Individual Risk Tolerance and Herding Behaviors in Financial Forecasts

Christoffersen, Jeppe; Stæhr, Simone

Link to publication in CBS Research Portal

Individual Risk Tolerance and Herding Behaviors in Financial Forecasts

Jeppe Christoffersen, Simone Stæhr

Journal article (Accepted manuscript*)

Individual Risk Tolerance and Herding Behaviours in Financial Forecasts[1]

Jeppe Christoffersen

Simone Stæhr

Department of Accounting and Auditing

Copenhagen Business School

Solbjerg Plads 3, Frederiksberg 2000

Denmark

Abstract

Financial analysts tend to demonstrate herding behaviour, which sometimes compromises accuracy. A number of explanations spanning rational economic logic, cognitive biases, and social forces have been suggested. Relying on an experimental setting where participants forecast future earnings from a rich information set, we posit and obtain

support for individual risk tolerance (or lack thereof) as an explanatory variable for herding behaviours. Specifically, less risk tolerant individuals forecast with less boldness and instead issue forecasts in agreement with the consensus forecast. The results are argued to be at least partially a product of cognitive biases and an intuitive reaction to uncertainty.

1 Introduction

'Who herds? Who doesn't?' The questions have been implicitly asked during many years of research on financial analysts' earnings forecasts and recently Huang, Krishnan, Shon, & Zhou, (2017) very explicitly posed the questions in the title of their article in a leading accounting journal. Although the questions are explicit, the answers remain elusive, primarily because explanations have related to the *circumstances* under which individuals herd more or less rather than on the *characteristics* of these individuals. In this paper, we seek to complement the current literature by expanding the potential explanations of variation in herding behaviours to also include variation in the personal characteristics of the analysts.

An example of a paper claiming but not aiming to answer the question of who herds is the above-cited study by Huang et al. (2017). The authors very explicitly answered the questions in their title ('Who Herds? Who Doesn't?') with the finding that herding propensity varies across time, forecast horizon, analyst following, broker/employer size

and forecast order. While these factors are very specific to the task and content of conducting financial forecasts, they relate very little to the characteristics of the persons producing these forecasts or their cognitive processes. A similar lack of focus on individuals is found in other contributions explaining variation in herding behaviour with reputational reasons (Holmes et al.*,* 2013), releases of macro data (Galariotis et al.*,* 2015), regulation limiting selectively disclosed information (Hahn & Song, 2013; Mensah & Yang, 2008), opaqueness of firms' information environment (Leece & White, 2017), analysts' access to information (Christensen et al.*,* 2017), analyst affiliation (Xue, 2016), brokerage size (Clement & Tse, 2005; Jegadeesh & Kim, 2010), the number of industries followed (Clement & Tse, 2005), firm earnings uncertainty (Song et al., 2009), dispersion across recommendations (Jegadeesh & Kim, 2010), forecast revision frequency (Jegadeesh & Kim, 2010), forecast horizon (De Bondt & Forbes, 1999), analyst's prior accuracy (Clement & Tse, 2005), fee concerns (Trueman, 1994), career concerns (Clarke & Subramanian, 2006; Clement & Tse, 2005; Hong et al., 2000; Nolte et al., 2014), experience (Clement & Tse, 2005; Hong et al., 2000; Youssef & Rajhi, 2010), and trust building (Kadous et al., 2009).

It is clear from the above that many of the previous studies focus on variables that differ between each individual, such as prior forecast errors and experience; however, these variables describe the individuals in their job situation and do not relate to their personalities more generally. There are two notable exceptions not listed above. One is Kumar (2010), who found that female analysts issue bolder forecasts. However, the explanation given is that only the best potential female analysts enter the profession due to a perception of discrimination in the analyst labour market; that is, the results were

argued to be an effect of biased self-selection rather than traits relating to typical female personalities. The other exception is Jiang et al. (2017), who found that analysts supporting the Republican Party are less likely to issue bold forecasts, that is, they are more likely to exhibit herding behaviour. The authors' key theoretical argument was that analysts who are most strongly aligned with the Republican Party are likely to possess conservative traits and thus prefer the status quo. This makes them cautious about interpreting new information and about updating their beliefs and, therefore, they do not stray from the consensus.

Our study is related to that of Jiang et al. (2017), since we focus on a personal characteristic closely related to conservatism, that is, individuals' general risk tolerance.[2] We hypothesize and find that the least risk-tolerant analysts – most likely the most conservative – issue less bold forecasts. However, we arrive at our hypotheses using theoretical arguments with a strong link to the psychological mechanisms of the individuals. This allows us to provide a detailed account of why some individuals will *not* be cautious in relation to new information, as implied by Jiang et al. (2017), but could, in fact, rely heavily on information from 'the pack'. Using students as analysts in an experimental setting that aims to reflect important dimensions of real-world contexts, this paper is the first to theorize about and empirically support the importance of the so-called herd instinct (Hirshleifer & Hong, 2003, p. 25) often hinted at in the literature on analyst earnings forecasts.

---

[2] Risk tolerance refers to participants' perceived willingness to take risks based on a 10-point Likert scale (from not willing to take risks to very willing to take risks).

We define herding behaviour as 'the propensity of analysts to follow the consensus' (Welch, 2003, p. 374, l 2 fb.). We also refer to making bold forecasts or, rather, to the boldness in forecasts, as being the opposite of forecasting close to the consensus estimate (herding). Relying on theories and positions in social psychology, such as group conformity (Asch, 1951, 1952, 1956), social comparison theory (Festinger, 1954), and partly behavioural decision theory (Einhorn & Hogarth, 1981; Kahneman 2011), we argue that individuals with lower risk tolerance feel a greater need for safety in their decision making. A search for safety could have a number of consequences for individuals forecasting future earnings. They may rely heavily on past earnings to predict future earnings (since these are objective), current developments (since these point towards the future), or management forecasts (since management may have inside knowledge). However, by drawing on arguments from the social psychology literature, we suggest that individuals will rely heavily on forecasts made by their peers. In other words, the lower their risk tolerance, the more likely they are to follow their peers rather than make bold forecasts based on their own processing of information about the company. This hypothesis is the paper's first and primary hypothesis.

This paper also seeks to add depth by developing theory regarding the differences in the strength of this relation depending on the strength of external stimuli. Specifically, considering findings in the psychology literature presenting the "general principle across a broad range of psychological phenomena' that, in individuals' minds, 'bad is stronger than good' (Baumeister et al., 2001, p. 323), the characteristics of news (bad or good) asymmetrically likely affect the relation stated in the first hypothesis. Therefore, it is

argued that the positive relation between risk tolerance and boldness in forecasts is moderated so it is amplified by bad news. This is our second hypothesis.

In line with other recent top journal contributions studying psychology-related mechanisms relevant to the behaviour of accounting and finance professionals (e.g. Chang et al., 2016; Emett & Nelson, 2017; Kumar, 2010, 2015; Rennekamp et al., 2015), we test and find support for our two hypotheses via an experimental setting using 289 business graduate students. Although the participants are not professional financial analysts, they do possess important analyst characteristics, such as knowledge of financial accounts and the limitations pertaining to earnings predictability. Furthermore, we believe that our theory and tests are relevant to the study of analysts conducting forecasts in a real-world setting, since our hypotheses rely on arguments related to deeply rooted mechanisms in individuals conducting earnings forecasts rather than related to market specificities. Even if these mechanisms have different degrees of importance in a real-world setting for professional analysts, since a multitude of other factors also play important roles (see above), we believe that the mechanisms of our hypotheses still apply to some extent. And possibly, to an even greater extent in a real-world setting, where the stakes are higher, potentially increasing the compulsion to seek safety by leaning on one's peers.

The experimental method allows us to provide controlled stimuli and directly observe the resulting forecasts. The experiment follows a 2×2 between-subjects design. All participants receive a rich information set. In one treatment, half of the participants additionally receive a consensus estimate before revising forecasts. In the other treatment,

half of the participants receive bad news about the company before revising forecasts while the other half receive good news.

We presume, in accordance with Kahneman's (2011) two-system theory (System 1 being the fast, intuitive decision making system and System 2 being the slower, more elaborate decision making system), that financial analysts, like the rest of us, have limited attentional resources. Therefore, we find it likely that financial analysts do not spend cognitive efforts to deliberate on the usefulness of herding in all forecasting situations. Our main theoretical contribution lies in exploring the mechanisms related to the herd instinct often hinted at but not yet investigated in prior literature. Specifically, we develop the notion that individuals in a stressful, ambiguous situation seek safety by following their peers and that this tendency is most pronounced for individuals who are generally less risk tolerant. We follow Jiang et al. (2017) in advancing theory pertaining to the influence of innate (rather than trained or experience-based) individual characteristics on herding behaviour. In doing so, we respond to reviews on financial analyst forecasts (FAFs) that have called for more research on the role of individual analysts' personal characteristics in forecasting behaviours (e.g. Bradshaw, 2011; Ramnath et al., 2008). The academic implications of this emerging strand of research are that herding behaviour cannot be explained solely using strictly rational arguments of information cascades or intentional herding for strategic and reputational reasons. This, in turn, implies that, in practice, measures taken by brokerage houses and public authorities to curtail herding should reflect the fact that emotions and intuition can make analysts respond in unanticipated ways to regulation and financial incentives. Hence, more customized interventions directed at financial analysts to change their instinctive behaviours could be

necessary if changes in behaviour are sought. Alternatively, biased forecasts could –
given the suggested inherency – be accepted as a *fait accompli* and published along with
analyst-specific data, allowing investors to adjust for such biases. This will further allow
recruiters to add this dimension to their hiring strategy to the extent that they see fit.

In arriving at our theoretical contribution, we also provide a methodological contribution
by developing a methodology for studying earnings forecast herding in an experimental
setting. We are inspired by Welch (2000, p. 394), who, in the closing comment of a much
cited paper, argued that, with access to 'information sets available to individual analysts
when making decisions, one could discriminate between true herding and information
that is simply received simultaneously and interpreted likewise by two analysts'.
Specifically, we conduct a first round of earnings forecasts where all the participants
receive identical information and, then, in a second round, we give half the participants a
consensus with which they can herd. We use the other half of the participants to establish
estimates of how those given a consensus would have forecasted in its absence. Each
individual receiving a consensus is thus compared to one or more individuals in the
control group who behaved similarly in the first round. The design requires many
participants but provides significant internal validity benefits and should, therefore, be of
interest to researchers embarking into experimental research on herding.

2 Background, theory, and hypothesis development

2.1 Herding behaviour

The phenomenon of herding behaviour has a long history and is, broadly, the tendency of
individuals to follow others when making decisions as a reaction to uncertainty (Keynes,

1930/1982). Because people generally seek to avoid uncertainty[3] (e.g. Hogarth, 1975), herding behaviour is widespread in many contexts. However, Keynes' implication of herding behaviour being associated with risk avoidance strategies as a hard-wired instinct of survival is not directly tested in modern world tasks, where survival is a given.

In modern world tasks in the financial domain, the previous literature has provided a wide range of explanations for why people herd, both in financial markets in general and in the specific context of FAFs (for reviews, see, e.g. Baddeley, 2010; Bikhchandani & Sharma, 2000; Cont & Bouchaud, 2000; Devenow & Welch, 1996). The explanations can be divided roughly into three groups differing in terms of the degree to which herding is perceived as being rational, bounded rational, or irrational.

One group of explanations holds as a central tenet that market participants are strictly rational and make decisions that are rational, purposive, and oriented towards a goal of economic welfare. This group includes economists who have generally presented herding as a rational result of a sophisticated logic that follows the principles of Bayesian updating and information cascades (Anderson & Holt, 1997; Bikhchandani et al., 1998; Scharfstein & Stein, 1990). Consistent with a rational view on herding, financial analysts might herd because they rely on other analysts to have superior information or superior skills (e.g. Clement & Tse, 2005; Guttman, 2010; Kim et al., 2011). This tendency is sometimes referred to as informational herding and is broadly considered rational.

---

[3] However, prospect theory (Kahneman & Tversky, 1979) suggests a general tendency of people to behave (more) riskily in the loss domain.

A second group of explanations is in agreement with Simon's (1955, 1979) idea of bounded rationality. Here, the view that individuals act in ways that are not always optimal in an economic sense but may be rational in a wider context is acknowledged. This could be because they seek other goals, such as pleasure or satisfaction, but these cannot be achieved because of cognitive constraints (Stanovich, 1999). In the grey area between perfect rationality and this view entailing bounded rationality lies the concept referred to as strategic herding. Here, external factors are argued to cause analysts to have strategic incentives, such as career concerns, which in some cases make a forecast close to the consensus the better choice (Hong et al., 2000). Asch (1951, 1952, 1956) relied on a less narrow view of rationality, which is the most pronounced in the psychology literature. Asch's series of experiments showed that most people conform to a group, even when they realize the group's answer is wrong. Asch argued that, even when the choice to follow a group creates mental conflicts in the individual (because the individual realizes that the group's answer is wrong), this conflict is resolved by reasoning that the choice of a majority should be correct. Furthermore, this argument is coherent with Festinger's (1954) social comparison theory, which suggests that people generally compare themselves to others to evaluate themselves, especially in situations with high uncertainty. Many years later, this view is still significant in the literature and Turner's (1991) work on social influences has argued that numbers are less influential than the consensus. Similar ideas have been reflected in the literature on financial markets, in which Shiller (1995) argued that, even after people make all the correct calculations along the way, social influence alone can result in their making the wrong choices at the end by herding towards the group.

While Asch, Festinger, Turner, and Shiller explained herding as a phenomenon that can be seen as (bounded) rational and, in fact, optimal if one considers non-economic social goals, recent experimental studies have suggested explanations that can be considered suboptimal. These explanations have been argued to result from cognitive biases, causing the information value of consensus to be weighted too heavily (Libby & Rennekamp, 2012; Seybert & Bloomfield, 2009). Some studies have further sought to explain herding with reference to emotions, leading to heuristics (see, e.g. De Bondt, 1999, who touches upon this idea in the context of financial analysts' herding behaviour). According to this view, the fear of standing out from the crowd seems to dominate the benefit of staying true to one's own beliefs. In the literature on evolutionary biology, it has been argued that one reason for the extent of the fear of standing out could be that, historically, it was life-threatening to not be part of a group (e.g. Keynes, 1930/1982; Simon, 1990). This finding implies that herding is just as likely to be an intuitive reaction to uncertainty as it is to be the result of conscious choice. Thus, a consensus estimate in the context of financial forecasts is viewed not only as a strategic reference point or an aggregated information signal but also as a clue that one can intuitively follow to make a choice that feels safer.

This paper supports the view that no single theory, explanation, or rationality paradigm can cover all aspects of herding behaviour. Thus, we aim to provide a targeted contribution that complements the focus of prior literature by adding a tenet of possible explanations of herding behaviour more related to the intuitive System 1 then the elaborate System 2. Therefore, this paper focuses on a rather limited subset of explanations based on social psychology and, in particular, cognitive biases, as well as the notion that herding occurs partly as an intuitive reaction to feeling uncertain about the

decision one is about to make. As Baddeley, Burke, Schultz, & Tobler (2010, p.7, l. 9 fb.) argued, 'When people are influenced by social information then this may reflect an interaction between a deliberative learning process and a more instinctive, affective, emotional response'. While neuroscientific tools are probably needed to accurately distinguish the effect of instinctive response from conscious intentions, we argue in the following that individuals' risk tolerance is likely an important antecedent of unconscious herding behaviours.

2.1.1 Individual risk tolerance and herding behaviour

Risk tolerance is considered an individual characteristic that influences how people make decisions under (risk and) uncertainty in various contexts (Brockhaus, 1980). Although it is generally agreed that individuals dislike risk (e.g. Hogarth, 1975), the level of risk tolerance differs among individuals and hence the same stimulus can lead to different reactions by different individuals. Furthermore, how individuals perceive and react to risk appears to be a very ingrained personal trait sometimes attached to emotions (Loewenstein et al., 2001; Slovic, 1999). Some studies even suggest that risk tolerance is genetically influenced, because it varies between individuals but remains somewhat stable over time (e.g. Andersen et al., 2008) and even over generations (Zyphur et al., 2009).

In a setting with rich information (including management forecasts, prior earnings, or recent developments), various mental shortcuts, such as representativeness bias, the availability heuristic, as well as anchoring and adjustments (Tversky & Kahneman, 1974), can simultaneously affect the decision making of individuals to various degrees, depending on each individual's personal characteristics, such as risk tolerance. Research

on humans' herding behaviour in general (including research on herding in financial markets), shows widespread evidence that herding is at least partly a response to feeling insecure about how to process the presented stimuli (Baddeley, 2010). Given this aversion towards the unknown – described by Shefrin (2002) as fear of the unknown in relation to ambiguity – individuals are likely to use the availability heuristic, which entails relying intuitively on information that is more available to the mind, including (but not limited to) information perceived to be salient because it comes from a group to which one belongs (Gilovich et al., 2002; Kahneman, 2011; Tversky & Kahneman, 1974). This tendency suggests that individuals who most dislike uncertainty and risk are the most likely to herd, because they use the availability heuristic more and rely on information that is salient and thus easily available to the mind.

This idea is formally stated in the first hypothesis, as follows.

> *H1: Individuals' risk tolerance is positively associated with the boldness of their forecasts.*

2.2 Asymmetries caused by (bad/good) news

Some studies have observed asymmetric patterns in FAF behaviour created by different reactions to negative versus positive market information (e.g. Easterwood & Nutt, 1999). Studies relying on archival data (e.g. Basu, 1997, Elliot et al., 1995) found that financial analysts use information inadequately more often in market upturns than in market downturns. Along the same lines, other studies (e.g. Beyer, 2008; Conrad et al., 2006) concluded that financial analysts are more likely to downgrade when prices decrease than to upgrade when prices increase. This asymmetric behaviour has often been argued to

occur because management delays bad news relative to good news and, hence, a stronger reaction to bad news than to good news is rationally expected in FAFs, since the forecasts contain more information.

Nonetheless, we suggest that at least part of the explanation is that individuals process positive and negative news differently and consequently react differently. This assertion is based on research across various fields finding that 'bad' is stronger than 'good' and that people are usually unaware of this asymmetry within themselves (for a review, see Baumeister et al., 2001). For instance, experimental studies have found that negative information is more processed, weighted more strongly, and provokes larger amplitudes in brain responses compared to positive information (Abele, 1985; Anderson, 1965; Ito et al., 1998; Peeters & Czapinski, 1990). This asymmetry is observed even in (mundane) everyday events (Nezlek & Gable, 2001), in which the magnitude of the events seems to play only a modest role (Hajcak et al., 2006). These latter findings suggest that individuals' asymmetric reactions to negative information relative to positive information is unconsciously generated and perhaps evolutionarily adaptive (Baumeister et al., 2001). Along the same lines, Kahneman (2003) suggested that the brain decides quickly whether something should be perceived as good or bad, supporting the idea that intuition rather than more elaborate thinking is provoked.

Based on the aforementioned good/bad asymmetry, we expect that an individual will feel more exposed if underreacting to bad news in an earnings forecast context than if underreacting to good news. This is because overlooking obviously negative information about the future of a firm is likely to be intuitively perceived as careless behaviour, whereas overlooking positive information is likely perceived as being careful. Therefore,

people who receive bad news prior to a forecast revision are likely to perceive that they are exposed to more risk or uncertainty than people who receive good news. Thus, we expect that the type of news (good or bad) related to the company in question moderates the relation between individual risk tolerance and boldness in forecasts so that the difference in reactions between more and less risk tolerant individuals is more obvious when bad news is presented than when good news is presented. This idea forms the second hypothesis of this paper, as follows.

> *H2: The positive relation between risk tolerance and boldness in forecasts is stronger when bad news is presented prior to forecast revisions (compared to good news).*

3 Methodology

The hypotheses are tested on data collected via an experiment. This method allows us to provide exact stimuli and directly observe individuals' forecasting outcomes (for examples of when experiments are particularly beneficial in financial accounting, see, e.g. Libby et al., 2002).

3.1 Design and Tasks

The participants are given a two-step task of forecasting earnings per share (EPS) for a given firm (see Appendixes A to D for the experimental materials). In step 1, all of the participants receive the same information about the firm in focus and are asked to make a one-year-ahead EPS forecast. This forecast in step 1 is called F1 and forms the baseline. In step 2, the participants are told that two quarters have passed since their first forecast

and they must now revise the EPS forecast based on newly released information. This revised forecast in step 2 is called F2. The newly released information varies along two treatment dimensions: consensus versus no-consensus and bad news versus good news. In other words, half of the participants receive a consensus estimate while the other half, the control group, do not. Furthermore, half of the participants receive good news about the firm's outlook while the other half receive bad news. Thus, the participants are randomly split into one of the four groups, as illustrated in Figure 1.

(FIGURE 1)

FAFs are conducted under ambiguous conditions, in the sense that various types of information are available and must be processed. The aim is to reflect similar characteristics of ambiguity in the design of a forecasting task. Therefore, roughly two A4 pages with both qualitative and quantitative information are presented in step 1. Closely simulating a real-world task, we largely copy actual information for an existing listed firm[4]. The information contains various publicly available information, such as historical financial performance, strategy, a chart illustrating historical stock prices, expectations about the future of the industry, and a statement that the firm's management expects EPS in the range of 2.25–2.35 for year 1. In other words, the experiment contains a rich information set and thus trades off some simplicity in design for higher external validity.

Given the 2×2 design, four groups receive slightly different information in step 2, with some information in common. All the participants are informed that the first quarterly

---

[4] The firm's identity is hidden to avoid potential bias (e.g., representativeness heuristic or confirmation bias). No participant gave the impression of knowing the identity of the underlying firm, neither in the written answers nor in debriefings.

report for the firm was roughly as expected and that the second quarterly report was just released. Furthermore, some highlights from last year's second quarterly report are presented together with a statement from the CEO clarifying that the firm maintains its expectation for year 1, that is, an EPS of 2.25–2.35. This management estimate is kept constant for all participants, regardless of the treatment, to isolate the treatment effect, meaning that the groups vary only along two treatment dimensions: consensus versus no consensus and good news versus bad news. We seek to vary elements as little as possible, since even variations that should theoretically have no economic effect on the outcome of the person making the forecast sometimes have considerable influence (e.g. Andersson & Hellman, 2007).

3.1.1 Treatments

Consensus versus no consensus: Half of the participants receive a consensus estimate (2.14) and are told that this estimate is the result of a revision performed after the release of new information. The other half of the participants do not receive the consensus estimate and are merely a control group to establish the effect of the consensus estimate, that is, to obtain a measure of forecast boldness. Thus, hypothesis tests are only performed on the half of the participants who receive a consensus estimate, since this group is the only one able to herd towards the consensus estimate.

Good news versus bad news: All of the financial numbers in this year's second quarterly report are manipulated equally upwards (good news group) or downwards (bad news group) in absolute terms compared to last year's second quarterly report. Furthermore, the qualitative information (statement from the CEO) is either positive and indicates a

positive earnings outlook (good news group) or negative and indicates a negative earnings outlook (bad news group). Although the numerical information of the quarterly reports can be manipulated without substantial asymmetry, we recognize that perceptions of the written text (CEO statement) can be subjective and can therefore vary in perceived magnitude among the participants. To minimize this variation, we allow only a few words to differ between the two groups in the news treatment.

3.2 Participants

The participants are 289 business graduate students enrolled in a course on financial statement analysis and valuation. On average, the participants are 25 years old (with also a median of 25 years), 27% are female, 38% invest in stocks, and 67% have a part-time job (57% in investment/trading, accounting, or corporate finance). Libby et al. (2002) argued that students with sufficient knowledge of the task can and should be used as participants. Students in the financial course cited should be quite familiar with the task of forecasting EPS and with the biases generally present in financial markets.[5]

It could still be argued, however, that students are less sophisticated users of financial information than professional analysts. Therefore, students are likely to trust the consensus estimate more if it is framed as having been conducted by professional analysts, because students likely expect professionals to have superior skills and information. This assumption is consistent with findings in social psychology that

---

[5] See Koonce and Mercer (2005) for further discussions of students as surrogates for sophisticated participants. See also Whitecotton (1996), in which professional analysts are found to have greater optimistic bias than MBA students.

individuals are more likely to conform to the behaviour of a group if it is viewed as more skilled (Bandura, 1965; Festinger, 1954). We strive to alleviate this potential issue by framing the information so that the participants perceive the consensus estimate as coming from 'colleagues', that is, people at the same level as themselves. This choice also aligns with the theoretical perspective of herding towards members of one's group. Untabulated results of tests of the data collected indicate that the students most akin to novices (based on relevant courses and experience) do not forecast significantly more closely to the consensus forecast than the more skilled students. This finding indicates that, even if our participants do perceive the consensus as coming from people with superior forecasting skills, they do not weight it more heavily than those students who should consider themselves more akin to professionals. Nonetheless, we recognize that at least some of the students might still consider the consensus estimate to be the work of professionals but we believe that our approach minimizes such connotations.

In an ideal world, we would have used professional analysts as participants in the experiment. However, with limited access to professional analysts, we chose to use a large number of students who are expected to be adequate for two reasons. First, their education and employment make them share certain characteristics with professional analysts. Second, the hypotheses investigated pertain to cognitive biases, which are likely to pertain to professionals as well as students. As a supplement to the robustness checks, we find that the hypotheses hold equally well for those students most akin to professionals (those investing in stocks and holding jobs within domains related to financial analysts) as for those least akin to them.

3.3 Procedure

The experiment was completed on paper in class. Only pens and calculators were allowed as aids. The participants were given limited time to complete the experiment.[6] Following standard guidelines in economic experiments, very little verbal introduction was provided prior to the actual experiment. In the experimental material, a more substantial written introduction was included (see Appendix D). At that point, it was again stipulated that participation was on a voluntary basis, together with other standard experiment rules where noncompliance would lead to exclusion from the experiment and losing the opportunity to receive monetary compensation.

Participation in the experiment included potential monetary gains in line with tradition in experimental economics. This reward is expected to enhance participants' efforts in the experiment and thus help capture natural behaviours. Inspired by Davis and Holt (1993, Chap. 2), we included a lottery offering few but relatively high stakes to motivate people to make an effort. The lottery chances are based on performance, the choices made throughout the experiment, and luck combined. Therefore, the participants were not given any direct monetary incentives to play strategically with their forecasts, that is, to stand out from the other participants or to follow management or the consensus estimate. Nor were they directly incentivized to be accurate with their forecasts, although the intention of the payment structure is to imply, without explicitly stating so, that accuracy can be beneficial.

---

[6] Pre-tests indicated how to set the time limit to balance the time pressure (as professional analysts also experience) but still obtain predominantly complete answers. Answers from pretests are excluded from the final sample.

We follow the experimental instructions of Libby and Rennekamp (2012, p. 208, footnote 13) and denote all earnings from the experiment in an experimental currency ('coins'), where the exchange rate (together with the choice of lottery) is unknown until the end of the experiment. Participating in the experiment offers a chance to earn 2,000 coins convertible into lottery tickets. All the lottery prizes are $286.

The week after executing the experiment, debriefings and lottery draws were (transparently) conducted in the classroom. In total, 10 participants won the lottery and each winner received $286, corresponding to approximately 62% of their average self-reported monthly disposable income ($395). Because Andersen et al. (2008) found risk tolerance to sometimes vary with personal finances, we checked but found no income effect.

At the debriefing, no one expressed any concerns about the specific calculation of coins; nor did anyone question whether the best solution to the forecasting task was also the most accurate.

3.4 Measurements and descriptive statistics

3.4.1 Boldness

According to Bernhard, Campello, & Kutsoati (2006, p. 658, l. 11 fb.), 'herding amounts to biasing a forecast away from an analyst's best estimate, toward the consensus forecast of earlier analysts; while anti-herding amounts to biasing a forecast away from that consensus'. We construct a measure of boldness (anti-herding) with which we aim to reflect how much each participant is affected by the consensus estimate given. First, for

each individual in the consensus group, we start by identifying all the participants in the control group who submitted a similar F1. We then compare F2 for each participant in the consensus group with the average F2 of the matched participants from the control group, who, apart from not receiving the consensus estimate, received identical information. This enables us to estimate the direct effect of the consensus estimate and thus allows for the recognition of boldness in forecasts.

We define similar forecasts as forecasts that differ by less than 0.10.[7] For example, if a participant in the consensus group stated an F1 of 2.10, this participant is matched with all participants from the control group who stated an F1 within the interval 2.00–2.20. We perform this step separately for each participant in the consensus group.

The measurement of boldness for each individual in the consensus group measures the degree to which $F2_{actual}$ is close to $F2_{pure}$ relative to the consensus, as follows:

Boldness = |(F2$_{actual}$ – Consensus)|/|(F2$_{pure}$ – Consensus)|

where

      F2 $_{actual}$ = actual revised forecast for each individual in the consensus group

      Consensus = consensus estimate of 2.14 given to the consensus group

F2 $_{pure}$ = average of the revised forecast for the matched participants in the control group

---

[7] The results generally persist if the interval is changed. We tried intervals of 0.05 and 0.15. The results are available on request.

Furthermore, if *Boldness > 1* or if $F2_{pure} < Consensus < F2_{actual}$ or $F2_{pure} > Consensus > F2_{actual}$, we assign the boldness measure a value of 1.

Because we cap the formulas at a value of one, the measure of boldness in forecasts can only have values between zero and one. If F2 for the consensus group is exactly the consensus (i.e. $F2_{actual} = Consensus$), the formula assigns a zero, reflecting the least bold forecasts (i.e. most herding). If the value of F2 for the consensus group moves away from that of the matched participants in the control group and towards consensus (i.e. $F2_{pure} < F2_{actual} < Consensus$ or $F2_{pure} > F2_{actual} > Consensus$), the boldness measure takes on a value between zero and one, depending on the extent to which it moves. Hence, a value of 0.5 reflects a value of F2 for the consensus group exactly between that of the matched participants in the control group and consensus. Furthermore, if the value of F2 for the consensus group moves away from that of the matched participants and away from the consensus estimate (i.e. $F2_{actual} < F2_{pure} < Consensus$ or $F2_{actual} > F2_{pure} > Consensus$), it is assigned a value of one by the formula, reflecting the boldest forecasts. Finally, a group of participants (42 observations) moves away from the value of F2 of their matched participants and states an F2 value on the other side of the consensus estimate (i.e. $F2_{pure} < Consensus < F2_{actual}$ or $F2_{pure} > Consensus > F2_{actual}$). In line with Clement and Tse (2005), we perceive this group as having extreme boldness[8] and thus assign these a value of one.

---

[8] If we do not assign these 42 observations a value of one but, instead, use the formula, the general results still hold but decrease slightly in strength.

A small number of participants (15 observations of the aforementioned 42) also move in a direction contradictory to the news signal received (i.e. $F2_{pure} < Consensus < F2_{actual}$ for bad news or $F2_{pure} > Consensus > F2_{actual}$ for good news). We are sceptical about whether their behaviour should be perceived as ignorant instead of extremely bold and therefore drop these 15 observations.[9]

Because our measure of boldness in forecasts requires at least one observation in the control group to have a difference of 0.1 or less from the F2 of each participant in the consensus group, we lose 17 observations that do not meet this requirement. In all, we havea total sample of 115 observations with a measure of boldness.

3.4.2 Risk tolerance

The measure of risk tolerance is based on answers to the following question from the Socio-Economic Panel Study (SOEP, e.g. Dohmen et al., 2005; Wagner et al., 2007):[10]

> How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please write a number between 1 and 10, where 1 means 'not willing to take risks' and 10 means 'very willing to take risks'.

---

[9] If these 15 observations are included, the general results still hold, except for the regression analyses in the robustness checks, where the difference between coefficients for *Risk Tolerance* in the two news groups becomes nonsignificant. This finding is further demonstrated and discussed in Section 4.

[10] The original measurement is an 11-point scale, on which zero is also an option. However, we reduced it to a 10-point scale (removing the zero option) to exclude a midpoint option.

This measure has been widely used to determine individual risk tolerance as a personal trait to predict behaviour (e.g. Caliendo et al., 2014; Weber et al., 2012). Although sometimes criticized by the economics literature for being too subjective and not directly incentivizing participants to respond according to their true preferences, this measurement was recently validated by a number of studies comparing risk measures across literatures and domains. The authors concluded that this measure is the best all-round predictor of risky behaviour across various domains (Dohmen et al., 2005, 2011; Wagner et al., 2007). In the experiment of Dohmen et al. (2011), the measure was compared to other popular risk measures, including revealed preference measures such as lottery choices, and was proven to outperform other measures in terms of predicting behaviour. These results were further supported by Lönnqvist et al. (2015). Additionally, Charness et al. (2013) argued that this measure might be more applicable in experimental settings to reflect real world behaviour than measures frequently used in economic experiments.

We drop one participant who did not respond to this question, for a total of 128 observations for risk tolerance (allowing a varying sample size among variables by not performing list-wise deletions). Descriptive statistics and correlations (including also the control variables used in various robustness checks) are shown in Table 1, Panel A.

(TABLE 1)

3.5 Manipulation and other checks

After the participants completed the main task, we checked the overall understanding of the experiment. A total of 78.2% of the participants wrote that the experiment was about behavioural finance, risk perception, expectations, decision making, or similar topics, reflecting good general attention and understanding. However, six participants indicated that they might have understood what theory and hypothesis the experiment is designed to test by writing keywords such as *influence from others* or *adjustments to news*. Rerunning the main analyses without these participants shows largely the same results (available on request), so we keep these observations to maintain the sample size.

3.5.1 Treatment effects

To test the effects of the treatments and to validate the measurement of boldness, we investigate how the participants respond to the consensus treatment. We generally expect the consensus group – those who in addition to the company information also receive a consensus estimate of 2.14 – revise their forecasts significantly closer to 2.14 than the control group. In agreement with our expectations, the results of a *t*-test support that participants in the consensus group revise their forecasts significantly more closely to the consensus estimate with a mean distance to consensus of 0.33, compared to a mean distance to consensus of 0.54 for the control group (*t*-stat. = 4.25, two tailed). Furthermore, we find that the absolute difference between the F2 value in the consensus group and that of their individually matched participants in the control group is significantly different from zero (*t*-stat. = 5.57). This finding suggests that the consensus

estimate causes a behavioural difference between the two group's F2 values. We therefore conclude that the consensus treatment provides additional information to the consensus group, as expected, and that the measurement of boldness in forecasts reliably captures the intended effect.

Furthermore, we want to test whether the characteristics of news (good vs. bad) received prior to the revision affect the participants' forecasting behaviour. In this test, the control group (the group that does not receive a consensus estimate) is excluded. As reported in Table 1, Panel B, the mean of F2 for the good (bad) news group is 2.15 (1.92), just above (slightly below) the disclosed consensus estimate. This difference is statistically significant (*t*-stat. = -2.25, two tailed), suggesting that the news treatment has the anticipated effect.

As expected, the measure of risk tolerance does not vary significantly with any treatment in the forecasting task, confirming the randomization effect. The results are available on request.

3.6 Analyses and results

As stated in H1, an individual's risk tolerance is expected to be positively associated with boldness in forecasts. This hypothesis is tested on the 115 consensus group participants for whom *Boldness* can be measured. To allow for *t*-tests, the participants are split into two groups based on the median value of *Risk Tolerance*, where a value greater than six is assigned a one (more risk tolerant) and zero otherwise (less risk tolerant).

(TABLE 2)

In Table 2, the results of the *t*-test (first column, total sample) show that the least risk-tolerant group has a mean *Boldness* value of 0.58, whereas the most risk-tolerant group has a mean *Boldness* value of 0.81. The two means are significantly different at the 1 % level (*t*-stat. = -3.14, two tailed). The results support H1.[11] Additional tests of sensitivity to this procedure of splitting the sample by the value of six on the risk tolerance scale largely reveal the same results. Specifically, splitting at a value of five yields the same results at a 5% significance level (*t*-stat. = -2.26), while splitting at a value of seven yields the same results at a 10% significance level (*t*-stat. = -1.78). Thus, the results support the hypothesis that individuals who are generally more willing to take risks are less likely to herd towards the consensus when making earnings forecasts.

Hypothesis H2 predicts that the type of news (good vs. bad) received prior to forecast revisions moderates the relation between individual risk tolerance and the boldness in forecasts. The relation is expected to be stronger for the bad news group than for the good news group. To test this assumption, the sample is split based on the news treatment and the tests of H1 are rerun. The results (second and third columns in Table 2, bad news and good news, respectively) show a significant difference in the degree of boldness between the forecasts for the least risk-tolerant group and those of the most risk-tolerant group in the subsample of participants receiving bad news prior to forecast revision (*t*-stat. = -3.76, two tailed). However, the significance disappears when the same test is conducted for the subsample receiving good news prior to forecast revision (*t*-stat. = -0.99, two tailed).

---

[11] The same results are found using nonparametric (Wilcoxon signed rank) tests.

These results are generally in agreement with H2, in that the relation between individual risk tolerance and the boldness in forecasts is more pronounced when bad news is received prior to forecast revision.

4 Robustness checks

In this section, we test the robustness of the main results using regression analyses to allow for control variables and to further understand whether the coefficients of the two news groups are statistically significantly different from each other. We also check whether the main results of our hypotheses are robust to alternative methods of measuring boldness and other methods of analysis. Finally, the main tests are rerun based on only a subsample, which allows for the subtraction of one dimension of confidence that could interact with individual risk tolerance. This step is to ensure that confidence is not driving the results (rather than risk tolerance).

4.1 Alternative measure of boldness

Studies relying on archival data cannot isolate the informational effect of a consensus estimate on individuals' forecasting behaviour as the experimental method in this study allows us to do. Therefore, a common way of measuring boldness in earnings forecasts in the FAF literature is by the ranked absolute distance from each individual's (revised) forecast to the consensus estimate (e.g. Hong et al., 2000). Inspired by these studies, we measure the distance of the revised forecast (F2) to the consensus estimate (2.14) for the consensus group and rank each distance relative to the other participants in the consensus group (i.e. the control group is excluded). This measure is called *Boldness2* and is treated as a continuous variable.

The *t*-tests related to hypothesis testing are rerun and all the results (available on request) still hold, suggesting support for both hypotheses, even with a less nuanced measurement of boldness in forecasts. In the next section, we include control variables and rerun regression analyses separately for both measurements of boldness. Descriptive statistics and correlations for *Boldness2* and all the control variables are included in Table 1.

4.2 Regression analyses

We further test the hypotheses using ordinary least squares regressions. In addition to allowing for control variables, this model allows for a more sophisticated statistical test of H2. The regression results are presented in Table 3.

4.2.1 Controls

Experience and skills can influence boldness in forecasts (Clarke & Subramanian, 2006; Jegadeesh & Kim, 2009; Jiang & Verardo, 2013; Leone & Wu, 2007). To control for aspects of experience and skills, we follow the use of proxies in prior literature on FAFs (e.g. Dearman & Shields, 2005) and include four variables: the first, *Course,* equals one if a participant already passed one or more accounting/finance courses at the graduate level and zero otherwise; the second variable, *Grade,* reflects the participant's grade point average from bachelor's studies on a five-point scale, with three being around the median for the year and a higher number reflecting a better grade point average; the third variable, *Invest*, takes the value of one if the participant invests in stocks and zero otherwise; and the fourth variable, *Job*, equals one for participants with a relevant part-time job and zero otherwise.

The variables *Boldness* (*Boldness2*) and *RiskTolerance* have a positive correlation of 20% (23%). None of the control variables are significantly correlated with any of the main variables.

4.2.2 Results

As stated in H1, we expect individual risk tolerance to be positively associated with boldness in forecasts. This hypothesis is tested on the 115 participants in the consensus group, using an ordinary least squares regression with *Boldness* as the dependent variable and *Risk Tolerance* as the main independent variable. Further, the four control variables described above are included. Hereafter, the test is rerun but including *Boldness2* as the dependent variable. Table 3, Panel A, shows the results of the following models:

Model 1a, $Boldness = \alpha_0 + \beta_{1\ RiskTolerance} + \beta_{2\ Job} + \beta_{3\ Invest} + \beta_{4\ Grade} + \beta_{5\ Course} + \varepsilon_i$

Model 1b, $Boldness = \alpha_0 + \beta_{1\ RiskTolerance} + \beta_{2\ News} + \beta_{1\ RiskTolerance}*_{News} + \beta_{2\ Job} + \beta_{3\ Invest} + \beta_{4\ Grade} + \beta_{5\ Course} + \varepsilon_i$

The full model nonsignificantly explains 8% of *Boldness* in the forecasting task ($F_{(5, 100)} = 1.80$, $p < 0.12$). The variable *Risk Tolerance* explains *Boldness* significantly with a coefficient of 0.05, suggesting that more risk-tolerant individuals are more likely to make bold forecasts (away from the consensus estimate) than are less risk-tolerant individuals.[12] This finding is in agreement with our predictions in H1. However, all the

---

[12] As often implied in the literature, males are generally more risk tolerant than females. Thus, as expected, additional analyses reveal that *Risk Tolerance* is significantly correlated with gender (22%) in our data as well. We therefore implement an interaction term between *Risk Tolerance* and gender in the regression model, but this interaction term turns out to be highly nonsignificant in explaining *Boldness* (coefficient = 0.01, *t*-stat. = 0.25, *p*-value = 0.80). Therefore, the relations between individual risk tolerance and gender do not seem to be driving the results.

controls are nonsignificant, suggesting that the model might improve if some variables were dropped. Using backwards selection, we apply different combinations but the other versions of the full model decrease R^2, whereas the significance of *Risk Tolerance* persists. Forward selection yields the same results. A variation inflation factor test reveals no sign of multicollinearity, with the highest (mean) variation inflation factor being 1.10 (1.04) and, thus, below the recommended ceiling of 10 (Kutner et al., 2004).[13]

(TABLE 3)

We rerun the regressions with *Boldness2* as the dependent variable. Table 3, Panel B, shows the results. The results still generally support H1; however, this model is statistically significant, unlike the model in which *Boldness* is included as the dependent variable. This finding suggests that we might improve the explanation of what determines the boldness in forecasts by including controls only when the measurement of boldness in forecasts cannot isolate the behaviour from other factors. For both measures of the dependent variable (*Boldness* and *Boldness2*), *Risk Tolerance* is only significant in the group receiving bad news. Although the difference between the coefficients for *Risk Tolerance* in the two subsamples (bad vs. good news) is significant for the model including *Boldness*, is it nonsignificant when *Boldness2* is included as the dependent variable. This finding suggests that the results of H2 are less robust, whereas the results for H1 seem very robust.

---

[13] All the main variables (but none of the controls) are treated as normal, based on a Shapiro–Wilk test for normality and a skewness and kurtosis test for normality. Furthermore, *Risk Tolerance* shows no sign of endogeneity based on a Durbin–Wu–Hausman test.

4.3 Rerunning the main tests on subsamples

4.3.1 Sophisticated sample

In the methodology section, we discuss the potential consequences of using students as substitutes for financial analysts. Although we would need a sample of professional analysts to conduct a direct test, we can still perform additional tests to address this potential concern further.

First, we test whether participants who herd more towards the consensus forecast have a tendency to be less sophisticated than the rest. If this is the case, then it is very likely that we would find the same difference between (sophisticated) students and professional analysts. This also addresses the question of whether our attempt to frame the consensus forecast in the experiment as a forecast made by equally skilled individuals is successful. If the students perceive the consensus forecast as having been made by professionals (with greater forecasting skills), we should expect the students to strongly follow the consensus forecast. However, only 20 participants in the total sample (less than 7%) made a forecast of 2.14, equal to the consensus forecast. Furthermore, we do not find any statistically significant difference between the skills and experience of this group compared to the rest of the sample, as measured by the four controls (listed in Section 3). Running a regression analysis with these four independent variables, we do not find any *t*-value greater than 0.92 for any of the variables.

Second, we rerun the tests (*t*-tests and regression analysis) of the main hypothesis on different subsamples representing the group of the most sophisticated students, again based on our four proxies for experience and skills. All the tests support the hypothesis

that individual risk tolerance is negatively related to herding behaviour. We test all 16 possible combinations of the sophisticated subsamples and find only one in which the significance decreases to the 10% level, which could partly be due to the considerably reduced sample size. Based on the discussion in section 3 and the results from the robustness checks addressed here,[14] we are confident that the use of graduate business students as participants is suitable for the purpose of this study.

4.3.2 Confidence

Prior studies have suggested that risk tolerance and confidence are overlapping constructs that can be difficult to separate (e.g. Loewenstein et al., 2001; Slovic, 1999). Furthermore, Moore et al. (1999) and Menkhoff et al. (2006) found that some aspects of confidence[15] affect herding behaviour in fund managers' trading decisions. We therefore retest the hypotheses on a subsample while controlling for a dimension of confidence sometimes referred to as the better-than-average effect (Svenson, 1981).

Immediately after the forecasting task is completed, the participants are asked to evaluate their own performance in the task on a five-point scale relative to how they expect others performed.[16] Inspired by the methods of Niederle and Vesterlund (2007), we rerun our tests but include only the 49 participants who evaluate their own performance as 'around average' (three on the five-point scale).

---

[14] The results are available on request.

[15] Referred to as the illusion of control and positive illusions.

[16] Since this task is performed immediately after the forecasting task, the participants' perceptions of their own performance could be affected by the treatment group to which they were assigned in the forecasting. We therefore test for this effect but find none on this measurement of better-than-average effects.

(TABLE 4)

The results from *t*-tests in Table 4 on the relevant subsample are roughly the same as for the full sample, supporting H1. When the sample is split according to the news treatment, the significance persists for the bad news group but disappears for the good news group, suggesting support for H2. We believe that this attempt to remove a potentially confounding dimension from the main results adds credibility to the findings from the primary analyses.

5 Conclusions

This study concentrates on decision making under uncertainty, which is comparable to the environment of financial analysts in the process of conducting (earnings) forecasts. The literature on FAFs is concerned with so-called biases, such as herding behaviours, since these can decrease the usefulness of forecasts by compromising accuracy. Because herding behaviour is a common phenomenon, improved understanding of how personal characteristics such as individual risk tolerance can impact it is important.

In this paper, we predict that less risk-tolerant individuals are more likely to herd towards consensus with their forecasts than more risk-tolerant individuals are. Furthermore, this relation is expected to be more pronounced when bad news is received prior to a forecast revision than when good news is. We rely on theories mainly in the behavioural and social psychology literature. We find support for our predictions by testing them based on a 2×2 between-subjects experimental setting using 289 business graduate students.

While other elements, such as informational herding (expecting others to have superior information or skills) or strategic herding (e.g. reputational advantages) are very likely to be at play in a real-world setting, this experiment holds such effects constant, at zero. Therefore, our results demonstrate that the herding behaviour of individuals conducting financial forecasts can also be connected to mental shortcuts and behavioural biases, leading to a more intuitive reaction based on personal characteristics such as individual risk tolerance. While prior literature focuses on explanations of herding that arise from elaborate thinking (System 2), such as in strategic herding, we focus on herding that stems from more intuitive behaviours (System 1) and thus add to the literature by broadening the perspective of herding behaviour in financial forecasts.

This paper contributes to a growing literature that generally suggests that personal characteristics, traits, and emotions affect decision making to a much greater extent than reflected in the literature. Thus, expanding the view beyond perfectly rational agents and relying more on bounded rationality and suboptimal decisions in the pursuit of understanding the herding behaviours of financial analysts could push the literature in the desired direction (e.g. the call for more research of Bradshaw, 2011). Therefore, future research should be concerned with the degree to which observed biases in FAFs (e.g. herding behaviour, optimism- or conservatism bias) at the individual level can arise from more unconscious mechanisms, such as intuition, cognitive biases, and emotions. Extension of our knowledge on this matter could serve as important input to improve the usefulness of FAFs as proxies for future movements in the stock market.

Furthermore, this paper speaks to brokerage houses and public authorities aiming to curtail herding. If intuition plays a more pronounced role than previously assumed,

analysts exhibiting herding behaviour are likely to continue on that path, even if external incentives change. Hence, more customized interventions directed at financial analysts to change their instinctive behaviours could be necessary if changes in behaviour are sought. In this respect, measures inspired by nudging theory and the related choice architecture literature designed to alter unconscious actions (Thaler et al., 2014) could prove more fruitful than measures aimed at changing behaviour assumed to follow individual analysts' private profit maximization.

Alternatively, given their prominence suggested in previous contributions and the inherency suggested and demonstrated in this study, biased forecasts could be accepted as something that is not possible to affect in predictable ways. The implications of such an attitude of surrender would be that measures should be taken not to change analyst behaviour but, rather, to provide investors with rich analyst-related information, enabling them to adjust and weigh in on the estimates of analysts on an informed basis. Similar to IQ and personality tests that companies use today to evaluate potential new employees, information about the individual risk tolerance of candidates for financial analyst positions could be helpful in adding a secondary but strategic dimension for recruiters.

The experimental methods in this paper have the advantage of controlling information and isolating herding behaviours from confounding factors and thus more directly associating herding behaviours with individual risk tolerance. Although we carefully designed the task to mirror the high information complexity that professional analysts must process, our design discounts the opportunity to directly observe and test the effects of strategic and reputational herding. In other words, we traded the opportunity to observe such effects for the greater internal validity of our results. We consider internal validity

pivotal because, untraditionally, we seek to investigate mechanisms on intuition that make herding prevail even when there is no direct incentive to herd. Future research is needed to broaden our understanding about the relation between herding due to circumstances (e.g. strategic herding) and herding due to personal characteristics (e.g. risk tolerance causing intuitive reactions to uncertainty). The effect that this paper finds by focusing on System 1 thinking could be a moderating effect in studies focusing on System 2. However, this remains unknown until future research can at least answer the questions of when and how a financial analyst determines that it is worth the effort of System 2 thinking to complete the task of earnings forecasting.

A clear limitation of this study is that the experimental setting relies on forecasts from a single-round setting (including two forecasts for each participant) without feedback on performance. It could be interesting to see whether the results persist in a setting with multiple rounds and immediate feedback, where learning effects are likely to be present. Nevertheless, it was assessed that the implementation of multiple rounds could bias the results through spillover effects, for example. Although we do not find any effects of our results on the level of experience of the participants, we cannot directly test whether a learning effect would decrease or increase individuals' tendency to rely on intuition. A decrease could occur because solving the task for the first time might demand more cognition than solving it for the second or third time. On the other hand, an increase could occur due to intuitive thinking increasing with experience of the task, because it can result in more comfort and hence decrease individuals' perceived need for more elaborate analyses. In the latter case, the results of this paper are likely more generalizable than expected, but future studies in this area are essential to make conclusions on this matter.

Appendix A

Forecasting Tas

On the next page you are asked to state your **one-year-ahead earnings per share (EPS) forecast** for a particular company. Today is year 0 and you are asked to forecast EPS for year 1.

EPS = Net Income / Number of Shares

The company is an actual listed company but in this experiment you will just know it as "The Firm".

Company info:

The Firm was founded in 1892 and operates today through 842 stores in the United States, and 157 stores in Canada, Europe, Asia, and Australia. The Firm is within the retail industry with clothes as the main product line. The Firm's products are considered to be medium-price. The sales are highly seasonal and usually peak in the spring and fall. Approximately 25 % of The Firm's total revenue comes from online sales. According to The Firm: *"The brand is our lifestyle, our focus—the value of having a great brand is far-reaching and cannot be overstated—it's a snowball effect".*

Industry info:

The industry is known by its intense competition between its many players which is characterized by very volatile earnings and volatile stock prices. Within the last five years The Firm has experienced many new competitors which can be of future threat according to The Firm itself: *"In light of the competitive challenges we face, we may not be able to compete successfully in the future. Further increase in competition could reduce our sales and harm our operating results and business".*

This figure illustrates the development in stock prices for The Firm in the last ten years:
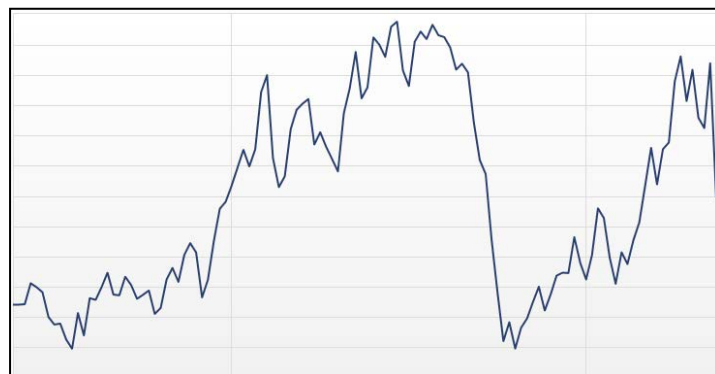
The Firm's Income Statement from the annual report.

(year 0 = current, year -3 = three years ago)

(in thousand $)	Year -3	Year -2	Year -1	Year 0
Revenue	3,468,777	4,158,058	4,510,805	4,116,897
Cost of Revenue	-1,251,348	-1,607,834	-1,694,096	-1,541,462
Gross Profit	2,217,429	2,550,224	2,816,709	2,575,435
Sales, General etc.	-1,980,249	-2,260,818	-2,435,069	-2,366,397
Non-Recurring Items	0	-68,022	-7,407	-128,215
Operating Income/EBIT	237,180	221,384	374,233	80,823
Interest Expense	-3,362	-3,577	-7,288	-7,546
Earnings Before Tax/EBT	233,818	217,807	366,945	73,277
Income Tax	-78,109	-74,669	-129,934	-18,649
Net Income	155,709	143,934	237,011	54,628
EPS ($)	1.67	1.47	2.85	1.45
Stock price ($)	45.64	58.70	49.61	46.20
Total Outstanding Shares (thousand, Year 0 and 1)	--------	--------	--------	37,693

Expectations for the industry:

The National Retail Foundation (NRF) says its retail sales forecast for year 1 predicts a growth of 4.1% (3.7 % year 0). Online sales are expected to increase between 9% and 12% this year.

The NRF President and CEO said: "Improvements in economic growth combined with positive expectations for continued consumer spending will put the retail industry in a relatively good place in year 1. Though headwinds from the looming debates about the debt ceiling, increased health care costs, and regulatory concerns still pose a risk for both consumers and retailers".

Expectations of The Firm:

The Firm partly explains the EPS level at year 0 by unusual expenses due to restructuring costs as well as an ongoing lawsuit about management ethics. The Firm expects little or no restructuring costs and no impact from the lawsuit in year 1. The Firm forecasts EPS for year 1 to be in the range of $2.25 to $2.35.

Appendix B

AVAILABLE INFORMATION TO THE PARTICIPANTS IN THE CONSENSUS-TREATMENT GROUP (ADDITIONALLY RECEIVING A CONSENSUS ESTIMATE) THAT RECEIVED BAD NEWS PRIOR TO STATING THEIR REVISED FORECAST (F2)[17]

New information in the market

The first quarterly report for year 1 was roughly as expected. The second quarterly report for year 1, which just got released, reveals the following numbers.

Highlights from the second quarter report, year 0 and year 1, in thousand $:

	Year 0, Q2	Year 1, Q2
Revenue	945,698	890,605
Gross profit	604,122	552,956
Operating income/EBIT	19,165	17,493
Net Income	11,371	10,877

After the release of the second quarterly report the CEO of The Firm said in a press release: "Because of a continued challenging environment, our sales for the second quarter were somewhat below plan." However, The Firm does not revise their total expectations for year 1.

On the other hand, based on 14 estimates from your colleagues, expected EPS for The Firm at year 1 is downgraded to an average (consensus) of $2.14.

Please revise your EPS forecast for The Firm for year 1 (If you do not deem a revision necessary, please write your forecast from the prior task again).

---

[17] The items of information that vary across treatments are underlined (but there is no underlining in the original materials available to the participants).

Appendix C

AVAILABLE INFORMATION TO THE PARTICIPANTS IN THE CONTROL GROUP (NOT RECEIVING A CONSENSUS ESTIMATE) THAT RECEIVED GOOD NEWS PRIOR TO STATING THEIR REVISED FORECAST (F2)[18]

New information in the market

The first quarterly report for year 1 was roughly as expected. The second quarterly report for year 1, which just got released, reveals the following numbers.

Highlights from the second quarter report, year 0 and year 1, in thousand $:

	Year 0, Q2	Year 1, Q2
Revenue	945,698	1,000,791
Gross profit	604,122	655,288
Operating income/EBIT	19,165	20,837
Net Income	11,371	11,865

After the release of the second quarterly report the CEO of The Firm said in a press release: Despite a continued challenging environment, our sales for the second quarter were somewhat above plan. However, The Firm does not revise their total expectations for year 1.

---

[18] The items of information that vary across treatments are underlined (but there is no underlining in the original materials available to the participants).

Please revise your EPS forecast for The Firm for year 1 (If you do not deem a revision necessary, please write your forecast from the prior task again).

Appendix D

WRITTEN INSTRUCTIONS TO THE EXPERIMENT

In-class experiment

This experiment takes approx. **40 minutes** and all your answers will be treated anonymously.

You can earn real money from this experiment. **You have a possible pre-tax payoff of DKK 2,000.**

If you decide to participate you need to remain seated until the experiment is over in approx. 40 minutes. If you do not wish to participate please exit the room now.

Paper, pen and calculator (e.g. on cell-phone) are required. Please place everything else by the wall (bag, jacket, computer, books etc.).

The number attached to the front page is your identification number. Detach one snip with the number on and leave the other snip with the number on attached to the front page. You will have to show the number in order to get your earnings so make sure you can find it again.

Answer all questions in chronological order. You are **not allowed to communicate** in any way during the experiment or to look at others' answers. After giving an answer you are not allowed to flip back and change it. If these rules are not obeyed, you may be excluded from the experiment and are not allowed to receive any potential earnings from the experiment.

Please read instructions carefully before you make your answers!

References

Abele, A. (1985). Thinking about thinking: Causal, evaluative, and finalistic cognitions about social situations. *European Journal of Social Psychology, 15,* 315–332.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. (2008). Lost in state space: Are preferences stable? *International Economic Review*, 49, 1091–1112.

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Personality and Social Psychology, 2,* 1–9.

Anderson, L. R., & Holt, C. A. (1997). Information cascades in the laboratory. *The American Economic Review*, 87, 847–862.

Andersson, P., & Hellman, N. (2007). Does pro forma reporting bias analyst forecasts? *European Accounting Review, 16,* 277–298.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In: Guetzkow, H. (ed.), *Groups, leadership, and men: Research in human relations.* Oxford, UK: Carnegie Press.

Asch, S. E. (1952). *Social psychology.* New York, NY: Prentice-Hall.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*, 1–70.

Baddeley, M. (2010). Herding, social influence and economic decision-making: Socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 281–290.

Baddeley, M., Burke, C., Tobler, P., & Schultz, W. (2010). Impacts of personality on herding in financial decision-making. Working paper. Cambridge Working Papers in Economics (CWPE)

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology*, 1, 589–595.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics*, 24, 3–37.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5, 323–370.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics*, 80, 657–675.

Beyer, A. (2008). Financial analysts' forecast revisions and managers' reporting behavior. *Journal of Accounting and Economics*, 46, 334–348.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, 12, 151–170.

Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff Papers*, 47, 279–310.

Bradshaw, M. T. (2011). Analysts' forecasts: What do we know after decades of work? Working paper.

Brockhaus, R. H. (1980). Risk taking propensity of entrepreneurs. *Academy of Management Journal*, 23, 509–520.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the 'black box' of sell-side financial analysts. *Journal of Accounting Research*, 53, 1–47.

Caliendo, M., Fossen, F., & Kritikos, A. S. (2014). Personality characteristics and the decisions to become and stay self-employed. *Small Business Economics,* 42, 787–814

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization*, 87, 43–51.

Christensen, D. M., Mikhail, M. B., Walther, B. R., & Wellman, L. A. (2017). From K Street to Wall Street: Political connections and stock recommendations. *Accounting Review*, 92, 87–112.

Clarke, J., & Subramanian, A. (2006). Dynamic forecasting behavior by analysts: Theory and evidence. *Journal of Financial Economics*, 80, 81–113.

Clement, M. B., & Tse, S. Y. (2005). Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance*, 60, 307–341.

Conrad, J., Cornell, B., Landsman, W. R., & Rountree, B. R. (2006). How do analyst recommendations respond to major news? *Journal of Financial and Quantitative Analysis*, 41, 25–49.

Cont, R., & Bouchaud, J. P. (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4, 170–196.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton, NJ: Princeton University Press.

Dearman, D. T., & Shields, M. D. (2005). Avoiding accounting fixation: Determinants of cognitive adaptation to differences in accounting method. *Contemporary Accounting Research*, 22, 351–384.

De Bondt, W. F. (1999). Herding in analyst earnings forecasts: Evidence from the United Kingdom. *European Financial Management*, 5, 143–163.

De Bondt, W. F. M., & Forbes, W. P. (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. *Journal of Finance*, 71, 267–302.

De Bondt, W. F. M., & Forbes, W. P. (1999). Herding in analyst earnings forecasts: Evidence from the United Kingdom. *European Financial Management*, 5, 143–163.

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European Economic Review*, 40, 603–615.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2005). Individual risk attitudes: New evidence from a large, representative, experimentally validated survey. Institute for the Study of Labor, discussion paper 1730.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9, 522–550.

Easterwood, J. C., & Nutt, S. R. (1999). Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *The Journal of Finance*, 54, 1777–1797.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, 32, 53–88.

Elliot, J. A., Philbrick, D. R., & Wiedman, C. I. (1995). Evidence from archival data on the relation between security analysts' forecast errors and prior forecast revisions. *Contemporary Accounting Research*, 11, 919–938.

Emett, S. A., & Nelson, M. W. (2017). Reporting accounting changes and their multi-period effects. *Accounting, Organizations & Society*, 57, 52–72.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.

Galariotis, E. C., Rong, W., & Spyrou, S. I. (2015). Herding on fundamental information: A comparative study. *Journal of Banking & Finance*, 50, 589–598.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Guttman, I. (2010). The timing of analysts' earnings forecasts. *The Accounting Review*, 85, 513–545.

Hahn, T., & Song, M. (2013). Regulation Fair Disclosure and analysts' reliance on earnings announcements. *Journal of Accounting & Public Policy*, 32, 26–43.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological psychology*, 71, 148–154.

Hirshleifer, D., & Hong Teoh, S. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, 9, 25–66.

Hogarth, R. M. (1975). Cognitive processes and the assessment of subjective probability distributions. *Journal of the American Statistical Association*, 70, 271–289.

Holmes, P., Kallinterakis, V., & Ferreira, M. L. (2013). Herding in a concentrated market: a question of intent. *European Financial Management*, 19, 497–520.

Hong, H., Kubik, J. D., & Solomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics*, 1, 121–144.

Huang, R., Krishnan, M., Shon, J., & Zhou, P. (2017). Who herds? Who doesn't? Estimates of analysts' herding propensity in forecasting earnings. *Contemporary Accounting Research*, 34, 374–399.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75, 887–900.

Jegadeesh, N., & Kim, W. (2009). Do analysts herd? An analysis of recommendations and market reactions. *The Review of Financial Studies*, 23, 901–937.

Jiang, D., Kumar, A., & Law, K. N. (2016). Political contributions and analyst behavior. *Review of Accounting Studies*, 21, 37–88.

Jiang, H., & Verardo, M. (2013). Does herding behavior reveal skill? An analysis of mutual fund performance. Working paper.

Kadous, K., Mercer, M., & Thayer, J. (2009). Is there safety in numbers? The effects of forecast accuracy and forecast boldness on financial analysts' credibility with investors. *Contemporary Accounting Research*, 26, 933–968.Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American Psychologist*, 58, 697.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Keynes, J. M. (1982). A treatise on money. In: *The collected writings of John Maynard Keynes, Vols. V and VI*. London, UK: Cambridge University Press (original work published 1930).

Kim, W., & Jegadeesh, N. (2010). Do analysts herd? An analysis of recommendations and market reactions. *Review of Financial Studies*, 23, 901–937.

Kim, Y., Lobo, G. J., & Song, M. (2011). Analyst characteristics, timing of forecast revisions, and analyst forecasting ability. *Journal of Banking and Finance*, 35, 2158–2168.

Koonce, L., & Mercer, M. (2005). Using psychological theories in archival financial accounting research. *Journal of Accounting Literature*, 24, 175–214.

Kumar, A. (2010). Self-selection and the forecasting abilities of female equity analysts. *Journal of Accounting Research*, 48, 393–435.

Kutner, M., Nachtsheim, C., & Neter, J. (2004). *Applied linear regression models*. New York, NY: McGraw-Hill/Irwin.

Leece, R., & White, T. (2017). The effects of firms' information environment on analysts' herding behavior. *Review of Quantitative Finance & Accounting*, 48, 503–525.

Leone, A. J., & Wu, J. S. (2007). What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts. Working paper.

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27, 775–810.

Libby, R., & Rennekamp, K. (2012). Self-serving attribution bias, overconfidence, and the issuance of management forecasts. *Journal of Accounting Research*, 50, 197–231.

Lönnqvist, J. E., Verkasalo, M., Walkowitz, G., & Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison. *Journal of Economic Behavior and Organization*, 119, 254–266.

Loewenstein, G., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286.

Markov, S., & Tan, M. Y. (2006). Loss function asymmetry and forecast optimality: Evidence from individual analysts' forecasts. Working paper.

Menkhoff, L., Schmidt, U., & Brozynski, T. (2006). The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European Economic Review*, 50, 1753–1766.

Mensah, Y. M., & Yang, R. (2008). An empirical evaluation of analysts' herding behavior following Regulation Fair Disclosure. *Journal of Accounting & Public Policy*, 27, 317–338.

Moore, D. A., Kurtzberg, T. R., Fox, C. R., & Bazerman, M. H. (1999). Positive illusions and forecasting errors in mutual fund investment decisions. *Organizational Behavior and Human Decision Processes*, 79, 95–114.

Nezlek, J. B., & Gable, S. L. (2001). Depression as a moderator of relationships between positive daily events and day-to-day psychological adjustment. *Personality and Social Psychology Bulletin*, 27, 1692–1704.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122, 1067–1101.

Nolte, I., Nolte, S., & Vasios, M. (2014). Sell-side analysts' career concerns during banking stresses. *Journal of Banking & Finance*, 49, 424–441.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1, 33–60.

Ramnath, S., Rock, S., & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting*, 24, 34–75.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *American Economic Review*, 80, 465–479.

Seybert, N., & Bloomfield, R. (2009). Contagion of wishful thinking in markets. *Management Science*, 55, 738–751.

Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Cambridge, MA: Harvard Business School Press.Shiller, R. J. (1995). Conversation, information, and herd behavior. *The American Economic Review*, 85, 181–185.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99–118.

Simon, H. A. (1979). Rational decision making in business organizations. *The American Economic Review*, 69, 493–513.

Simon, H. A. (1990). A mechanism for social selection and successful altruism. *Science*, 250, 1665–1668.

Slovic, P. (1999). Trust, emotion, sex, politics, and science: Surveying the risk-assessment Battlefield. *Risk Analysis*, 19, 689–701.

Song, M., Kim, D., & Won, C. (2009) Earnings uncertainty and analyst forecast herding. *Asia-Pacific Journal of Financial Studies*, 38, 545–574.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47, 143–148.

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2014). Choice architecture. In: Shafir, E. (ed.), *The behavioral foundations of public policy*. Princeton, NJ: Princeton University Press.

Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies*, 7, 97–124.

Turner, J. C. (1991). *Social influence*. Milton Keynes, UK: Open University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German socio-economic panel study (SEOP) -- scope, evolution and enhancements. *Schmollers Jahrbuch: Zeitschrift fur Wirtschafts- und Sozialwissenschaften/Journal of Applied Social Science Studies*, 127, 139–169.

Weber, M., Weber, E. U., & Nosić, A. (2012). Who takes risks when and why: Determinants of changes in investor risk taking. *Review of Finance*, 17, 847–883.

Welch, I. (2000). Herding among security analysts. *Journal of Financial Economics*, 58, 369–396.

Whitecotton, S. M. (1996). The effects of experience and a decision aid on the slope, scatter, and bias of earnings forecasts. *Organizational Behavior and Human Decision Processes*, 66, 111–121.

Xue, H. (1992). Independent and affiliated analysts: Disciplining and herding. *Accounting Review*, 92, 243–267.

Youssef, M., & Rajhi, T. H. (2010). Does herding behavior affect analysts' earnings forecasts? A study of French listed firms. *International Journal of Economics and Finance*, 2, 125–136.

Zyphur, M. J., Narayanan, J., Arvey, R. D., & Alexander, G. J. (2009). The genetics of economic risk preferences. *Journal of Behavioral Decision Making*, 22, 367–377.

Figure 1

An Illustration of the Steps and how Information is Distributed between Participants in the Forecasting Task.
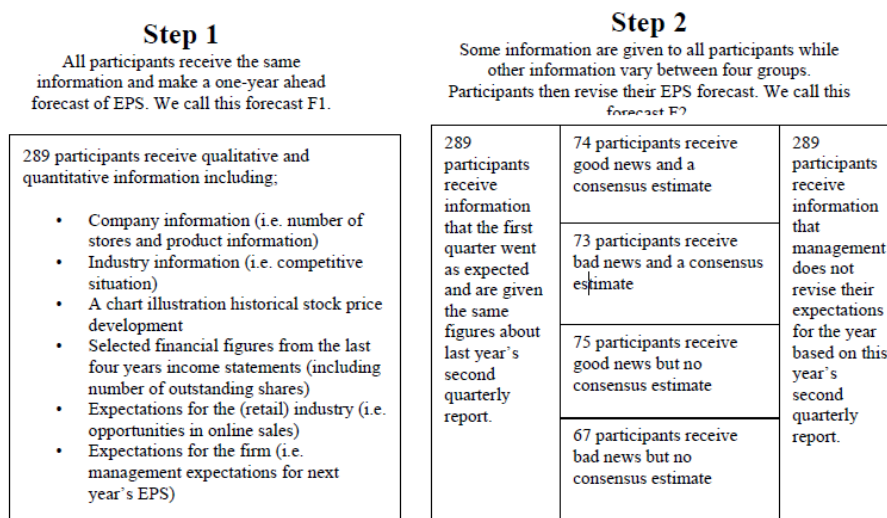

Step 1
All participants receive the same information and make a one-year ahead forecast of EPS. We call this forecast F1.

289 participants receive qualitative and quantitative information including;

- Company information (i.e. number of stores and product information)
- Industry information (i.e. competitive situation)
- A chart illustration historical stock price development
- Selected financial figures from the last four years income statements (including number of outstanding shares)
- Expectations for the (retail) industry (i.e. opportunities in online sales)
- Expectations for the firm (i.e. management expectations for next year's EPS)

Step 2
Some information are given to all participants while other information vary between four groups. Participants then revise their EPS forecast. We call this forecast F2.

289 participants receive information that the first quarter went as expected and are given the same figures about last year's second quarterly report.

74 participants receive good news and a consensus estimate

73 participants receive bad news and a consensus estimate

75 participants receive good news but no consensus estimate

67 participants receive bad news but no consensus estimate

289 participants receive information that management does not revise their expectations for the year based on this year's second quarterly report.

Table 1. Descriptive Statistics

Panel A. Overall descriptive statitstics and correlations

This table presents descriptive statistics and pairwise correlations (Pearson's- and Spearman's correlations generally reveal the same results). Panel A presents those for all observations while Panel B presents decriptives for the subsample. *Boldness* refers to the level of boldness (anti-herding) a participant acts with by their revised forecast (F2). The measure matches a participant from the treatment group (those receiving a consensus estimate in the experimental materials) with participants from the control group (those not receiving a consensus estimate in the experimental materials) with a similar baseline (stating the same first forecast (F1)), to asses the level of influence the consensus estimate has. *Boldness* is a continuous measure taking values between 0 and 1 with larger numbers meaning a greater level of boldness (less herding). *Risk Tolerance* refers to the participants' perceived willingness to take risks based on a 10-point Likert scale (*not willing to take risks* to *very willing to take risks*). Larger numbers means more willing to take risk. *Job* is a dummy variable where 1 is assigned if a participant has a relevant part-time job (i.e. within the areas of accounting or finance) and zero otherwise. *Invest* is a dummy variable where 1 is assigned if a participant actively invests in stocks and zero otherwise. *Grade* reflects the participants' grade point average from their bachelor degrees on a five-point scale where 3 reflects a grade point average around the median for the year. Larger numbers means better grades. *Course* is a dummy variable where 1 is assigned if a participant already passed a master-level accounting or finance course and zero otherwise. The first two variables (*Boldness* and *Risk Tolerance*) are considered the main variables of the study whereas all other variables are used as control variables for robustness checks. All variables include only the treatment group (approximately half the total sample) and n varies due to no list-wise deletion. Stars indicate significance on level: *** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$.

	n	Mean	Median	SD	Min	Max	1	2	3	4	5	6
1. Boldness	115	0.69	1.00	0.40	0.00	1.00						
2. Risk Tolerance	128	6.07	6.00	1.76	1.00	10.00	0.20**					
3. Boldness2	124	72.90	75.00	41.50	10.50	142	0.74***	0.23**				
4.	1	0.	1.00	0.	0.	1.	0.	0.	0.			

Job	29	76	0	53	00	00	10	03	12			
5. Invest	127	0.40	0.00	0.49	0.00	1.00	-0.08	0.04	-0.14	-0.01		
6. Grade	126	3.85	4.00	0.83	1.00	5.00	0.03	-0.05	-0.10	0.08	0.00	
7. Course	119	0.62	1.00	0.49	0.00	1.00	-0.07	0.01	0.03	-0.23**	-0.22**	-0.10

Panel B. Detailed descritive statistics for F1 and F2

This table presents descriptive statistics (n, mean, median and standard deviation) for F1 and F2 for the Control Group and the Treated Group, respectively. Furthermore, in *italics*, descriptive statistics for F2 and are reported for the sub-group of the treated group for which there was a near-match on F1 in the control group, giving rise to observations with a meaningful boldness measure. All statistics are reported for the total sample and seperately for the two groups in the news treatment, bad news and good news respectively.

		All				Bad news				Good news			
		n	Mean	Median	SD	n	Mean	Median	SD	n	Mean	Median	SD
Control Group	F1	140	2.50	2.30	1.13	66	2.47	2.24	1.10	74	2.53	2.30	1.71
	F2	137	2.09	2.00	1.06	65	2.09	2.00	0.10	72	2.10	1.83	1.19

Treated Group	F1	145	2.50	2.25	1.33	71	2.68	2.30	1.66	74	2.31	2.25	0.87
	F2	142	2.21	2.14	0.78	68	2.16	2.05	0.85	74	2.25	2.14	0.71
Treated Group observations with meaningful boldness measure	*F2*	*115*	*2.1*	*2.10*	*0.58*	*50*	*1.9*	*1.93*	*0.57*	*65*	*2.2*	*2.14*	*0.53*

Table 2. T-tests of differences in *Boldness* between the two groups of *Risk Tolerance*

Notes: Results of t-tests of differnces in means of *Boldness* between two groups of *Risk Tolerance*. *Boldness* and R*isk Tolerance* are defined in Table 1. *Less risk tolerant* and *more risk tolerant* refers to the variable *Risk Tolerance* split by median, where *more risk tolerant (less risk tolerant)* refers to the group with an above (below or at) median answer. Results are presented for the total sample and for the two groups in the news-treatment respectively. The total sample is split in two groups where the *Bad news (Good news)* group refers to those participants that are presented with a negative (positive) outlook about the future of the firm in the experimental materials prior to making a forecast revision (F2)). T-statistics (t-tests) and z-statistics (Mann-Whitney tests) are reported for two-tailed tests. Stars indicate significance on level: *** p < 0.001, ** p < 0.05, * p < 0.1.

	Total sample		Bad news		Good news	
	n	Mean	n	Mean	n	Mean
Less risk tolerant	61	0.58	28	0.47	33	0.68
More risk tolerant	54	0.81	22	0.86	32	0.77
t-stat.		-3.14 ***		-3.76 ***		-0.99

z-stat.　　　　-2.73 ***　　　　-3.14 ***　　　　-0.83

Table 3. Regression analyses of *Risk Tolerance* on *Boldness* including controls

This table presents regression results of Risk Tolerance on Boldness. All variables are defined in Table 1. Panel A shows regression results where *Boldness* is the dependent variable for the following two models; Model 1a: *Boldness* = α0 + β1 *RiskTolerance* + β2 *Job* + β3 *Invest* + β4 *Grade* + β5 *Course* + εi; Model 1b: *Boldness* = α0 + β1 *RiskTolerance* + β2 *News* + β1 *RiskTolerance*News* + β2 *Job* + β3 *Invest* + β4 *Grade* + β5 *Course* + εi. Panel B shows regression results where *Boldness2* is the dependent variable for the following two models; Model 2a: *Boldness2* = α0 + β1 *RiskTolerance* + β2 *Job* + β3 *Invest* + β4 *Grade* + β5 *Course* + εi; Model 2b: *Boldness2* = α0 + β1 *RiskTolerance* + β2 *News* + β1 *RiskTolerance*News* + β2 *Job* + β3 *Invest* + β4 *Grade* + β5 *Course* + εi. *Boldness2* is an alternative way to capture the level of boldness (anti-herding) of each participants revised forecast (F2). *Boldness2* refers to the absolute distance between a participants revised forecast (F2) and the consensus estimate given in the experimental materials (an estimated average of 2.14 based on the forecasts of 11 others) ranked relatively to the other participants. Larger numbers means a greater level of boldness (less herding). Results are presented for the total sample and for the two groups in the news-treatment respectively. The total sample is split in two groups where *Bad news* (*Good news*) refers to those participants that are presented with a negative (positive) outlook about the future of the firm in the experimental materials prior to making a forecast revision (F2)). Model statistics are reported seperately for each model at the end of the table. Tests of difference between coefficients (*Bad news* vs. *Good news*) are two-sided. Stars indicate significance on level: *** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$.

PANEL A: *Boldness* as dependent variable

	Model 1a		Model 1b		Bad news		Good news		Bad vs. Good	
	β	t	β	t	β	t	β	t	chi	z
Risk Tolerance	0.05 **	2.20	0.11 ***	2.89	0.11 ***	2.82	0.00	0.10	4.14 **	2.03 **
News			0.74 ***	2.2						

This article is protected by copyright. All rights reserved.

	β	t	β	t	β	t
Risk Tolerance * News			-2.12	**	-2.21	
Job	0.04	0.41	0.05	0.58	0.04	0.25
Invest	-0.10	-1.25	-0.06	-0.74	-0.08	-0.67
Grade	0.02	0.34	0.00	0.10	-0.01	-0.10
Course	-0.08	-1.01	-0.09	-0.14	-0.01	-0.13
n	106		106		46	
R^2	0.08		0.14		0.25	
F	1.80		2.00	*	2.58	**

PANEL B: *Boldness2* as dependent variable

	Model 2a		Model 2b		Bad news		Good news		Bad vs. Good	
	β	t	β	t	β	t	β	t	chi	z

Risk Tolerance	5.61	**	2.50	8.96	2.17	**	8.53	**	2.15	2.69	0.98	1.14	1.07
News				37.31	1.15								
Risk Tolerance * News				-5.92	-1.16								
Job	8.83		0.95	9.17	0.94		9.68		-0.57	20.29	1.83		
Invest	-14.93	*	-1.98	-13.11	-1.66	*	-14.86		-1.24	-9.16	-0.86		
Grade	-6.58		1.56	-7.11	1.70	*	-2.14		0.33	-9.27	1.82		
Course	-5.90		0.74	-5.94	0.76		-4.15		-0.36	-2.80	-0.25		
n	113			113			49			64			
R^2	0.12			0.14			0.19			0.15			
F	3.44	***		2.39	*		2.71	*		2.18	*		

Table 4. T-tests of differences in *Boldness* between the two groups of *Risk Tolerance* tested on a subsample

This table presents results of t-tests of differnces in means of *Boldness* between two groups of *Risk Tolerance. Boldness* and R*isk Tolerance* are defined in Table 1. *Less risk tolerant* and *more risk tolerant* refers to the variable *Risk Tolerance* split by the median, where *more risk tolerant (less risk tolerant)* refers to the group with an above (at or below) median answer. All tests are based on a subsample including only those participants that consider their own performance in the forecasting task around average compared to how they think the others did in the same task. To elicit the assesment of the participants own performance, they make answers on a five-point scale with the middlepoint (3) being that they consider their own performance around average. Thus, *Total sample* in this table includes only those 49 participants who consider their own performance around average. Results are presented for the total sample and for the two groups in the news-treatment respectively. The total sample is split in two groups where the *Bad news (Good news)* group refers to those participants that are presented with a negative (positive) outlook about the future of the firm in the experimental materials prior to making a forecast revision (F2)). T-statistics (t-tests) and z-statistics (Mann-Whitney tests) are reported for two-tailed tests. Stars indicate significance on level: *** p < 0.001, ** p < 0.05, * p < 0.1.

| | Total sample | | | Bad news | | | Good news | |
	n	Mean		n	Mean		n	Mean
Less risk tolerant	22	0.46		11	0.42		9	0.51
More risk tolerant	27	0.78		9	0.91		13	0.73
t-stat.		-2.96	***		-3.22	***		-0.41
z-stat.		-2.64	***		-2.36	**		-0.48