Expert Forecasts: Fast, Frugal, Flawed[§]

Markus Glaser [*] Zwetelina Iliewa [**]

This draft: January 15, 2014

Abstract

We conduct a natural field experiment with financial market professionals and document that they rely too often on intuition as opposed to deliberate reasoning, resulting in a significant bias in their expert forecasts. There is a significant difference between professionals' expectations when they are asked to forecast stock price levels or stock returns - two questions used interchangeably in real-world surveys. The difference constitutes a significant violation of the procedure invariance assumption of normative decision theory. The reason for professionals to rely too often on their flawed intuition lies in an extreme confidence, a characteristic of intuitive responses.

Keywords: procedure invariance, framing effect, dual-processing theory, natural field experiment, professional forecasters, judgmental forecasting

JEL Classification: G1, C93

[§] We thank Edgar Erdfelder, Lena Jaroszek, Doron Kliger, Lukas Menkhoff and Michael Ungeheuer for valuable suggestions. We further thank participants at the 2013 World Meetings of the Economic Science Association (ESA), the Behavioral Finance Working Group (BFWG) Conference 2013 as well as participants at the Research in Cognitive Psychology Workshop at the University of Mannheim for helpful comments. We also thank Anna-Lena Huthmacher for excellent research assistance. We gratefully acknowledge financial support by the German Research Foundation (DFG).

[*] Markus Glaser is from the Institute for Capital Markets and Corporate Finance, Ludwig-Maximilians-Universität Munich, Schackstr. 4, 80539 Munich, e-mail: glaser@bwl.lmu.de

[**] Corresponding author: Zwetelina Iliewa, Centre for European Economic Research (ZEW) Mannheim, L7.1, 68161 Mannheim and the Institute for Capital Markets and Corporate Finance, Ludwig-Maximilians-Universität Munich, e-mail: iliewa@zew.de

Expert Forecasts: Fast, Frugal, Flawed

Abstract

We conduct a natural field experiment with financial market professionals and document that they rely too often on intuition as opposed to deliberate reasoning, resulting in a significant bias in their expert forecasts. There is a significant difference between professionals' expectations when they are asked to forecast stock price levels or stock returns - two questions used interchangeably in real-world surveys. The difference constitutes a significant violation of the procedure invariance assumption of normative decision theory. The reason for professionals to rely too often on their flawed intuition lies in an extreme confidence, a characteristic of intuitive responses.

Keywords: procedure invariance, framing effect, dual-processing theory, natural field experiment, professional forecasters, judgmental forecasting

JEL Classification: G1, C93

1 Introduction

Which of the two is more important for an economic expert, intuition or econometrics?[1]

`Undoubtedly: intuition.`
Andreas Rees, Chief Economist Germany, UniCredit

People cannot fly and we therefore only have the capacity for accurate spatial orientation on the ground, not in the air. Nevertheless, when sitting in an airplane our inner senses already give us some impression of our relative position to the ground, although sometimes a dangerously erroneous one. For example, when flying in dense clouds or at night over the ocean a pilot may be certain, based on his own perception, that he is flying straight ahead when he is actually spinning or he may even feel he is flying up when he is actually headed down.[2] Because we are aware of the fact that under certain conditions our inner senses are erroneous we have developed navigation systems to assist pilots' decisions. Pilots are trained to rely on them disregarding their inner senses. Financial market forecasting is another instance of a task for which our inner senses cannot be of any value for particular reasons (see Kahneman and Klein, 2009). Nevertheless, as our evidence shows, many professional forecasters insist on "flying the plane" on their intuition. In this study we show just how erroneous the results can be.

Without making any strong assumptions on the information available to professional forecasters, which is unobservable in the real world, we test the hypothesis that financial market forecasts are unbiased by testing for a necessary condition for unbiasedness. In particular we analyze whether the expectations of professionals are independent of superficial features such as the exact form in which they are elicited, i.e. whether they satisfy the procedure invariance assumption of normative decision theory. For this purpose we conduct a natural field experiment with 191 financial market professionals from leading financial companies (banks, insurance companies, large corporates) in Germany. We examine the difference in stock market expectations between two logically identical elicitation forms - asking subjects

1 The quote is translated from German from an interview "Who is actually ... Andreas Rees, Chief Economist Germany, UniCredit" for Wirtschaftswunder. The interview in German is available under following link.

2 These examples are borrowed from Sharot (2011), Chapter I, "Which Way is Up?"

to forecast future stock price levels versus asking them to forecast stock returns. We are interested in these particular forms of elicitation, because they are used interchangeably in real-world surveys (see, e.g., Duke/CFO Magazine Business Outlook Survey for return forecasts; see, e.g., Livingston Survey of the Federal Reserve Bank of Philadelphia for a price level forecasts). Experimental evidence by Glaser, Langer, Reynders, and Weber (2007) challenges the information content of such surveys by examining these particular elicitation forms and showing in a lab experiment with students that the resulting expectations are predictably affected by the way they are elicited.

A growing body of literature challenges traditional normative theory of choice which assumes that people have clear preferences and expectations and especially that these are invariant with respect to the manner in which information is represented (description invariance) and the way they are elicited (procedure invariance). Violations to either of the invariance assumptions are generally referred to as framing effect. Existing literature provides evidence on violations to procedure invariance in elicited risk preferences (see e.g. Slovic and Lichtenstein, 1983, for evidence on preference reversals), risk perception (see e.g. Soll and Klayman, 2004) and inflation expectations (see Bruine de Briun, Klaauw, and Topa, 2011). Among the more widely studied violations to description invariance are the valence framing effect, the difference in revealed preferences when information is presented in the domain of gains or losses (see Levin, Gaeth, and Schreiber, 2002, for a literature overview), and scale variability of risk perception (see e.g. Kaufmann, Weber, and Haisley, 2013; Lawrence and O'Connor, 1993).

The study which is most closely related to ours is an experimental study by Simmons, Nelson, Galak, and Frederick (2011), which examines a violation of procedure invariance in sport betting. The study points to a fundamental difference in the responses of very experienced sports fans to logically identical questions - asking them to predict whether a team will win or lose against the point spread and asking them to predict the exact outcome of the game. The authors conclude that the difference is related to intuitive thinking, especially to the extreme subjective confidence which is characteristic for intuitive responses. De Martino, Kumaran, Seymour, and Dolan (2006) provide neurobiological evidence on the more widely studied violation of description invariance - valence framing. In a discussion of this study, Kahneman and Frederick (2006) ascribe the occurrence of the valence framing effect to intuition, better yet, its basic inability to disregard superficial features such as the exact representation of information or the exact wording of the question. Several recent experimental and empirical studies evidence an impact of intuitive thinking on biases in financial market expectations and investment decisions of retail investors and students. Grinblatt, Keloharju, and Linnainmaa (2011) and Grinblatt, Keloharju, and Linnainmaa (2012) analyze the impact of IQ, a proxy for the high tendency to trust one's gut feeling (see e.g., Stanovich and West, 2002), on the investment decisions of private investors and find that a low IQ score results in a higher susceptibility to behavioral biases and inferior portfolio allocation. A conceptually similar proxy is used by Kumar and Korniotis (2013) who analyze the impact of cognitive abilities on investment performance. The authors thereby approximate cognitive abilities with correlated demographic factors. Kempf, Merkle, and Niessen-Ruenzi (2012) link affect to stock market return expectations in an experimental setting and provide evidence on biased risk-return expectations resulting from intuitive judgment. According to the experimental evidence by Glaser and Walther (2013) subjects make inferior investment decisions when their intuition is in play. To the best of our knowledge this is the first study to use data on financial market professionals to document a link between intuitive thinking and a judgmental bias.

The main results of our study can be summarized as follows: Firstly, we provide evidence that the stock market forecasts of professionals violate the procedure invariance assumption of normative decision theory. The difference in the forecasts resulting from the two elicitation forms - price levels and returns - is both economically and statistically significant. Secondly, we show that the bias is driven by intuitive thinking as measured by both self-reported measures and reaction time. Hence, the use of judgmental forecasting, which makes use of fast and frugal, but in financial markets evidently not necessarily helpful heuristics, is particularly relevant even for the real-world forecasts of financial market professionals. Thirdly, our results suggest that the reason for professionals to rely on their intuition is the extreme confidence that comes with it. We also draw attention to the elicitation form which is strongly preferred by professionals, namely price levels. According to our evidence it is exactly the domain of price levels which drives the framing effect and in which participants are more prone to rely on their flawed intuition.

2 Data

2.1 The Panel of the ZEW Financial Market Survey

The ZEW Financial Market Survey is a monthly survey conducted since December 1991 among roughly 350 financial market professionals. The panel of participants covers a heterogeneous sample of financial market practitioners: active (e.g., portfolio managers) and passive (e.g., professional forecasters); sell-side, i.e. participants from large German companies, and buy-side participants, i.e. investors and investment advisors. Participants are initially selected by ZEW and invited to participate in the panel for their occupation as financial market professionals at leading financial institutions (banks, insurance companies) and large industrial companies in Germany. According to their occupation, participants are categorized as follows: treasurer (26.25%), economist or asset researcher (20%), portfolio manager (13.75%), advisor (8.13%), trader (7.5%), and other (e.g., corporate executives, wealth manager etc.). The participants are almost exclusively male (only 6 female participants). Within the scope of the survey questionnaire participants are generally asked about their medium term forecasts (6 months ahead) on macroeconomic variables, interest rates, international stock market indices and foreign exchange. Furthermore, the survey questionnaire always contains a Special Questions section on diverse financial market topics and participants are used to responding to it whenever the topic is in their area of expertise.

The participants are not provided with any monetary incentives. However, all participants regularly and timely receive the press releases containing the survey results of the most recent wave. Timely and detailed information on the survey results is evidently valuable to the participants because of the comparably high overall market attention to scheduled releases of two indicators calculated from the survey responses - "ZEW Economic Sentiments" for Germany and "ZEW Current Situation" for Germany. Bloomberg ranks the scheduled release indicators based on their relevance as approximated by the number of release alert subscriptions of Bloomberg users. As of October 2012 the Bloomberg relevance indicators for ZEW Economic Sentiments and ZEW Current Situation were 98.24 points and 94.74 points, comparable to the relevance index of the European Central Bank Interest Rates Announcements (97.67 points) and higher than the relevance indicators for macroeconomic announcements such as the German Consumer Price Index (75.44 points) and the unemployment rate in Germany (84.21 points). As the results are evaluated and published strictly anonymously there are no incentives for strategic response behavior such as rational herding. There are no mandatory questions or any other restrictions of the response behavior. Moreover, participants are encouraged to respond strictly according to their area of expertise.

In September 2012 we collected background information on the methods used by experiment participants when conducting short-term DAX forecasts - the main focus of our experiment. An overview of the results is included in Figure 7 in the Appendix. Technical analysis is by far the most intensively used forecasting tool - 64% of the participants indicate that it is of great importance for their short-term forecasts. This result is in line with recent evidence by Menkhoff (2010) on the wide usage of technical analysis of fund managers in Germany, especially for an investment horizon of several weeks. Another study by Hoffmann and Shefrin (2011) surveys the methods used by online investors and shows that technical analysis is the second most preferred method after intuition. The consistency with both studies indicates a high representativeness of the ZEW panel of financial market professionals. Further factors which play a role for the short-term DAX forecasts are experience, fundamental analysis and intuition with respectively 43%, 30% and 22% of the participants ranking them as highly important. In contrast, the majority of participants consider econometric models and simulations of little importance for their short-term DAX forecasts.

2.2 Experimental Design

In the period from September 2012 until September 2013 we conducted a quarterly repeated field experiment with the professional stock market forecasters from the ZEW Financial Market Survey. Within the scope of the field experiment we asked the online participants - round 260 professionals - to forecast the German stock market performance index DAX one month ahead (point forecast) and to provide their subjective 90% confidence intervals. In September 2013 they were additionally asked to provide their one-year ahead DAX forecasts and 90% confidence intervals. The experimental questions were included as a quarterly repeated part of the regular survey; this is why no additional incentives for participation were necessary, other than the usual non-monetary incentives to participate in the ZEW panel (see Section 2.1). The questions were included at the end of the survey questionnaire within a flexible part called Special Questions. Prior to the first wave of the field experiment the subjects were randomly divided into two equally sized groups - a return group and a level group (randomized between-subject design). The participants did not receive any indication that an experiment was being conducted and they were unlikely to expect it because experiments had never been conducted within the scope of the ZEW Financial Market Survey before. The experimental design is as close to a real-world forecasting task as can be and constitutes a natural field experiment in the terms of Harrison and List (2004). Our experiment combines the advantages of laboratory experiments (randomization) and natural experiments (realism).

The exact wording of the questions for the level group and the return group are given as follows:

I expect the DAX in 1 month at … points. With a probability of 90% the DAX will then lie between … and … points.

Within 1 month I expect a DAX return (monthly percentage change) of … %. With a probability of 90% the DAX return will then lie between … % and … %.

In order to retain high ecological validity the questions for both groups are adopted from existing regular surveys. The questions for the level group are adopted from a regular question within the ZEW Financial Market Survey on mid-term DAX forecasts (6 months ahead), which has been a part of the survey since 2003 (see Deaves, Lüders, and Schröder, 2010). The exact wording of the questions for the return group is adopted from the Duke/CFO Magazine Business Outlook Survey (see Ben-David, Graham, and Harvey, 2013) and adjusted to the shorter forecasting horizon of our experiment. We include a definition of returns in brackets, "monthly percentage change", in order to specify that non-annualized returns are required. Feedback to a beta version of our questionnaire indicated that financial market professionals usually understand "returns" as "annualized returns" unless stated otherwise. The questions on subjective confidence intervals in the return group and the level group are held comparable in spite of the deviation from the exact wording of Ben-David, Graham, and Harvey (2013) in order to avoid well-documented differences in the width of subjective confidence intervals resulting from the elicitation format (see e.g., Soll and Klayman, 2004).

[Insert Figure 1 about here]

In each wave the respondents had roughly two weeks to submit their response: 31.08.-17.09.2012 (September 2012 wave), 26.11.-10.12.2012 (December 2012 wave), 04.03.-18.03.2013 (March 2013 wave), 03.06.-17.06.2013 (June 2013 wave) and 02.09.-14.09.2013 (September 2013 wave). The survey waves always begin and end on a Monday. Participants receive an email with the link to the questionnaire at the beginning of the survey wave, an email reminder on the last Friday before the deadline for submission and a phone reminder on the last day of the survey wave. The deadline for submission is always 17:00h CET. Overall, 191 professionals participated in the experiment with an average of 3.4 responses, which resulted in a total of 650 responses. In each wave we have received from 117 to 147 responses. The response rate to the experimental question was from 63.6% (in September 2013) to 73.5% (in December 2012). We measure response rate to the experimental question as the percentage of responses over all participants who have potentially obtained the experimental question in this month, i.e. all online participants who have clicked on the link to the ZEW survey.

As illustrated in Figure 1 the period covered an overall bullish market phase on the German stock market. The survey questionnaire does not contain any information on the German stock market or any other information sources. However, since we restrict the experiment to the online participants of the ZEW panel, it is plausible to assume that they have timely access to all publicly available information and are well informed about the current level of the assets they frequently track (including DAX).

The forecasts of the level group are converted into return forecasts using the DAX daily open on the day of the response (calculation method "daily open"). Whenever the response was submitted on a Saturday, Sunday or an official holiday the DAX daily close of the last working day prior to the day of the response is applied. Daily data on DAX is downloaded from Datastream. Since there is a clustering of responses before noon (50% of the responses are submitted before 12 o'clock), using the daily opening price as a proxy for the actual DAX level is connected to a rather small measurement error in most of the cases. Additionally, in order to account for a potential measurement error resulting from intraday trends, we have conducted robustness checks using the average between the DAX daily open and the DAX daily close on the day of the response (calculation method "daily avg"). We exclude responses for which the forecast lies outside the subjective 90% confidence interval (consistency check).

During the experiment the participants did not receive any feedback on their own DAX one-month forecasts. Furthermore, participants have not received any feedback on the aggregate responses to the experimental questions, as the results have not been published in any form.

3 Framing Effect and Intuitive Thinking Outside the Lab

The study by Glaser, Langer, Reynders, and Weber (2007) provides experimental evidence for the particular framing effect in stock market expectations - asking subjects to forecast future stock price levels or asking them to forecast stock returns - and casts serious doubt on the information content of real-world surveys. The authors argue for external validity of their laboratory experiment by providing an overview of the literature on subjective stock market forecasting. In existing studies with students and professionals alike, the findings of almost each single existing study concerning the expectation formation process are systematically related to the exact forecasting domain (price levels vs. returns) which is applied. The argument that finance professionals are not immune to judgmental biases altogether is proclaimed by a handful of studies: Northcraft and Neale (1987) document a significant use of the anchoring-and-adjustment heuristic by real estate experts; Roszkowski and Snelbecker (1990) provide evidence on the valence framing effect in a study with financial planners. More recent studies even go one step further and alert that professionals may even be more susceptible to judgmental biases than students: Haigh and List (2005) show that traders on the Chicago Board of Trade exhibit higher myopic loss aversion than students; Gilad and Kliger (2008) demonstrate that investment advisors in large commercial banks and accountants in CPA firms are more prone to priming than students. Simmons, Nelson, Galak, and Frederick (2011) suggest that under some conditions violations to procedure invariance are more likely to be present in the forecast of experts as opposed to novices without any prior knowledge. According to the authors this is likely to be the case whenever experts' prior knowledge results in a strong tendency to rely on misleading intuition.

The opposite view is supported by Schoorman, Mayer, Douglas, and Hetrick (1994), who cast serious doubt on the real-world relevance of framing effects altogether. The authors argue that real-world relevant business decisions are connected to a large amount of available information and information evidently mitigates the valence framing effect, which they examine. Furthermore, in some environments and for some tasks framing effects are evidenced to diminish with the relevance of the task (see e.g., Schoorman, Mayer, Douglas, and Hetrick, 1994; McElroy and Seta, 2003), expertise in the task and statistical knowledge (see e.g., Bless, Betsch, and Franzen, 1998). These arguments would be particularly valid reasons not to expect any framing effect in professionals' forecasts if professionals could be assumed to be able to acquire actual skills and expertise in forecasting, and if they could generally be assumed to make use of statistical forecasting tools. However, there is evidence against both assumptions: firstly, Kahneman and Klein (2009) argue that the acquisition of expertise is impossible on financial markets altogether; secondly, the use of statistical methods by professionals should not be taken for granted, as evidently even they rely on other forecasting methods instead (see Section 2.1).

In a first step we therefore examine whether expectations of professional stock market forecasters are susceptible to the particular framing effect under study. We formulate the following hypothesis:

H1: *The stock market forecasts of professionals violate procedure invariance (framing effect).*

The study by Simmons, Nelson, Galak, and Frederick (2011), which examines a violation of procedure invariance in sports betting, concludes that the bias is related to the reliance on intuitive thinking, especially to the extreme subjective confidence which is characteristic for intuitive responses. McElroy and Seta (2003) provide experimental evidence that the more intensively studied valence framing effect is determined by intuitive thinking. Furthermore, De Martino, Kumaran, Seymour, and Dolan (2006) provide neurobiological evidence that the valence framing effect is determined by emotional processes - a result which is interpreted by Kahneman and Frederick (2006) in terms of dual-processing theory. The psychological literature on dual processing hypothesizes that judgments and decisions are generally a result of two processes of thoughts - a fast, effortless, automatic and emotional intuition, labeled System 1, and a slower and more effortful deliberate reasoning, labeled System 2 (see e.g., Kahneman, 2003; Stanovich and West, 2002). The way intuition works is by subconsciously making use of fast and frugal heuristics. Glaser, Langer, Reynders, and Weber (2007) suggest a particular heuristic which explains the sign of the framing effect in their experiment - the representativeness heuristic. In general, heuristics may or may not lead to invariant and unbiased judgment and decisions and whether agents can develop skilled intuition depends by and large on particular properties of the decision environment itself (see Kahneman and Frederick, 2006; Kahneman and Klein, 2009). For the particular task of financial market forecasting Kahneman and Klein (2009) argue that skilled intuition is impossible and agents should thus refrain from trusting their gut feelings. Against this backdrop it seems preposterous that financial market practitioners admit to relying on their intuition and claim it superior to deliberate forecasting techniques (see

Section 2.1). A recent study by Hoffmann and Shefrin (2011) surveys private investors and documents that they admit to primarily use their own intuition when making investment decisions. Meyler and Rubene (2009) provide evidence on professional forecasters from the ECB Survey of Professionals Forecasters, who admit to using judgment more often than econometric or fundamental analysis when conducting macroeconomic forecasts. Northcraft and Neale (1987), Mussweiler and Schneller (2003) and Campbell and Sharpe (2009) provide evidence that the expectations of professionals are compatible with the use of fast and frugal heuristics.

We therefore hypothesize that the difference between return and price level forecasts of professionals is driven by reliance on (flawed) intuition and vanishes when deliberate forecasting strategies are applied.

`H2a`: *Intuitive thinking intensifies the framing effect.*

`H2b`: *Deliberate forecasting strategies diminish the framing effect.*

3.1 Are Professionals Prone to Framing?

In the following we test hypothesis `H1`. Table 1 shows that the framing effect is significant, both economically and statistically, in expectations of professional forecasters. Asking professionals about the expected monthly return as opposed to the index level results in significantly more optimistic expectations.

[INSERT TABLE 1 ABOUT HERE]

We report coefficients from panel regressions and cluster-robust standard errors.[3] The independent variables are given and coded as follows: dummy coding for the treatment fixed effect (D^{ret}= 1 if return group); unweighted effect coding for the wave fixed effects (C^{t_j}= 1 for $t = t_j$, C^{t_j}= -1 for $t = t_1$ and C^{t_j}=0 else, where $j = 2 \dots 5$). The application of effect coding allows comparing the framing effect in each individual wave with the global

[3] We have also estimated mixed models with individual, wave and treatment random effects for all regressions included in the paper. All reported results are robust to the exact model specification.

average framing effect. According to Table 1, regression (1), the average difference in optimism over all individuals and all waves is as high as 1.50 percentage points. A monthly expected return of 1.31 percent implies a strong investment recommendation on the part of the return group. Following the average forecast of the level group, in contrast, an investor would be advised to refrain from short-term investments in the DAX as such an investment is connected to an expected, albeit insignificant, loss of minus 0.19 percent.

Over the period the experiment has been conducted, there are no statistically significant fluctuations in the scope of the framing effect. Figure 2 plots the average forecasts of the return group and the level group in each wave, which can be derived from the regression coefficients in Table 1. All results are robust to alternative methods of calculation of return expectations from price level forecasts.

[Insert Figure 2 about here]

Given the overall positive trend of the DAX performance index during the five waves of our experiment, the evidence is consistent with the hypothesized use of the representativeness heuristic (see Andreassen, 1988; Glaser, Langer, Reynders, and Weber, 2007). However, the fact that there has not been any negative trend in the DAX index during the experiment does not allow us to test this hypothesis against alternatives. For instance, it can be hypothesized that there is a general optimism bias caused by the return domain but not the price level domain. Another possible explanation could be that the return domain makes the use of statistical methods more obvious to the forecaster and therefore results in responses which are closer to the base rate (see Kahneman and Frederick, 2002).[4]

A strand of literature on questionnaire design alerts that the framing effect may only be an apparent bias and rather results from fundamental differences in the information content of allegedly identical survey questions (see e.g., Bradburn, 1982; Bruine de Briun, Klaauw, and Topa, 2011). While in real-world surveys on stock market expectations the questions on future stock price levels and future stock returns are used interchangeably,

[4] The average monthly return of the DAX performance index since December 1964 lies at 0.65 percent and is closer to the average forecast of the return group in more recent samples (e.g., average return of 0.97 percent since January 2003; average return of 1.11 percent since January 2009).

there is no evidence that practitioners understand them the same way. Firstly, it can be argued that differences in the expectations arise from different treatments of dividends, especially if practitioners understand returns as dividend-adjusted returns. This criticism does not apply to the German stock market index because the default index is already a total return index ("performance index"), which means that both the level group and the return group should be taking the dividends into account. Furthermore, the companies in the German DAX index make dividend decisions on a yearly basis, mostly in May and June, hence differences in the treatment of dividends cannot be an explanation for the bias in monthly return expectations in four of the five survey waves. Secondly, feedback to a demo version of the questionnaire has indicated that unless stated otherwise participants understand returns as annualized returns instead of monthly percentage change (see Section 2.2). Although we have explicitly specified that in the questionnaire, it may be argued that some participants overlook the specification and report annualized returns in spite of it. In order to address this criticism we extended the experiment in September 2013 to include DAX one-year ahead forecasts. The exact wording of the questions for both groups was analogous to the question for DAX one-month ahead forecasts.

The results for two alternative methods for calculation are displayed in Table 1, regressions (3) and (4). The DAX forecasts of the level group are significantly positive with an average of 3.39 percent. In line with the evidence on one-month forecasts, the return group reports significantly higher expectations. The magnitude of the framing effect is as high as 2.67 percentage points for one-year ahead forecasts and is significant at the 10% significance level.

3.2 The Impact of Intuition on the Framing Effect

In the following we test hypothesis `H2a` on the impact of intuitive thinking on the framing effect. We measure intuitive thinking firstly by self-reported measures and secondly by means of reaction times. In line with hypothesis `H2a`, the framing effect in our sample is driven mainly by participants who put relatively great weight on their intuition and rely on it at least as much as on analytical methods. The scope of the framing effect in this group of participants is namely more than three times higher than in the group of more analytical forecasters. Fast responses also exhibit significantly larger framing effect as compared to responses which are better thought through, which is also in line with our hypothesis.

A. Self-reported Measures of Intuitive Thinking

In September 2012 we included a question on the factors important for the conduction of one's own forecasts. Among other things, participants were asked to rate the importance of intuition on a three-point Likert-scale with the categories "low", "medium", "high". We collected data from 123 of the participants in our experiment and the group sizes are given as follows: low importance (34 participants), medium importance (62 participants), high importance (27 participants). From this question we construct two alternative measures of intuitive thinking which capture the absolute importance of intuition and its relative importance as compared to analytical methods. The absolute importance of intuition reflects "Faith in Intuition" - the participants' tendency to trust their initial gut feeling in forecasting tasks (see Epstein, Pacini, Denes-Raj, and Heier, 1996, for a description and factor analysis of the short Rational-Experiential Inventory).

Our measure of the relative importance of intuition is constructed analogously to the Intuitive-Analytical Score introduced by Sjöberg (2003). We use participants' responses to an exhaustive list of analytical factors - technical analysis, fundamental analysis, econometric models, simulations, inhouse forecasts and consensus forecast - and compare the most important one among them with the importance of intuition (within-subject, between-factor comparison). Based on the relative importance of intuition, participants are divided into three groups and the group sizes are given as follows: intuition is much less important than analytical methods (27 participants), intuition is less important than analytical methods (56 participants), intuition is of comparable importance or more important than analytical methods (39 participants).

Table 2 and Figure 3 show that the scope of the framing effect is largest in the group of intuitive forecasters as measured by both Faith-in-Intuition and Intuitive-Analytical Score. We report coefficients from panel regressions with cluster-robust standard errors.

Asking participants with the highest Faith-in-Intuition score to forecast returns instead of price levels results in more optimistic forecasts by 2.11 percentage points, which is higher compared to the scope of the framing effect in the more analytical groups (1.18 percentage points in the medium Faith-in-Intuition group; 0.99 percentage points in the low Faith-in-Intuition group). It should be noted that the framing effect, although statistically insignificant at the 10% significance level in the low Faith-in-Intuition group, does not diminish to zero. This can be due to a self-reporting bias in the proxy for intuitive thinking, which results partly from the subconscious nature of intuition and partly from a potential social-desirability bias. The difference in the scope of the framing effect between the three groups is economically significant but not statistically significant at the 10% significance level, which is partly due to the small group size of the high Faith-in-Intuition group.

[INSERT TABLE 2 ABOUT HERE]

[INSERT FIGURE 3 ABOUT HERE]

As can be seen from Table 2, regressions (3) and (4), there are large differences in the scope of the framing effect between the three Intuitive-Analytical Score groups. The magnitude of the framing effect in the most intuitive group is more than three times larger than the magnitude of the bias in the other two groups. The difference between the scope of the framing effect in the two largest group (high and medium Intuitive-Analytical Score) is also statistically significant at the 10% significance level.

B. Reaction Time

Reaction times are generally used to distinguish between intuitive and deliberate approaches to decision-making (see e.g. Rubinstein, 2007). For our purposes we make use of the characteristic of intuition to provide fast and effortless responses and assume that short reaction times reflect intuitive responses whereas longer reaction times are more likely to reflect analytical responses. Reaction times have a great advantage against survey-based measures of intuition because they are not prone to any self-reporting biases.

In March we measured the time participants take to respond to the Special Questions section (reaction time). We measure the reaction time from the moment the participants see the first question of the Special Questions section (including our experimental questions at its very beginning) until the moment the Special Questions section is completed. Few participants have returned to the previous parts of the questionnaire after entering the Special Questions section. These participants are excluded from the analysis, because their reaction time is biased upwards.

The experimental part of the Special Questions section in March 2013 contained the usual short-term DAX forecasts (point estimates and subjective confidence intervals). The remaining Special Questions addressed participants' estimates on the gold price, assessment of the fair value of both DAX and gold, an assessment of the correlation between DAX and three other asset classes and a question on participants' correlation forecasting practices.[5] We acknowledge, that extremely short reaction times as measured for the whole Special Questions section are mostly due to reluctance to respond to the non-experimental part of the Special Questions section. That is why as a robustness check we conduct a separate analysis on the subsample of participants who have filled out all questions from the Special Questions section. Focusing on this subsample of participants who have responded to all Special Questions is also particularly interesting, because these participants are likely to have better statistical knowledge, as indicated by their willingness to provide correlation estimates.

Reaction times are measured in seconds and rescaled in minutes. Figure 4 shows the results of a polynomial fit over the forecasts of both treatment groups over reaction time. As the distribution of the reaction time variable is highly positively skewed we use logarithmized reaction times for the purposes of the subsequent regressions. We run simple OLS regressions with an interaction term between the treatment dummy and log reaction time.

[Insert Table 3 about here]

5 We have also measured the reaction time for the Special Questions in June 2013 and September 2013. However, the non-experimental part of the Special Questions in these waves was more extensive than in March 2013 and completely disconnected from the experimental questions which introduced severe noise in the measured reaction times. The exact wording of the non-experimental Special Questions is available upon request.

[Insert Figure 4 about here]

The regression results are displayed in Table 3, columns 1-4. According to our results the scope of the framing effect significantly decreases with reaction time at the 5% significance level.[6] The results from our reaction time proxy for intuitive thinking are consistent with the results of the self-reported proxies for intuitive thinking - Faith-in-Intuition and Intuitive-Analytical Score.

3.3 Deliberate Forecasting Strategies as a Remedy for the Framing Effect

In the following we test hypothesis `H2b` on the advantage of deliberate forecasting strategies as a remedy for the framing effect. We approximate deliberate forecasting strategies by means of the self-reported use of technical analysis, because technical analysis is the only strictly deliberate forecasting method among the methods that are prevailingly used by the professionals in our sample. In line with hypothesis `H2b`, relying heavily on technical analysis diminishes the framing effect significantly at the 1% significance level.

In September 2012 we asked participants to assess the importance of several factors for the conduction of their one-month DAX forecasts. An overview of the responses is included in the Appendix (see Figure 7). Participants were asked to assess the importance on a Likert-type scale with three categories - "low", "medium" and "high". We collected data for 123 participants in the field experiment. Among other things we have asked participants to assess the importance of deliberate forecasting methods such as econometric models, simulations and technical analysis. Technical analysis provides deliberate technical trading rules based exclusively on historical price levels. In contrast to the high importance of technical analysis, only a minority of 6.61% of the participants consider econometric models of high relevance for their short-term forecasts. Analogously, simulations are considered highly important by a minority of 8.47% of the participants. The evidence is consistent with the survey evidence provided by Menkhoff (2010), who documents that, for forecasting horizons of several weeks, technical analysis is the predominant method of choice of

[6] Unreported regressions with median-centered reaction time instead of logarithmized reaction time show that the framing effect is highly significant at the median reaction time.

international fund managers. Therefore, we restrict the following analysis on the impact of deliberate forecasting methods to the use of technical analysis. Table 4 reports the results from panel regressions with cluster-robust standard errors.

[Insert Table 4 about here]

The results indicate that the framing effect in the group of technical analysts is not significant at the 10% significance level. Moving from the group which considers technical analysis to be of merely medium importance to the group of forecasters who rely heavily on technical analysis decreases the framing effect by 3.04 percentage points. The difference in the scope of the framing effect is significant at the 1% significance level. We therefore conclude that the use of deliberate forecasting methods such as technical analysis can be a remedy against the framing effect.

It should be stressed at this point that the fact that a particular forecasting method is deliberate does not mean that it is of any value for forecasting. Nor does it mean that the use of this forecasting method is rational. It merely indicates that the resulting forecasts are consciously thought through. In this sense satisfying the assumption of procedure invariance and not being prone to framing effects is merely a necessary but not a sufficient condition for rational expectations.

4 Why Do Professional Forecasters Rely on Their Flawed Intuition?

In the previous section we have shown that reliance on intuition is connected to one particular violation of the procedure invariance assumption of normative decision theory. But why would stock market professionals rely on their intuition if it provides flawed responses? What factors determine whether subjects count on intuition instead of deliberate reasoning for a particular task in a particular situation at a particular point of time? Dual-processing models differ in the answers they provide to this question. Some dual-processing models assume that the tendency to decide intuitively is a personal trait that is stable over time and over decision situations (see Evans, 2008, for a categorization). We have implicitly adopted this assumption in Section 3.2 when we categorized participants according to the Faith-in-Intuition measure and the Intuitive-Analytical-Score. Another set of models assumes that the stronger reliance on intuition depends on situation-specific factors such as low level of motivation and low cognitive capacity. Experimental evidence supports the claim that subjects tend to trust their gut feelings more often when they are unmotivated, cognitively busy or depleted (see e.g., Schoorman, Mayer, Douglas, and Hetrick, 1994; Gilbert, Pelham, and Krull, 1988; Baumeister, Bratslavsky, Muraven, and Tice, 1998). These arguments are relevant in our setting because our participants work in a highly dynamic and stressful environment and we do not provide any monetary incentives for accurate responses. Therefore we formulate and test following hypothesis:

H3: *The framing effect is intensified by a lack of motivation or a lack of cognitive resources.*

A third category of models assumes an internal dialog between both systems. These models assume that intuition (i.e. System 1) provides a fast and effortless assessment in almost all situations and thereby serves as a decision default. Subsequently, reasoning (i.e. System 2) sometimes monitors and eventually corrects the intuitive assessment. The activation of System 2 thereby depends on characteristics of the decision default itself, such as subjective confidence. An example of a more general model in this category is provided by the Parallel-Constraint-Satisfaction (see e.g. Glöckner, 2008). When making a decision or a judgment (i.e. a forecast) subjects seek to maximize the fit between the separate pieces of information that instantly come to mind. They are particularly fast and confident in their responses when the pieces of information easily fit together. In contrast, subjects become slow, less confident and employ cognitive effort when conflicts occur, which is more likely to be the case the more informed they are in general. The idea that conflicts result in more deliberate reasoning is supported by neurobiological evidence (see e.g. Botvinick, Braver, Barch, Carter, and Cohen, 2001). Simmons and Nelson (2006) provide experimental evidence that the tendency to switch from intuitive to analytical responses is negatively correlated with intuitive confidence. Given that intuitive responses are flawed (see Section 3) we hypothesize a relationship exactly opposite to what can be expected from well-calibrated forecasters, namely, that they will be most confident whenever they are most biased.

`H4`: *(Flawed) intuitive responses are more confident.*

It can further be hypothesized that the elicitation form may influence the tendency to suppress intuitive responses itself. Firstly, Kahneman and Frederick (2002) suggest that certain formulations can make the applicability of statistical methods more apparent or draw attention to relevant information. Secondly, if the response to a certain question requires information which cannot be activated from memory and a deliberate information search in required, following Glöckner and Betsch (2008) the resulting responses are expected to be deliberate, not intuitive. Following the experimental evidence by Andreassen (1988) we would then hypothesize that individuals forecast more deliberately when they are asked to forecast returns. Andreassen (1988) provides evidence that return charts are more difficult to recall than price level charts.

Experimental studies on other violations of either of the invariance assumptions oftentimes identify one domain that decreases the respective bias under study. For instance it has been argued for the case of judgmental biases in subjective probability estimates, that asking subjects to estimate frequencies as opposed to probabilities results in better estimates (see e.g. Kahneman and Frederick, 2002). For the related case of subjective risk perception Soll and Klayman (2004) identify an elicitation form that makes individuals less susceptible to the overconfidence bias. Simmons, Nelson, Galak, and Frederick (2011) examine sports gambling and provide evidence of a significant bias in the betting behavior of high-involvement sports fans when asked to predict whether a team will win or loose against the exogenously determined point spread. Changing the elicitation form and asking participants to predict the exact point difference between both teams, diminishes the bias - a result which is not achieved by any other methods such as explicit warnings and experience with the betting task. The study also provides direct evidence that the results in the inferior elicitation mode are driven by (evidently flawed) intuitive thinking, better yet by the characteristic for intuition extreme subjective confidence. Therefore we hypothesize:

`H5`: *The tendency to rely on intuition differs between the two forecasting domains.*

In the following we provide several tests on hypotheses H3, H4 and H5.

4.1 Intuitive Confidence

We have already shown in Section 3.2 that the framing effect prevails in fast and intuitive responses. In the following we test whether the fast and biased responses are also connected to extreme confidence as stated in H4. Taking the results from all measures of intuitive thinking together - both self-reported measures and reaction time - our results for the level group are in line with the predictions of H4 - intuitive and evidently flawed forecasts are provided with high confidence whenever participants are asked to forecast price levels. In contrast, the return group displays a confidence pattern closer to what would be expected from well-calibrated forecasters, that is a a positive relation between subjective confidence and information content of subjective forecasts, according to one of our measures for intuitive thinking - reaction time.

We measure subjective confidence based on the subjective volatility estimates σ_{it} of each individual i in each wave t, which are derived from the upper and lower bounds of the subjective confidence intervals applying the approach by Pearson and Tukey (1965).

Table 5 reports the regression results for both self-reported proxies for intuitive thinking - Faith-in-Intuition and Intuitive-Analytical-Score. We report the coefficients from panel regressions with cluster-robust standard errors. An illustration of the regression results is displayed in Figure 5 and Figure 6. The results for the level group are in line with H4 for both proxies for intuitive thinking - subjective confidence is lower (subjective volatility estimates are higher) in the group with the lowest tendency towards intuitive thinking as compared to the group with the highest tendency towards intuitive thinking. The increase in subjective confidence with intuitive thinking as measured by Faith-in-Intuition (Intuitive-Analytical Score) is economically and statistically significant at the 5% level (10% level). Additional regressions with a dummy variable for the level group, instead of the return group, show that the subjective confidence in the return group does not change significantly with intuitive thinking.[7]

[7] It can be noted that the average confidence intervals in the most analytical group are more narrow in

The results from OLS regressions of the subjective volatility estimates on reaction time based on the sample in March 2013 are displayed in Table 3, columns 5-8. An illustration of the regression results is displayed in Figure 5 and Figure 6. The coefficient β_2 describes the development of the subjective volatility estimates of the level group with reaction time. The positive coefficients indicate that subjective confidence decreases the more time the participants take to respond. The decrease is statistically significant at the 10% significance level for the specifications in columns 6 and 8. The confidence of the return group, however, develops in the opposite direction. The results of regressions analogous to Table 3 with a dummy variable for the level group, instead of the return group, are included in the Appendix, Table 8. Against the predictions of H4, the subjective confidence of the return group becomes larger with reaction time, and the increase is statistically significant at the 10% significance level (one-sided test).

4.2 The Domain of Price Levels

Against the backdrop of the asymmetry between the return group and the level group in terms of the correlation of subjective confidence with intuitive thinking, documented in Section 4.1, it should be noted that there is also an apparent asymmetry in the development of the forecasts of both groups over reaction time. Table 3 reveals that the evidenced decrease of the framing effect with reaction time is due to significant change in the level forecasts with reaction time, not the return forecasts. This indicates that it is the fast and flawed forecasts of the level group, not the return group, that drive the framing effect. Together with the observation that the fast and flawed forecasts of the level group are correlated with high confidence, this gives rise to the hypothesis that return forecasts are generally more thought through and level forecasts are on average more intuitive. A comparison of the mean reaction time of both groups provides support for this claim. On average it takes additional 1.57 minutes to respond to the return question, a difference which

the return domain than in the level domain. This is in line with the experimental results by Glaser, Langer, Reynders, and Weber (2007). There are several potential explanations for this observation (e.g. numerosity effect) other than the participants being truly more confident in the return treatment. Analyzing the absolute difference between the level of confidence in both treatments is an interesting research question itself, however, it is outside the scope of this study.

is significant at a 10% significance level according to a non-parametric Mann-Whitney test.[8]

The evidenced tendency to forecast more deliberately in the return treatment as opposed to the level treatment is in line with the hypothesis that some forecast domains, in this case the return domain, make the applicability of more time-consuming statistical methods more apparent (see Kahneman and Frederick, 2002). It is also in line with the hypothesis, that the return formulation requires information that participants cannot recall and therefore have to search deliberately (see Glöckner and Betsch, 2008), namely information on past returns. Andreassen (1988) makes a strong case for this hypothesis by providing experimental evidence that price charts are easier to recall than return charts. Even if both price level and return charts are available, subjects appear to pay more attention to the price level charts. Prior to the beginning of the field experiment, in December 2011, we asked professionals which numbers they consider when they gather information about the stock market indices they regularly track. Among the responses we included the responses "index level" and "index return".[9] We collected data from 142 participants. Only a minority of 16.9% participants indicated that they inform themselves of the index return. Most of them do so in addition to getting informed about the current index level and only 2 participants inform themselves only on the index returns. The evidence that participants do not pay attention to stock returns in the first place, makes a strong case for the hypothesis that they cannot recall relevant information when they are asked to forecast returns (e.g. historical average monthly return) but rather have to search for it deliberately, which in turn induces a deliberate forecasting strategy.

Before the beginning of the experiment, in March 2012 we also examined participants' preferences towards price levels as a forecast domain. We asked participants for their DAX forecasts one-month ahead and thereby allowed them to provide either a return forecast or a price level forecast. From 136 professionals, who submitted a response, only 6 submitted a return forecast, which indicates a strong preference towards responding in the price levels domain. In March 2013 we also asked the survey participants about the exact forecast

[8] Similar results are obtained from a Wald-test on the log reaction times.

[9] Other response categories were given as follows: high-low spread, historical volatility, implied volatility, trading volume.

domain of their forecasts outside the scope of the ZEW survey. From 79 participants who indicated that they are conducting explicit forecasts outside the scope of the ZEW survey, 49.4% indicated that they are usually conducting level forecasts and only 13.9% - return forecasts. The evidence raises another concern, which highlights the relevance of research on judgmental forecasting for real-world situations - participants appear to be most biased in the domain they are most comfortable with and use most often.

4.3 Lack of Motivation and Cognitive Resources

In this section we test hypothesis H3. In a first step we measure the impact of lack of motivation on the scope of the framing effect. We measure motivation by the relative relevance of the ZEW survey to the other daily activities of the participants. Following Schoorman, Mayer, Douglas, and Hetrick (1994) participants perceive tasks as relevant when their contribution is relevant for the final decision. The absolute relevance of the survey is equal for all participants - the survey forecasts are mainly used for the calculation of sentiment indicators which eventually have an impact on the decisions of others. However, for the group of participants whose daily activity consists of providing forecasts to serve the decisions of others, the relevance of the survey is comparable to their everyday activity. In June 2013 we therefore ask participants whether they conduct regular forecasts and if so for which purposes. Of 139 respondents 56.8% indicate that they conduct forecasts regularly. We use this group of participants (”Forecasters”) as a proxy for high relative relevance of the ZEW survey. Thereof, 50.7% conduct forecasts to serve the decisions of clients and 79.7% report that their forecasts are used for the inhouse trading strategy.

[Insert Table 6 about here]

Table 6, columns 1-2, reports the results of panel regressions with cluster-robust standard errors. The results indicate a significant framing effect (1.67 percentage points, significant at the 1% level) in the group of Forecasters. Against the prediction of hypothesis H3, the framing effect is even lower by 0.52 percentage points in the group of participants who do not conduct forecasts outside the scope of the ZEW survey, although the difference is not statistically significant at the 10% significance level. Hence, based on our evidence we cannot conclude that the framing effect is amplified by lack of motivation. The result, however, can be due to a smaller than assumed heterogeneity between the levels of motivation in the group of Forecasters and non-Forecasters.

In a second step we test for the impact of lack of cognitive resources on the scope of the framing effect. We approximate cognitive busyness by the exact timing of the submission of the response. We hypothesize that the framing effect should be intensified on Fridays driven by distraction from the upcoming weekend (as evidenced by DellaVigna and Pollet, 2009). The exact timing of the response is an appropriate measure of cognitive busyness because it is largely exogenously determined by the timing of emails and reminders by the ZEW team to the participants. We argue that this is the case for the following reasons: Firstly, the majority of the responses are submitted on a Monday or a Friday as illustrated in the Appendix, Figure 9, which can be explained by the scheduled beginning and end of the survey waves on Mondays and scheduled reminders on Fridays. Secondly, there is a clustering of responses in the morning, especially on Mondays and Fridays, which can be explained by the fact that all e-mails to the participants are sent out before noon. Nevertheless, we acknowledge that our results may underestimate the impact of cognitive busyness. In situations of extreme cognitive overload, participants are unlikely to begin the survey. Hence, our results rather reflect the impact of different variations of moderate cognitive busyness on the framing effect. On the positive side, the proxy has the advantage of being available for each observation in the sample. Furthermore, the time of response is not prone to any self-reporting biases.

Table 6, columns 3-4, displays the regression results from panel regressions with cluster-robust standard errors. Against the predictions of `H3`, the framing effect does not increase, but rather decreases by 0.44 percentage points for responses submitted on a Friday, although the decrease is not statistically significant at the 10% level.

5 Conclusion

Our experiment so far allows conclusions on at least one channel through which the violation to procedure invariance influences the development of financial markets, that is through the recommendations of analysts who, as a group, are particularly prone to the judgmental bias. Womack (1996) and Barber, Lehavy, McNichols, and Trueman (2001) provide evidence that the forecasts of analysts are indeed relevant for the decisions of traders and have a significant influence on the development of the market price. Given that these forecasts are influenced by a superficial aspect such as the domain in which they are conducted, it may well be the case that incorporating them actually decreases instead of increases the information efficiency of the market price. Further analysis of this issue is particularly relevant in light of our evidence that professionals are rather homogeneous in their preferences for one particular forecast domain, hence the impact of the framing effect cannot be averaged out by considering analysts' forecasts in the aggregate (for instance sentiment indicators).

Moreover, the forecast domain that professionals use most widely is exactly the one that makes them more prone to intuitive and confident responses. So far this has been evidenced for stock market forecasts and for sporting bets (Simmons, Nelson, Galak, and Frederick, 2011) alike. There are two possible explanation for this observation. Firstly, there might be specific properties of particular forecast domains that make us rely more heavily on our intuition and we might tend to use these "intuitive" forecast domains more often because this way we process our tasks more easily and we feel more comfortable about our decisions (without being aware of the potential compromises on quality that we are making). Secondly, it might be just the fact that a particular domain happens to be the default domain (for exogenous reasons) and we are very familiar with it, which makes us very often respond automatically (i.e. intuitively). Future research should analyze how mental representations of the decision environment (e.g. forecasting domains, forms of information representation etc.) evolve endogenously particularly in such decision environments, like financial markets, in which skilled intuition is not feasible.

References

Andreassen, P. B. (1988): "Explaining the Price-Volume Relationship: The Difference between Price Changes and Changing Prices," *Organizational Behavior and Human Decision Processes*, 41, 371–389.

Barber, B., R. Lehavy, M. McNichols, and B. Trueman (2001): "Can Investors Profit from the Prophets? Security Analysts Recommendations and Stock Returns," *The Journal of Finance*, 61(2), 531–563.

Baumeister, R. F., E. Bratslavsky, M. Muraven, and D. M. Tice (1998): "Ego Depletion: Is the Active Self a Limited Resource?," *Journal of Personality and Social Psychology*, 74(5), 1252–1265.

Ben-David, I., J. Graham, and C. Harvey (2013): "Managerial Miscalibration," Quarterly Journal of Economics, forthcoming.

Bless, H., T. Betsch, and A. Franzen (1998): "Framing the framing effect: The impact of context cue on solutions to the 'Asian disease' problem," *European Journal of Social Psychology*, 28, 287–291.

Botvinick, M. M., T. Braver, S. Barch, C. Carter, and J. Cohen (2001): "Conflict Monitoring and Cognitive Control," *Psychological Review*, 108(3), 624–652.

Bradburn, N. (1982): "Question-Wording Effects in Surveys," in *Question Framing and Response Consistency*, ed. by R. M. Hogarth. Jossey-Bass Inc., Publishers, San Francisco.

Bruine de Briun, W., W. Klaauw, and G. Topa (2011): "Expectaitons in Inflation: The biasing Effect of Thoughts about Specific Prices," *Journal of Economic Psychology*, 32, 834–845.

Campbell, S. D., and S. A. Sharpe (2009): "Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices," *Journal of Financial and Quantitative Analysis*, 44(2), 369–390.

De Martino, B., D. Kumaran, B. Seymour, and R. Dolan (2006): "Frames, Brains, and Rational Decision-Making in the Human Brain," *Science*, 313, 684–687.

Deaves, R., E. Lüders, and M. Schröder (2010): "The Dynamics of Overconfidence: Evidence from Stock Market Forecasters," *Journal of Economic Behavior & Organization*, 75, 402–412.

DellaVigna, S., and J. M. Pollet (2009): "Investor Inattention and Friday Earnings Announcements," *The Journal of Finance*, 64(2), 709–749.

Epstein, S., R. Pacini, V. Denes-Raj, and H. Heier (1996): "Individual Differences in Intuitive-Experiential and Analytical-Rationall Thinking Styles," *Journal of Personality and Social Psychology*, 71(2), 390–405.

Evans, J. (2008): "Dual-Processin accounts of reasoning, judgment, and social cognition," *Annual Review of Psychology*, 59, 255–278.

Gilad, D., and D. Kliger (2008): "Priming the Risk Attitudes of Professionals in Financial Decision Making," *Review of Finance*, 12, 567–586.

Gilbert, D. T., B. Pelham, and D. Krull (1988): "On Cognitive Busyness: When Person Perceivers Meet Persons Perceived," *Journal of Personality and Social Psychology*, 54(5), 733–740.

Glaser, M., T. Langer, J. Reynders, and M. Weber (2007): "Framing Effects in Stock Market Forecasts: The Difference Between Asking for Prices and Asking for Returns," *Review of Finance*, 11(2), 325–357.

Glaser, M., and T. Walther (2013): "Run, Walk, or Buy? Financial Literacy, Dual-Process Theory, and Investment Behavior," Working Paper.

Glöckner, A. (2008): "How Evolution Outwits Bounded Rationality: The Efficient Interaction of Automatic and Deliberate Processes in Decision Making and Implications for Institutions," in *Better than Conscious? Decision Making, the Human Mind, and Implications for Institutions*, ed. by C. Engel, and W. Singer. The MIT Press, Cambridge.

Glöckner, A., and T. Betsch (2008): "Multiple-reason decision making based on automatic processing," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1055–1075.

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2011): "IQ and Stock Market Participation," *The Journal of Finance*, 66(6), 2121–2164.

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2012): "IQ, trading behavior, and performance," *Journal of Financial Economics*, 104, 339–362.

Haigh, M. S., and J. A. List (2005): "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," *The Journal of Finance*, 60(1), 523–534.

Harrison, G. W., and J. A. List (2004): "Field Experiments," *Journal of Economic Literature*, 42, 1009–1055.

Hoffmann, A., and H. Shefrin (2011): "Online Investors: What They Want, What They Do, And How Their Portfolios Perform," SSRN Working Paper.

Kahneman, D. (2003): "A Perspective on Judgment and Choice. Mapping Bounded Rationality," *American Psychologist*, 58(9), 697–720.

Kahneman, D., and S. Frederick (2002): "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in *Heuristics and Biases. The Psychology of Intuitive Judgment*, ed. by T.Gilovich, D. Griffin, and D. Kahneman, pp. 49–81. Cambridge University Press, Cambridge, United Kingdom.

——— (2006): "Frames and brains: elicitaiton and control of response tendencies," *Trends in Cognitive Sciences*, 11(2), 45–46.

Kahneman, D., and G. Klein (2009): "Conditions for intuitive expertise: A failure to disagree," *American Psychologist*, 64(6), 515–526.

Kaufmann, C., M. Weber, and E. Haisley (2013): "The Role of Experience Sampling and Graphical Displays on One's Investment Risk Appetite," *Management Science*, 59, 323–340.

Kempf, A., C. Merkle, and A. Niessen-Ruenzi (2012): "Low Risk and High Return-Affective Attitudes and Stock Market Expectations," *European Financial Management, forthcoming.*

Kumar, A., and G. M. Korniotis (2013): “Do Portfolio Distortions Reflect Superior Information or Psychological Biases?,” *Journal of Financial and Quantitative Analysis*, FirstView, 1–80.

Lawrence, M., and M. O’Connor (1993): “Scale, Variability, and the Calibration of Judgmental Prediction Intervals,” *Organizational Behavior and Human Decision Processes*, 56, 441–458.

Levin, I., G. Gaeth, and J. Schreiber (2002): “A New Look at Framing Effects: Distribution of Effect Sizes, Individual Differences, and Independence of Types of Effects,” *Organizational Behavior and Human Decision Processes*, 88(1), 411–429.

McElroy, T., and J. Seta (2003): “Framing effects: An analytic-holistic perspective,” *Journal of Experimental and Social Psychology*, 30, 4415–4430.

Menkhoff, L. (2010): “The use of technical analysis by fund managers: International evidence,” *Journal of Banking and Finance*, 34, 2573–2586.

Meyler, A., and I. Rubene (2009): “Results of a special questionnaire for participants on the ECB Survey of Professional Forecasters (SPF),” MPRA Archive.

Mussweiler, T., and K. Schneller (2003): “What goes up must come down - How charts influence decisions to buy and sell stocks,” *Journal of Behavioral Finance*, 4, 121–130.

Northcraft, G. B., and M. Neale (1987): “Experts, Amateur, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions,” *Organizational Behavior and Human Decision Processes*, 39, 84–97.

Pearson, E., and J. Tukey (1965): “Approximate Means and Standard Deviations Based on Distances Between Percentage Points of Frequency Curves,” *Biometrika*, 52(3-4), 533–546.

Roszkowski, M. J., and G. Snelbecker (1990): “Effects of ”Framing” on Measures of Risk Tolerance: Financial Planners Are Not Immune,” *The Journal of Behavioral Economics*, 19(3), 237–246.

Rubinstein, A. (2007): "Instinctive and Cognitive Reasoning: A Study of Response Times," *The Economic Journal*, 117, 1243–1259.

Schoorman, F., R. Mayer, C. Douglas, and C. Hetrick (1994): "Escalation of Commitment and the Framing Effect: An Empirical Investigation," *Journal of Applied Social Psychology*, 24(6), 509–528.

Sharot, T. (2011): *The Optimism Bias: A Tour of the Irrationally Positive Brain*. Pantheon Books, New York.

Simmons, J. P., and L. D. Nelson (2006): "Intuitive Confidence: Choosing Between Intuitive and Nonintuitive Alternatives," *Journal of Experimental Psychology: General*, 135(3), 409–428.

Simmons, J. P., L. D. Nelson, J. Galak, and S. Frederick (2011): "Intuitive Biases in Choice versus Estimation: Implications for the Wisdom of Crowds," *Journal of Consumer Research*, 38, 1–15.

Sjöberg, L. (2003): "Intuitive vs. analytical decision making: which is preferred?," *Scandinavian Journal of Management*, 19, 17–29.

Slovic, P., and S. Lichtenstein (1983): "Preference Reversals: A Broader Perspective," *American Economic Review*, 73(4), 596–605.

Soll, J. B., and J. Klayman (2004): "Overconfidence in Interval Estimates," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 299–314.

Stanovich, K., and R. West (2002): "Individual Differences in Reasoning: Implications for the Rationality Debate?," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. by T.Gilovich, D. Griffin, and D. Kahneman, pp. 421–440. Cambridge University Press, Cambridge, United Kingdom.

Womack, K. L. (1996): "Do Brokerage Analysts' Recommendations Have Investment Value?," *Journal of Finance*, 51, 137–167.

Table 1: Framing Effect over Time and over Forecast Horizons

Table 1 reports regression coefficients for the framing effect in DAX one-month and DAX one-year forecasts. Level forecasts are converted into return forecasts by means of DAX daily open on the day of the response (calculation method "daily open") or by means of DAX daily average on the day of the response (calculation method "daily avg"). Daily data on DAX is downloaded from Datastream. Reported are coefficients from panel regressions with cluster-robust standard errors. The independent variables are given and coded as follows: dummy coding for the treatment fixed effect (D^{ret}= 1 if return group); effect coding for wave fixed effects. The interpretation of the regression coefficients is indicated below.

$$r_{it} = \beta_0 + \beta_1 D^{ret} + \sum_{j=2}^{5} \beta_i C^{t_j} + \sum_{j=2}^{5} \beta_{4+j} D^{ret} C^{t_j} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)
calculation method	daily open	daily avg	daily open	daily avg
forecast horizon	1 mth	1 mth	1 year	1 year
$\beta_1 = \overline{r}^{ret} - \overline{r}^{lev}$	0.0150***	0.0158***	0.0267*	0.0274*
	(0.0034)	(0.0034)	(0.0149)	(0.0150)
$\beta_2 = r^{lev,t_2} - \overline{r}^{lev}$	0.0046	0.0053*		
	(0.0029)	(0.0029)		
$\beta_3 = r^{lev,t_3} - \overline{r}^{lev}$	0.0088***	0.0083***		
	(0.0028)	(0.0027)		
$\beta_4 = r^{lev,t_4} - \overline{r}^{lev}$	0.0019	0.0031		
	(0.0030)	(0.0029)		
$\beta_5 = r^{lev,t_5} - \overline{r}^{lev}$	-0.0067	-0.0067		
	(0.0045)	(0.0045)		
$\beta_6 = (r^{ret,t_2} - r^{lev,t_2}) - (\overline{r}^{ret} - \overline{r}^{lev})$	-0.0007	-0.0014		
	(0.0041)	(0.0041)		
$\beta_7 = (r^{ret,t_3} - r^{lev,t_3}) - (\overline{r}^{ret} - \overline{r}^{lev})$	-0.0053	-0.0048		
	(0.0034)	(0.0034)		
$\beta_8 = (r^{ret,t_4} - r^{lev,t_4}) - (\overline{r}^{ret} - \overline{r}^{lev})$	-0.0035	-0.0046		
	(0.0042)	(0.0042)		
$\beta_9 = (r^{ret,t_5} - r^{lev,t_5}) - (\overline{r}^{ret} - \overline{r}^{lev})$	0.0055	0.0054		
	(0.0049)	(0.0049)		
$\beta_0 = \overline{r}^{lev}$	-0.0019	-0.0027	0.0339***	0.0333***
	(0.0022)	(0.0022)	(0.01000)	(0.0100)
N	650	650	114	114
N_g	191	191	-	-

Cluster-robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 2: Framing Effect and Self-Reported Level of Intuitive Forecasting

Table 2 reports regression coefficients for intergroup differences in the scope of the framing effect in DAX one-month forecasts. Level forecasts are converted into return forecasts by means of DAX daily open on the day of the response (calculation method "daily open") or by means of DAX daily average on the day of the response (calculation method "daily avg"). Daily data on DAX is downloaded from Datastream. Reported are coefficients from panel regressions with cluster-robust standard errors. In columns (1) and (2) intuition refers to Faith-in-Intuition as measured by self-reported importance of intuition for the conduction of short-term (1 month ahead) stock market forecasts - "low", "medium" or "high". In columns (3) and (4) intuition refers the Intuitive-Analytical Score - relative importance of intuition as compared to the importance of the most important among the analytical methods. The independent variables are given and coded as follows: return group dummy (D^{ret}= 1 if return group); intuition approximated by Faith-in-Intuition and Intuitive-Analytical-Score respectively ($i2$ indicates the middle category of intuitive thinking, $i1$ indicates the lowest category of intuitive thinking, $i3$ indicates the highest category of intuitive thinking).

$$r_{it} = \beta_0 + \beta_1 D^{ret} + \beta_2 D^{i2} + \beta_3 D^{i1} + \beta_4 D^{ret} D^{i2} + \beta_5 D^{ret} D^{i1} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)
calculation method	daily open	daily avg	daily open	daily avg
proxy for intuitive thinking	absolute	absolute	relative	relative
$\beta_1 = r^{ret,i3}$	0.0211**	0.0219**	0.0251***	0.0255***
	(0.0084)	(0.0085)	(0.0068)	(0.0069)
$\beta_2 = r^{lev,i2} - r^{lev,i3}$	0.0039	0.0040	0.0071	0.0068
	(0.0061)	(0.0062)	(0.0058)	(0.0059)
$\beta_3 = r^{lev,i1} - r^{lev,i3}$	0.0080	0.0077	0.0110	0.0102
	(0.0071)	(0.0072)	(0.0069)	(0.0070)
$\beta_4 = (r^{ret,i2} - r^{lev,i2}) - (r^{ret,i3} - r^{lev,i3})$	-0.0093	-0.0094	-0.0172*	-0.0169*
	(0.0105)	(0.0106)	(0.0095)	(0.0095)
$\beta_5 = (r^{ret,i1} - r^{lev,i1}) - (r^{ret,i3} - r^{lev,i3})$	-0.0112	-0.0108	-0.0175	-0.0167
	(0.0113)	(0.0114)	(0.0111)	(0.0111)
$\beta_0 = r^{lev,i3}$	-0.0035	-0.0043	-0.0050	-0.0054
	(0.0049)	(0.0050)	(0.0045)	(0.0046)
N	479	479	478	478
N_g	123	123	122	122

Cluster-robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 3: Development of Framing Effect and Subjective Confidence with Reaction Time

Table 3 reports regression coefficients for the development of the framing effect in DAX one-month forecasts and subjective volatility estimates with reaction time in the subsample of March 2013. Level forecasts are converted into return forecasts by means of actual DAX daily open on the day of the response (calculation method "daily open") and daily average on the day of the response (calculation method "daily avg"). DAX daily data is downloaded from Datastream. Columns (1), (3), (5) and (7) report the results for regressions on the whole sample and columns (2), (4), (6) and (8) report the results on the participants who have filled in the entire Special Questions section (group with sound statistical knowledge). Dummy coding is used for the treatment fixed effect (D^{ret}= 1 if return group). The reaction time T_i^* is measured in seconds, rescaled in minutes and logarithmized.

$$r_i = \beta_0 + \beta_1 D^{ret} + \beta_2 T_i^* + \beta_3 D^{ret} T_i^* + \epsilon_i$$
$$\sigma_i = \beta_0 + \beta_1 D^{ret} + \beta_2 T_i^* + \beta_3 D^{ret} T_i^* + \epsilon_i$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations on DAX 1m [percentage points]				Subjective volatility estimates on DAX 1m			
	daily open	daily open	daily avg	daily avg	daily open	daily open	daily avg	daily avg
β_1	0.0209***	0.0494***	0.0219***	0.0492***	0.0042	0.0309	0.0042	0.0309
	(0.0068)	(0.0134)	(0.0068)	(0.0134)	(0.0083)	(0.0219)	(0.0083)	(0.0219)
β_2	0.00942**	0.0259***	0.00893**	0.0245***	0.0054	0.0221*	0.0054	0.0221*
	(0.0037)	(0.0071)	(0.0037)	(0.0071)	(0.0043)	(0.0116)	(0.0043)	(0.0116)
β_3	-0.0120**	-0.0284***	-0.0115**	-0.0270***	-0.0112*	-0.0304**	-0.0112*	-0.0304**
	(0.0052)	(0.0087)	(0.0052)	(0.0087)	(0.0062)	(0.0142)	(0.0061)	(0.0142)
β_0	-0.0019	-0.0257**	-0.0029	-0.0255**	0.0242***	0.0022	0.0242***	0.0022
	(0.0043)	(0.0104)	(0.0043)	(0.0103)	(0.0050)	(0.0169)	(0.0050)	(0.0169)
N	118	46	118	46	111	46	111	46

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 4: Framing Effect and Deliberate Reasoning

Table 4 reports regression coefficients for intergroup differences in the scope of the framing effect in DAX one-month forecasts. Level forecasts are converted into return forecasts by means of DAX daily open on the day of the response (calculation method "daily open") or by means of DAX daily average on the day of the response (calculation method "daily avg"). Daily data on DAX is downloaded from Datastream. Reported are coefficients from panel regressions with cluster-robust standard errors. We measure deliberate reasoning by means of the self-reported importance of technical analysis for the conduction of short-term (1 month ahead) stock market forecasts - "low", "medium" or "high". The independent variables are given and coded as follows: return group dummy (D^{ret}= 1 if return group); technical analysis dummies for the categories "low" and "medium" ($ta2$ indicates medium importance, $ta1$ indicates low importance, $ta3$ indicates high importance of technical analysis).

$$r_{it} = \beta_0 + \beta_1 D^{ret} + \beta_2 D^{ta2} + \beta_3 D^{ta1} + \beta_4 D^{ret} D^{ta2} + \beta_5 D^{ret} D^{ta1} + \epsilon_{it}$$

	(1)	(2)
calculation method	daily open	daily avg
$\beta_1 = r^{ret,ta3}$	0.0073	0.0081
	(0.0055)	(0.0055)
$\beta_2 = r^{lev,ta2} - r^{lev,ta3}$	-0.0115*	-0.0115*
	(0.0060)	(0.0060)
$\beta_3 = r^{lev,ta1} - r^{lev,ta3}$	0.0003	0.0005
	(0.0060)	(0.0060)
$\beta_4 = (r^{ret,ta2} - r^{lev,ta2}) - (r^{ret,ta3} - r^{lev,ta3})$	0.0304***	0.0304***
	(0.0088)	(0.0088)
$\beta_5 = (r^{ret,ta1} - r^{lev,ta1}) - (r^{ret,ta3} - r^{lev,ta3})$	0.0013	0.0010
	(0.0113)	(0.0113)
$\beta_0 = r^{lev,ta3}$	0.0027	0.0020
	(0.0035)	(0.0035)
N	479	479
N_g	123	123

Cluster-robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 5: Subjective Confidence/Uncertainty and Self-Reported Level of Intuitive Forecasting

Table 5 reports regression coefficients for intergroup differences in the subjective confidence. Subjective confidence is derived from the subjective 90% confidence interval for one's own DAX one-month forecast using the approach by Pearson and Tukey (1965). Subjective confidence intervals of the level group are converted into return intervals by means of DAX daily open on the day of the response (calculation method "daily open") or by means of DAX daily average on the day of the response (calculation method "daily avg"). Daily data on DAX is downloaded from Datastream. We report coefficients from panel regressions with cluster-robust standard errors. In columns (1) and (2) intuition refers to Faith-in-Intuition as measured by self-reported importance of intuition for the conduction of short-term (1 month ahead) stock market forecasts - "low", "medium" or "high". In columns (3) and (4) intuition refers the Intuitive-Analytical Score - relative importance of intuition as compared to the importance of the most important among the analytical methods. The independent variables are given and coded as follows: return group dummy (D^{ret}= 1 if return group); intuition approximated by Faith-in-Intuition and Intuitive-Analytical-Score respectively ($i2$ indicates the middle category of intuitive thinking, $i1$ indicates the lowest category of intuitive thinking, $i3$ indicates the highest category of intuitive thinking).

$$\sigma_{it} = \beta_0 + \beta_1 D^{ret} + \beta_2 D^{i2} + \beta_3 D^{i1} + \beta_4 D^{ret} D^{i2} + \beta_5 D^{ret} D^{i1} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)
calculation method	daily open	daily avg	daily open	daily avg
proxy for intuitive thinking	absolute	absolute	relative	relative
$\beta_1 = \sigma^{ret,i3}$	0.0010	0.0010	0.0008	0.0008
	(0.0066)	(0.0066)	(0.0049)	(0.0049)
$\beta_2 = \sigma^{lev,i2} - \sigma^{lev,i3}$	0.0019	0.0019	0.0059	0.0059
	(0.0037)	(0.0037)	(0.0039)	(0.0039)
$\beta_3 = \sigma^{lev,i1} - \sigma^{lev,i3}$	0.0209**	0.0208**	0.0205*	0.0204*
	(0.0090)	(0.0090)	(0.0105)	(0.0105)
$\beta_4 = (\sigma^{ret,i2} - \sigma^{lev,i2}) - (\sigma^{ret,i3} - \sigma^{lev,i3})$	0.0016	0.0016	0.0031	0.0031
	(0.0080)	(0.0080)	(0.0074)	(0.0074)
$\beta_5 = (\sigma^{ret,i1} - \sigma^{lev,i1}) - (\sigma^{ret,i3} - \sigma^{lev,i3})$	-0.0137	-0.0136	-0.0185	-0.0185
	(0.0121)	(0.0121)	(0.0124)	(0.0124)
$\beta_0 = \sigma^{lev,i3}$	0.0222***	0.0222***	0.0217***	0.0217***
	(0.0028)	(0.0028)	(0.0024)	(0.0024)
N	456	456	455	455
N_g	121	121	120	120

Cluster-robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 6: Framing Effect and Lack of Motivation and Cognitive Resources

Table 6 reports regression coefficients for intergroup differences in the framing effect in DAX one-month forecasts. Level forecasts are converted into return forecasts by means of DAX daily open on the day of the response (calculation method ”daily open”) and daily average on the day of the response (calculation method ”daily avg”) . Daily data on DAX is downloaded from Datastream. We report coefficients from panel regressions with cluster-robust standard errors. The independent variables are given and coded as follows: return group dummy (D^{ret}= 1 if return group); a proxy for motivation based on occupation as a professional forecaster ($D^{noFcster}$=1 for ”low” motivation group consisting of participant who do not conduct on a regular basis forecasts outside the scope of the ZEW Financial Market Survey); a proxy for availability of cognitive resources based on the timing of the submission of the responses (D^{Friday}=1 for ”low” cognitive resources due to the distraction from the upcoming weekend).

$$r_{it} = \beta_0 + \beta_1 D^{ret} + \beta_2 D^{noFcster} + \beta_3 D^{ret} D^{noFcster} + \epsilon_{it}$$
$$r_{it} = \beta_0 + \beta_1 D^{ret} + \beta_2 D^{Friday} + \beta_3 D^{ret} D^{Friday} + \epsilon_{it}$$

	(1)	(2)	(3)	(4)
calculation method proxy for motivation/cognitive ressources	daily open Forecaster	daily avg Forecaster	daily open no Friday	daily avg no Friday
$\beta_1 = r^{ret,high}$	0.0167***	0.0174***	0.0163***	0.0176***
	(0.0044)	(0.0045)	(0.0036)	(0.0036)
$\beta_2 = r^{lev,low} - r^{lev,high}$	0.0022	0.0024	0.0040	0.0055
	(0.0048)	(0.0049)	(0.0039)	(0.0039)
$\beta_3 = (r^{ret,low} - r^{lev,low}) - (r^{ret,high} - r^{lev,high})$	-0.0052	-0.0054	-0.0044	-0.0058
	(0.0076)	(0.0077)	(0.0053)	(0.0052)
$\beta_0 = r^{lev,high}$	-0.0027	-0.0034	-0.0031	-0.0044*
	(0.0031)	(0.0031)	(0.0022)	(0.0022)
N	547	547	650	650
N_g	139	139	191	191

Cluster-robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Figure 1: Market Phases Covered by the Field Experiment

Daily data on DAX performance index was taken from Datastream. The reference lines correspond to the respective timing of the waves in September 2012, December 2012, March 2013, June 2013 and September 2013.

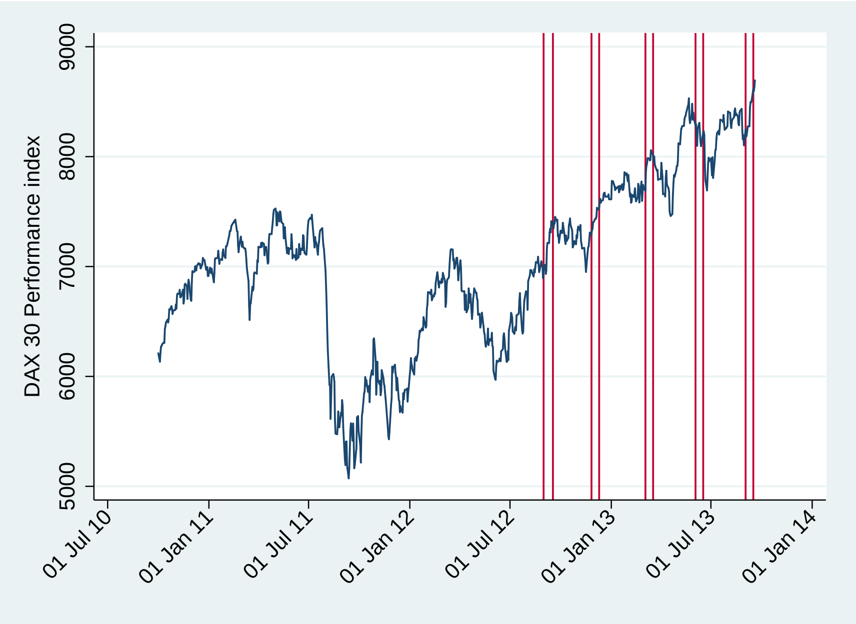


Figure 2: Framing Effect Over Time

Figure 2 displays the difference between the return forecasts by the return group and the level group on average and over all waves of the experiment. Return forecasts of the level group are calculated from the level forecasts and DAX daily open on the day of the response. The corresponding regressions are displayed in Table 1, column (1).

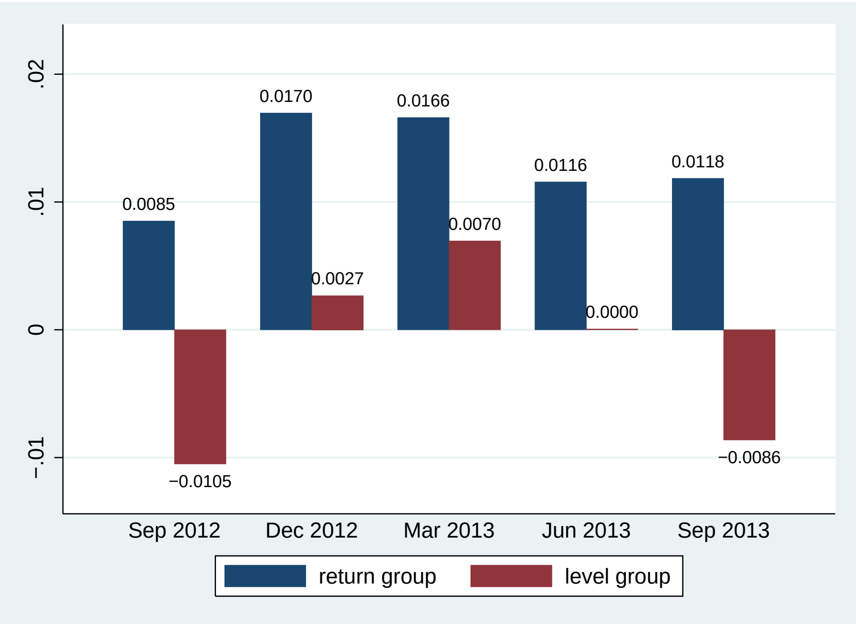


Figure 3: Impact of Intuitive Thinking on Framing Effect

Figure 3 displays differences in the scope of the framing effect depending on self-reported level of intuitive thinking. Absolute (relative) importance of intuition if measured by "Faith in Intuition" ("Intuitive-Analytical Score"). The corresponding regressions are displayed in Table 2, columns (1) and (3).

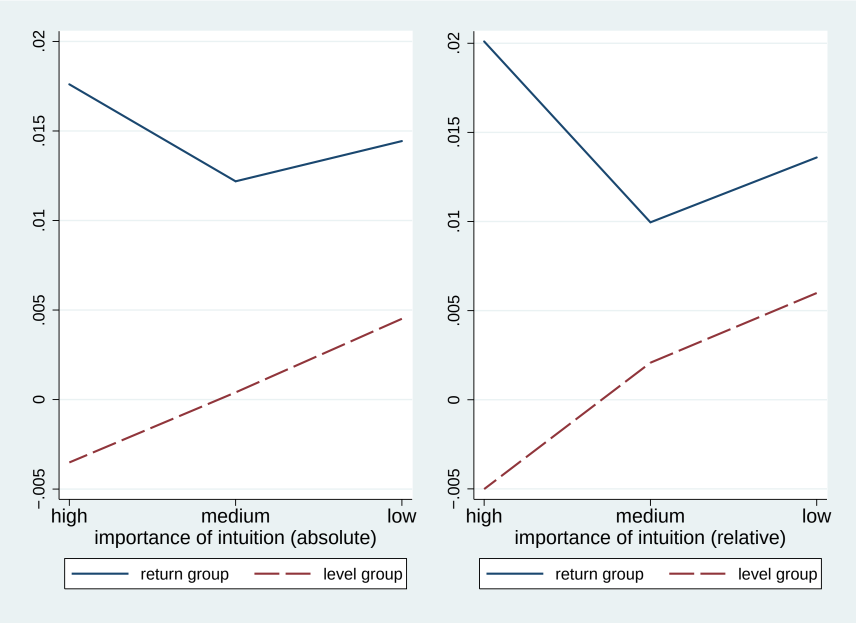


Figure 4: Impact of Reaction Time on Framing Effect

Figure 4 displays forecasts of the return group and the level group depending on the reaction time. Fractional polynomials are used for the predicted lines. The left part contains all responses, the right part displays the results only for the participants who have responded to the whole Special Questions section (participants with sound statistical knowledge). For the calculation of return forecasts of the level group DAX daily open on the day of the response is applied. Vertical reference lines indicate respective median reaction times.

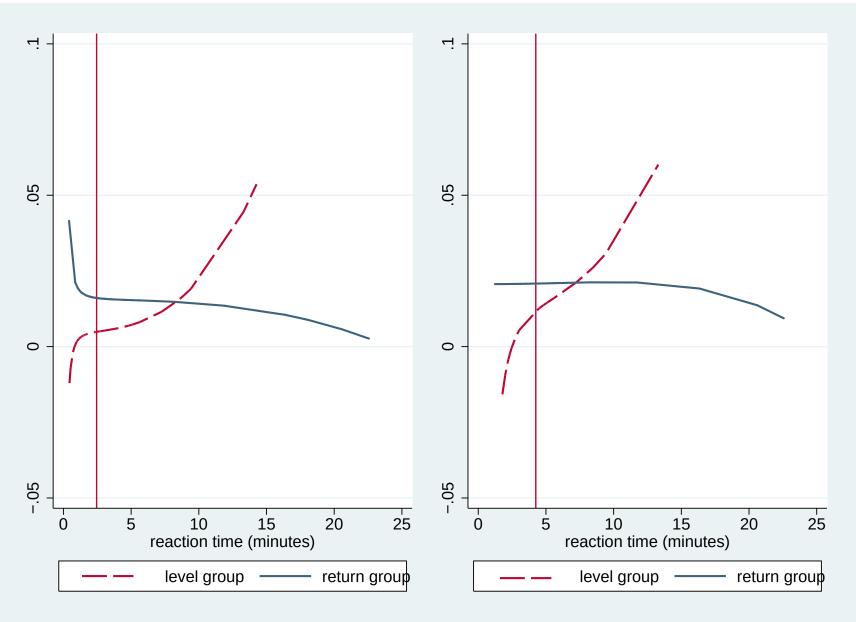


Figure 5: Impact of Intuitive Thinking on Subjective Confidence of the Level Group

Figure 5 displays differences in the subjective volatility estimates depending on the level of intuitive thinking as measured by both self-reported measures (absolute and relative) and reaction time. The corresponding regressions are displayed in Table 5, columns (1) and (3) and Table 3, columns (5)-(8).

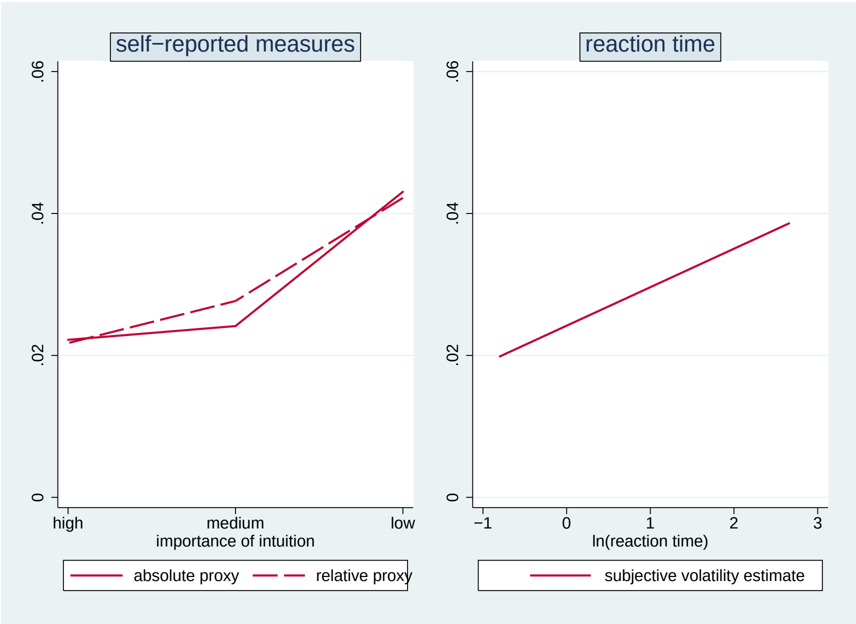
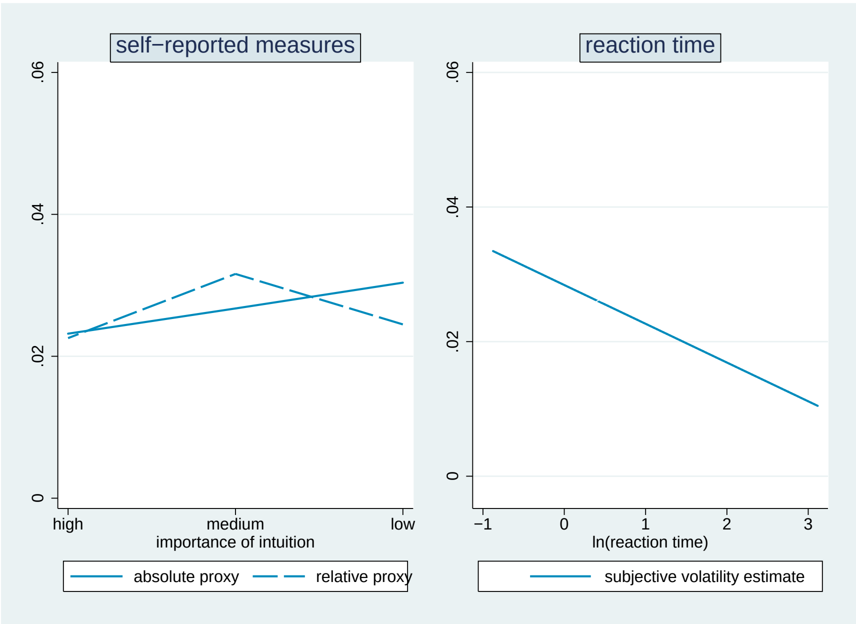


Figure 6: Impact of Intuitive Thinking on Subjective Confidence of the Return Group

Figure 6 displays differences in the subjective volatility estimates depending on the level of intuitive thinking as measured by both self-reported measures (absolute and relative) and reaction time. The corresponding regressions are displayed in Table 5, columns (1) and (3) and Table 8, columns (5)-(8).



6 Appendix

Table 7: Variable Definitions

Variable	Description
TREATMENT GROUPS	
A. POINT FORECAST	
- Description:	Return and level treatment, randomized, between-subject design
- Question wording:	[return treatment] *Within 1 month I expect a DAX return (monthly percentage change) of . . . %.*
	[level treatment] *I expect the DAX in 1 month at . . . points.*
- Timing:	Sep 2012-Sep 2013
- Coding:	$D^{ret} = 1$ & $D^{lev} = 0$ for return treatment
	$D^{ret} = 0$ & $D^{lev} = 1$ for level treatment
B. SUBJECTIVE CONFIDENCE	
- Description:	Return and level treatment, randomized, between-subject design
- Question wording:	[return treatment] *With a probability of 90% the DAX return will then lie between . . . % and . . . %.*
	[level treatment] *With a probability of 90% the DAX will then lie between . . . and . . . points.*
IMPORTANCE OF DIVERSE FORECASTING METHODS	
- Description:	Self-reported importance of diverse methods for the conduction of own forecasts (short-term forecasts and mid-term forecasts are elicited separately);
- Time of elicitation:	Special Question in September 2012
- Question wording:	*How important are following factors for your [e.g., short-term (1 month ahead)] DAX forecasts? . . . technical analysis, fundamental analysis, econometric models, intuition, experience, consensus forecasts, inhouse forecasts, simulations*
- Response categories:	"low", "medium", "high"
INTUITIVE FORECASTING	
A. FAITH IN INTUITION	
- Description:	Self-reported importance of intuition for the conduction of own short-term DAX forecasts
- Time of elicitation:	Special Question in September 2012

Continued on next page

Table 7 – continued from previous page

Variable	Description
- Question wording:	*How important are the following factors for your [short-term (1 month ahead)] DAX forecasts?*
	... intuition
- Response categories:	"low", "medium", "high"
- Coding:	$D^{i1} = 1$ if importance of intuition is "low"
	$D^{i2} = 1$ if importance of intuition is "medium"
	$D^{i2} = 0$ & $D^{i1} = 0$ if importance of intuition is "high"
B. Intuitive-Analytical Score	
- Description:	Self-reported importance of intuition for the conduction of own short-term DAX forecasts compared to the self-reported importance of the most important analytical method; Analytical methods: technical and fundamental analysis, econometric tools, simulations, inhouse and consensus forecasts
- Time of elicitation:	Special Question in September 2012
- Coding:	$D^{i1} = 1$ if intuition is much less (by 2 categories) important than analytical methods
	$D^{i2} = 1$ if intuition is less (by 1 category) important than analytical methods
	$D^{i2} = 0$ & $D^{i1} = 0$ if intuition is of comparable or higher importance
C. Reaction Time	
- Description:	The time between the beginning of the Special Questions section and the submission of the questionnaire is measured in seconds and converted into minutes. The Special Questions section in March 2013 contains questions on DAX and gold price forecasts, correlation assessments between DAX and other asset classes and a comments field; Excluded are participants who submitted a comment and participants who returned to the main questionnaire after entering the Special Questions section
- Time of elicitation:	March 2013
Cognitive Busyness	
Day of Week	
- Description:	Cognitive busyness is approximated by the timing of the

Continued on next page

Table 7 – continued from previous page

Variable	Description
	response (day of week); Response timing is measured by the time of the submission of the questionnaire
- Time of elicitation:	every wave
- Variable definition:	$D^{Fri} = 1$ if the questionnaire is submitted on a Friday
MOTIVATION	
FORECASTER	
- Description:	Self-reported information on regularly conducted forecasts outside the scope of the ZEW Financial Market Survey
- Time of elicitation:	Special Question in March 2013
- Question wording:	*What is the usual type of your regular forecasts outside the scope of the ZEW Financial Market Survey?*
- Response categories:	"level forecasts", "return forecasts", "range forecast" "directional forecast", "probability estimate", "other" "I do not conduct any explicit forecast"
- Coding:	$D^{noFcster} = 1$ if does not conduct any explicit forecasts outside the scope of the ZEW Financial Market Survey $D^{noFcster} = 0$ if participant conducts forecasts regularly

Table 8: Development of Framing Effect and Return Forecasts with Log Reaction Time

Table 8 reports regression coefficients for the development of the framing effect in DAX one-month forecasts and subjective volatility estimates with reaction time in the subsample of March 2013. Level forecasts are converted into return forecasts by means of actual DAX daily open on the day of the response (calculation method "daily open") and daily average on the day of the response (calculation method "daily avg"). DAX daily data is downloaded from Datastream. Columns (1), (3), (5) and (7) report the results for regressions on the whole sample and columns (2), (4), (6) and (8) report the results on the participants who have filled in the entire Special Questions section (group with sound statistical knowledge). Dummy coding is used for the treatment fixed effect (D^{lev}= 1 if level group). The reaction time T_i^* is measured in seconds, rescaled in minutes and logarithmized.

$$r_i = \beta_0 + \beta_1 D^{lev} + \beta_2 T_i^* + \beta_3 D^{lev} T_i^* + \epsilon_i$$
$$\sigma_i = \beta_0 + \beta_1 D^{lev} + \beta_2 T_i^* + \beta_3 D^{lev} T_i^* + \epsilon_i$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations on DAX 1m [percentage points]				Subjective volatility estimates on DAX 1m			
	daily open	daily open	daily avg	daily avg	daily open	daily open	daily avg	daily avg
β_1	-0.0209***	-0.0494***	-0.0219***	-0.0492***	-0.0042	-0.0309	-0.0042	-0.0309
	(0.0068)	(0.0134)	(0.0068)	(0.0134)	(0.0083)	(0.0219)	(0.0083)	(0.0219)
β_2	-0.0026	-0.0025	-0.0026	-0.0025	-0.0058*'	-0.0083	-0.0058*'	-0.0083
	(0.0037)	(0.0050)	(0.0036)	(0.0050)	(0.0044)	(0.0082)	(0.0044)	(0.0082)
β_3	0.0120**	0.0284***	0.0115**	0.0270***	0.0112*	0.0304**	0.0112*	0.0304**
	(0.0052)	(0.0087)	(0.0052)	(0.0087)	(0.0062)	(0.0142)	(0.0061)	(0.0142)
β_0	0.0189***	0.0237***	0.0189***	0.0237***	0.0284***	0.0332**	0.0284***	0.0332**
	(0.0053)	(0.0086)	(0.0053)	(0.0085)	(0.0066)	(0.0140)	(0.0066)	(0.0140)
N	118	46	118	46	111	46	111	46

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01, *' p(one-sided test)<0.1

Figure 7: Methods Used for Short-Term DAX Forecasts

Figure 7 displays the number of respondents who have indicated that a particular method has a high, medium or low importance for the conduction of their DAX forecasts 1 month ahead.

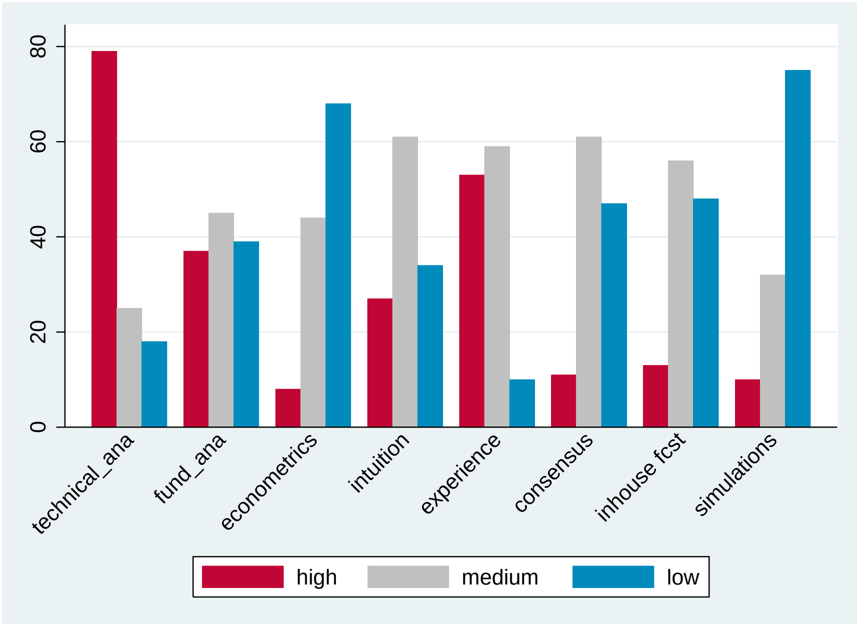


Figure 8: Usual Forecast/Investment Horizon and Preferences towards Response Mode

Figure 8 displays the sample distribution of usual forecast/investment horizons among the participants (left graph). Participants' preferences for a particular response mode were elicited in March 2012.

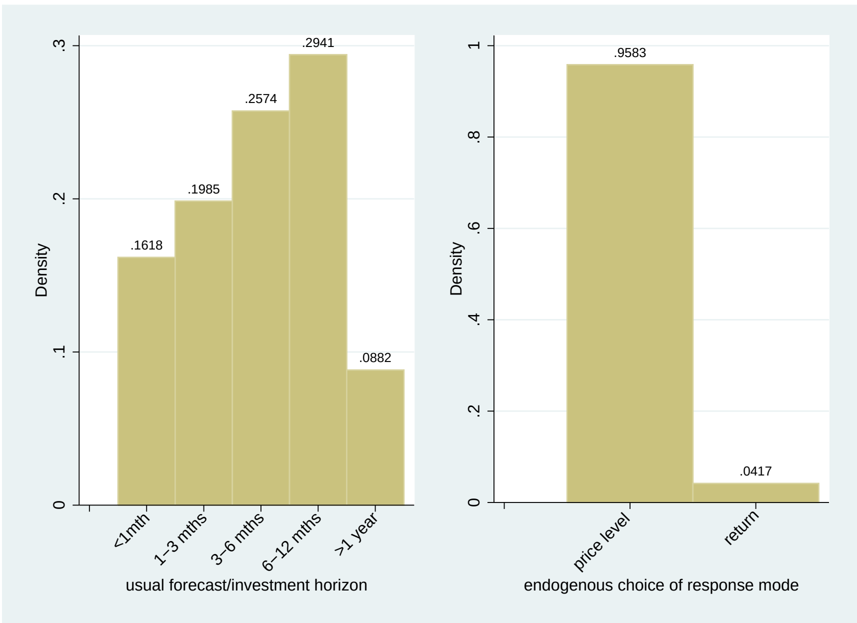


Figure 9: Response Timing