000	0.000
Royal B. of Canada	272	0.771	o	0.249	o	0.114	0.000	0.000	0.298	o	0.107	o	0.576	0.000	0.000
Suncorp	212	0.436	-	0.000	o	0.489	0.000	0.000	0.038	o	0.155	+	0.057	0.000	0.000
UBS	449	0.791	-	0.005	o	0.973	0.000	0.000	0.382	-	0.023	-	0.006	0.000	0.000
Westpac	450	0.798	-	0.000	+	0.031	0.000	0.000	0.455	-	0.042	o	0.316	0.000	0.000
Consensus Forec.	504	0.806	-	0.004	o	0.887	0.000	0.000	0.430	-	0.031	o	0.323	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; Sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 4: Results of the Measurement of Forecast Quality for Australia (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
			DM Test		Sign Acc. Test		Unbiasedness F Test	Unbiasedness DW		DM Test		Sign Acc. Test		Unbiasedness F Test	Unbiasedness DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
ANZ	468	0.904	o	0.471	+	0.000	0.000	0.000	0.377	o	0.180	+	0.035	0.000	0.000
BIS Shrapnel	465	0.897	o	0.159	+	0.001	0.000	0.000	0.490	o	0.215	+	0.001	0.000	0.000
BT Financial Group	468	0.886	o	0.432	+	0.000	0.109	0.000	0.427	o	0.847	+	0.002	0.000	0.000
Centre of Policy St.	219	0.780	o	0.393	o	0.107	0.001	0.000	0.015	-	0.096	o	0.663	0.000	0.000
Citigroup	276	0.854	o	0.889	o	0.156	0.002	0.000	0.239	o	0.235	o	0.379	0.000	0.000
Commonwealth B.	470	0.912	o	0.759	o	0.290	0.000	0.000	0.378	o	0.190	o	0.372	0.000	0.000
Deloitte Acc. Econ.	450	0.874	o	0.592	o	0.205	0.000	0.000	0.430	o	0.407	o	0.063	0.000	0.000
Deutsche Bank	124	0.560	o	0.867	o	0.358	0.000	0.000	0.044	o	0.089	o	0.839	0.000	0.000
Goldman Sachs	227	0.913	o	0.162	+	0.000	0.006	0.000	0.457	o	0.351	+	0.000	0.409	0.000
HSBC	272	0.953	o	0.502	o	0.109	0.000	0.000	0.713	o	0.172	+	0.028	0.000	0.000
JPMorgan Chase	410	0.877	o	0.450	+	0.008	0.001	0.000	0.387	o	0.902	o	0.387	0.000	0.000
Macquarie	411	0.900	o	0.537	+	0.005	0.000	0.000	0.368	o	0.388	o	0.482	0.000	0.000
Merrill Lynch	325	0.914	o	0.633	+	0.006	0.008	0.000	0.523	o	0.723	o	0.892	0.000	0.000
Moody's Analytics	188	0.906	o	0.457	+	0.002	0.028	0.000	0.364	o	0.779	o	0.660	0.000	0.000
Nation. Australia B.	432	0.896	o	0.366	+	0.046	0.843	0.000	0.452	o	0.970	o	0.066	0.001	0.000
Nomura	352	0.817	o	0.627	o	0.887	0.000	0.000	0.140	o	0.573	o	0.087	0.000	0.000
Royal B. of Canada	272	0.865	o	0.121	+	0.000	0.000	0.000	0.383	o	0.612	o	0.366	0.036	0.000
Suncorp	212	0.619	-	0.083	o	0.922	0.796	0.000	0.044	o	0.668	o	0.646	0.000	0.000
UBS	472	0.874	o	0.707	+	0.031	0.006	0.000	0.363	o	0.576	o	0.401	0.000	0.000
Westpac	470	0.930	o	0.506	+	0.008	0.000	0.000	0.503	o	0.234	+	0.002	0.000	0.000
Consensus Forec.	504	0.897	o	0.308	+	0.000	0.001	0.000	0.427	o	0.737	o	0.206	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 5: Results of the Measurement of Forecast Quality for China (1-Year Base Lending Rate)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Bank of China	258	0.671	o	0.317	+	0.000	0.000	0.000	0.037	o	0.623	+	0.000	0.000	0.000
Barclays Capital	101	0.841	o	0.733	o	0.317	0.000	0.000	0.030	o	0.386	+	0.000	0.000	0.000
BNP Paribas	108	0.888	o	0.830	+	0.009	0.000	0.000	0.441	o	0.343	+	0.000	0.000	0.000
Citigroup	100	0.599	o	0.229	+	0.001	0.000	0.047	0.007	o	0.179	+	0.000	0.000	0.000
Daiwa Capital	115	0.897	o	0.766	+	0.000	0.000	0.952	0.675	o	0.359	+	0.000	0.000	0.000
Deutsche Bank	158	0.702	o	0.240	+	0.036	0.002	0.000	0.000	o	0.445	+	0.000	0.000	0.000
Hang Seng Bank	107	0.918	o	0.532	+	0.001	0.001	0.000	0.547	o	0.885	+	0.590	0.000	0.000
HSBC Economics	212	0.665	o	0.332	+	0.000	0.000	0.000	0.045	o	0.530	+	0.000	0.000	0.000
IHS Economics	278	0.735	o	0.408	+	0.000	0.000	0.000	0.026	o	0.441	+	0.000	0.000	0.000
ING	216	0.574	o	0.298	+	0.002	0.000	0.000	0.036	o	0.889	+	0.000	0.000	0.000
JPMorgan Chase	96	0.580	o	0.271	+	0.000	0.001	0.001	0.000	o	0.324	+	0.002	0.000	0.098
Morgan Stanley	93	0.808	o	0.632	+	0.008	0.000	0.000	0.013	o	0.895	+	0.000	0.000	0.000
Nomura	158	1.017	o	0.860	+	0.000	0.001	0.000	1.432	o	0.751	+	0.000	0.000	0.000
Oxford Economics	232	0.695	-	0.089	+	0.000	0.000	0.000	0.153	o	0.618	+	0.000	0.000	0.000
Consensus Forec.	300	0.718	o	0.564	+	0.000	0.000	0.000	0.040	o	0.738	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 6: Results of the Measurement of Forecast Quality for Hong Kong (Prime Lending Rate)

	Forecast Horizon 4 Months								Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Bank of China	334	0.967	+	0.025	+	0.000	0.026	0.000	0.574	+	0.058	+	0.000	0.000	0.000
Bank of East Asia	483	0.986	+	0.024	+	0.000	0.003	0.000	0.738	o	0.257	+	0.000	0.000	0.000
C. Pacific-Yamaichi	168	0.985	o	0.264	+	0.000	0.004	0.000	0.287	o	0.131	+	0.012	0.000	0.000
Credit Suisse	168	0.957	o	0.211	+	0.000	0.010	0.000	0.479	o	0.203	+	0.020	0.000	0.000
Daiwa Research I.	273	0.956	+	0.044	+	0.000	0.121	0.000	0.488	+	0.076	+	0.000	0.000	0.000
Deutsche Bank	264	0.967	+	0.031	+	0.000	0.159	0.000	0.633	o	0.418	+	0.000	0.000	0.000
FAZ Institute	132	0.939	o	0.208	+	0.000	0.829	0.000	0.449	o	0.161	+	0.001	0.000	0.000
Goldman Sachs	156	0.972	o	0.703	+	0.001	0.011	0.000	0.785	o	0.697	+	0.000	0.000	0.000
Hang Seng Bank	363	0.994	+	0.021	+	0.000	0.002	0.000	0.778	+	0.047	+	0.000	0.006	0.000
HSBC Economics	341	0.999	o	0.682	+	0.000	0.083	0.000	0.787	o	0.197	+	0.000	0.349	0.000
IHS Economics	288	0.957	+	0.075	+	0.000	0.003	0.000	0.527	o	0.201	+	0.000	0.000	0.000
Sakura Institute	139	0.378	o	0.946	+	0.000	0.169	0.000	0.153	+	0.009	+	0.000	0.000	0.000
S. Chartered Bank	209	0.970	+	0.036	+	0.000	0.023	0.000	0.441	o	0.668	+	0.000	0.000	0.000
UBS	132	0.930	+	0.098	+	0.000	0.863	0.001	0.253	o	0.408	+	0.045	0.455	0.000
Consensus Forec.	504	0.975	+	0.023	+	0.000	0.042	0.000	0.720	+	0.043	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 7: Results of the Measurement of Forecast Quality for Hong Kong (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
			DM Test		Sign Acc. Test		Unbiasedness F Test	Unbiasedness DW		DM Test		Sign Acc. Test		Unbiasedness F Test	Unbiasedness DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Bank of China	336	0.923	+	0.088	+	0.003	0.080	0.000	0.760	o	0.152	+	0.000	0.000	0.000
Bank of East Asia	484	0.906	o	0.121	+	0.000	0.010	0.000	0.804	o	0.163	+	0.001	0.000	0.000
Citigroup	339	0.967	o	0.125	+	0.000	0.848	0.000	0.768	o	0.390	+	0.001	0.000	0.000
C. Pacific-Yamaichi	168	0.981	o	0.628	+	0.000	0.000	0.000	0.374	o	0.317	+	0.001	0.000	0.000
Credit Suisse	185	0.887	o	0.231	o	0.091	0.292	0.000	0.573	o	0.577	o	0.284	0.000	0.000
Daiwa Research	275	0.818	+	0.096	o	0.191	0.064	0.000	0.481	o	0.418	o	0.301	0.000	0.000
Deutsche Bank	297	0.921	o	0.900	o	0.067	0.012	0.000	0.744	o	0.736	+	0.014	0.000	0.000
FAZ Institute	132	0.921	o	0.456	o	0.392	0.061	0.000	0.374	o	0.931	+	0.030	0.000	0.000
Goldman Sachs	375	0.862	o	0.214	+	0.000	0.081	0.000	0.795	o	0.223	+	0.000	0.001	0.000
Hang Seng Bank	363	0.913	o	0.143	+	0.023	0.595	0.000	0.832	+	0.041	+	0.000	0.011	0.000
HSBC	342	0.962	o	0.318	o	0.244	0.082	0.000	0.827	o	0.182	+	0.000	0.002	0.000
ING	405	0.920	o	0.202	+	0.000	0.339	0.000	0.794	o	0.368	+	0.000	0.000	0.000
Morgan Stanley	100	1.013	o	0.411	o	0.401	0.240	0.761	0.818	o	0.744	+	0.008	0.000	0.000
Nomura	196	0.947	o	0.251	o	0.474	0.000	0.000	0.868	o	0.336	+	0.001	0.000	0.000
Sakura Institute	139	0.126	o	0.291	o	0.098	0.123	0.001	0.002	o	0.301	o	0.077	0.044	0.000
Societe Generale	117	0.834	o	0.464	o	0.184	0.024	0.094	0.676	o	0.327	o	0.149	0.004	0.000
S. Chartered Bank	212	0.782	o	0.415	o	0.521	0.000	0.002	0.392	o	0.897	+	0.013	0.000	0.000
UBS	131	0.808	o	0.638	o	0.674	0.001	0.001	0.026	o	0.767	o	0.510	0.000	0.000
Consensus Forec.	504	0.909	o	0.154	+	0.000	0.001	0.000	0.778	o	0.140	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 8: Results of the Measurement of Forecast Quality for India (10-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	190	0.182	o	0.110	O	0.273	0.007	0.000	0.025	o	0.725	+	0.001	0.004	0.000
Confed of Indian I.	144	0.949	-	0.037	o	0.073	0.000	0.032	0.684	o	0.563	o	0.093	0.000	0.000
Deutsche Bank	101	0.932	-	0.011	+	0.020	0.090	0.000	0.907	o	0.329	+	0.030	0.008	0.000
Dresdner Bank	84	0.747	o	0.919	+	0.011	0.021	0.790	0.653	o	0.287	+	0.037	0.183	0.829
FERI	156	0.389	-	0.058	o	0.270	0.000	0.000	0.003	o	0.245	o	0.123	0.000	0.000
Hindustan Lever	176	0.897	-	0.063	o	0.601	0.000	0.000	0.718	o	0.112	o	0.500	0.000	0.000
HSBC Securities	272	0.849	o	0.126	o	0.080	0.015	0.000	0.812	o	0.299	+	0.000	0.000	0.000
HIS Economics	186	0.607	o	0.170	+	0.028	0.076	0.000	0.317	o	0.140	-	0.046	0.000	0.000
NCAER	214	0.881	-	0.005	o	0.214	0.000	0.000	0.712	-	0.003	o	0.688	0.000	0.000
Nomura	224	0.951	o	0.171	o	0.106	0.000	0.000	0.951	o	0.537	+	0.000	0.000	0.000
Tata Services	327	0.922	-	0.073	+	0.001	0.033	0.000	0.813	o	0.163	+	0.000	0.000	0.000
UBS	138	0.905	o	0.178	o	0.128	0.000	0.000	0.730	o	0.518	+	0.013	0.000	0.000
Consensus Forec.	504	0.934	o	0.167	+	0.001	0.000	0.000	0.843	o	0.944	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 9: Results of the Measurement of Forecast Quality for India (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	221	0.862	o	0.316	o	0.558	0.000	0.000	0.104	o	0.669	+	0.001	0.000	0.000
Confed of Indian I.	166	0.908	o	0.621	+	0.001	0.000	0.009	0.656	o	0.624	+	0.028	0.000	0.002
Deutsche Bank	151	0.905	o	0.304	+	0.011	0.109	0.000	0.443	o	0.413	+	0.000	0.000	0.000
Dresdner Bank	232	0.808	NA	NA	+	0.010	0.001	0.000	0.373	o	0.296	+	0.000	0.018	0.000
Goldman Sachs	98	0.113	o	0.389	+	0.026	0.002	0.000	0.000	o	0.553	o	0.835	0.000	0.000
Hindustan Lever	168	0.864	NA	NA	+	0.011	0.000	0.001	0.752	-	0.001	+	0.000	0.000	0.002
HSBC Securities	234	0.502	o	0.172	o	0.585	0.000	0.000	0.167	o	0.513	+	0.000	0.000	0.000
JPMorgan	108	0.484	o	0.315	o	0.793	0.000	0.707	0.452	o	0.249	+	0.001	0.004	0.594
NCAER	212	0.809	o	0.335	+	0.000	0.006	0.000	0.336	o	0.927	o	0.090	0.000	0.000
Tata Services	325	0.854	o	0.487	+	0.000	0.000	0.000	0.180	o	0.870	+	0.000	0.000	0.000
UBS	136	0.666	o	0.158	o	0.200	0.001	0.000	0.049	o	0.186	o	0.133	0.122	0.002
Consensus Forec.	504	0.799	o	0.393	+	0.002	0.001	0.000	0.223	o	0.502	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 10: Results of the Measurement of Forecast Quality for India (10-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Bahana Securities	194	0.916	o	0.990	o	0.108	0.000	0.000	0.569	o	0.553	-	0.033	0.000	0.000
Bank Danamon	98	0.809	NA	NA	o	0.301	0.000	0.023	0.464	o	0.367	+	0.043	0.000	0.000
Castle Asia	108	0.787	-	0.034	o	0.433	0.360	0.000	0.440	o	0.513	+	0.003	0.000	0.000
Citigroup	228	0.912	o	0.505	o	0.273	0.000	0.615	0.443	o	0.707	o	0.722	0.000	0.000
Danareksa S.	375	1.001	NA	NA	o	0.422	0.000	0.000	0.844	o	0.273	o	0.317	0.156	0.000
GK Goh	110	1.107	o	0.347	o	0.053	0.276	0.000	0.969	o	0.864	o	0.542	0.000	0.000
HSBC Economics	272	0.916	o	0.270	o	0.983	0.000	0.000	0.590	o	0.281	o	0.601	0.000	0.000
ING	199	0.937	o	0.353	o	0.191	0.008	0.000	0.747	o	0.519	o	0.983	0.000	0.000
Nomura	134	0.867	o	0.338	o	0.866	0.000	0.095	0.419	o	0.308	-	0.004	0.000	0.000
Societe Generale	124	0.822	-	0.002	NA	NA	0.450	0.000	0.653	o	0.222	NA	NA	0.031	0.000
S. Chartered Bank	174	0.861	o	0.111	o	0.676	0.144	0.000	0.358	o	0.660	o	0.971	0.018	0.000
Consensus Forec.	504	0.926	o	0.265	o	0.884	0.000	0.000	0.590	o	0.501	o	0.547	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 11: Results of the Measurement of Forecast Quality for Malaysia (Base Lending Rate)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
AMSecurities	212	0.890	o	0.226	+	0.000	0.000	0.000	0.493	o	0.318	o	0.309	0.000	0.000
CIBD-CIMB	157	1.002	+	0.096	+	0.003	0.001	0.000	0.612	o	0.272	+	0.017	0.000	0.000
Citigroup	94	0.848	o	0.305	+	0.007	0.000	0.000	0.105	o	0.821	+	0.013	0.000	0.000
Deutsche Bank	92	0.996	o	0.261	NA	NA	0.022	0.000	1.016	o	0.222	+	0.000	0.000	0.010
Goldman Sachs	118	0.792	-	0.049	+	0.000	0.000	0.000	0.388	o	0.301	+	0.000	0.000	0.000
HSBC Securities	149	0.551	-	0.056	o	0.178	0.000	0.000	0.005	o	0.441	+	0.001	0.000	0.000
JM Sassoon	150	0.918	o	0.218	+	0.000	0.055	0.000	0.746	o	0.158	+	0.001	0.270	0.000
Kanega Research	109	0.954	o	0.131	+	0.009	0.014	0.000	0.142	o	0.155	+	0.012	0.000	0.000
Kay Hian Research	261	0.951	o	0.249	+	0.001	0.495	0.000	0.918	o	0.231	+	0.001	0.368	0.000
Maybank	145	0.791	o	0.390	NA	NA	0.082	0.000	0.037	-	0.086	o	0.201	0.000	0.000
MIER	276	0.862	o	0.277	+	0.000	0.095	0.000	0.636	o	0.475	o	0.697	0.204	0.000
RHB Research	400	1.021	o	0.287	+	0.000	0.000	0.000	0.676	o	0.169	+	0.000	0.045	0.000
Societe Generale	104	0.911	o	0.216	+	0.000	0.000	0.000	1.102	o	0.800	+	0.008	0.000	0.000
S. Chartered Bank	165	1.051	o	0.240	+	0.000	0.003	0.000	0.654	o	0.555	o	0.059	0.000	0.000
Consensus Forec.	480	0.939	o	0.130	+	0.000	0.005	0.000	0.606	o	0.160	+	0.000	0.180	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 12: Results of the Measurement of Forecast Quality for Malaysia (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
AMSecurities	211	0.896	o	0.261	o	0.051	0.706	0.000	0.364	o	0.400	o	0.161	0.000	0.000
Baring- ING	427	0.882	o	0.499	+	0.000	0.072	0.000	0.603	o	0.920	+	0.000	0.000	0.000
BofA-Merrill Lynch	111	1.050	o	0.643	o	0.068	0.052	0.000	0.924	o	0.650	o	0.536	0.000	0.000
CIBD-CIMB	156	0.962	o	0.622	o	0.350	0.028	0.000	0.588	o	0.184	o	0.148	0.000	0.000
Citigroup	383	0.923	o	0.100	+	0.004	0.011	0.000	0.534	o	0.474	+	0.002	0.000	0.000
Deutsche Bank	127	0.996	o	0.546	o	0.071	0.006	0.000	0.884	o	0.153	+	0.000	0.000	0.014
Goldman Sachs	464	0.877	-	0.081	o	0.663	0.000	0.000	0.573	NA	NA	+	0.017	0.000	0.000
HSBC Economics	195	0.819	-	0.051	o	0.060	0.000	0.000	0.138	o	0.479	+	0.004	0.000	0.000
JM Sassoon	150	0.917	o	0.180	+	0.000	0.472	0.000	0.689	o	0.248	+	0.002	0.064	0.000
Kanega Research	118	0.819	o	0.113	+	0.005	0.098	0.000	0.077	o	0.179	+	0.000	0.047	0.000
Kay Hian Research	120	0.841	o	0.575	+	0.001	0.009	0.001	0.497	o	0.276	o	0.940	0.000	0.000
Maybank	141	0.876	o	0.110	o	0.699	0.008	0.000	0.200	o	0.144	o	0.248	0.000	0.000
MIER	282	0.915	o	0.243	+	0.002	0.006	0.000	0.604	o	0.424	o	0.482	0.000	0.000
RHB Research	404	0.935	o	0.245	+	0.000	0.527	0.000	0.501	o	0.263	+	0.000	0.000	0.000
Societe Generale	118	1.019	o	0.938	+	0.000	0.436	0.001	1.540	o	0.810	o	0.095	0.000	0.000
S. Chartered Bank	238	0.963	o	0.237	+	0.001	0.028	0.000	0.581	o	0.445	o	0.238	0.025	0.000
UOB Kay Hian	220	0.908	o	0.274	o	0.189	0.654	0.000	0.352	o	0.294	+	0.007	0.000	0.000
Consensus Forec.	504	0.927	o	0.199	+	0.000	0.154	0.000	0.588	o	0.291	+	0.000	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 13: Results of the Measurement of Forecast Quality for New Zealand (10-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
ANZ Bank	488	0.832	-	0.010	o	0.286	0.000	0.000	0.575	-	0.053	+	0.007	0.000	0.000
ASB Bank	274	0.798	-	0.011	o	0.849	0.000	0.000	0.393	-	0.072	o	0.630	0.000	0.000
Bank of NZ	480	0.779	-	0.038	o	0.806	0.000	0.000	0.426	-	0.047	o	0.423	0.000	0.000
BERL	340	0.483	-	0.024	o	0.951	0.000	0.000	0.231	-	0.086	o	0.355	0.043	0.000
Credit Suisse FB	114	0.297	o	0.263	o	0.602	0.002	0.000	0.239	o	0.983	o	0.638	0.213	0.000
Deutsche Bank NZ	468	0.831	-	0.006	+	0.032	0.000	0.000	0.481	-	0.001	-	0.004	0.000	0.000
First NZ Capital	348	0.830	-	0.000	o	0.988	0.000	0.000	0.565	o	0.145	o	0.198	0.000	0.000
Goldman Sachs NZ	144	0.516	-	0.059	o	0.955	0.043	0.000	0.093	o	0.308	-	0.015	0.000	0.000
HSBC Economics	216	0.867	-	0.003	o	0.550	0.000	0.000	0.490	-	0.053	o	0.129	0.000	0.000
Infometrics	498	0.762	-	0.002	o	0.469	0.000	0.000	0.426	-	0.021	o	0.549	0.000	0.000
JPMorgan Chase	331	0.754	-	0.000	o	0.444	0.019	0.000	0.361	-	0.035	o	0.833	0.000	0.000
Macquarie	172	0.776	-	0.003	o	0.271	0.000	0.000	0.341	-	0.050	o	0.472	0.000	0.000
National Bank NZ	212	0.465	o	0.663	o	0.107	0.529	0.000	0.000	o	0.180	o	0.455	0.000	0.000
NZIER	440	0.764	-	0.000	o	0.257	0.000	0.000	0.572	-	0.060	+	0.002	0.000	0.000
Ord Minnett S.	128	0.234	-	0.005	o	0.567	0.008	0.000	0.367	o	0.783	+	0.000	0.000	0.000
UBS	476	0.779	-	0.001	-	0.038	0.000	0.000	0.393	-	0.084	-	0.040	0.000	0.000
Westpac	482	0.799	-	0.003	o	0.663	0.000	0.000	0.497	-	0.092	o	0.359	0.000	0.000
Consensus Forec.	504	0.793	-	0.005	o	0.377	0.000	0.000	0.504	o	0.118	o	0.075	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 14: Results of the Measurement of Forecast Quality for New Zealand (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
ANZ Bank	488	0.939	o	0.287	+	0.000	0.063	0.000	0.694	o	0.383	o	0.446	0.000	0.000
ASB Bank	274	0.998	o	0.243	+	0.000	0.018	0.000	0.684	o	0.654	o	0.434	0.000	0.000
Bank of NZ	480	0.937	o	0.152	+	0.000	0.068	0.000	0.758	o	0.341	+	0.003	0.000	0.000
BERL	340	0.770	o	0.275	o	0.173	0.000	0.000	0.403	o	0.213	o	0.128	0.038	0.000
Credit Suisse FB	114	0.742	o	0.215	+	0.012	0.490	0.000	0.042	o	0.236	+	0.000	0.043	0.000
Deutsche Bank NZ	468	0.947	+	0.076	+	0.000	0.064	0.000	0.630	o	0.497	+	0.007	0.000	0.000
First NZ Capital	348	0.960	o	0.161	+	0.000	0.124	0.000	0.637	o	0.394	o	0.299	0.000	0.000
Goldman Sachs NZ	146	0.825	o	0.265	+	0.000	0.000	0.000	0.079	o	0.464	+	0.008	0.000	0.000
HSBC Economics	198	0.972	o	0.294	+	0.001	0.000	0.000	0.805	o	0.419	+	0.038	0.000	0.000
Infometrics	498	0.925	o	0.411	+	0.000	0.000	0.000	0.472	o	0.969	o	0.363	0.000	0.000
JPMorgan Chase	327	0.959	o	0.328	o	0.071	0.115	0.000	0.435	o	0.861	o	0.743	0.000	0.000
Macquarie	176	1.002	o	0.320	+	0.001	0.021	0.000	0.806	o	0.324	o	0.379	0.000	0.000
National Bank NZ	212	0.807	o	0.130	+	0.000	0.124	0.000	0.270	o	0.361	+	0.000	0.000	0.000
NZIER	440	0.891	o	0.997	o	0.067	0.015	0.000	0.493	o	0.860	o	0.258	0.000	0.000
Ord Minnett S.	128	0.772	o	0.537	+	0.009	0.078	0.000	0.685	o	0.193	+	0.001	0.860	0.000
UBS	478	0.927	o	0.412	+	0.001	0.496	0.000	0.558	o	0.571	+	0.028	0.000	0.000
Westpac	482	0.946	o	0.137	+	0.000	0.041	0.000	0.542	o	0.611	+	0.039	0.006	0.000
Consensus Forec.	504	0.929	o	0.150	+	0.000	0.562	0.000	0.603	o	0.291	+	0.009	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 14: Results of the Measurement of Forecast Quality for Singapore (Prime Lending Rate)

	Forecast Horizon 4 Months								Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	94	0.580	o	0.604	NA	NA	0.004	0.000	0.308	o	0.163	+	0.000	0.000	0.000
Credit Suisse	95	0.847	-	0.000	NA	NA	0.000	0.000	0.355	-	0.000	NA	NA	0.000	0.002
Deutsche Bank	164	0.996	-	0.000	NA	NA	0.000	0.000	0.920	-	0.000	NA	NA	0.000	0.000
Goldman Sachs	151	0.792	-	0.003	o	0.696	0.000	0.000	0.871	o	0.196	+	0.003	0.000	0.000
HSBC	322	1.090	-	0.000	NA	NA	0.000	0.000	2.542	-	0.000	NA	NA	0.000	0.000
JM Sassoon	164	1.144	o	0.457	o	0.119	0.000	0.000	2.295	o	0.956	-	0.030	0.000	0.000
Kay Hian Research	122	0.784	o	0.280	o	0.055	0.000	0.000	1.378	o	0.951	+	0.001	0.000	0.000
Morgan Stanley	95	0.818	-	0.000	NA	NA	0.000	0.000	0.704	-	0.010	o	0.338	0.000	0.000
Nomura	183	0.973	-	0.000	o	0.505	0.000	0.000	1.054	-	0.008	+	0.003	0.000	0.000
OCBC Bank	194	0.970	-	0.000	NA	NA	0.000	0.000	0.868	-	0.000	NA	NA	0.000	0.000
Sakura Institute	144	0.749	o	0.430	+	0.007	0.000	0.000	0.398	o	0.903	+	0.003	0.000	0.000
S. Chartered Bank	168	0.978	-	0.061	o	0.107	0.000	0.000	0.863	o	0.395	o	0.064	0.000	0.000
U. Overseas Bank	190	0.937	-	0.013	o	0.058	0.000	0.000	1.050	o	0.542	+	0.004	0.000	0.000
UOB Kay Hian	143	9.037	-	0.000	NA	NA	0.000	0.000	2.053	-	0.000	NA	NA	0.000	0.000
Consensus Forec.	392	0.946	-	0.000	o	0.090	0.000	0.000	0.749	-	0.002	-	0.039	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 16: Results of the Measurement of Forecast Quality for Singapore (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	382	0.847	o	0.196	+	0.017	0.002	0.000	0.431	o	0.189	o	0.229	0.353	0.000
Credit Suisse	154	0.922	o	0.246	o	0.058	0.270	0.000	0.424	o	0.531	o	0.264	0.000	0.000
DBS Bank	208	0.892	o	0.554	+	0.001	0.000	0.000	0.379	o	0.349	-	0.032	0.000	0.000
Deutsche Bank	240	0.916	o	0.999	+	0.003	0.017	0.047	0.761	o	0.844	+	0.049	0.000	0.000
Goldman Sachs	454	0.818	o	0.111	+	0.007	0.000	0.000	0.522	o	0.557	+	0.007	0.000	0.000
HSBC	335	0.926	o	0.553	o	0.103	0.000	0.000	0.501	o	0.948	-	0.002	0.000	0.000
IHS	212	0.891	o	0.119	o	0.990	0.451	0.000	0.311	-	0.098	o	0.622	0.000	0.000
ING	279	0.697	o	0.319	+	0.028	0.024	0.000	0.097	o	0.442	+	0.011	0.000	0.000
JM Sassoon	166	0.541	o	0.494	+	0.041	0.149	0.163	0.064	o	0.495	o	0.914	0.626	0.000
Kay Hian Research	236	0.572	-	0.057	o	0.375	0.059	0.000	0.188	o	0.426	+	0.023	0.000	0.000
Merrill Lynch	127	0.767	o	0.313	o	0.855	0.470	0.027	0.371	o	0.229	o	0.874	0.000	0.001
Nomura	252	0.663	o	0.571	o	0.385	0.202	0.000	0.426	o	0.319	+	0.031	0.000	0.000
OCBC Bank	323	0.831	o	0.554	o	0.291	0.000	0.000	0.527	o	0.750	+	0.002	0.000	0.000
Sakura Institute	128	0.315	-	0.052	o	0.796	0.001	0.444	0.001	o	0.391	o	0.492	0.000	0.028
S. Chartered Bank	243	0.658	o	0.484	+	0.016	0.049	0.002	0.098	o	0.823	o	0.066	0.000	0.000
UBS	122	0.625	o	0.951	o	0.435	0.146	0.002	0.355	o	0.492	o	0.411	0.000	0.000
U. Overseas Bank	206	0.751	o	0.412	+	0.026	0.176	0.000	0.615	o	0.240	+	0.000	0.031	0.000
UOB Kay Hian	143	0.837	o	0.434	o	0.184	0.051	0.000	0.342	o	0.719	+	0.000	0.000	0.000
Consensus Forec.	504	0.770	o	0.511	+	0.000	0.000	0.000	0.436	o	0.403	+	0.002	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

Table 17: Results of the Measurement of Forecast Quality for South Korea (3-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Credit Suisse	112	0.295	-	0.032	o	0.790	0.009	0.000	0.403	-	0.021	o	0.340	0.000	0.000
Daewoo Securities	208	0.849	-	0.090	o	0.874	0.000	0.000	0.586	o	0.465	o	0.172	0.000	0.000
Daishin Economics	133	0.634	-	0.098	o	0.466	0.000	0.000	0.164	-	0.000	-	0.001	0.001	0.000
Dresdner Bank	179	0.724	-	0.002	o	0.277	0.000	0.000	0.643	-	0.006	o	0.473	0.000	0.000
HSBC Economics	94	0.838	-	0.009	o	0.848	0.002	0.000	0.634	o	0.463	o	0.108	0.000	0.000
Hyundai Securities	228	0.714	-	0.005	o	0.364	0.000	0.000	0.216	-	0.056	o	0.166	0.000	0.000
ING Baring	94	0.322	o	0.149	-	0.046	0.000	0.000	0.382	-	0.000	-	0.040	0.001	0.000
LG Group	211	0.856	-	0.056	o	0.283	0.000	0.000	0.338	-	0.001	-	0.001	0.000	0.000
Samsung ER	196	0.923	-	0.000	o	0.198	0.000	0.000	0.830	-	0.021	-	0.025	0.000	0.000
Sakura	143	0.581	-	0.013	o	0.252	0.000	0.000	0.020	-	0.031	o	0.107	0.372	0.000
Shinhan Securities	144	0.640	-	0.002	NA	NA	0.000	0.000	0.138	-	0.000	NA	NA	0.000	0.000
Societe Generale	92	0.841	-	0.000	o	0.204	0.000	0.000	0.573	o	0.725	o	0.123	0.000	0.002
UBS	101	0.912	-	0.018	o	0.073	0.001	0.000	0.820	-	0.038	o	0.781	0.000	0.005
Consensus Forec.	278	0.834	-	0.047	o	0.795	0.000	0.000	0.485	-	0.008	-	0.014	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 18: Results of the Measurement of Forecast Quality for Taiwan (10-Year Government Bond Yield)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	176	0.325	o	0.116	o	0.791	0.000	0.043	0.000	-	0.036	o	0.645	0.000	0.294
HSBC	192	0.629	-	0.004	o	0.681	0.000	0.000	0.127	-	0.094	o	0.163	0.000	0.000
IHS	234	0.645	-	0.023	-	0.034	0.000	0.000	0.222	-	0.018	o	0.222	0.000	0.000
ING	224	0.675	o	0.119	o	0.752	0.000	0.005	0.208	o	0.120	o	0.941	0.000	0.000
Nomura	167	0.635	-	0.004	o	0.444	0.001	0.000	0.001	-	0.032	o	0.391	0.000	0.000
Polaris Research	134	0.721	-	0.000	o	0.601	0.000	0.000	0.339	-	0.003	o	0.314	0.000	0.000
Taiwan Institute R.	126	0.719	o	0.227	+	0.024	0.000	0.000	0.222	o	0.102	NA	NA	0.000	0.000
Consensus Forec.	236	0.672	-	0.045	o	0.453	0.000	0.000	0.201	-	0.065	o	0.234	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test; NA = not available.

Table 19: Results of the Measurement of Forecast Quality for Thailand (3-Month Interest Rates)

		Forecast Horizon 4 Months							Forecast Horizon 13 Months						
							Unbiasedness							Unbiasedness	
			DM Test		Sign Acc. Test		F Test	DW		DM Test		Sign Acc. Test		F Test	DW
Institution	#	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value	TOTA	Res	P-Value	Res	P-Value	P-Value	P-Value
Citigroup	298	0.823	o	0.366	+	0.003	0.000	0.000	0.468	o	0.570	+	0.000	0.000	0.000
Deutsche Bank	82	0.972	o	0.762	+	0.006	0.064	0.000	0.797	o	0.320	o	0.908	0.298	0.000
Goldman Sachs	377	0.762	o	0.231	o	0.297	0.063	0.000	0.522	o	0.368	o	0.366	0.000	0.000
HSBC Economics	346	0.865	o	0.349	+	0.000	0.164	0.000	0.361	o	0.332	o	0.060	0.000	0.000
ING	400	0.820	o	0.328	+	0.006	0.481	0.000	0.344	o	0.297	+	0.000	0.000	0.000
Kasikornbank	390	0.761	o	0.148	+	0.001	0.000	0.000	0.391	o	0.401	+	0.007	0.000	0.000
Merrill Lynch	155	0.865	o	0.270	o	0.659	0.000	0.002	0.538	o	0.524	-	0.025	0.000	0.000
Morgan Stanley	85	0.963	o	0.379	o	0.804	0.194	0.027	0.131	o	0.509	o	0.095	0.030	0.095
Nomura	146	0.791	o	0.487	+	0.000	0.000	0.421	0.579	o	0.912	+	0.006	0.000	0.007
Phatra Thanakit S.	334	0.850	o	0.370	+	0.007	0.000	0.000	0.554	o	0.382	+	0.001	0.000	0.000
Siam C. Bank	175	0.899	o	0.254	+	0.006	0.000	0.008	0.725	o	0.391	o	0.506	0.000	0.000
S. Chartered Bank	206	0.841	o	0.838	+	0.000	0.000	0.000	0.398	o	0.474	+	0.047	0.000	0.000
Consensus Forec.	504	0.841	o	0.437	+	0.000	0.008	0.000	0.477	o	0.379	o	0.186	0.000	0.000

= number of observations; TOTA = TOTA coefficient; DM test = Diebold-Mariano test; Res = result; o = no significant result; - = significantly worse than a naïve or random walk forecast; + = significantly better than a naïve or random walk forecast; sign acc. test = sign accuracy test; unbiasedness = test for unbiasedness; DW = Durbin-Watson test.

A further factor is that it is obviously more difficult to forecast market interest rates than those which are set or controlled by governments. The base lending rate in China, the prime lending rate in Hong Kong, the base lending rate and the three-month interest rate in Malaysia as well as the prime lending rate in Singapore are set directly by the respective administration or – at least partly – managed by it. The success rate for the relevant forecast time series is relatively high: 75% of the forecast time series for these interest rates predict the future interest rate trend (rising or falling) significantly better than a random walk forecast.

5 Conclusion

We analyzed interest rate forecasts for the Asia-Pacific region in the period from 1990-2016. To do so, we examined individual interest rate forecasts from Australia, China, Hong Kong, India, Indonesia, Malaysia, New Zealand, Singapore, South Korea, Taiwan and Thailand. As a basis we used forecasting data which had been published in the journal Asia Pacific Consensus Forecasts on a monthly basis. We did not limit ourselves to the analysis of consensus forecasts, however: We also evaluated all of the forecast time series issues by banks, investment companies, consulting firms, associations and industrial companies. Overall we assessed 532 forecast time series with a total of 85,264 individual interest rate forecasts. The variety of procedures which we used to measure the quality of forecasts enabled us to create a comprehensive evaluation of forecasting performance in the Asia-Pacific region. We carried out a comparison to naïve forecasts. We examined the forecast time series for evidence of topically-orientated trend adjustments. In addition, we deployed the sign accuracy test and the unbiasedness test.

The results are very sobering in part. 95.9% of all forecast time series are characterized by the phenomenon of topically-oriented trend adjustments. This means that the overwhelming majority of all forecast time series reflect the present rather than the future. In total, 99.4% of all forecast time series proved to be biased. Given that topically-orientated trend adjustments usually lead to the error term u_t not being distributed randomly, the result of the unbiasedness test is not surprising.

Only a small proportion of the forecast time series (3.6%) reflected the future interest rate trend significantly more precisely than a naïve forecast. The only forecast whose success went beyond rare individual cases was that for the prime lending rate in Hong Kong. 46.7% of these forecast time series predict the future interest rate trend significantly better than a random walk forecast.

However, some of the results of the study are also surprisingly positive. The sign accuracy test reveals that in 248 out of 513 forecast time series (48.3%), the future trend (rising or falling interest rates) has been grasped significantly better than by a random walk forecast. In this context, at least part of the forecasts for Australia, China, Hong Kong, India, Malaysia, New Zealand, Singapore and Thailand proved to be particularly successful.

Overall it can be stated that – at least in some countries and for some forecast horizons – forecasts of future interest rate trends in the Asia-Pacific region are significantly more successful than those made in other parts of the world. This has consequences for portfolio management: Active portfolio management strategies have no prospects of success in many financial markets because the necessary forecasting competence is simply not there. However, this is different, for example, in the case of the Indian bond market. 61.5% of the forecast time series on the interest rates of Indian state bonds with ten years remaining to maturity (forecast horizon: 13 months) predict the future interest rate trend (rising or falling) significantly better than a random walk forecast. This should suffice in order to achieve systematic excess returns with active portfolio management strategies.

Acknowledgments

We would especially like to thank Matthias Noack, Benjamin Schulz, Petrik Runst, Harm Alhusen and Tim Schneider for helpful comments and support. We would also like to thank the participants of the doctoral research seminar hosted by Kilian Bizer at the University of Göttingen in 2017.

References

Albrecht, T. (2000): Zur Eignung professioneller Zinsprognosen als Entscheidungsgrundlage, Ein Vergleich der Zinsprognosen deutscher Banken mit der Zinserwartung „des Marktes“, Sofia – Diskussionsbeiträge zur Institutionenanalyse, 00-7, 1-29.

Andres, P. / Spiwoks, M. (1999): Prognosequalitätsmatrix – Ein methodologischer Beitrag zur Beurteilung der Güte von Kapitalmarktprognosen, Jahrbücher für Nationalökonomie und Statistik, 219(5/6), 513-542.

Baghestani, H. (2005): Improving the Accuracy of Recent Survey Forecasts of the T-Bill Rate, Business Economics, 40, 36-40.

Baghestani, H. / Arzaghi, M. / Kaya, I. (2015): On the Accuracy of Blue Chip Forecasts of Interest Rates and Country Risk Premiums, Applied Economics, 47(2), 113-122.

Baghestani, H. / Danila, L. (2014): Interest Rate and Exchange Rate Forecasting in the Czech Republic: Do Analysts Know Better than a Random Walk?, Czech Journal of Economics and Finance, 64(4), 282-295.

Baghestani, H. / Jung, W. / Zuchegno, D. (2000): On the Information Content of Futures Market and Professional Forecasts of Interest Rates, Applied Financial Economics, 10, 679-684.

Baghestani, H. / Marchon, C. (2012): An Evaluation of Private Forecasts of Interest Rate Targets in Brazil, Economics Letters, 115, 352-355.

Bates, J. M. / Granger, C. W. J. (1969): The Combination of Forecasts, Operational Research Quarterly, 20, 451-468.

Beechey, M. / Österholm, P. (2014): Policy Interest-Rate Expectations in Sweden: A Forecast Evaluation, Applied Economics Letters, 21(13/15), 984-991.

Belongia, M. T. (1987): Predicting Interest Rates: A Comparison of Professional and Market-Based Forecasts, Federal Reserve Bank of St. Louis Review, 69(3), 9-15.

Benke, H. (2004): Kapitalmarktprognosen auf dem Prüfstand, Stiftung/Sponsoring, 4, 25-28.

Benke, H. (2006): Was leisten Kapitalmarktprognosen? – die Sicht eines Stiftungsmanagers, Zeitschrift für das gesamte Kreditwesen, 59(17), 902-906.

Brooks, R. / Gray, J. B. (2004): History of the Forecasters, Journal of Portfolio Management, 31(1), 113-117.

Butter, den F. A. G. / Jansen, P. W. (2013): Beating the Random Walk: A Performance Assessment of Long-Term Interest Rate Forecasts, Applied Financial Economics, 23, 749-765.

Cho, D. W. (1996): Forecasts Accuracy: Are Some Business Economists Consistently Better than Others? Business Economics, 31, 45-49.

Chortareas, G. / Jitmaneeroj, B. / Wood, A. (2012): Forecast Rationality and Monetary Policy Frameworks: Evidence from UK Interest Rate Forecasts, Journal of International Financial Markets, Institutions and Money, 22(1), 209-231.

Chun, A. L. (2009): Forecasting Interest Rates and Inflation: Blue Chip Clairvoyants or Econometrics?, EFA 2009 Bergen Meeting Paper, SSRN, 946667.

Diebold, F. X. / Lopez, J. A. (1996): Forecast Evaluation and Combination, Maddala, G. S. / Rao, C. R. (Ed.), Handbook of Statistics, 14, Amsterdam.

Diebold, F. X. / Mariano, R. S. (1995): Comparing Predictive Accuracy, Journal of Business and Economic Statistics, 13, 253-263.

Domian, D. L. (1992): Money Market Mutual Fund Maturity and Interest Rates, Journal of Money, Credit and Banking, 24(4), 519-527.

Dua, P. (1988): Multiperiod Forecasts of Interest Rates, Journal of Business and Economic Statistics, 6(3), 381-384.

Francis, J. (1991): Management Anticipation of Interest Rates: The Case of Commercial Banks, Journal of Business Finance and Accounting, 18, 675-695.

Friedman, B. M. (1980): Survey Evidence on the 'Rationality' of Interest Rate Expectations, Journal of Monetary Economic, 6, 453-465.

Goodhart, C. A. E. / Lim, W. B. (2008): Interest Rate Forecasts: A Pathology, IJCB Discussion Paper, 612.

Gosnell, T. F. / Kolb, R. W. (1997): Accuracy of International Interest Rate Forecasts, Financial Review, 32(3), 431-448.

Granger, C. W. J. / Newbold, P. (1973): Some Comments on the Evaluation of Economic Forecasts, Applied Economics, 5(1), 35-47.

Greer, M. (2003): Directional Accuracy Tests of Long-Term Interest Rate Forecasts, International Journal of Forecasting, 19(2), 291-298.

Gubaydullina, Z. / Hein, O. / Spiwoks, M. (2011): The Status Quo Bias of Bond Market Analysts, Journal of Applied Finance and Banking, 1(1), 31-51.

Hafer, R. W. / Hein, S. E. (1989): Comparing Futures and Survey Forecasts of Near-Term Treasury Bill Rates, Federal Reserve Bank of St. Louis Review, 5, 33-42.

Hafer, R. W. / Hein, S. E. / MacDonald, S. S. (1992): Market and Survey Forecasts of the Three-Month Treasury-Bill Rate, Journal of Business, 65(1), 123-138.

Henriksson, R. D. / Merton, R. C. (1981): On Market Timing and Investment Performance, II. Statistical Procedures for Evaluating Forecasting Skills, Journal of Business, 54, 513-533.

Ilmanen, A. (1996): Market Rate Expectations and Forward Rates, Jounal of Fixed Income, 6(2), 8-22.

Jongen, R. / Verschoora, W. F. C. / Wolff, C. C. P. (2011): Time-Variation in Term Premia: International Survey-Based Evidence, Journal of International Money and Finance, 30, 605-622.

Joutz, F. / Stekler, H. O. (2000): An Evaluation of the Predictions of the Federal Reserve, International Journal of Forecasting, 16, 17-38.

Knüppel, M. / Schultefrankenfeld, G. (2013): The Empirical (Ir)relevance of the Interest Rate Assumption for Central Bank Forecasts, Deutsche Bundesbank Discussion Paper, 11/2013.

Kolb, R. A. / Stekler, H. O. (1996): How Well do Analysts Forecast Interest Rates?, Journal of Forecasting, 15(5), 385-394.

Kunze, F. / Gruppe, M. (2014): Performance of Survey Forecasts by Professional Analysts: Did the European Debt Crisis Make it Harder or Perhaps Even Easier? Social Sciences, 3, 128-139.

Kunze, F. / Gruppe, M. / Wendler, T. (2015): Interest Rate Forecasting and the Financial Crisis: A Turning Point in More than Just one Way, International Journal of Financial Engineering and Risk Management, 2(1), 1-16.

Kunze, F. / Kramer, J. / Rudschuck, N. (2014): Interest Rate Forecasts in Times of Financial Crisis: What Might be Interesting to Know? European Journal of Political Economy, 34, 45-52.

Kunze, F. / Wegener, C. / Bizer, K. / Spiwoks, M. (2017): Forecasting European Interest Rates in Times of Financial Crisis – What Insights do we get from International Survey Forecasts?, Journal of International Financial Markets, Institutions and Money, 48, 192-205.

Laplace, P. S. (1814): Essai Philisophique Sur les Probabilities, Paris.

Merton, R. C. (1981): On Market Timing and Investment Performance, I. An Equilibrium Theory of Value for Market Forecasts, Journal of Business, 54, 363-406.

Miah, F. / Khalifa, A. A. / Hammoudeh, S. (2016): Further Evidence on the Rationality of Interest Rate Expectations: A Comprehensive Study of Developed and Emerging Economies, Economic Modelling, 54, 574-590.

Mincer, J. / Zarnowitz, V. (1969): The Evaluation of Economic Forecasts, Mincer, J. (Ed.), Economic Forecasts and Expectation, Columbia University Press, New York, 3-46.

Mitchell, K. / Pearce, D. K. (2007): Professional Forecasts of Interest Rates and Exchange Rates: Evidence from the Wall Street Journal's Panel of Economists, Journal of Macroeconomics, 29(4), 840-854.

Mose, J. S. (2005): Expert Forecasts of Bond Yields and Exchange Rates, Danmarks National Bank, Monetary Review, 2005, 91-95.

Oliver, N. / Pasaogullari, M. (2015): Interest Rate Forecasts in Conventional and Unconventional Monetary Policy Periods, Federal Reserve Bank of Cleveland Economic Commentary, 5.

Pierdzioch, C. (2015): A Note on the Directional Accuracy of Interest-Rate Forecasts, Applied Economics Letters, 22(13), 1073-1077.

Scheier, J. / Spiwoks, M. (2006): Aktives Portfoliomanagement am britischen Anleihenmarkt, Meyer-Bullerdiek, F. / Spiwoks, M. (Ed.), Bank- und Finanzwirtschaft, 4, Frankfurt am Main.

Schwarzbach, C. / Kunze, F. / Rudschuck, N. / Windels, T. (2012): Asset Management in the German Insurance Industry: The Quality of Interest Rate Forecasts, Zeitschrift für die gesamte Versicherungswissenschaft, 101(5), 693-703.

Simon, D. P. (1989): The Rationality of Federal F/s Rate Expectations: Evidence from a Survey, Journal of Money, Credit, and Banking, 21, 388-393.

Spiwoks, M. (2003): Qualität der Zinsprognosen deutscher Banken – eine empirische Analyse, Kredit und Kapital, 36(3), 289-308.

Spiwoks, M. / Hein, O. (2007): Die Währungs-, Anleihen- und Aktienmarktprognosen des Zentrums für Europäische Wirtschaftsforschung – Eine empirische Untersuchung des Prognoseerfolges von 1995 bis 2004, AStA-Wirtschafts- und Sozialstatistisches Archiv, 01, 43-52.

Spiwoks, M. / Bedke, N. / Hein, O. (2008): Forecasting the Past: The Case of US Interest Rate, Financial Markets and Portfolio Management, 22(4), 357-379.

Spiwoks, M. / Bedke, N. / Hein, O. (2009): The Pessimism of Swiss Bond Market Analysts and the Limits of the Sign Accuracy Test – An Empirical Investigation of their Forecasting Success Between 1998 and 2007, International Bulletin of Business Administration, 4, 6-19.

Spiwoks, M. / Bedke, N. / Hein, O. (2010): Topically Orientated Trend Adjustment and Autocorrelation of the Residuals – An Empirical Investigation of the Forecasting Behavior of Bond Market Analysts in Germany, Journal of Money, Investment and Banking, 14, 16-35.

Spiwoks, M. / Gubaydullina, Z. / Hein, O. (2015): Trapped in the Here and Now – New Insights into Financial Market Analyst Behavior, Journal of Applied Finance and Banking, 5(1), 35-56.

Tabak, B. M. / Feitosa, M. A. (2008): How Informative are Interest Rate Survey-based Forecasts?, Brazilian Administrative Review, 5(4), 304-318.

Thiele, J. (1993): Kombination von Prognosen, Physica Verlag, Heidelberg.

Throop, A. W. (1981): Interest Rate Forecasts and Market Efficiency, Economic Review of the Federal Reserve Bank of San Francisco, Spring, 29-43.

Zarnowitz, V. / Braun, P. (1992): Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisions of Forecasting Performance, NBER Working Paper, 3965.

Appendix

Table 20: Overview of Studies on the Accuracy of Survey-Based Interest Rate Forecasts

Study	Countries Analyzed	Interest Rates Analyzed	Data Source	Period Considered	Methods Used	Result
Friedman (1980)	USA	Fed Funds Rate, 3-Month and 12-Month Bills, 6-Month Eurodollars, Utility Bonds, Municipal Bonds	Goldsmith-Nagan Bond and Money Market Letter	1969 - 1977	Unbiasedness Test, Efficiency Test, Consistency Test	Negative
Throop (1981)	USA	3-Month Treasury Bill Rate	Goldsmith-Nagan Bond and Money Market Letter	1970 - 1979	MSE, RMSE	Positive
Belongia (1987)	USA	3-Month Treasury Bill Rate	Wall Street Journal	1981 - 1986	Direction of Change, MAE, RMSE	Negative
Dua (1988)	USA	3-Month and 12-Month Treasury Bill Rate, Fed Funds Rate	Goldsmith-Nagan Bond and Money Market Letter / Federal Reserve Bulletin / Bond Buyer	1972 - 1985	MAE, RMSE, Theil´s U	Mixed
Simon (1989)	USA	Fed Funds Rate	Money Market Services	1984 - 1987	MAE, MSE	Negative
Hafer / Hein (1989)	USA	3-Month Treasury Bill Rate	Bond and Money Market Letter	1969 - 1989	Bias Tests, Market Efficiency Tests	Negative
Francis (1991)	USA	Various Bank Interest Rates in Pennsylvania	Call Reports	1983 - 1986	Mann-Whitney Test	Negative
Zarnowitz / Braun (1992)	USA	3-Month Treasury Bill Rate	ASA-NBER Quarterly Survey	1968 - 1990	ME, MAE, RMSE	Mixed
Hafer / Hein / MacDonald (1992)	USA	3-Month Treasury Bill Rate	Bond and Money Market Letter / Wall Street Journal	1977 - 1988	Unbiasedness Test, ME, MAE, RMSE, Theil´s U	Negative
Domian (1992)	USA	3-Month Treasury Bill Rate	IBC / Donoghue's Money Fund Report	1982 - 1990	Granger Causality	Negative
Ilmanen (1996)	USA	3-Month Treasury Bill Rate and 30-year Government Bond Yield	Wall Street Journal	1981 - 1994	Yield Change Predictions Compared to Forwards and No-Change	Negative

Study	Countries Analyzed	Interest Rates Analyzed	Data Source	Period Considered	Methods Used	Result
Kolb / Stekler (1996)	USA	3-Month Treasury Bill Rate and 30-Year Government Bond Yield	Wall Street Journal	1982 - 1990	Compared to No-Change, Random Walk measured by Skillings-Mack, Fisher's Exact	Negative
Cho (1996)	USA	3-Month Treasury Bill Rate and 30-Year Government Bond Yield	Wall Street Journal	1989 - 1994	Rank Consistency Test	Mixed
Gosnell / Kolb (1997)	GER, JPN, CH, GB, USA	3-Month Euromarket Rate	Risk	1990 - 1992	Measured against No-Change Model and Forward Rate Forecast	Mixed
Baghestani / Jung / Zuchegno (2000)	USA	3-Month Treasury Bill Rate	ASA-NBER Quarterly Survey	1983 - 1995	Unbiasedness Test	Negative
Albrecht (2000)	GER	3-Month Rate, 10-Year Government Bond Yield	Finanzen	1991 - 1997	ME	Negative
Spiwoks (2003)	GER	10-Year Government Bond Yield	Consensus Forecasts	1989 - 1999	Theil's U, TOTA Coefficient	Negative
Greer (2003)	USA	30-Year Government Bond Yield	Wall Street Journal	1984 - 1998	Binomial Test, Directional Accuracy Test, Institutional Affiliation Test	Mixed
Brooks / Gray (2004)	USA	30-Year and 10-Year Government Bond Yield	Wall Street Journal	1982 - 2002	Simplified Sign Accuracy Test, Simplified Unbiasedness Test	Negative
Benke (2004)	GER	10-Year Government Bond Yield	Handelsblatt	1991 - 2003	Simplified Sign Accuracy Test	Negative
Mose (2005)	GER, USA	10-Year Government Bond Yield	Consensus Forecasts	1989 - 2005	MAE	Negative
Baghestani (2005)	USA	3-Month Treasury Bill Rate	Survey of Professional Forecasters (SPF)	2001 - 2003	ME, MAE, RMSE	Negative
Scheier / Spiwoks (2006)	GB	10-Year Government Bond Yield	Consensus Forecasts	1989 - 2004	Theil's U_2, TOTA Coefficient	Mixed

Study	Countries Analyzed	Interest Rates Analyzed	Data Source	Period Considered	Methods Used	Result
Benke (2006)	GER	10-Year Government Bond Yield	Handelsblatt	1992 - 2005	Simplified Sign Accuracy Test	Negative
Spiwoks / Hein (2007)	FRA, GER, ITA, JPN, GB, USA	10-Year Government Bond Yield	ZEW-Finanzmarktreport	1995 - 2004	RMSE, MARE	Negative
Mitchell / Pearce (2007)	USA	3-Month Treasury Bill Rate and 30-Year Government Bond Yield	Wall Street Journal	1982 - 2002	Unbiasedness Test	Negative
Tabak / Feitosa (2008)	BRA	Short Term Interest Rate	Selic / Bloomberg and Central Bank of Brazil	1982 - 2002	MSE, Diebold-Mariano	Positive
Goodhart / Lim (2008)	NZ, GB	3-Month Official Cash NZ, Official Bank Rate UK (Libor)	RBNZ and BoE Interest Rate Forecasts	NZ 2000 - 2006 UK 1992 - 2004	Unbiasedness Test	Mixed
Spiwoks / Bedke / Hein (2008)	USA	10-Year Government Bond Yield and 3-Month Treasury Bill Rate	Consensus Forecasts	1989 - 2004	Unbiasedness Test, Sign Accuracy Test, Efficiency Test	Mixed
Spiwoks / Bedke / Hein (2009)	CH	3-Month Interest Rate and 10-Year Government Bond Yield	Consensus Forecasts	1998 - 2007	Unbiasedness Test, Sign Accuracy Test, TOTA Coefficient, Efficiency Test	Negative
Chun (2009)	USA	Fed Funds Rate, Short, Medium and Long Maturity Yield	Blue Chip Financial Forecasts	1993 - 2011	Compared against Time-series Models, Parametric Yield Curve Models and Futures Prices	Mixed
Spiwoks / Bedke / Hein (2010)	GER	3-Month Interest Rate and 10-Year Government Bond Yield	Consensus Forecasts	1989 - 2006	Unbiasedness Test, TOTA Coefficient, Efficiency Test, Sign Accuracy Test, Modified Diebold-Mariano Test, Theil's U_2	Mixed

Study	Countries Analyzed	Interest Rates Analyzed	Data Source	Period Considered	Methods Used	Result
Gubaydullina / Hein / Spiwoks (2011)	CAN, CH, ESP, FRA, GER, ITA, JPN, NLD, NOR, SWE, GB, USA	10-Year Government Bond Yield and 3-Month Interest Rate	Consensus Forecasts	1989 - 2009	TOTA Coefficient	Negative
Jongen / Verschoor / Wolff (2011)	23 countries inter alia AUS, HK, IDN, MYS, NZ, SGP, TWN	3-Month Interest Rates	Consensus Forecasts	1995 - 2009	Dickey-Fuller Unit Root Test, Expectations Hypothesis Tests	Mixed
Schwarzbach / Kunze / Rudschuck / Windels (2012)	GER	10-Year Government Bond Yield	Bloomberg, Reuters	1999 - 2011	Augmented Dickey Fuller Test (ADF Test), Johansen Approach, Granger Causality	Negative
Chortareas / Jitmaneeroj / Wood (2012)	GB	3-Month Interest Rate and 10-Year Government Bond Yield	Consensus Forecasts	1989 - 2006	Unbiasedness Test, Orthogonality Test	Negative
Baghestani / Marchon (2012)	BRA	Central Bank of Brazil Selic Interest Rate Target	Central Bank of Brazil Online Survey	2003 - 2011	Unbiasedness Test	Positive
Butter / Jansen (2013)	GER, GB, JPN, NLD, USA	10-Year Government Bond Yield	Consensus Forecasts	2003 - 2008	Successful Forecasts as a Percentage of Total Forecasts	Negative
Kunze / Kramer / Rudschuk (2013)	EUR	3-Month EURIBOR	Bloomberg/Reuters Professional Survey Forecasts	1998 - 2011	Granger Causality	Mixed
Knüppel / Schultefrankenfeld (2013)	BRA, GB	Interest Rates Central Bank	COPOM, IBGE	1999 - 2011	RMSE	Positive
Kunze / Gruppe (2014)	EUR	3-Month EURIBOR	Consensus Forecasts	1998 - 2013	Quandt-Andrews Breakpoint Test, Theil´s U	Mixed
Baghestani / Danila (2014)	CZE	2-Week Repo Rate and 12-Month Interbank Interest Rate (PRIBOR)	Czech National Bank (CNB)	2005 - 2012	Theil´s U, Diebold-Mariano Test, Fisher´s Exact Test	Mixed

Study	Countries Analyzed	Interest Rates Analyzed	Data Source	Period Considered	Methods Used	Result
Beechay / Österholm (2014)	SWE	Government Bond Yield, Forward Rate and Interest-Rate Swaps	Prospera, Swedish Financial Markets	2002 - 2012	Unbiasedness Test, Efficiency Test, Modified Diebold Mariano Test, RMSE	Mixed
Kunze / Gruppe / Wendler (2015)	EUR	3-Month EURIBOR	Consensus Forecasts	1998 - 2013	Sign Accuracy Test, Turning Point Analysis, RMSE	Mixed
Spiwoks / Gubaydullina / Hein (2015)	CAN, CH, ESP, GER, FRA, GB, ITA, JPN, NLD, NOR, SWE, USA	10-Year Government Bond Yield	Consensus Forecasts	1989 - 2009	TOTA Coefficient	Negative
Oliver / Pasaogullari (2015)	USA	Fed Funds Rate, 1-Year, 5-Year and 10-Year Bond Yield	Blue Chip Financial Forecasts	1990 - 2012	RMSE	Mixed
Baghestani / Arzaghi / Kaya (2015)	AUS, CAN, CH, EUR, GB, JPN, USA	3-Month Eurocurrency Rate and 10-Year Government Bond Yield	Blue Chip Financial Forecasts	1999 - 2008	Unbiasedness Test, Theil’s U, ME, MAE, Sign Accuracy Test, Rationality Test	Mixed
Pierdzioch (2015)	USA	Prime Interest Rate, Treasury Bill Rate and T-Bond Rate	Livingston Survey	1981 - 2013	ROC (Relative Operating Characteristic) Curves Analysing Techniques	Positive
Miah / Khalifa / Hammoudeh (2016)	30 Countries inter alia CHN, HK, IND, KOR, PHL, SGP, THAI, TWN	Long-Term and Short-Term Interest Rates	Fx4casts.com	2001 - 2012	Unbiasedness Test, Efficiency Test, Unit Root Test	Mixed
Kunze / Wegener / Bizer / Spiwoks (2017)	GER, GB	3-Month Interbank Rate and 10-Year Government Bond Yield	Consensus Forecasts	1993 - 2014	RMSE, Theil´s U, Diebold-Mariano Test	Negative

Chapter 6

OVERCONFIDENCE: THE INFLUENCE OF POSITIVE AND NEGATIVE AFFECT

Submitted to:
Journal of Neuroscience, Psychology, and Economics

Published in:
Wolfsburg Working Papers, 17-02, March 2017.
Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse, 17-01, January 2017.

Abstract

The consequences of overconfidence affect many spheres of economic life. As yet, few factors are known that determine the extent of possible overconfidence. There are also few studies concerning the influence of positive and negative emotions on self-assessment. It has not yet been examined whether emotions can affect learning effects regarding self-assessment, wherefore the present study addresses this research question.

In a real-effort-task experiment the participants are presented with tasks over the course of 5 rounds. After each round, they are asked to assess their own performance. They are then given feedback on their actual performance, thereby allowing for learning effects. Their mood is induced by positive (treatment “positive”), negative (treatment “negative”) and neutral (treatment “neutral”) movie clips. There are no significant differences in the three treatments regarding absolute and relative overconfidence. However, the participants’ moods differed with regard to the occurrence of learning effects. Obvious learning effects can be established in a neutral mood when examining absolute overconfidence. These learning effects cannot be detected in positive and negative moods.

Keywords

Overconfidence, emotions, self-assessment, feedback, learning effect

JEL Classification

D81, D83, D84

1 Introduction

In psychological research, the overconfidence bias is a widely-known phenomenon of individual behavior. In economic research, this phenomenon is regarded as a reason for inefficient markets (Proeger and Meub, 2013). Different groups of people such as investors, managers, bankers and other economic actors have been proven to show overconfidence (Barber and Odean, 2001; Menkhoff et al., 2013; Ifcher and Zarghamee, 2014). Considering the numerous psychological and economic studies that have been concerned with the phenomenon of overconfidence,[1] the following literature review is limited to those studies that directly address the present research question:

The study by Allwood and Bjorhag (1991) did not reveal any findings that indicate the influence of negative affect on the extent of overconfidence. A positive mood could not be induced successfully which is why the influence of positive affect could not be examined.

Allwood et al. (2002) compared the effect of positive and negative emotions on the extent of overconfidence without taking into account a neutral treatment (control group). Their study showed that the participants are liable to overconfidence when being in a positive or in a negative mood. A significant difference regarding overconfidence between the two treatments could not be detected.

Kuvaas and Kaufmann (2004) published a similar study. They also compared the influence of positive and negative emotions on overconfidence – and did not consider a neutral treatment (control treatment) either. They concluded that there are no differences between the overconfidence shown in both treatments.

De Paola et al. (2014) assessed the effect of superstition and positive and negative emotions on overconfidence. Their examination was conducted as a field experiment with approx. 700 Italian students who were randomly allocated numbered seats before a written exam. Moods were induced by lucky numbers, unlucky numbers and neutral numbers. De Paola et al. (2014) ascertain that the students generally overestimate themselves systematically and that their overconfidence increases due to the lucky numbers. Unlucky numbers, by contrast, have a cushioning effect on the extent of overconfidence.

The study by Ifcher and Zarghamee (2014) is of great importance for the present study and is therefore presented in detail in the following: In two experiments, Ifcher and Zarghamee examined if positive, negative or neutral (control treatment) moods affect self-assessment.

In the first experiment, they examined the effects of positive and neutral moods on the extent of overconfidence. The moods were successfully induced with the help of movie clips. Those participants with a positive mood show more absolute overconfidence as well as more relative overconfidence than those participants with a neutral mood.[2] Overestimation is a common phenomenon relating to this behavior. 72% of the participants showed absolute

[1] For an overview see Moore and Healy (2008), Adel and Mariem (2013), Ifcher and Zarghamee (2014).

[2] Research literature distinguishes between two kinds of overconfidence: "absolute overconfidence" (AOC), a form of self-evaluation in absolute numbers, and "relative overconfidence" (ROC), when the participants assess their own success in comparison to other participants (see also Ifcher and Zarghamee (2014), De Paola et al. (2014)). The following tables and figures will contain the abbreviations AOC and ROC.

overconfidence, while 62% showed relative overconfidence. On average, the participants overestimated their performance by 18.29%. The difference between positive and neutral moods, however, proved to be insignificant regarding absolute overconfidence.

In the second experiment, Ifcher and Zarghamee examined the effect of negative and neutral moods on the extent of overconfidence. The moods were successfully induced using movie clips. Those participants with a negative mood showed both more absolute overconfidence and more relative overconfidence than those participants with a neutral mood. These differences, however, are statistically not relevant. The participants overestimated their performance by averagely 9.2%.

Both positive and negative mood induction increases overconfidence in comparison to participants with a neutral mood. Comparing the treatments "positive" and "negative", the overconfidence in the "negative" treatment is lower than in the "positive" treatment. First, the results of these previous studies will be reviewed. Therefore, the first hypothesis reads as follows:

Hypothesis 1:

H1a: Positive or negative emotions influence overconfidence to a larger extent than a neutral mood does.

H1b: Overconfidence is stronger when feeling positive emotions rather than negative emotions.

The above-mentioned studies have not examined the development of overconfidence over time but only conducted the experiments in one session. However, the present study will focus on the long-term influence of positive and negative affect on overconfidence.

Conducting the experiment in more than one session enables the researcher to examine if the participants have any learning effects when assessing their own performance. Since the experiment consists of 5 rounds and since the participants receive feedback on their performances after each round, they can reflect on their self-assessment for the next round and thereby learn to easier assess their own capabilities. This makes it possible to analyze the influence of emotions on learning effects. We therefore consider the possibility that moods can influence possible learning effects that result from repeated self-assessment with individual feedback.

There are now some studies that investigate overconfidence within the framework of games with multiple periods containing feedback:

Clark and Friesen (2009) carried out a real-effort-task experiment in two rounds. After the first round, the participants estimated the number of the tasks that they completed correctly and were then given feedback on the actual number of correct answers. For the second round, the participants reconsidered their self-assessment and indeed performed better in the second round than in the first round. Hence, they experienced some learning effects.

Grossmann and Owens (2012) conclude that small learning effects could be achieved through self-assessment. However, they did not have any statistically relevant success.

Proeger and Meub (2014) performed a real-effort-task experiment. The participants had to solve 10 simple calculations with three fixed variables and one random variable. The participants were then asked to assess their own performance. They were asked to estimate how many problems they solved correctly. Each participant received some feedback. It can be concluded that the participants achieved a learning effect through self-assessment. In each round (3 rounds in total), the participants' self-assessment improved.

The previous studies did not examine the influence of emotions on learning effects. It remains to be examined whether learning effects are influenced by emotions. Therefore, it must be researched whether repeated self-assessment with individual feedback can lead to individuals breaking away from overconfidence even when they are in a positive or in a negative mood. Consequently, hypothesis 2 reads as follows:

Hypothesis 2:

H2a: Learning effects are triggered by feedback on repeated self-assessment in a neutral mood.

H2b: Learning effects that are attained by repeated self-assessment with individual feedback are not affected by positive or negative moods.

The present paper is structured in 4 paragraphs. The following paragraph describes the experimental design. The results are then presented in the subsequent paragraph. Finally, the last paragraph summarizes the most striking results of the study.

2 Experimental Design

The experimental design follows the design by Ifcher and Zarghamee (2014) who took the following approach:

First, the participants take part in a quiz that consists of 30 questions (20 general knowledge questions (Moore and Small, 2007) and 10 mathematical tasks (Niederle and Vesterlund, 2007)), which they have 15 minutes to answer. They will receive $0.50 for each correct answer. Secondly, the mood is induced. For mood induction, the participants are shown movie clips that are supposed to trigger positive or negative emotions. Meanwhile, the control group looks at a screen saver or watches neutral movie clips. In the next step, the participants assess their performance in the quiz. They estimate the number of the tasks that they completed correctly (absolute overconfidence) and assess the quality of their performance in comparison to their fellow gamblers (relative overconfidence). They receive $5.00 for each correct estimation. The fourth step includes the manipulation check of the participants' mood induction, using PANAS[3]. In the fifth step, they answer questions regarding demographic and personal characteristics. The average profit is $15.00.

[3] PANAS stands for Positive and Negative Affect Schedule.

The experimental design of the present study is structured as follows: After the participants have read the detailed instructions, they must answer four control questions[4] correctly. Before the start of the experiment, their mood is then scaled using the following question:

How are you feeling now? Please mark the adequate number!		
	1—2—3—4—5—6—7—8—9—10	
very bad		very good

Each round starts with a real-effort task. There are 25 general knowledge questions and 25 mathematical tasks. The general knowledge questions were taken from Moore and Small (2007) and supplemented with five similar questions. The mathematical tasks by Niederle and Vesterlund (2007) were also used and supplemented with 15 similar tasks. Hence, the experiment consists of 50 tasks and is structured into 5 rounds. Each round contains 5 general knowledge tasks and 5 mathematical tasks. The participants are allowed 45 seconds to complete the real-effort task. They receive a material incentive to motivate them to answer the questions correctly. 2 points are awarded for each correct answer. They can receive 100 points in total if all questions are answered correctly. This is equivalent to a payout of €15.00.

Moods are then induced using short movie clips. Emotional movie clips are one of the most effective methods to trigger emotions.[5] Movie clips are often used in economic experiments to evoke certain moods (see e.g. Allwood et al., 2002; Kirchsteiger et al., 2006; Rottenberg et al., 2007; Ifcher and Zarghamee, 2014; Oswald et al., 2015).

The movie clips have been chosen from the study by Schaefer et al. (2010), which analyzes more than 70 movie clips for their effect on mood induction. The movie clips are categorized into positive, negative and neutral effects on a person's mood and ranked according to their effectivity.

The present experiment uses the following movie clips to induce a positive mood: (1) Benny and Joon (122 seconds): Benny (Johnny Depp) clowns around in a café. (2) Life is Beautiful (266 seconds): Mother and son are reunited after World War II. (3) Dead Poets Society (163 seconds): The students mount their desks to express their solidarity with Mr. Keating (Robin Williams). (4) Forrest Gump (121 seconds): Father (Tom Hanks) meets son. (5) Dinner for Schmucks (101 seconds): Complex comic scene.

The following movie clips were used to induce a negative mood: (1) Saving Private Ryan (327 seconds): A combat scene in World War II. (2) The Piano (42 seconds): One of the main characters has a finger chopped off with an ax. (3) The Blair Witch Project (232 seconds): Final scene when the main characters are obviously killed. (4) Schindler's List (76 seconds): Corps-

[4] The control questions are used to check if the participants understood the instructions of the experiment.
[5] There are different methods of mood induction, including real situations, memories and imaginations, noises and music, gifts, movie clips or the Velten technology. See Westermann et al. (1996).

es are burnt in a concentration camp. (5) City of Angels (257 seconds): Maggie (Meg Ryan) dies in Seth's (Nicolas Cage) arms.

The neutral movie clips were the following: (1) The Lover (43 seconds): Marguerite (Jane March) gets into a car. She is taken to a house in a busy street where she knocks on a door. A Chinese man opens the door and she enters the house. (2) Blue (40 seconds): A man clears the drawers of his desk. A woman passes through an alley and salutes another woman on the way. (3) Train Ride (58 seconds): A train crosses a green countryside.[6] (4) Blue (25 seconds): A woman holding a box goes up an escalator. (5) Blue (16 seconds): A person throws a piece of foil out of the window of a car.

Those movie clips inducing negative emotions are shown in the treatment "negative", while those clips triggering positive emotions are shown in the treatment "positive" and the neutral clips are presented in the treatment "neutral" (control group).

In each round, a manipulation check was conducted after the participants watched the movie clip. The participants were asked the following question:[7]

Which emotions did you experience while watching the movie clip? Please mark one number accordingly! 1—2—3—4—5—6—7—8—9—10 very negative very positive

Afterwards, the participants' self-assessment is captured by the following questions:

- How many of the 10 tasks did you complete correctly?
- How many tasks did you complete correctly compared to the other participants, i.e. how many more or fewer tasks compared to the average number of tasks completed by the other participants?

To motivate the participants to assess their performance as accurately as possible, they receive 8 points for each overlap of the estimated and the actual performance. They can receive 80 points in total, which equates to a payout of €12.00.

The investigation of the self-assessment follows the approach by Ifcher and Zarghamee (2014): Absolute overconfidence is captured and relative confidence is considered.

The absolute overconfidence is the difference between the assumed number of correctly completed tasks and the actual number of correctly solved tasks. If, for example, a participant assumes that they completed 10 (4) tasks correctly but only 7 tasks were actually solved, their absolute overconfidence would be +3 (-3).

[6] The movie clip "Train Ride" is similar to the clip chosen by Gendolla and Krüsken (2002).

[7] Similar manipulation checks were conducted in the studies by Kirchsteiger et al. (2006), Rottenberg et al. (2007), Lahav and Meer (2012), Andrade et al. (2015).

Relative overconfidence results from the difference between the assumed and the actual relative success in comparison to the other participants. For example, a participant assumes that they averagely solved 4 tasks more (4 tasks less) than the other participants. In fact, they only solved 2 more tasks correctly than the other participants. Hence, relative overconfidence is +2 (-6).

After each of the five rounds, the participants are given feedback on the success of their self-assessment. Each participant is told how many tasks they completed correctly (absolute overconfidence) and how they performed in relation to the other participants (relative overconfidence). Over the course of the five rounds, the participants can thereby learn from their experience in the previous rounds and progressively assess their own performance in a more realistic way.

With the exception of presented movie clips, the experimental process is the same for the treatments "negative", "positive" and "neutral". Table 1 provides an overview on the research method.

Table 1: Overview on the Treatments of the Experiment

Treatment	Real-effort Task	Mood Induction	Manipulation Check	Self-assessment	Feedback
Negative	Yes	Yes	Yes	Yes	Yes
Neutral	Yes	Yes	Yes	Yes	Yes
Positive	Yes	Yes	Yes	Yes	Yes

The treatments "positive" and "negative" lasted approx. 45 minutes, while the experiment lasted approx. 35 minutes for the control group. This can be attributed to the length of the movie clips, which are considerably shorter for the induction of a neutral mood than for the induction of a positive or negative mood.

The participants are remunerated for their performance. The total number of points awarded to each participant (180 points are possible to achieve in total) is converted into a sum of money in euros. Each point equals €0.15. Each participant furthermore receives a show-up fee of €2.50. The participants can earn €29.50 in total. They earned €12.81 on average. The minimum payout was €2.50; the maximum payout was €19.60.

The experiment was programmed in z-Tree (Fischbacher, 2007). The instructions and screenshots are given in the appendices of this paper.

The experiment was conducted between 30 March and 22 April 2015 with students of the Ostfalia University of Applied Sciences. 104 participants took part in the 22 sessions of the experiment. 45 participants study at the Faculty of Business (43.3%), 45 participants study at the Faculty of Automotive Engineering (43.3%) and 14 participants study at the Faculty of Public Health Services (13.5%). 28 women (26.9%) and 76 men (73.1%) participated in the experiment. The participants were assigned to the treatments as follows: 34 participants (32.7%) played the treatment "negative", 32 participants (30.8%) the treatment "positive" and 38 participants (36.5%) played the treatment “neutral”. The average age of the participants was 23.7 years.

3 Results

3.1 Mood Induction

First, I will determine if the mood induction was successful. The average mood of the participants in treatment “negative” was 3.28 (SD 1.24). The average mood of the participants in treatment “neutral” was 5.52 (SD 0.95). The average mood of the participants in treatment “positive” was 7.03 (SD 1.50). Figure 1 gives an overview on the participants’ moods in each treatment and round.

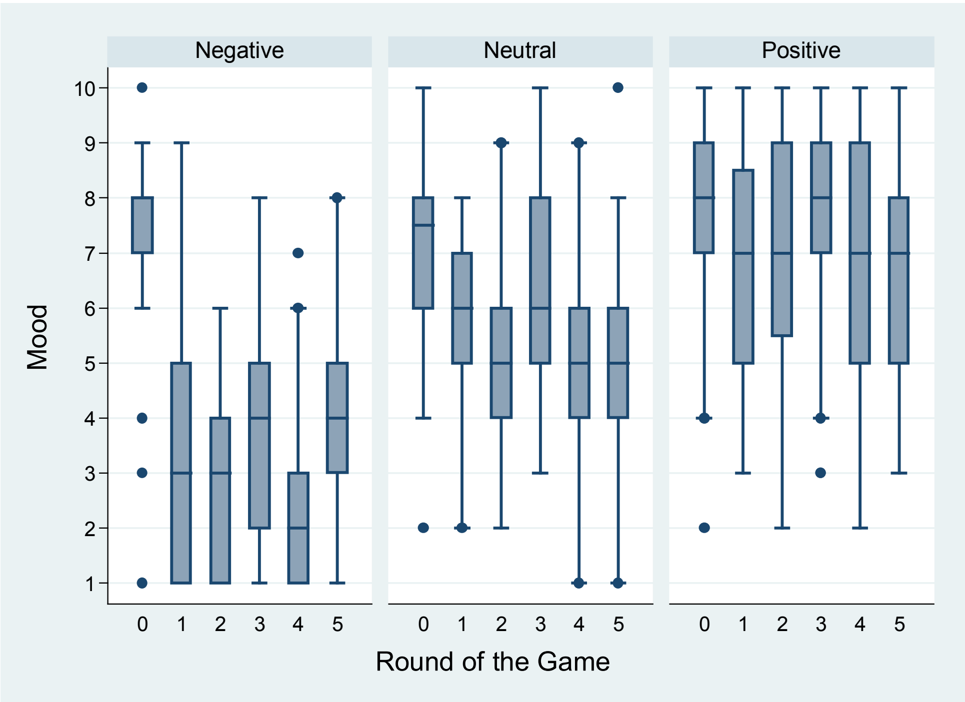


Figure 1: Participants’ Moods before the Experiment and in the Five Rounds of the Game

Figure 1 shows that the mood induction in the individual treatments was successful. Before the start of the treatments, the participants were in a positive mood (see the boxplots of round 0). After the mood induction, the moods of the participants in the three treatments disperse (rounds 1-5). Figure 2 summarizes the different moods of the participants in the three treatments of the five rounds. This highlights the success of the mood induction. The participants’ moods before the start and in each round of the experiment as well as the standard deviations are shown in Table 2.

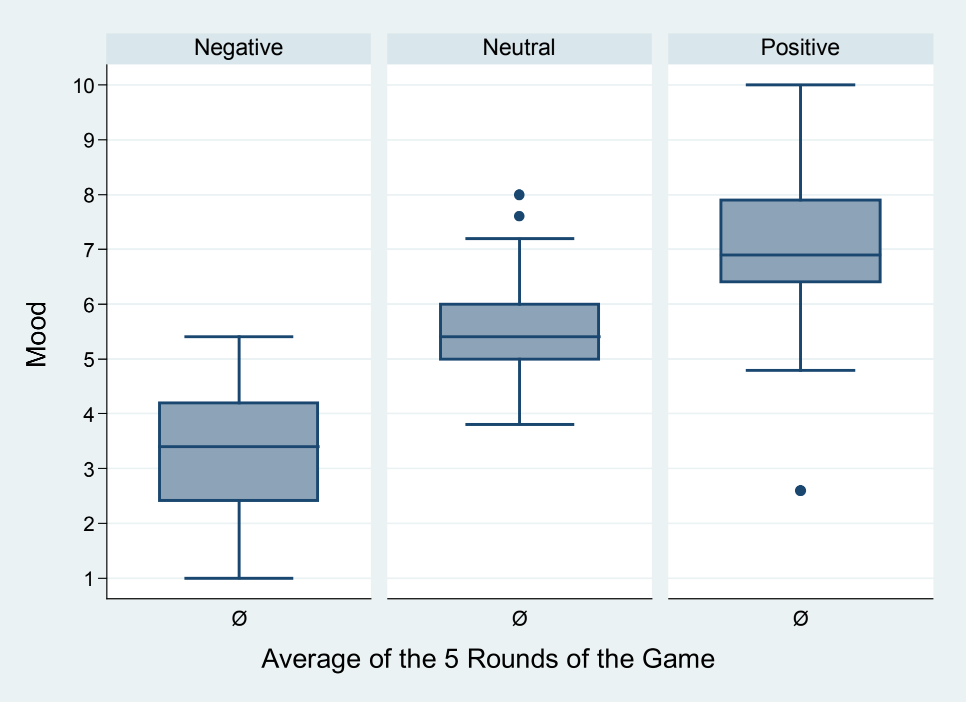


Figure 2: Participants' Average Moods in the Five Rounds of the Game (Excluding the Mood before the Experiment)

Table 2: Participants' Average Moods per Round

Treatment	#	Average Mood per Round (Standard Deviation)						
		Before the Experiment	1	2	3	4	5	**Ø**[a]
Negative	34	7.21 (1.75)	3.32 (2.10)	2.65 (1.47)	3.85 (1.88)	2.53 (1.67)	4.06 (1.54)	**3.28*** (1.24)**
Neutral	38	7.16 (1.84)	5.68 (1.44)	5.26 (1.54)	6.26 (1.98)	5.37 (1.63)	5.00 (2.10)	**5.52*** (0.95)**
Positive	32	7.69 (1.82)	6.88 (2.10)	7.09 (1.96)	7.53 (1.93)	7.13 (2.24)	6.53 (2.27)	**7.03*** (1.50)**

Please note: The significant values are printed in bold (*** p<0.01; ** p<0.05; * p<0.1).

[a]To calculate the average values, the values of the five rounds of the game were taken into account; the values before the start of the experiment were not considered.

Comparing the treatments (Table 2), the treatment "negative" shows significantly lower values than the treatment "positive" (z= -6.561, p= 0.0000; Mann-Whitney U Test). The treatment "negative" also shows considerably lower values when compared to the treatment "neutral" (z= -6.382, p= 0.0000; Mann-Whitney U Test). Furthermore, the treatment "positive" shows significantly higher values than the treatment "neutral" (z= 4.570, p= 0.0000; Mann-Whitney U Test). Mood induction was therefore successful.

3.2 Absolute Overconfidence

This subchapter presents the results of absolute overconfidence. Figure 3 portrays the results of the three treatments and the five rounds of the game. It can clearly be observed that the treatments “negative”, “neutral” and “positive” do not differ greatly regarding absolute overconfidence. The median is on the level of 0 for a total of seven times (treatment “negative”: Rounds 4 and 5; treatment “neutral”: Rounds 3 and 4; treatment “positive”: Rounds 2, 3 and 4).

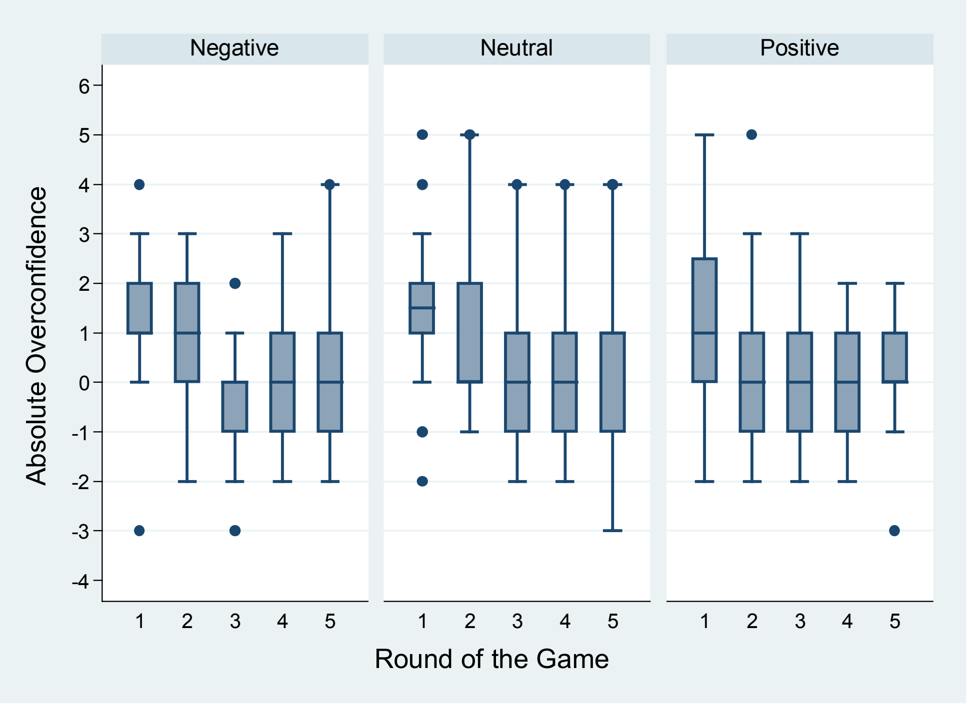


Figure 3: Participants’ Absolute Overconfidence in the Three Treatments and in the Five Rounds of the Game

Figure 4 shows the average absolute overconfidence in the five rounds. It can clearly be seen how close the absolute overconfidence in the three treatments lie together. The lower whiskers are all between 0 and -1. The upper whiskers are all between 1 and 2. The median for all three treatments is between 0 and 1. In the treatments “neutral” and “positive” 25% of the values are lower than 0 and 75% of the values are higher than 0. In the treatment “negative” 80% of the values are higher than 0 and 20% of the values are lower than 0. Thus, all three treatments show a clear tendency towards overconfidence.

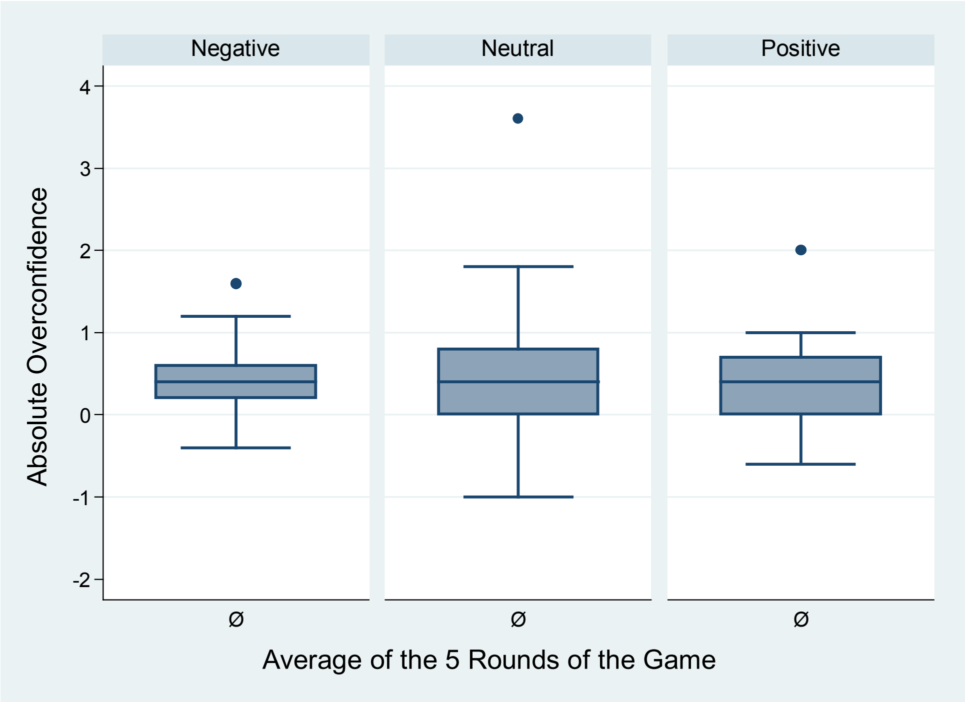


Figure 4: Participants' Absolute Overconfidence per Treatment

Table 3 shows the participants' self-assessment regarding absolute overconfidence in each treatment. It can be observed that overestimation is predominant. In the treatment "negative", 46.5% of participants overestimate themselves, in the treatment "positive" 43.1% of participants overestimate themselves and in the treatment "neutral" 51.1% of participants overestimate themselves. Underestimation occurs in the treatment "negative" with 25.9% of participants, with 26.3% in the treatment "positive" and with 26.8% in the treatment "neutral". Accurate self-assessment was detected with 27.6% of participants in treatment "negative", with 30.6% in the treatment "positive" and with 22.1% in the treatment "neutral".

Table 3: Participants' Self-assessment (Absolute Overconfidence) per Treatment

Treatment	#	Absolute Overconfidence in %		
		Underrating	Adequate Self-Assessment	Overconfidence
Negative	34	25.9	27.6	46.5
Neutral	38	26.8	22.1	51.1
Positive	32	26.3	30.6	43.1

Table 4 shows a summary of all the numbers on overconfidence. Surprisingly, the average absolute values for overconfidence are highest in the treatment "neutral" with 0.55 (SD 0.79). This number is followed by 0.41 (SD 0.46) in the treatment "negative" and by 0.40 (SD 0.61) in the treatment "positive".

Table 4: Participants' Absolute Overconfidence per Round

Treatment	#	Average AOC Values per Round (Standard Deviation)					
		1	2	3	4	5	**Ø**
Negative	34	1.29 (1.27)	0.71 (1.34)	-0.47 (1.31)	0.38 (1.37)	0.15 (1.31)	**0.41** **(0.46)**
Neutral	38	1.50 (1.43)	0.71 (1.51)	0.21 (1.44)	0.03 (1.24)	0.32 (1.65)	**0.55** **(0.79)**
Positive	32	1.38 (1.79)	0.31 (1.75)	0.00 (1.32)	0.09 (1.00)	0.22 (1.26)	**0.40** **(0.61)**

Please note: The significant values are printed in bold (*** p<0.01; ** p<0.05; * p<0.1).

In the treatment "negative", the participants assumed that they averagely solved 5.39 (SD 1.59) tasks correctly. They actually completed only 4.98 (SD 1.68) tasks correctly. The difference is 0.41. Hence, the participants overestimated their own performance by 8.23%. The same is true for the treatment "positive". On average, the participants assumed that they solved 5.36 (SD 1.56) tasks correctly. They actually only completed 4.96 (SD 1.63) tasks correctly. The difference is 0.40. The participants therefore overestimated their own performance by 8.06%. In the treatment "neutral", the participants assumed that they averagely solved 5.16 (SD 1.34) tasks correctly. They actually only completed 4.61 (SD 1.75) tasks correctly. The difference is 0.55, wherefore they overestimated their own performance by 11.93%.

The Wilcoxon-Rank-Sum Test does not reveal any significant differences between the treatments "positive" or "negative" and the treatment "neutral" (treatment "negative" vs. treatment "neutral": z=-0.705, p=0.4805; treatment "positive" vs. treatment "neutral": z=-0.706, p=0.4801). Therefore, hypothesis 1a for absolute overconfidence must be discarded. The overconfidence of the participants in the treatments "negative" and "positive" is not significantly higher than in the treatment "neutral". The results of the study by Ifcher and Zarghamee (2014) are hereby confirmed.

Hypothesis 1b must also be rejected for absolute overconfidence since the values of the treatment "positive" are not significantly higher than the values of the treatment "negative" (Wilcoxon-Rank-Sum Test: z=0.045, p=0.9638).

3.3 Relative Overconfidence

This subchapter presents the results on relative overconfidence. Figure 5 gives an overview on the relative overconfidence of the participants in the respective rounds and treatments. This overview, similarly, reveals only few differences among the treatments. The interquartile ranges are larger for relative overconfidence than for absolute overconfidence.

One reason for this is likely the difficulty that the participants experience in assessing the other participants' performance in the respective round in order to estimate their own success in comparison to the other participants.

For the large part, the boxes stretch below 0. This indicates that the participants rather underestimate than overestimate their own relative performance. It is remarkable that 9 out of 15 medians are at the level of 0 and that 12 of the 15 quartiles are on the level of 1.

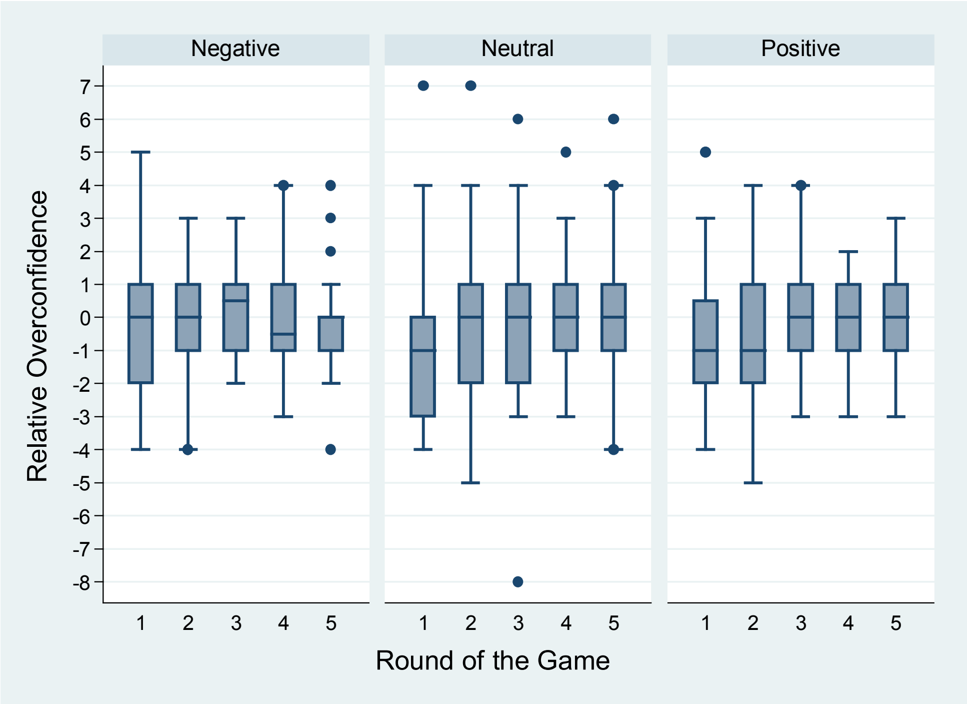


Figure 5: Participants' Relative Overconfidence in the Rounds and Treatments

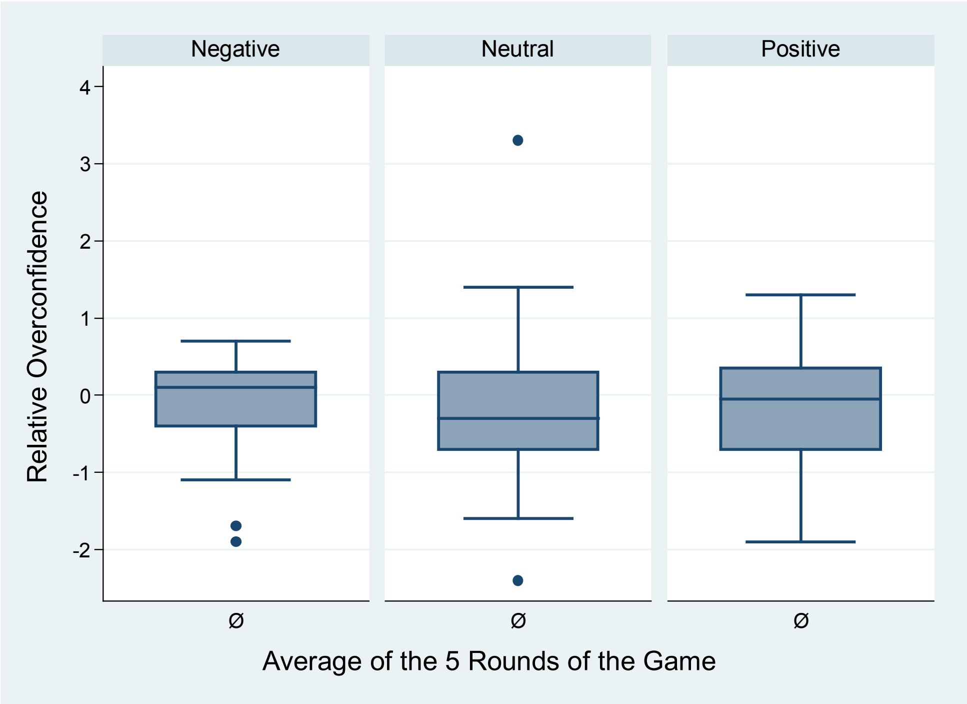


Figure 6: Participants' Relative Overconfidence per Treatment

Figure 6 shows the average values of the five rounds per treatment concerning relative overconfidence. As is the case for the findings on absolute overconfidence, the values for relative overconfidence are similar. It can be clearly seen that 80% of the boxes are below 0. Only 20% are above 0.

Different to absolute overconfidence, the participants overestimate their own performance with regard to relative overconfidence (Table 5). In the treatment “negative” 38.2% underestimate their relative success, with 45.6% in the treatment “positive” and 44.2% in the treatment “neutral”. Overestimation can also be observed. In the treatment “negative” 35.9% overestimate their success relative to the other participants’ performance, with 33.8% overestimation in the treatment “positive” and 29.5% in the treatment “neutral”. A correct self-assessment was given by 25.9% of the participants in the treatment “negative”, by 20.6% in the treatment “positive” and by 26.3% in the treatment “neutral”.

Table 5: Participants’ Self-assessment Relative to the Other Participants (Relative Overconfidence) per Treatment

Treatment	#	Relative Self-Assessment in %		
		Underestimation	Accurate Self-Assessment	Overestimation
Negative	34	38.2	25.9	35.9
Neutral	38	44.2	26.3	29.5
Positive	32	45.6	20.6	33.8

Table 6: Participants’ Relative Overconfidence per Round

Treatment	#	Average ROC Values per Round (Standard Deviation)					
		1	2	3	4	5	**Ø**
Negative	34	-0.38 (1.94)	-0.18 (1.78)	0.29 (1.22)	-0.12 (1.74)	-0.18 (1.53)	**-0.11 (0.64)**
Neutral	38	-0.84 (2.26)	-0.26 (2.24)	-0.13 (2.42)	-0.08 (1.63)	0.11 (2.08)	**-0.24 (0.96)**
Positive	32	-0.69 (2.05)	-0.66 (2.21)	0.00 (1.48)	-0.09 (1.40)	0.13 (1.50)	**-0.26 (0.74)**

Please note: The significant values are printed in bold (*** $p<0.01$; ** $p<0.05$; * $p<0.1$).

Table 6 shows the values for relative overconfidence. The average values of the five rounds are striking because they are negative in the three treatments. It can be concluded that, on average in the five rounds, the participants underestimate their own performance relative to the other participants’ performance in the session.

In the treatment “negative”, the participants assumed that they averagely completed 0.23 (SD 1.87) fewer tasks correctly than the average of the other participants. In fact, they fulfilled only 0.12 (SD 1.96) fewer tasks successfully than the average of the other participants. The difference is 0.11. In the treatment “positive”, the participants assumed that they averagely completed 0.26 (SD 1.71) fewer tasks correctly than the average of the other participants. However, it turned out that they accomplished exactly as many tasks as the average of the other participants (0.00 (SD 1.68)). The difference is 0.26. In the treatment “neutral”, the participants assumed that they averagely completed 0.34 (SD 1.71) fewer tasks correctly than the average of the other participants. In fact, they fulfilled only 0.10 (SD 1.81) fewer tasks successfully than the average of the other participants. This makes a difference of 0.24.

To answer the question whether there are significant differences between the treatments, the Wilcoxon-Rank-Sum Test was used (treatment “negative” vs. treatment “positive”: z= 0.779, p= 0.4361; treatment “negative” vs. treatment “neutral”: z= 1.459, p= 0.1445; treatment “positive” vs. treatment “neutral”: z= 0.579, p= 0.5627). No significant differences between the three treatments can be established. Hypothesis 1a must be discarded for relative overconfidence. The participants in the treatments “negative” and “positive” do not show a significantly higher relative overconfidence than the participants in the treatment “neutral”. These results are contradictory to the results by Ifcher and Zarghamee (2014), who detected a stronger manifestation of relative overconfidence in positive and negative moods than in a neutral mood. Hypothesis 1b must also be neglected for relative overconfidence because the relative overconfidence in the treatment “positive” was not significantly higher than the relative overconfidence in the treatment “negative”.

3.4 Learning Effects

This subchapter will analyze if the participants experienced any learning effects. To assess hypotheses 2a and 2b the values of absolute overconfidence in the first three rounds were compared to the values of absolute overconfidence in the last two rounds. This is reasonable because the participants might need more than one feedback on their performance to improve their self-assessment, or in order to experience a learning process.

To compare the first three rounds of the game to the last two rounds, the Wilcoxon-Signed-Rank Test was administered.

In the treatment “neutral”, significant learning effects could be established (z= 3.187, p= 0.0014). Those participants taking part in the treatment “neutral” were obviously able to use the feedbacks of the first round to improve their self-assessment over time.

Therefore, hypothesis 2a cannot be neglected. The results by Clark and Friesen (2009) and by Proeger and Meub (2014) can be confirmed.

The results for the treatments “positive” and “negative” are fundamentally different. The Wilcoxon-Signed-Rank Test shows that the results of the last two rounds do not significantly differ from the results of the first three rounds (treatment “positive” (z= 1.600, p= 0.1096), treatment “negative” (z= 1.301, p= 0.1934)).

This indicates that both positive and negative moods prevent the participants from having any learning effects and from achieving more realistic self-assessment.

The same approach was taken to assess relative overconfidence. The values of the relative overconfidence in the first three rounds were compared to the values of the absolute overconfidence in the last two rounds. No learning effects can be detected in neither the treatment “neutral” nor in the treatments “positive” or “negative”. The relative self-assessment in the first three rounds does not significantly differ from the relative self-assessment in the last two rounds (Wilcoxon-Signed-Rank Test for the treatment “neutral”: z= -1.588, p= 0.1122; for the treatment “positive”: z= -1.562, p= 0.1183; for the treatment “negative”: z= 0.342, p= 0.7323). Regarding relative overconfidence, hypothesis 2a must therefore be discarded. The assessment of hypothesis 2b for learning effects concerning relative overconfidence is thereby omitted.

The fact that the assessment of relative overconfidence has not returned any learning effects even in the treatment “neutral” is probably owing to the complexity of assessing one’s own performance in comparison to the other participants’ achievements (relative overconfidence) as opposed to assessing one’s own performance (absolute overconfidence). More time is needed to reflect on the other participants’ performance in the session in order to assess one’s own relative capability.

4 Conclusion

The present study examines the phenomenon of overconfidence and addresses two research questions: (1) The study examines the influence of positive and negative emotions on self-assessment. (2) The study also addresses the question if any learning effects through self-assessment are influenced by positive or negative emotions.

The economic experiment was conducted with 122 students of the Faculties of Business, Automotive Engineering and Public Health Services of the Ostfalia University of Applied Sciences.

Positive, negative and neutral movie clips were used for mood induction. Absolute and relative overconfidence were equally assessed.

The study produced the following results:

1. The participants’ moods were successfully induced using positive, negative and neutral movie clips. The average moods in the three treatments are significantly different. The treatment “negative” shows an average mood value of 3.28, the treatment “neutral” an average mood value of 5.52 and the treatment “positive” an average mood value of 7.03.
2. The participants’ overconfidence in the treatments “negative” and “positive” is not significantly higher than their overconfidence in the control group (treatment “neutral”). This is equally true for the absolute overconfidence and the relative overconfidence. Therefore, hypothesis 1a must be discarded.
3. Furthermore, the treatment “positive” does not present a significantly increased tendency towards overconfidence when compared to the treatment “negative”. This is equally true for the absolute overconfidence and the relative overconfidence. Hence, hypothesis 1b must also be rejected.
4. Participants with a neutral mood (control group) achieve striking learning effects regarding absolute overconfidence. In the last two rounds, they assess their own performance

significantly more accurately than in the first three rounds. This is why hypothesis 2a cannot be rejected.

5. Participants with a positive or a negative mood (treatment “positive” and treatment “negative”) do not achieve any considerable learning effects regarding absolute overconfidence. They do not assess their performance more accurately in the last two rounds than in the first three rounds. It can hence be established that both positive and negative emotions can influence possible learning effects. Therefore, hypothesis 2b must be rejected.
6. Regarding relative overconfidence, no learning effects could be detected. In neither of the treatments “neutral”, “positive” or “negative”, the participants can forecast their relative performance over the course of the game more accurately than in any other. In all three treatments, the relative overconfidence of the first three rounds does not significantly differ from the relative overconfidence in the last two rounds.

Acknowledgements

I would especially like to thank Markus Spiwoks, Kilian Bizer, Till Pröger, Lukas Meub, Julia Brüggemann, John Jermyn, Antje Fitschen-Lischewski, Matthias Noack, Varun Vruddhula and Kai-Christian Zellemann, as well as various seminar participants at Göttingen and Hamburg for helpful comments and support.

References

Adel, B. / Mariem, T. (2013): The Impact of Overconfidence on Investors' Decisions, Business and Economic Research, 3(2), 53-75.

Alicke, M. / Klotz, M. / Breitenbecher, D. / Yurak, T. / Vredenburg, D. (1995): Personal Contact, Individuation, and the Better-Than-Average Effect, Journal of Personality and Social Psychology, 68(5), 804-825.

Allwood, C. M. / Bjorhag, C. G. (1991): Mood and Realism of Confidence Judgements of one's own Answers to General Knowledge Questions, Scandinavian Journal of Psychology, 32(4), 358-371.

Allwood, C. M. / Granhag, P. A. / Jonsson, A.-C. (2002): Does Mood Influence the Realism of Confidence Judgements? Scandinavian Journal of Psychology, 43(3), 253-260.

Andrade, E. B. / Odean, T. / Lin, S. (2015): Bubbling with Excitement: An Experiment, Review of Finance, 20(2), 1-20.

Barber, B. M. / Odean, T. (2001): Boys will be Boys: Gender, Overconfidence, and Common Stock Investment, Quarterly Journal of Economics, 116(1), 261-292.

Clark, J. / Friesen, L. (2009): Overconfidence in Forecasts of own Performance: An Experimental Study, Economic Journal, 119(534), 229-251.

De Paola, M. / Gioia, F. / Scoppa, V. (2014): Overconfidence, Omens and Gender Heterogeneity: Results from a Field Experiment, Journal of Economic Psychology, 45, 237-252.

Fischbacher, U. (2007): Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments, Experimental Economics, 10(2), 171-178.

Gendolla, G. / Krüsken, J. (2002): Informational Mood Impact on Effort-Related Cardiovascular Response: The Diagnostic Value of Mood Counts, Emotion, 2(3), 251-262.

Grossmann, Z. / Owens D. (2012): An Unlucky Feeling: Overconfidence and Noisy Feedback, Journal of Economic Behavior and Organization, 84(2), 510-524.

Ifcher, J. / Zarghamee, H. (2014): Affect and Overconfidence: A Laboratory Investigation, Journal of Neuroscience, Psychology, and Economics, 7(3), 125-150.

Kirchsteiger, G. / Rigottii, L. / Rustichini, A. (2006): Your Morals Might be Your Moods, Journal of Economic Behavior and Organization, 59(2), 155-172.

Kuvaas, B. / Kaufmann, G. (2004): Impact of Mood, Framing, and Need for Cognition on Decision Makers' Recall and Confidence, Journal of Behavioral Decision Making, 17(1), 59-74.

Lahav, Y. / Meer, S. (2012): The Effect of Induced Mood on Prices in Asset Markets – Experimental Evidence, Working Paper Ben-Gurion University of the Negev.

Menkhoff, L. / Schmeling, M. / Schmidt, U. (2013): Overconfidence, Experience, and Professionalism: An Experimental Study, Journal of Economic Behavior and Organization, 86, 92-101.

Moore, D. A. / Healy, P. J. (2008): The Trouble with Overconfidence, Psychological Review, 115(2), 502-517.

Moore, D. A. / Small, D. A. (2007): Error and Bias in Comparative Judgement: On Being Both Better and Worse than We Think We Are, Journal of Personality and Social Psychology, 92(6), 972-989.

Niederle, M. / Vesterlund, L. (2007): Do Women Shy away from Competition? Do Men Compete Too Much?, Quarterly Journal of Economics, 122(3), 1067-1101.

Oswald, A. / Proto, E. / Sgroi D. (2015): Happiness and Productivity, Journal of Labor Economics, 33(4), 789-822.

Pröger, T. / Meub, L. (2014): Overconfidence as a Social Bias: Experimental Evidence, Economics Letters, 122(2), 203-207.

Rottenberg, J. / Ray, R. D. / Gross, J. J. (2007): Emotion Elicitation Using Films, Coan, J. A. / Allen, J. J. B. (Ed.), The Handbook of Emotion Elicitation and Assessment, London, 9-28.

Schaefer, A. / Nils, F. / Sanchez, X. / Philippot, P. (2010): Assessing the Effectiveness of a Large Database of Emotion-Eliciting Films: A New Tool for Emotion Researchers, Cognition and Emotion, 24(7), 1153-1172.

Westermann, R. / Spies, K. / Stahl, G. / Hesse, A. (1996): Relative Effectiveness and Validity of Mood Induction Procedures: A Meta-Analysis, European Journal of Social Psychology, 26(4), 557-580.

Appendix A: English Translation of the Instructions

The Game

In each part of this game you will be given 10 tasks. You have 45 seconds to complete each task. There are 5 rounds in total. In the following, you are presented two examples:

- Question: What is the capital of the federal country of Saarland?
  Answer: **Saarbrücken**

- Task: Please add the five numbers given below and enter your result into the input field:

  26 16 86 05 41

  Answer: **174**

First, you will complete the tasks. Then you will watch a short movie clip that is shorter than 5 minutes. After that, you will be asked to assess your own performance by answering the following two questions:

- How many tasks did you complete correctly?
- How many tasks did you complete correctly in comparison to the other participants? How many more or less?
  - Example 1: I think that I gave three correct answers less than the average of the participants. Hence, you enter -3.
  - Example 2: I think that I gave three correct answers more than the average of the participants. Hence, you enter +3.

You have 45 seconds to complete the self-assessment. After each round of the game, you will receive feedback on your actual performance.

Award for Points

For each correct answer, you will receive **2 points**.

For each correct self-assessment, you will receive **8 points**.

You can be awarded 180 points in total.

The Payout

The basic payout is €2.50. For each point, you will receive €0.15. You can earn up to €29.50 in total.

Please note

Please keep quiet during the experiment!

Please do not look at your seatmate's monitor!

You are **not** allowed to use any auxiliary devices (calculator, smartphone etc.). All electronic devices must be switched off!

Please note the timing given in the upper right hand corner of the monitor. If you do not enter an answer in the given time, you will not be awarded any points for the respective task.

Appendix B: Original Instructions in German

Das Spiel

In diesem Spiel werden Sie in jedem Spielabschnitt 10 Aufgaben bekommen. Für jede Aufgabe haben Sie 45 Sekunden Zeit. Insgesamt gibt es 5 Spielabschnitte. Im Folgenden finden Sie zwei Beispielaufgaben:

– Aufgabe: Wie lautet die Landeshauptstadt von Saarland?

  Antwort: **Saarbrücken**

– Aufgabe: Bitte addieren Sie die unten angegebenen fünf Zahlen und tragen Sie Ihr Ergebnis in die Ergebniszeile ein.

  26 16 86 05 41

  Antwort: **174**

Zunächst werden Sie die Aufgaben lösen. Anschließend werden Sie einen kurzen Filmausschnitt zu sehen bekommen, der nicht länger als 5 Minuten dauert. Im Anschluss werden Sie Ihre Leistung einschätzen, indem Sie die beiden folgenden Fragen beantworten:

– Wie viele von den 10 Aufgaben haben Sie korrekt gelöst?
– Wie viele Aufgaben haben Sie im Vergleich zum Durchschnitt der anderen Teilnehmer korrekt gelöst? Wie viele mehr oder wie viele weniger?

  - Beispiel 1: Ich glaube, ich habe drei korrekte Lösungen weniger als der Durchschnitt der anderen Teilnehmer. Also geben Sie -3 ein.
  - Beispiel 2: Ich glaube, ich habe drei korrekte Lösungen mehr als der Durchschnitt der anderen Teilnehmer. Also geben Sie +3 ein.

Für die Selbsteinschätzung haben Sie 45 Sekunden Zeit. Nach jedem Spielabschnitt bekommen Sie ein Feedback über Ihre tatsächliche Leistung.

Die Punktevergabe

Für jede richtig gelöste Aufgabe erhalten Sie **2 Punkte**.

Für jede richtige Selbsteinschätzung erhalten Sie **8 Punkte**.

Insgesamt können Sie bis zu 180 Punkte erreichen.

Die Auszahlung

Sie erhalten eine Grundauszahlung von 2,50 Euro. Darüber hinaus erhalten Sie für jeden erreichten Punkt 0,15 Euro. Insgesamt können Sie bis zu 29,50 Euro verdienen.

Hinweise

Bitte verhalten Sie sich während des Experiments ruhig!

Bitte schauen Sie Ihren Nachbarn nicht auf den Bildschirm!

Es sind **keine** Hilfsmittel (Taschenrechner, Smartphones etc.) zugelassen. Alle elektronischen Geräte bleiben ausgeschaltet!

Bitte beachten Sie die jeweiligen Zeitangaben oben rechts am Bildschirm. Wenn Sie innerhalb dieser Zeit keine Eingabe machen, erhalten Sie keine Punkte für die jeweilige Aufgabe.

Appendix C: Screenshot of the Experiment with Z-Tree (Reconstructed in Order to Improve Readability)

Screenshot 1: Measurement of Mood before the Experiment (Reconstructed in Order to Improve Readability)

How are you feeling now? Please mark the adequate number!

very bad ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ very good

1–2–3–4–5–6–7–8–9–10

O.K.

Figure A-1: Measurement of Mood before the Experiment

Screenshot 2: Manipulation Check after the Attempt to Influence Mood (Reconstructed in Order to Improve Readability)

Which emotions did you experience while watching the movie clip?

Please mark one number accordingly!

very negative ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ ʘ very positive

1–2–3–4–5–6–7–8–9–10

O.K.

Figure A-2: Manipulation Check after the Attempt to Influence Mood