Essays in Behavioural Finance

Jonas Frey

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2023

To my parents,
for always supporting me on my academic journey

Acknowledgements

I could not have written this thesis without the support of many people. First and foremost, I would like to thank my advisors, Pedro Bordalo and Martin Schmalz, as well as my co-author Markus Dertwinkel-Kalt, who was like a third advisor to me, for their guidance and support. I am eternally grateful to have been allowed to learn how to become a researcher from you. I also want to thank my co-author Axel Ockenfels from whom I learned a lot during the work on our joined project.

I also profited tremendously from the academic community at Oxford. Special thanks go to professors Renée Adams, Bige Kahraman, Jiri Knesl, Alan Morrison, Ken Okamura and Mungo Wilson, who all gave me valuable feedback and guidance during my doctoral studies and to my fellow (former) DPhil students, especially Johan Cassel Pegelow, Laura Fritsch, Beata Gafka, Ryan Lee, Alexander Montag, Magdalena Plesa, Dean Ryu and Qirong Song, who both helped me improve my research and were the reason that I had a great time in Oxford.

Moreover, I want to acknowledge financial support from the Saïd Business School Foundation and Green Templeton College.

Finally, I want to thank my family and friends, who always supported me and gave me the energy to finish this thesis. I am particularly grateful to my parents, Torsten and Bettina Frey who made possible everything I did to become the person and researcher I am today. I also want to thank my partner Maja Geulig, for all the good times and her endless support. Last but not least, I would like to thank Kai Kortus for the good times and many interesting discussions.

Abstract

This thesis studies the role of contrast effects and biased expectations in financial decision-making and financial markets. The first study explores the role of skewness preferences in dynamic decision-making at the hands of salience theory. Previous research suggests that otherwise risk-averse people are willing to take risks if the outcome distribution is positively skewed. Salience theory can explain this by assuming that states with high contrasts between the outcomes attract attention and their probability is overestimated. Skewness preferences are particularly important in dynamic setups because these allow agents to endogenously create skewness through the choice of their stopping strategy. I extend salience theory to a dynamic setup and show that it predicts that agents will take gambles if the expected value is not too negative. Moreover, if they gamble they choose stopping strategies that yield positively skewed outcome distributions. These predictions differ both from expected utility theory and other behavioural models. I test the predictions experimentally and find broad support.

In the second study, I examine whether the earnings forecasts of analysts after a firm announces its earnings depend on the earnings surprises of companies that announced shortly before the firm. Evidence from a plethora of domains suggests that the interpretation of information depends on how it compares to contrasting information. Thus, the earnings of a given firm might look worse the better other firms perform. I find that positive earnings surprises of other firms make analysts revise their forecast of a firm's earnings upwards but, at the same time, make their forecast more pessimistic relative to the true earnings. This result is in line with a positive news channel in combination with a contrast effect channel of the other firms' earnings surprise on the analysts' forecasts.

In the third study, I develop a method to test if a given return predictor reflects mispricing rather than risk. Asset pricing research has uncovered hundreds of characteristics that can predict the cross-section of returns, but the nature of many of these remains elusive. If a predictor is linked to returns through risk, it should be unrelated to changes in the market's expectations about firm profits. Alternatively, return predictably can be explained by biased expectations. If a return predictor captures this form of mispricing, it should predict changes in expectations in addition to returns. I use the earnings forecasts of professional analysts as a proxy to the market's expectations and test for 173 return predictors if they can also predict forecast revisions. I find that around 40% of predictors can do so and, thus, reflect mispricing.

Contents

1 Introduction **1**

References 19

2 Optimal Stopping in a Dynamic Salience Model **25**

2.1 Introduction 27

2.2 Related Literature on Behavioral Stopping 31

2.3 A Dynamic Version of Salience Theory of Choice under Risk 33

2.3.1 Static Model 33

2.3.2 Dynamic Model 34

2.4 Stopping Behavior of a Naïve Salient Thinker 37

2.4.1 Motivating Example 37

2.4.2 Main Theoretical Result 37

2.4.3 Gambling an (Un)Fair Process with Stop-Loss and Take-Profit Strategies 40

2.5 An Experiment on Dynamic Gambling Behavior 44

2.5.1 Experimental Design 44

2.5.2 Implementation and Discussion of the Design 48

2.5.3 Experimental Predictions 53

2.6 Experimental Results on Dynamic Gambling Behavior 55

2.6.1 Data and Descriptive Statistics 55

2.6.2 Main Test of our Salience Predictions 57

2.6.3 On the Salience-Mechanism: Frequency and Direction of Strategy Adjustments 62

2.6.4 On the Disposition to Stop and the Role of Cognitive Skills 63

2.6.5 Limitations of the model in explaining the data 64

2.7 Discussion of Alternative Models 65

2.7.1 Expected Utility Theory 65

2.7.2 Cumulative Prospect Theory 65

2.7.3 Reference-Dependent Preferences without Probability Weighting . 66

2.7.4 Disappointment Aversion 68

2.8 Conclusion . . . 70
References . . . 71
2.A Proofs . . . 74
2.A.1 Preliminary Results on Arithmetic Brownian Motions . . . 74
2.A.2 Motivating Example . . . 81
2.A.3 Main Theoretical Result . . . 82
2.A.4 Additional Results on Stop-Loss and Take-Profit Strategies . . . 87
2.A.5 Salience and the Disposition Effect . . . 89
2.B Sophisticated Stopping Behavior Without Commitment . . . 90
2.B.1 Statement of the Result . . . 90
2.B.2 Proof of Proposition 4 for a Finite Expiration Date . . . 91
2.C Additional Details on the Main Experiment . . . 93
2.C.1 Parameters and Layout of the Static Choices . . . 93
2.C.2 Experimental Instructions . . . 93
2.D Salience Predictions on Static Skewness Preferences . . . 103
2.E Additional Figures . . . 106
2.F Stopping Behavior under Cumulative Prospect Theory . . . 109
2.G Salience-Driven Decoy Effects . . . 112
2.G.1 Overview of the Experimental Design and Results . . . 112
2.G.2 A More General Salience Model . . . 116
2.G.3 Salience Predictions on Decoy Effects . . . 117
2.G.4 Experimental Instructions . . . 119
2.H Stopping a Process Prematurely . . . 124

3 Do contrast effects influence earnings expectations? 127
3.1 Introduction . . . 129
3.2 Data . . . 134
3.2.1 Sources and Inclusion Criteria . . . 134
3.2.2 Calculation of Derived Variables . . . 135
3.2.3 Comparing Earnings per Share Across Companies . . . 137
3.2.4 Descriptive Statistics . . . 138
3.3 Methodology . . . 140
3.3.1 Questions & Hypothesis . . . 140
3.3.2 Regression Specification . . . 144
3.4 Expectations . . . 145
3.4.1 Main Result . . . 145
3.4.2 Persistence of the Effect . . . 149
3.4.3 The Link Between Errors and Revisions . . . 150
3.4.4 Heterogeneity in the Contrast Effect . . . 151

3.5 Returns . . . 155
3.6 Conclusion . . . 159
References . . . 160
3.A Robustness of the Effect of Contrast on Forecast Revisions and Errors . . . 164
3.B More Detailed Comparison with Hartzmark and Shue . . . 167
3.C Effect of the Own Surprise . . . 168
3.D Additional Return Results . . . 170

4 Which stock return predictors reflect mispricing? 174
4.1 Introduction . . . 176
4.2 Data & Methodology . . . 182
4.2.1 Motivation of the Empirical Specification . . . 182
4.2.2 Dataset and Variable Construction . . . 185
4.2.3 Multiple Hypothesis Testing . . . 189
4.2.4 Descriptive Statistics . . . 189
4.3 Mispricings . . . 191
4.3.1 Main Results . . . 191
4.3.2 Which Predictors are Mispricings? . . . 198
4.3.3 Detecting Spurious Predictors . . . 200
4.4 Build-up or Resolution? . . . 203
4.4.1 Main Results . . . 203
4.4.2 Which Predictors are Build-up and Which are Resolution? . . . 209
4.5 Comparison to other Classification Methods . . . 211
4.5.1 Mispricing Classification of the Original Authors . . . 211
4.5.2 Mispricing Classification of Holcblat et al. (2022) . . . 212
4.5.3 Build-up and Resolution Classification of van Binsbergen et al. (2023) 214
4.6 Conclusion . . . 216
References . . . 216
4.A Derivations of Equation 1 & 2 . . . 220
4.B Long-Term Growth Data . . . 222
4.C Additional Figures & Tables . . . 224
4.D Detailed Predictor Definitions . . . 234

5 Conclusion 245
References . . . 247

Chapter 1

Introduction

Researchers embarking on the task of understanding or modelling the decisions of an economic agent must understand two things about this agent: They need to know her preferences (what she wants to achieve) and her beliefs about the information relevant to the decision. This is a daunting task because humans often have a multitude of conflicting desires and two people who are presented with the same information rarely form the exact same beliefs. Early economists have acknowledged this complexity. In his *Theory of Moral Sentiment* Adam Smith discussed how behaviour is shaped both by rational deliberation and basic "passions" (Ashraf et al., 2005) and Keynes (1936) coined the term "Animal Spirits" to describe how behaviour is often impulsive rather than the result of careful optimisation. As economics came to rely more on mathematical models, formal theories of beliefs and preferences had to be adopted. Rather than attempting to model the "Animal Spirits", economists largely agreed to make two simplifying assumptions: agents' preferences can be represented by a (discounted) expected utility function (Bernoulli, 1738; Von Neumann and Morgenstern, 1947) and they form expectations rationally (Muth, 1961). These assumptions became the de-facto standard not only because of their mathematical convenience but also because they capture some important drivers of behaviour. Few people would argue that humans do not form beliefs about the consequences of potential decisions and then try to make the choice that leads to the preferred expected outcome, in line with expected utility theory. Similarly, it is natural to assume that their expectations are forward-looking and based on the information available to them, rather than just an extrapolation of past trends.

However, today a plethora of evidence from economics, psychology and neuroscience suggests that both expected utility maximisation and rational expectations are systematically violated and fail to capture central components of preferences and expectation formation. The challenge for behavioural models is, therefore, to retain the useful concepts of utility maximisation and forward-looking expectations, while at the same time describing observed behaviour more accurately than rational models. Finally, the models should explain behaviour through scientifically founded cognitive mechanisms and be applicable to different economic settings.

Behavioural economists have often turned to psychology research to explain why agents do not behave in line with standard economic theory. Psychologists have discovered a multitude of *cognitive biases*, that describe systematic mistakes in the reasoning of people. For example, when making probability judgements, people overweight information that easily comes to mind (*availability bias,* Tversky and Kahneman, 1973) and they conflate the *representativeness* of an outcome for its population with its probability (Kahneman and Tversky, 1972). Moreover, their beliefs are anchored and they adjust insufficiently when receiving new information (Tversky and Kahneman, 1974). Biases not only affect beliefs but also preferences. For example, a majority of people show *loss aversion*, that is they evaluate the outcome of a decision not in absolute terms but in terms of gains and losses relative to a reference point, whereby losses are overweighted (Tversky and Kahneman, 1981).

One cognitive mechanism that is of particular importance for the research in this thesis is the contrast effect. Evidence from a plethora of domains shows that humans do not process information in isolation, but relative to contrasting information that is available and that these relative comparisons distort absolute value judgements. Chapter 2 shows that the contrast effect can lead otherwise risk-averse individuals to behave as if they are risk-seeking in dynamic decision problems. Chapter 3 demonstrates that contrast effects lead to biases in the earnings expectations of professional stock market analysts. Chapter 4 shows that studying biases in expectations helps to understand the nature of stock return predictability. Before going into more detail on the research conducted in this thesis, I give an overview of the existing research on contrast effects and the

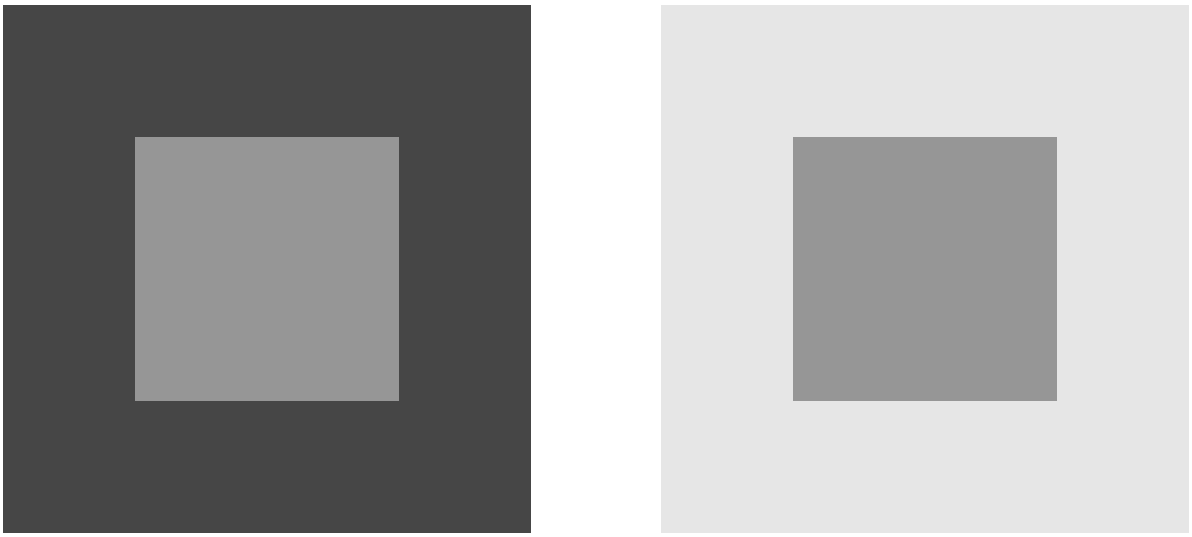
theoretical and empirical research on expectations in economics.

Contrast effects

In the context of visual perception, contrasts are important to focus attention on relevant objects. Humans are not able to fully interpret all the information in their field of vision. Instead, they focus on relevant objects. If an object visually stands out from its background it draws our attention and it is recognised. For instance, when driving a car, we notice the information on a red street sign because it stands out from the green background provided by the trees growing next to the road. At the same time, we do not notice if all the trees are of the same or of different kinds because the individual trees do not stand out. This effect has been demonstrated experimentally, for instance by Treisman and Gelade (1980). In their experiment, subjects were shown a set of objects in different colours. Some objects served as background distractions. They had two different shapes and two different colours. Subjects repeatedly saw a set of objects that either contained only the background objects or also a single object that stood out because it had a unique shape or colour. Their task was to determine as fast as possible whether a unique object was present. The study found that the response time of subjects was independent of how many background objects were presented, suggesting that subjects were immediately able to focus on the object that stood out and did not have to sequentially go over all the objects. Further evidence that contrast drives visual perception comes from the fact that computer models of visual perception based on the idea that contrasts attract attention perform similarly to humans in object recognition tasks (Itti et al., 1998).

However, this reliance on contrast in visual perception can also lead to biased information processing. For example, the two inner squares in Figure 1.1 are the same shade of grey, but the left one appears lighter because it is contrasted with a darker shade of grey whereas the right one is contrasted with a lighter shade of grey. Similarly, research demonstrates that objects tend to appear larger when they are contrasted with other objects that are small and/or close and smaller if they are contrasted with objects that are large and/or far away (Roberts et al., 2005).

Figure 1.1: Contrast Effect in Colour Perception



Self-created graphic based on a well-known phenomenon whose original discoverer is unclear. The two inner squares are the same shade of grey.

Contrast effects do not only affect visual perception but also the perception of temperature (Egeth et al., 1970), loudness (Holland and Lockhead, 1968) and weight (Sherif et al., 1958). Moreover, contrast effects have also been found in decisions made by non-human animals (Waite, 2001), as well as the firing rate of (animals') neurons (Wallis and Kennerley, 2011), suggesting that they are a fundamental component of perception and decision-making that stems directly from workings of the brain.

Given the evidence that contrasts shape our thinking on a deep level, researchers have studied if they also affect higher-level perception and decision-making. Initially, this research was mostly conducted by psychologists using laboratory experiments. Among others, this research showed that contrast effects influence perceptions of attractiveness (Cash et al., 1983; Kenrick and Gutierres, 1980) and crime severity (Kerr et al., 1982; Pepitone and DiNubile, 1976; Rodríguez and Blanco, 2016). Moreover, contrast affects influence loan quality judgements by professional auditors in a laboratory setting (O'Reilly et al., 2004). A meta-study of the evidence published in psychology journals up to 1999 also supports the existence of a contrast effect (Conway Dato-on and Dahlstrom, 2003).

More recently, field experiments demonstrated contrast effects in high-stakes real-world decision scenarios. Bhargava and Fisman (2014) find that participants in a speed

dating field experiment are less likely to declare romantic interest in someone if they rated the previous person they saw more favourably. Radbruch and Schiprowski (2020) find that the evaluations of candidates for a competitive scholarship are negatively related to the evaluation received by the previous candidate interviewed by the same assessor. Similar effects have been found in the decisions to approve R&D projects by senior employees of a large professional service firm (Criscuolo et al., 2021), the teaching evaluations of business school students (Cassar and Ko, 2023), and exam grades by university teachers (Goldbach et al., 2022). Moreover, people moving to the same city rent more expensive apartments if they previously lived in a city with high average rents compared to a city with cheap rents, even after controlling for potential confounds such as income, taste or beliefs about local housing costs (Simonsohn and Loewenstein, 2006).

There are also a few papers that provide evidence for contrast effects in the context of stock markets. Hartzmark and Shue (2018) find that the market reaction to a firm's earnings surprise is more positive the more negative the earnings surprises of firms announcing on the previous day were. He and Li (2020) split firms announcing on the same day by whether their earnings surprise is above or below yesterday's average earnings surprise and also find evidence of a contrast effect. Moreover, they show that the effect is larger for firms with high uncertainty.

Finally, contrast effects can also be used to explain common violations of expected utility theory in choice under risk. Specifically, they can explain that people show an affinity for risk with a positively skewed outcome distribution while at the same time disliking risks with a negatively skewed outcome distribution or being generally risk averse. Common examples of risks with positively skewed outcome distributions that are attractive to many humans are lottery games that offer a large positive payoff with a small probability while offering a small loss (the cost of the lottery ticket) with a large probability. Buying insurance for an expensive good is an example where agents show an aversion to negatively skewed risks. Under expected utility theory (EUT), agents are either risk-averse, risk-neutral or risk-seeking. Therefore EUT cannot explain why some people simultaneously buy insurance and play lottery games.

To explain this behaviour, Kahneman and Tversky (1979) and Tversky and Kahneman (1992) introduced prospect theory of choice under risk which assumes that people mechanically overweight small probabilities (1979) or probabilities of extreme outcomes (1992) when assessing the attractiveness of a prospect. More recently, Bordalo et al. (2012) proposed salience theory of choice under risk, in which the overweighting of probabilities is psychologically founded in the contrast effect. In this model, agents overweight states with extreme payoffs relative to the other states, because these states stand out due to their high contrast and attract attention. Salience theory can explain various puzzling behaviours in static choice under risk such as the Allais paradox (Allais, 1953). Moreover, it can also explain common asset pricing puzzles such as the equity premium puzzle and the fact that growth stocks underperform value stocks Bordalo et al. (2013).

Expectations

Today the rational expectations hypothesis is the standard assumption in economics and finance. It was originally formulated by Muth (1961) who assumed that the forecast of individual economic agents can deviate from the optimal forecast suggested by the underlying economic model, but only in a non-systematic way so that the aggregated forecast fit the model. A more recent formulation of rational expectations by Lucas (1972) and Lucas and Prescott (1971), assumes that expectations are also correct on an individual level in the sense that the expected distribution equals the actual distribution of the forecasted variable and this is now the more commonly used definition.

While the rational expectations hypothesis is the workhorse assumption today, it is predated both by the extrapolative expectations hypothesis, in which agents extrapolate past observed values into the future (Metzler, 1941), and the adaptive expectations hypothesis, in which agents update their expectations after observing a change in past data but anchor their new forecast to their previous level (Nerlove, 1958). As late as 1976 one of the earliest experimental papers on expectation formation considered extrapolative and adaptive expectations to be the leading hypothesis and did not even test the rational expectations hypothesis (Schmalensee, 1976). Subsequent experiments that did test the rational expectations hypothesis generally rejected it (Brennscheidt, 1993; Garner, 1982)

as cited in (Leitner and Leopold-Wildburger, 2011), (N. Harvey et al., 1994; Hey, 1994). These findings have been consistently confirmed by many other experiments (Leitner and Leopold-Wildburger, 2011). Thus, the shift towards rational expectations as the standard model of expectations was not driven by evidence. Instead, the reason for incorporating rational expectations into economic models was that they are forward-looking and thus allow the evaluation of different policies in a model. For instance, if a change in monetary policy fundamentally changes the money supply, it is unrealistic to assume that agents will not recognise this and continue to base their inflation expectations on the previous expectation level as they would under adaptive and extrapolative expectations.

Since forward-looking expectations are a crucial component of a model of expectation formation, any alternative to rational expectations must not only fit observed data but also retain this feature. One approach to improve the classical rational expectations model is to assume that agents are rational but face information frictions. Examples of such frictions are that not every agent receives new information immediately (Carroll, 2003; Mankiw and Reis, 2002), or that agents have limited attention and can only process parts of the available information (Gabaix, 2014). Moreover, models that relax the assumption of rationality and allow agents to be affected by biases or use simple heuristics have been proposed. In one class of models agents dynamically switch between different models of information processing based on their past performance and these models can include both rational and behavioural models (e.g. Anufriev and Hommes, 2012; Brock and Hommes, 1997), see also Hommes (2006) for a review. In other models, agents are overconfident and thus overestimate the precision of their private information (Daniel et al., 1998), see also Daniel and Hirshleifer (2015) for a review. Bordalo et al. (2018) propose a model of diagnostic expectations based on a formalisation of the representativeness heuristic (Kahneman and Tversky, 1972). A trait is said to be representative of a group if it is relatively more common in this group than in other groups. Applying this idea to expectations implies that agents update their expectations in the correct direction when they receive new information but they "overweight those future states whose likelihood increases the most in light of current news relative to what they know already." The reason is that the new information is most representative of these states. For example,

after a young firm has high earnings growth for a few years, investors think it will be "the next Google" because high earnings growth is representative of very successful firms. Nevertheless, most firms with high earnings growth never become superstar firms. The model also nests rational expectations as a special case and hence can be easily tested against the rational expectations hypothesis.

Early empirical tests of expectations relied mostly on experiments. While experiments are useful because they offer control over the data-generating process and the agents' information sets and thus allow for clean tests of models of expectation formation, they suffer from the issue that they are mostly conducted with student samples and have low stakes. Furthermore, especially in the early stages of experimental economics, they generally had small sample sizes. Therefore, it can be questioned whether these findings can be generalised to the behaviour of highly skilled professionals with high stakes whose decisions affect real economic outcomes.

Surveys of the beliefs of either households or professionals are an alternative to experiments. While they still suffer from the issue of being unincentivised, they eliminate the problem that students may form expectations differently than real economic actors. The concern of missing incentives is somewhat alleviated by the fact that survey expectations have been found to be linked to actual choices. For instance, investors' expectations are linked to mutual fund flows (Greenwood and Shleifer, 2014) and portfolio choice (Giglio et al., 2021b). Another advantage of surveys is that they can be merged with real-world outcome data to study how expectations are shaped by past and shape future economic outcomes.

Since the research in this thesis uses data on professional forecasters (as opposed to households), the following review will focus on those types of forecasts. Early studies using professional forecast survey data are La Porta (1996) who finds that stock markets analysts' long-term earnings forecasts overreact to news and La Porta et al. (1997), who show that the value premium can be explained by biased expectations which cause value stocks to have systematically more positive earnings surprises. These findings have recently been revisited under the lens of the diagnostic expectations model Bordalo et al. (2019, 2022) and the evidence supports overreaction to news in line with the model.

Moreover, the diagnostic expectations model can also explain errors in forecasts by professional forecasters for several different macroeconomic variables such as GDP, inflation and unemployment (Bordalo et al., 2020) as well as treasury bond yields (d'Arienzo, 2020).

Besides overreaction, there is also evidence for underreaction in professional forecasts. Coibion and Gorodnichenko (2012, 2015) study US inflation forecasts by professional forecasters and find evidence of information rigidity and Bouchaud et al. (2019) show that one year ahead earnings forecasts underreact to news.

Summaries of the research conducted in this thesis

Chapter 2 extends salience theory (Bordalo et al., 2012) to a dynamic setup and tests the model experimentally. As discussed above, a central feature of salience theory is that states with high contrast in their payoffs attract attention and are thus salient. Salient states then have their probabilities overweighted. The salience of a state is calculated using a salience function that needs to fulfil the following two properties.

1. Ordering. Let $x \geq y$. Then, for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$, we have

$$\sigma(x + \epsilon, y - \epsilon') > \sigma(x, y).$$

2. Diminishing sensitivity. For any $x > y$ and any $\epsilon > 0$, we have

$$\sigma(x + \epsilon, y + \epsilon) < \sigma(x, y).$$

These properties are sufficient if only positive payoffs are allowed. If we allow for negative payoffs a third property is necessary.

3. Reflection. For any $w, x, y, z \geq 0$, it holds that $\sigma(x, y) > \sigma(w, z)$ if and only if $\sigma(-x, -y) > \sigma(-w, -z)$.

The ordering property captures the idea of the contrast effect. All else equal, a state where the possible payoffs are further apart, i.e., have a higher contrast, is more salient. The diminishing sensitivity property captures the idea that relative contrasts

matter and the same absolute contrast is less salient at a higher level. This ensures that, for instance, the contrast between zero and one hundred stands out more and attracts more attention than the contrast between one million and one million and one hundred. Finally, the reflection property ensures that changing the sign of all payoffs of two states does not affect their salience ranking.

As discussed above, a central feature of the model is that it can explain a preference for positively skewed and an aversion to negatively skewed risk (*skewness preferences*) which has been documented in data on choices (e.g. Dertwinkel-Kalt and Köster, 2019). Intuitively, positive skewness means that the right tail of the distribution is longer than the left tail. There are several formal definitions of skewness. For example, it can be defined as the third standardised moment.

$$S[X] := \mathbb{E}\left[\left(\frac{X - \mathbb{E}[X]}{\sqrt{Var[X]}}\right)^3\right]. \tag{1.1}$$

To illustrate why salience theory generates skewness preferences it is easiest to consider binary lotteries. Consider an agent whose preferences are described by salience theory (a *salient thinker*), and who needs to choose between the following lotteries. For simplicity, let's assume that the salient thinker has a linear value function ($U(x) = x$).

$$L_1 : (\$20, 20\%; \$120; 80\%)$$
$$L_2 : (\$100, 100\%)$$

Note that L_2 is a degenerate lottery. It always has a payoff of \$100. The expected value of L_1 is also \$100 and it is negatively skewed. These lotteries yield two possible payoff states:

$$S_1 : (L_1 : \$120, L_2 : \$100, 80\%)$$
$$S_2 : (L_1 : \$20, L_2 : \$100, 20\%)$$

In this example S_2 is more salient. By the ordering property, salience increases in the difference (contrast) between the payoffs of the two lotteries, which is higher for S_2 and by the diminishing sensitivity property it decreases in the payoff level which is lower for S_2. Thus a salient thinker overweights the probability of S_2 which captures the downside of the lottery and since both the lottery and the safe option have the same expected value, she will choose the safe option.

Next, we consider the following lotteries. L_1^* is now positively skewed and its expected value is also \$100.

$$L_1^* : (\$180, 20\%; \$80; 80\%)$$
$$L_2 : (\$100, 100\%)$$

These lotteries lead to the following states:

$$S_1^* : (L_1^* : \$180, L_2 : \$100, 20\%)$$
$$S_2^* : (L_1^* : \$80, L_2 : \$100, 80\%)$$

In this case, it is not entirely clear which state is more salient. S_1^* has the higher contrast but also the higher level. The salience of the states thus depends on the strength of the contrast effect relative to the diminishing sensitivity effect. However, if we assume a sufficiently strong contrast effect, S_1^* will be more salient and the agent will choose L_1^* because the upside of the lottery is salient. Therefore, salience theory can explain skewness-seeking behaviour.

While skewness preferences are already crucial to explain behaviour in static choices, they are likely even more important for dynamic choices because those allow agents to create skewness through their stopping strategy. To illustrate this think of a casino that offers a gamble that is represented by a fair coin flip. If the coin comes up heads the gambler wins \$1 and if it comes up tails she loses \$1. This gamble is not inherently skewed. However, if the agent can take the gamble multiple times, she can, for instance, decide to keep gambling until she either lost \$10 or won \$100, which gives her a positively skewed payoff distribution. Barberis (2012) builds a model in which agents have preferences

described by Cumulative Prospect Theory (CPT) (Tversky and Kahneman, 1992) and can take symmetric gables with an expected value of zero multiple but not infinitely many times. He shows that agents with a wide range of CPT parameters start gambling despite being loss averse.

The preferences of these agents are time inconsistent. By further assuming that the agents are *naive*, that is unaware of their time inconsistency, Barberis shows that they often lose more money than they initially planned to. The reason is that after the agent incurred a small loss and reaches the point at which she initially planned to stop, she can formulate a new strategy where she again plans to play until she either lost a little more money or won a large amount. This strategy adjustment can occur multiple times, but due to the finite end date the agent eventually "runs out of skewness". Since the upside of a skewed strategy can only be reached by winning a considerable amount of times, as the final period approaches the possible upside gets smaller and smaller and the skewness of possible strategies decreases until, in the last period, only a symmetric risk is achievable. Ebert and Strack (2015) apply a model in which the agent has CPT preferences to a continuous time setup, which is relevant because many real-world dynamic decisions such as investing in the stock market or searching for an optimal product lack discrete time steps. The agent in the model needs to decide when to stop a stochastic process (specifically an Arithmetic Brownian Motion or ABM) without an expiration date. In this setup, a naive CPT agent without access to commitment will always gamble until bankruptcy regardless of negative the expected value of the underlying process is. In contrast, an agent who is *sophisticated* that is aware of her time inconsistency and lacks the ability to commit to her initial strategy will never start gambling regardless of how positive the expected value of doing so is (Ebert and Strack, 2018).

The reason for these extreme results is that the mechanical overweighting of the probabilities of extreme outcomes, and thus the skewness preferences, can get arbitrarily large in CPT. In contrast, probability weighting in salience theory depends on the payoffs of the states and is bounded. Thus salience theory is a natural candidate for a model that can not only explain choices under risk in static contexts but also make reasonable predictions for dynamic and continuous time contexts.

In Chapter 2, we first show that the extreme result of Ebert and Strack (2015) holds also with a finite expiration date. Deriving predictions for a finite time setup is important because it makes the model testable. We then derive a continuous time dynamic salience model and show that it predicts that agents might start gambling with an ABM that has a drift of zero or even a somewhat negative drift if their skewness preferences are sufficiently strong. However, regardless of the strength of their skewness preferences, there exists a threshold drift below which salient thinkers will not start to gamble.

If we restrict the strategies that a salient thinker can choose to *stop loss and take profit* strategies, we can analyse which types of strategies a salient thinker finds attractive. A *stop loss and take profit* is a strategy, where the salient thinker sets an upper and a lower bound and the process is stopped if it reaches either bound. If the salient thinker, is restricted to such strategies, she will always choose a strategy where the upper bound is further away from the current value of the process than the lower bound (a *loss exit strategy*) because in this case, the contrast of the state in which the process stops at the upper bound is larger than the contrast of the state where the process stops at the lower bound.

We test our model in an experiment in which agents have to stop a process resembling an ABM. We find that more than 90% of subjects start gambling a process with a drift of zero and fewer and fewer subjects are willing to gamble as the drift and thus the expected value of gambling becomes more negative. Moreover, conditional on starting to gamble subjects stop earlier for more negative drifts. Finally, if subjects gamble they are likely to choose *loss exit* strategies. For the median subject 73% of strategies chosen throughout the experiment are *loss exit* strategies.

In sum, Chapter 2 shows that, unlike CPT, salience theory makes reasonable predictions when applied to a dynamic and continuous context. This suggests that salience theory is a potential candidate for a unified model of choice under risk that incorporates skewness preferences.

Chapter 3 also studies contrast effects but in the context of beliefs rather than preferences. Despite the widespread evidence for contrast effects in perception, to the best of my knowledge, there has been no formal test of contrast effects in expectations so

far. Since contrasting information is widely available in contexts where expectations are formed, the existence of a contrast effect in expectation formation would have widespread consequences which motivates the research in Chapter 3. Moreover, showing that contrast effects can explain deviations both from the standard economic model of preferences as well as from the standard model of belief formation would support the ongoing trend in behavioural economics to find portable and psychologically founded mechanisms that can explain behaviour across a wide range of different contexts. This is crucial because in order to remain tractable, economic theory cannot rely on a large set of different cognitive biases that can each only explain behaviour in a few situations.

A setup previously studied by Hartzmark and Shue (2018) provides an ideal context to test for a contrast effect in expectations. As discussed above, Hartzmark and Shue find that the market's response to a firm's earnings surprise is more negative the more positive the earnings surprises of contrasting firms were. They interpret this as evidence of a "perceptual error rather than an expectational error" because it affects returns not at the time of the contrasting announcement but only after the firm's own earnings announcement. This rules out that the contrast surprise on the day prior to the announcement has an immediate effect on the expectations for the upcoming announcement. However, it does not rule out that contrast effects lead to biased expectations about future earnings announcements after the firm's announcement for the current quarter. Since any given quarterly earnings represent only a tiny fraction of a firm's value it is unlikely that there would be a strong effect of contrast surprises on returns if they only affected the perception of the currently announced earnings and not also expectations for future earnings.

I use earnings forecasts of professional analysts as a proxy for the market's expectations of a firm's future earnings and study if forecast revisions around an earnings announcement as well as forecast errors of the forecasts made after an announcement depend on the earnings surprise of contrasting firms. I use three types of firms for my contrast: firms that announced their earnings on the day before the firm announced, firms that announced on the same day but at an earlier time as the firm and firms that were covered by the same analyst and announced within 15 days before the firm. For the first two contrast, I require firms to be large, defined as having a market capitalisation

above the 90th percentile of stocks listed on the New York Stock Exchange. I do this because large firms receive more news coverage and are thus more likely to be in the mind of the analyst when she makes her forecast. Since a firm that the analyst recently covered herself is sure to be in her mind, I do not require the contrast firm to be large for this type of contrast.

I find that analyst forecast revisions are positively linked to all types of contrast surprises. However, forecast errors, defined as actual value minus forecasted value, are also positively linked to contrast surprises, which suggests that analysts are too pessimistic after good contrast surprises. A plausible interpretation for this result is that positive contrast surprises contain positive information about the firm's future prospects (e.g, about the state of the economy in general) to which analysts react, but they also provide a contrast for its own surprise, which makes it look less favourably and leads to more pessimism. I further show that subsequent forecast revisions are also positively linked to the contrast surprises, suggesting that the initial underreaction to these surprises gets corrected but that even forecasts made up to 75 days after the firm's announcement remain more pessimistic after a positive contrast surprise, suggesting that the upwards revisions of forecasts do not fully correct the bias. Finally, I study if the contrast effect in expectations can predict subsequent stock returns and find inconclusive evidence.

In Chapter 4, I also study the link between analyst earnings expectations and returns. Rather than focussing on a specific source of predictable returns and expectation biases as in Chapter 3, I use data on earnings expectations to understand the nature of stock market return predictability more broadly.

If markets are efficient, the only explanation for return predictability is non-diversifiable risk. If the marginal investor is risk averse, she will only hold riskier stocks if she is rewarded with higher returns. However, risks that are specific to individual assets can be eliminated through diversification and are thus not priced. This insight is captured by the canonical Capital Asset Pricing Model (CAPM) (Lintner, 1965; Sharpe, 1964) in which a stock's returns are driven by its covariance with the entire market. Under the assumptions of the model, this is the only priced risk factor. However, evidence suggests that other stock characteristics also systematically predict returns. For instance,

stocks with low market capitalisation earn higher returns than those which high market capitalisation and stocks with a high book-to-market value outperform those with a low book-to-market value (Fama and French, 1993). Moreover, there is momentum in stock returns: stocks that recently had more positive returns continue to earn higher returns (Jegadeesh and Titman, 1993). Today, hundreds of other characteristics have been found to predict returns (see e.g. C. R. Harvey et al., 2016). A common method to test if a characteristic predicts the cross-section of stock returns, is to sort stocks into percentiles (usually deciles) based on the characteristic and form portfolios out of extreme percentiles.

If one does not assume market efficiency, systematic mispricing is another explanation for return predictability. For a long time, many researchers have argued that mispricing cannot persist in the market because a few rational arbitrageurs would be sufficient to eliminate systematic mispricing. However, Kozak et al. (2018) show that even without limits to arbitrage, mispricing that is aligned with common factors can persist.

Given that the fundamental value of a stock reflects its discounted stream of future payments to the investor, mispricing can stem from two possible sources, it can either reflect biased expectations about the firms' future earnings or about the future required rate of return (i.e., the discount factor).

Understanding which return predictors are driven by risk and mispricing respectively, is important for several reasons. First, understanding how asset prices are formed and if they reflect fundamental values is at the core of academic asset pricing research. If prices do not reflect fundamental values this implies that capital is not allocated optimally which has negative implications for the real economy. Moreover, if a return predictor is driven by mispricing, investors may be able to achieve superior risk-adjusted returns by trading on this predictor, which is not the case for return predictability that is driven by risk.

Researchers who discover new return predictors often offer some interpretation on whether they think their return predictor is driven by mispricing or risk. Chen et al. (2022) study these original interpretations and find that "18% predictors are attributed to risk-based theory. 58% are attributed to mispricing and 24% have uncertain origins". However, they also find that predictors attributed to risk have no better out-of-sample

performance than those attributed to mispricing. This is surprising because mispricing-driven return predictability could disappear out of sample due to arbitrage trading, whereas risk factors should be persistent. Therefore, it is crucial to formally test the nature of each return predictor.

A few recent papers have developed such tests. Pukthuanthong et al. (2019) develop a protocol that classifies a return predictor as a risk factor if it is related to the principal components of the covariance matrix of returns, priced and has a reasonable Sharpe ratio. Notably, systematic mispricing that is aligned with common factors as discussed in Kozak et al. (2018) would pass this protocol and be considered a risk factor. Holcblat et al. (2022) develop a test based on the idea that if every risk-averse investor, regardless of her specific utility function, would rather invest in the long portfolio of a return predictor than in its short portfolio, then the return predictability cannot be explained by risk. Thus, they classify a return predictor as mispricing if they cannot reject this hypothesis. If they can reject it they consider a return predictor to be a potential risk factor. However, there can still be mispricings among the potential risk factors because a rejected null hypothesis only means there is some risk-averse utility function that rationalises the risk-return profile of the short portfolio but this might not be the utility function of the marginal or indeed any investor. Moreover, a lack of statistical power can cause the test to classify a risk factor as mispricing because the classification is based on failing to reject a null hypothesis. Gafka et al. (2021), decompose the abnormal returns that predictors earn into those earned on days on which macroeconomic news is announced to those they earn on other days. They argue that a return predictor is a risk factor if it earns most of its returns on announcement days. However, given the evidence that earnings expectations can both underreact and overreact to news, mispricing might also be linked to the announcement of news.

My method to test if a return predictor reflects mispricing relies on examining if there are systematic biases in expectations about firms' future earnings. If a predictor's abnormal returns are driven by biased expectations they should be accompanied by predictable changes in the market's earnings expectations. Specifically, the stocks that

predictably earn higher returns over a time period should also have predictably more positive revisions of earnings expectations over the same period.

I use analysts' earnings expectations as a proxy for the market's earnings expectations and obtain data on 173 significant return predictors from a dataset by Chen and Zimmermann (2021). The predictor data includes information on whether a given stock was part of a predictor's long or short portfolio in a given month as well as monthly stock returns. For each of the 173 predictors, I separately test if the stocks in the long portfolios have more positive earnings forecast revisions than those in the short portfolio and find significant evidence that this is the case for 40% of the predictors. Thus, these predictors can be classified as mispricing. The predictors in my sample can be grouped into categories that likely capture similar sources of underlying return predictability. I examine which categories contain a high share of mispricing predictors and find that momentum, volatility, accruals as well as predictors related to investment or investment growth are particularly likely to be mispricings.

My test differs from previous approaches by identifying mispricing that is driven by biased expectations, whereas existing tests group all types of mispricing together. Hence, it allows for a more detailed understanding of the type of mispricing in the market. This is important because different types of mispricing have different implications. For instance, mispricing driven by market frictions does not suggest attractive trading strategies whereas expectation bias driven mispricing does.

Moreover, my test is not based on the same return data that was used to establish return predictability in the first place. This is an advantage because recent evidence suggests that due to multiple hypothesis testing issues, many published return predictors may be false positives (Chordia et al., 2020; Giglio et al., 2021a; C. R. Harvey et al., 2016; Hou et al., 2020). If return predictability is just an in sample statistical artefact it is unlikely that analyst forecasts also show significant matching patterns by chance.

Using analyst expectations also comes with drawbacks: they are only a noisy proxy for the expectations of the marginal investor and they are updated infrequently and are thus not available for every firm at every point in time. Moreover, it is not possible to detect mispricings that are driven by biased expectations about the future required rate

of return or market frictions, using my data. Thus, my test can be seen as complementary to existing methods since it has orthogonal strengths and weaknesses.

As a second step, I further classify the mispricing predictors into those capturing the build-up and those capturing the resolution of mispricing. In a recent paper, van Binsbergen et al. (2023) argue that while most researchers implicitly assume that mispricing return predictors generally capture the convergence of prices back to their fundamental value (resolution of mispricing) a sizeable fraction actually captures the divergence of prices from their fundamental value (build-up of mispricing). Their test relies on calculating the fundamental value of stocks based on an assumption about the correct stochastic discount factor (SDF) and ex-post realised dividend data.

I provide an alternative and complementary approach based on forecast errors in analyst expectations. If a return preditor captures the resolution of mispricing, stocks that subsequently earn higher returns are initially undervalued relative to those with lower subsequent returns. If this undervaluation is driven by biased expectations, analyst forecasts should be more pessimistic as well. For predictors that capture the build-up of mispricing, stocks are initially priced correctly and the stocks with subsequent higher returns become overvalued. Therefore, analyst expectations should be too optimistic for those stocks after the abnormal returns were realised. Applying this test to my data, I can classify 45% of my mispricing predictors as resolution and 10% as build-up. The remaining predictors cannot be classified.

Overall, the results of Chapter 4 suggest that a large share of stock return predictability is driven by mispricing and that it is, therefore, crucial to formally test the nature of return predictors.

References

Allais, M. (1953). "Le Comportement de l'Homme Rationnel Devant Le Risque: Critique Des Postulats et Axiomes de l'Ecole Americaine". In: *Econometrica* 21.4, pp. 503–546. ISSN: 0012-9682. DOI: 10.2307/1907921.

Anufriev, Mikhail and Cars H. Hommes (2012). "Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments". In: *American Economic Journal: Microeconomics* 4.4, pp. 35–64. ISSN: 1945-7669. DOI: 10.1257/mic.4.4.35.

D'Arienzo, Daniele (2020). "Maturity Increasing Overreaction and Bond Market Puzzles". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3733056.

Ashraf, Nava, Colin F Camerer, and George Loewenstein (2005). "Adam Smith, Behavioral Economist". In: *Journal of Economic Perspectives* 19.3, pp. 131–145. ISSN: 0895-3309. DOI: 10.1257/089533005774357897.

Barberis, Nicholas (2012). "A Model of Casino Gambling". In: *Management Science* 58.1, pp. 35–51. ISSN: 0025-1909. DOI: 10.1287/mnsc.1110.1435.

Bernoulli, Daniel (1738). "Exposition of a New Theory on the Measurement of Risk". Trans. by Louise Sommer. In: *Econometrica* 22.1, pp. 23–36. ISSN: 0012-9682. DOI: 10.2307/1909829.

Bhargava, Saurabh and Ray Fisman (2014). "Contrast Effects in Sequential Decisions: Evidence from Speed Dating". In: *The Review of Economics and Statistics* 96.3, pp. 444–457. ISSN: 0034-6535. DOI: 10.1162/REST_a_00416.

Van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni (2023). "Dynamic Asset (Mis)Pricing: Build-up versus Resolution Anomalies". In: *Journal of Financial Economics* 147.2, pp. 406–431. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2022.11.005.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic Expectations and Stock Returns". In: *The Journal of Finance* 74.6, pp. 2839–2874. ISSN: 1540-6261. DOI: 10.1111/jofi.12833.

— (2022). *Belief Overreaction and Stock Market Puzzles*. preprint.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020). "Overreaction in Macroeconomic Expectations". In: *American Economic Review* 110.9, pp. 2748–2782. ISSN: 0002-8282. DOI: 10.1257/aer.20181219.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). "Salience Theory of Choice Under Risk". In: *The Quarterly Journal of Economics* 127.3, pp. 1243–1285. ISSN: 0033-5533. DOI: 10.1093/qje/qjs018.

— (2013). "Salience and Asset Prices". In: *American Economic Review* 103.3, pp. 623–628. ISSN: 0002-8282. DOI: 10.1257/aer.103.3.623.

— (2018). "Diagnostic Expectations and Credit Cycles". In: *The Journal of Finance* 73.1, pp. 199–227. ISSN: 1540-6261. DOI: 10.1111/jofi.12586.

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar (2019). "Sticky Expectations and the Profitability Anomaly". In: *The Journal of Finance* 74.2, pp. 639–674. ISSN: 1540-6261. DOI: 10.1111/jofi.12734.

Brennscheidt, Gunnar (1993). *Predictive Behavior*. Red. by G. Fandel and W. Trockel. Vol. 403. Lecture Notes in Economics and Mathematical Systems. Berlin, Heidelberg: Springer. ISBN: 978-3-540-57311-1 978-3-642-95718-5. DOI: 10.1007/978-3-642-95718-5.

Brock, William A. and Cars H. Hommes (1997). "A Rational Route to Randomness". In: *Econometrica* 65.5, pp. 1059–1095. ISSN: 0012-9682. DOI: 10.2307/2171879.

Carroll, Christopher D. (2003). "Macroeconomic Expectations of Households and Professional Forecasters". In: *The Quarterly Journal of Economics* 118.1, pp. 269–298. ISSN: 0033-5533. DOI: 10.1162/00335530360535207.

Cash, Thomas F., Diane Walker Cash, and Jonathan W. Butters (1983). ""Mirror, Mirror, on the Wall...?": Contrast Effects and Self-Evaluations of Physical Attractiveness". In: *Personality and Social Psychology Bulletin* 9.3, pp. 351–358. ISSN: 0146-1672. DOI: 10.1177/0146167283093004.

Cassar, Gavin and Taeho Ko (2023). "Peer Effects in Subjective Performance Evaluation". In: *Contemporary Accounting Research* n/a.n/a. ISSN: 1911-3846. DOI: 10.1111/1911-3846.12876.

Chen, Andrew Y., Alejandro Lopez-Lira, and Tom Zimmermann (2022). *Peer-Reviewed Theory Does Not Help Predict the Cross-section of Stock Returns*. DOI: 10.2139/ssrn.4308069. URL: https://papers.ssrn.com/abstract=4308069 (visited on 01/20/2023). preprint.

Chen, Andrew Y. and Tom Zimmermann (2021). *Open Source Cross-Sectional Asset Pricing*. SSRN Scholarly Paper 3604626. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3604626.

Chordia, Tarun, Amit Goyal, and Alessio Saretto (2020). "Anomalies and False Rejections". In: *The Review of Financial Studies* 33.5, pp. 2134–2179. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa018.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What Can Survey Forecasts Tell Us about Information Rigidities?" In: *Journal of Political Economy* 120.1, pp. 116–159. ISSN: 0022-3808. DOI: 10.1086/665662.

— (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". In: *American Economic Review* 105.8, pp. 2644–2678. ISSN: 0002-8282. DOI: 10.1257/aer.20110306.

Conway Dato-on, Mary and Robert Dahlstrom (2003). "A Meta-Analytic Investigation of Contrast Effects in Decision Making". In: *Psychology & Marketing* 20.8, pp. 707–731. ISSN: 1520-6793. DOI: 10.1002/mar.10093.

Criscuolo, Paola, Linus Dahlander, Thorsten Grohsjean, and Ammon Salter (2021). "The Sequence Effect in Panel Decisions: Evidence from the Evaluation of Research and Development Projects". In: *Organization Science* 32.4, pp. 987–1008. ISSN: 1047-7039. DOI: 10.1287/orsc.2020.1413.

Daniel, Kent and David Hirshleifer (2015). "Overconfident Investors, Predictable Returns, and Excessive Trading ". In: *Journal of Economic Perspectives* 29.4, pp. 61–88. ISSN: 0895-3309. DOI: 10.1257/jep.29.4.61.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam (1998). "Investor Psychology and Security Market Under and Overreactions". In: *Journal of Finance* 53.6, pp. 1839–1885. ISSN: 0022-1082. DOI: 10.1111/0022-1082.00077.

Dertwinkel-Kalt, Markus and Mats Köster (2019). "Salience and Skewness Preferences". In: *Journal of the European Economic Association*. DOI: 10.1093/jeea/jvz035.

Ebert, Sebastian and Philipp Strack (2015). "Until the Bitter End: On Prospect Theory in a Dynamic Context". In: *The American Economic Review* 105.4, pp. 1618–1633. ISSN: 0002-8282.

— (2018). *Never, Ever Getting Started: On Prospect Theory Without Commitment*. SSRN Scholarly Paper ID 2765550. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.2765550.

Egeth, Howard E., Arthur S. Kamlet, and Russell A. Bell (1970). "The Reversal of Classical Contrast in Temperature Perception". In: *Psychonomic Science* 19.2, pp. 96–96. ISSN: 0033-3131. DOI: 10.3758/BF03337438.

Fama, Eugene F. and Kenneth R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds". In: *Journal of Financial Economics* 33.1, pp. 3–56. ISSN: 0304-405X. DOI: 10.1016/0304-405X(93)90023-5.

Gabaix, Xavier (2014). "A Sparsity-Based Model of Bounded Rationality". In: *The Quarterly Journal of Economics* 129.4, p. 1661. ISSN: 0033-5533. DOI: 10.1093/qje/qju024.

Gafka, Beata, Pavel G. Savor, and Mungo Ivor Wilson (2021). *Sources of Return Predictability.* DOI: 10.2139/ssrn.3937343. URL: https://papers.ssrn.com/abstract=3937343 (visited on 06/08/2023). preprint.

Garner, C. Alan (1982). "Experimental Evidence on the Rationality of Intuitive Forecaster". In: *Research in Experimental Economics.* Greenwich. Conn, JAI Press.

Giglio, Stefano, Yuan Liao, and Dacheng Xiu (2021a). "Thousands of Alpha Tests". In: *The Review of Financial Studies* 34.7, pp. 3456–3496. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa111.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021b). "Five Facts about Beliefs and Portfolios". In: *American Economic Review* 111.5, pp. 1481–1522. ISSN: 0002-8282. DOI: 10.1257/aer.20200243.

Goldbach, Carina, Jörn Sickmann, and Thomas Pitz (2022). "Sequential Decision Bias Evidence from Grading Exams". In: *Applied Economics* 54.32, pp. 3727–3739. ISSN: 0003-6846. DOI: 10.1080/00036846.2021.1976390.

Greenwood, Robin Marc and Andrei Shleifer (2014). "Expectations of Returns and Expected Returns". In: *Review of Financial Studies* 27.3, pp. 714–746. ISSN: 0893-9454. DOI: 10.1093/rfs/hht082.

Hartzmark, Samuel M. and Kelly Shue (2018). "A Tough Act to Follow: Contrast Effects in Financial Markets". In: *The Journal of Finance* 73.4, pp. 1567–1613. ISSN: 1540-6261. DOI: 10.1111/jofi.12685.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (2016). " and the Cross-Section of Expected Returns". In: *The Review of Financial Studies* 29.1, pp. 5–68. ISSN: 0893-9454. DOI: 10.1093/rfs/hhv059.

Harvey, Nigel, Fergus Bolger, and Alastair McClelland (1994). "On the Nature of Expectations". In: *British Journal of Psychology* 85.2, pp. 203–229. ISSN: 2044-8295. DOI: 10.1111/j.2044-8295.1994.tb02519.x.

He, Wen and Yan Li (2020). "Comparing with the Average: Reference Points and Market Reactions to above-Average Earnings Surprises". In: *Journal of Banking & Finance* 117, p. 105824. ISSN: 0378-4266. DOI: 10.1016/j.jbankfin.2020.105824.

Hey, John D. (1994). "Expectations Formation: Rational or Adaptive or ?" In: *Journal of Economic Behavior & Organization* 25.3, pp. 329–349. ISSN: 0167-2681. DOI: 10.1016/0167-2681(94)90104-X.

Holcblat, Benjamin, Abraham Lioui, and Michael Weber (2022). *Anomaly or Possible Risk Factor? Simple-To-Use Tests.* DOI: 10.2139/ssrn.4064159. URL: https://papers.ssrn.com/abstract=4064159 (visited on 07/05/2022). preprint.

Holland, Morris K. and G. R. Lockhead (1968). "Sequential Effects in Absolute Judgments of Loudness". In: *Perception & Psychophysics* 3.6, pp. 409–414. ISSN: 1532-5962. DOI: 10.3758/BF03205747.

Hommes, Cars H. (2006). "Chapter 23 Heterogeneous Agent Models in Economics and Finance". In: *Handbook of Computational Economics.* Ed. by L. Tesfatsion and K. L. Judd. Vol. 2. Elsevier, pp. 1109–1186. DOI: 10.1016/S1574-0021(05)02023-X.

Hou, Kewei, Chen Xue, and Lu Zhang (2020). "Replicating Anomalies". In: *The Review of Financial Studies* 33.5, pp. 2019–2133. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy131.

Itti, L., C. Koch, and E. Niebur (1998). "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.11, pp. 1254–1259. ISSN: 1939-3539. DOI: 10.1109/34.730558.

Jegadeesh, Narasimhan and Sheridan Titman (1993). "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency". In: *The Journal of Finance* 48.1, pp. 65–91. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1993.tb04702.x.

Kahneman, Daniel and Amos Tversky (1972). "Subjective Probability: A Judgment of Representativeness". In: *Cognitive Psychology* 3.3, pp. 430–454. ISSN: 0010-0285. DOI: 10.1016/0010-0285(72)90016-3.

— (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–291. ISSN: 0012-9682. DOI: 10.2307/1914185.

Kenrick, Douglas T. and Sara E. Gutierres (1980). "Contrast Effects and Judgments of Physical Attractiveness: When Beauty Becomes a Social Problem". In: *Journal of Personality and Social Psychology* 38.1, pp. 131–140. ISSN: 1939-1315(Electronic),0022-3514(Print). DOI: 10.1037/0022-3514.38.1.131.

Kerr, Norbert L., Douglas L. Harmon, and James K. Graves (1982). "Independence of Multiple Verdicts by Jurors and Juries1". In: *Journal of Applied Social Psychology* 12.1, pp. 12–29. ISSN: 1559-1816. DOI: 10.1111/j.1559-1816.1982.tb00845.x.

Keynes, John Maynard (1936). *The General Theory of Employment, Interest, and Money.* Palgrave Macmillan. ISBN: 978-0-230-00476-4.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh (2018). "Interpreting Factor Models". In: *The Journal of Finance* 73.3, pp. 1183–1223. ISSN: 1540-6261. DOI: 10.1111/jofi.12612.

La Porta, Rafael (1996). "Expectations and the Cross-Section of Stock Returns". In: *The Journal of Finance* 51.5, pp. 1715–1742. ISSN: 0022-1082. DOI: 10.2307/2329535.

La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny (1997). "Good News for Value Stocks: Further Evidence on Market Efficiency". In: *The Journal of Finance* 52.2, pp. 859–874. ISSN: 0022-1082. DOI: 10.2307/2329502.

Leitner, Johannes and Ulrike Leopold-Wildburger (2011). "Experiments on Forecasting Behavior with Several Sources of Information A Review of the Literature". In: *European Journal of Operational Research* 213.3, pp. 459–469. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2011.01.006.

Lintner, John (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets". In: *The Review of Economics and Statistics* 47.1, pp. 13–37. ISSN: 0034-6535. DOI: 10.2307/1924119.

Lucas, Robert E (1972). "Expectations and the Neutrality of Money". In: *Journal of Economic Theory* 4.2, pp. 103–124. ISSN: 0022-0531. DOI: 10.1016/0022-0531(72)90142-1.

Lucas, Robert E and Edward C. Prescott (1971). "Investment Under Uncertainty". In: *Econometrica* 39.5, pp. 659–681. ISSN: 0012-9682. DOI: 10.2307/1909571.

Mankiw, N. Gregory and Ricardo Reis (2002). "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve". In: *The Quarterly Journal of Economics* 117.4, pp. 1295–1328. ISSN: 0033-5533. DOI: 10.1162/003355302320935034.

Metzler, Lloyd A. (1941). "The Nature and Stability of Inventory Cycles". In: *The Review of Economics and Statistics* 23.3, pp. 113–129. ISSN: 0034-6535. DOI: 10.2307/1927555.

Muth, John F. (1961). "Rational Expectations and the Theory of Price Movements". In: *Econometrica* 29.3, pp. 315–335. ISSN: 0012-9682. DOI: 10.2307/1909635.

Nerlove, Marc (1958). "Adaptive Expectations and Cobweb Phenomena". In: *The Quarterly Journal of Economics* 72.2, pp. 227–240. ISSN: 0033-5533. DOI: 10.2307/1880597.

O'Reilly, Dennis M., Robert A. Leitch, and Douglas H. Wedell (2004). "The Effects of Immediate Context on Auditors' Judgments of Loan Quality". In: *Auditing: A Journal of Practice & Theory* 23.1, pp. 89–105. ISSN: 02780380. DOI: 10.2308/aud.2004.23.1.89.

Pepitone, Albert and Mark DiNubile (1976). "Contrast Effects in Judgments of Crime Severity and the Punishment of Criminal Violators". In: *Journal of Personality* 33.4, pp. 448–459. ISSN: 0022-3514.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam (2019). "A Protocol for Factor Identification". In: *The Review of Financial Studies* 32.4, pp. 1573–1607. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy093.

Radbruch, Jonas and Amelie Schiprowski (2020). *Interview Sequences and the Formation of Subjective Assessments*. 045. University of Bonn and University of Cologne, Germany.

Roberts, Brian, Mike G Harris, and Tim A Yates (2005). "The Roles of Inducer Size and Distance in the Ebbinghaus Illusion (Titchener Circles)". In: *Perception* 34.7, pp. 847–856. ISSN: 0301-0066. DOI: 10.1068/p5273.

Rodríguez, Gabriel and Sara Blanco (2016). "Contrast Effect on the Perception of the Severity of a Criminal Offence". In: *Anuario de Psicología Jurídica* 26.1, pp. 107–113. ISSN: 1133-0740. DOI: 10.1016/j.apj.2016.02.001.

Schmalensee, Richard (1976). "An Experimental Study of Expectation Formation". In: *Econometrica* 44.1, pp. 17–41. ISSN: 0012-9682. DOI: 10.2307/1911378.

Sharpe, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*". In: *The Journal of Finance* 19.3, pp. 425–442. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1964.tb02865.x.

Sherif, Muzafer, Daniel Taub, and Carl I. Hovland (1958). "Assimilation and Contrast Effects of Anchoring Stimuli on Judgments". In: *Journal of Experimental Psychology* 55.2, pp. 150–155. ISSN: 0022-1015(Print). DOI: 10.1037/h0048784.

Simonsohn, Uri and George Loewenstein (2006). "Mistake: The Effect of Previously Encountered Prices on Current Housing Demand". In: *The Economic Journal* 116.508, pp. 175–199. ISSN: 0013-0133. DOI: 10.1111/j.1468-0297.2006.01052.x.

Treisman, Anne M. and Garry Gelade (1980). "A Feature-Integration Theory of Attention". In: *Cognitive Psychology* 12.1, pp. 97–136. ISSN: 0010-0285. DOI: 10.1016/0010-0285(80)90005-5.

Tversky, Amos and Daniel Kahneman (1973). "Availability: A Heuristic for Judging Frequency and Probability". In: *Cognitive Psychology* 5.2, pp. 207–232. ISSN: 0010-0285. DOI: 10.1016/0010-0285(73)90033-9.

— (1974). "Judgment under Uncertainty: Heuristics and Biases". In: *Science* 185, pp. 1124–1131. DOI: 10.1126/science.185.4157.1124.

— (1981). "The Framing of Decisions and the Psychology of Choice". In: *Science* 211.4481, pp. 453–458. DOI: 10.1126/science.7455683.

— (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty". In: *Journal of Risk and Uncertainty* 5.4, pp. 297–323. ISSN: 1573-0476. DOI: 10.1007/BF00122574.

Von Neumann, John and Oskar Morgenstern (1947). *Theory of Games and Economic Behavior, 2nd Rev. Ed.* Theory of Games and Economic Behavior, 2nd Rev. Ed. Princeton, NJ, US: Princeton University Press, pp. xviii, 641. xviii, 641.

Waite, Thomas A. (2001). "Background Context and Decision Making in Hoarding Gray Jays". In: *Behavioral Ecology* 12.3, pp. 318–324. ISSN: 1045-2249. DOI: 10.1093/beheco/12.3.318.

Wallis, Jonathan D. and Steven W. Kennerley (2011). "Contrasting Reward Signals in the Orbitofrontal Cortex and Anterior Cingulate Cortex". In: *Annals of the New York Academy of Sciences* 1239.1, pp. 33–42. ISSN: 1749-6632. DOI: 10.1111/j.1749-6632.2011.06277.x.

Chapter 2

Optimal Stopping in a Dynamic Salience Model

Abstract

We study dynamic choice under risk through the lens of salience theory. We derive predictions on salient thinkers' gambling decisions and strategy choices. We test our model experimentally and find support for all of our predictions. We also detect a strong correlation between static and dynamic choices, suggesting that salience theory can coherently explain risky choice in both static *and* dynamic contexts. Our results help to understand when people sell assets, stop gambling, enter the job market or retire.

2.1 Introduction

Dynamic decisions under risk are ubiquitous. Examples include the decisions when to sell an asset, when to stop gambling, when to enter the job market or to retire, when to buy a flight ticket or a durable good, and when to stop searching for a house or a spouse. An extensive behavioral literature of choice under risk — going back to Kahneman and Tversky (1979) — has singled out the skewness of the underlying probability distribution as one important driver of risk attitudes. In particular, people typically seek (positively) skewed and avoid negatively skewed risks. The corresponding literature on skewness seeking has traditionally focused on static decisions in order to explain, for instance, why people buy expensive insurance and lottery tickets at the same time. In dynamic problems, skewness seeking even matters when the underlying risk or stochastic process is symmetric (e.g., Barberis, 2012; Ebert, 2020; Ebert and Strack, 2015) because the decision maker can select a skewed return distribution through the choice of her stopping strategy. Hence, skewness seeking may play an even bigger role in dynamic than in static problems.

Since skewness seeking in static and dynamic decisions plausibly stems from the same cognitive mechanism, the existing literature has adapted static models of choice under risk — such as expected utility theory (EUT) or cumulative prospect theory (CPT; Amos Tversky and Daniel Kahneman, 1992) — to dynamic decision making. We follow this approach by studying, both theoretically and experimentally, the dynamic implications of Bordalo et al.'s (2012) salience theory. In salience theory, skewness seeking in static problems originates from the idea that outcomes, which are extreme relative to a reference point — such as the relatively large upside of a right-skewed risk — are particularly salient. These outcomes attract a disproportionate amount of attention and their probabilities are overweighted (Bordalo et al., 2012). Thus, salience theory exhibits the core prediction that agents seek skewness (Dertwinkel-Kalt and Köster, 2020). At the same time, the salience distortions are bounded, which places a tight limit on the degree of probability weighting and therefore also on skewness seeking. Models that imply much stronger skewness seeking than salience theory (such as CPT) could make reasonable predictions in static settings but make unreasonably extreme predictions in

some important dynamic settings (Ebert, 2015). Hence, salience theory represents a natural candidate for modelling skewness seeking in dynamic choices.

We therefore apply salience theory to dynamic contexts, and derive, test and support salience theory's predictions for dynamic choice under risk. We also document a strong relation between skewness seeking in static and dynamic setups. This suggests that there is a common mechanism that drives static and dynamic choice under risk, highlighting the importance of developing a model that can explain both static and dynamic decisions.

In Section 2.3, we adapt salience theory to analyze dynamic choice under risk at the hand of standard optimal stopping problems. We ask when a naïve[1] salient thinker stops an Arithmetic Brownian Motion (ABM) with a non-positive drift and a finite expiration date. Expected utility theory with a concave utility function cannot explain gambling when the agent loses money in expectation, that is, when the ABM's drift is negative. With a specific stopping strategy in mind, a *salient thinker*, however, inflates the probabilities of those realizations that "differ most" from his current wealth level (*contrast effect*), which might render gambling attractive. Adopting the naïve decision rule proposed in the literature (e.g., Ebert and Strack, 2015), we assume that the (naïve) salient thinker continues to gamble as long as he can find at least one stopping strategy that is more attractive to him than stopping.

In Section 2.4, we derive our theoretical results. Unlike an expected-utility agent with a concave utility function, a naïve salient thinker gambles even when he loses money in expectation. At the same time, unlike in CPT (Ebert and Strack, 2015), a naïve salient thinker stops any ABM with a sufficiently negative drift. In a next step, we restrict the choice set to all *stop-loss and take-profit strategies*[2] to learn more about how a näive

[1]Non-linear probability weighting implies that an agent's optimal strategy at time t might no longer be optimal at some later point in time (e.g., Machina, 1989). Optimal stopping behavior under salience theory and other behavioral models, thus, depends on whether the agent is aware of this time-inconsistency (i.e., the agent is sophisticated) or not (i.e., the agent is naïve). We assume that the agent is naïve (as Ebert and Strack, 2015), which is also supported by our experimental results.

[2]A stop-loss and take-profit strategy is characterized by a stop-loss threshold below the current value of the process and a take-profit threshold above the current value of the process at which the process will be stopped. These strategies are often proposed by retail banks to their customers (see., e.g., the brokerage data by Heimer et al., 2023) and have attracted much attention in the related literature (e.g., Ebert and Strack, 2015; Fischbacher et al., 2017; Heimer et al., 2023; Xu and Zhou, 2013). Important for us, these strategies allow agents to obtain skewed return distributions even if the underlying stochastic process is symmetric.

salient thinker plans to stop, and how he will revise this plan over time. These additional restrictions allow for more interesting experimental predictions of salience theory. First, a salient thinker chooses a particular subset of stop-loss and take-profit strategies, which give rise to a right-skewed distribution of returns; so-called *loss-exit strategies* (see, e.g., Barberis, 2012; Heimer et al., 2023). A loss-exit strategy is defined as a stop-loss and take-profit strategy for which the stop-loss threshold is closer to the current value of the process than the take-profit threshold, so that — by the contrast effect — stopping at a gain is more salient than stopping at a loss. Second, a naïve salient thinker does not necessarily follow his initial plan, but might instead revise his strategy over time. In particular, salience theory can rationalize stopping behavior that is consistent with the well-known disposition effect (e.g., Barberis, 2012; Odean, 1998; Shefrin and Statman, 1985).

Section 2.5 presents a laboratory experiment that is designed to test our salience-based predictions on stopping behavior. Participants in the experiment have to decide when to stop ABMs with different non-positive drifts. Subjects stop the process by defining an upper and a lower bound and the process is stopped if it reaches either bound. If a process is stopped, subjects can either sell it or restart it by moving the bounds. This design allows us to test whether subjects choose loss-exit strategies (i.e., strategies with a salient upside) and whether they revise their initial strategies as predicted by the model. We validate our approach of adapting the static salience model to an optimal stopping problem by further eliciting skewness seeking in static choices. Generalizing results from Dertwinkel-Kalt and Köster (2020), we show that, for a fixed expected value and variance, a salient thinker chooses a binary lottery over the safe option paying its expected value with certainty if and only if the lottery's skewness exceeds some threshold. If salience is indeed the psychological mechanism driving skewness seeking in general, it should coherently explain revealed attitudes toward skewness in such static choices as well as in the optimal stopping problems.

In Section 2.6, we present our experimental results. First of all, we find that subjects select skewed return distributions: for the median subject, more than 70% of all chosen strategies are loss-exit strategies. Furthermore, 93% of the subjects revise their initial strategy at least once, and actual behavior is reminiscent of the disposition effect.

We next examine how sensitive subjects stopping decisions are to the drift of the process. While it might seem obvious that fewer subjects should gamble with more negative drifts, both EUT and CPT predict that this is not the case. Risk averse EUT agents should not gamble for any non-positive drift. In contrast, subjects following the commonly used CPT-models as analyzed by Ebert and Strack (2015) should always gamble regardless of how negative the drift is. We find that most subjects start gambling a fair process with a drift of zero, but then stop before reaching the expiration date. Moreover, subjects stop the earlier, the more negative the drift of the process. Interpreted through the lens of our model, these results indicate heterogeneity in the strength of salience distortions of our subjects: Around 95% of the subjects reveal sufficiently strong salience distortions that they start the fair process, but for only 60% of the subjects salience distortions are so strong that also gambling with the most negative drift is attractive.

We also find a positive correlation between static and dynamic skewness seeking, which is both, statistically and economically, significant: subjects that reveal stronger skewness seeking in static choices also have a larger propensity to choose loss-exit strategies in the dynamic ones. Overall, our experimental results suggest that a good model of dynamic risk-taking without commitment should include moderately strong skewness seeking. Moreover, it should be based on a mechanism that can also explain skewness seeking in static choices. Our salience theory model fulfils these criteria and is therefore able to coherently explain choices in static and dynamic problems.

In Section 2.7, we show that several popular models of static choice under risk struggle to explain the dynamic evidence from our experiment because they either predict too strong skewness seeking or no skewness seeking at all. Arguably, alternative models with the right strength of skewness seeking that allows to explain our data can be developed, but we are not aware of any such model among the commonly used ones. Moreover, as salience theory can jointly explain static and dynamic data, we regard it as a prime candidate for a unified model of static and dynamic choice under risk.

We conclude in Section 5 by discussing further applications of our findings.

2.2 Related Literature on Behavioral Stopping

Our paper is related to a large theoretical (e.g., Barberis, 2012; Ebert and Strack, 2015, 2018; He et al., 2019; Henderson et al., 2017; Karni and Safra, 1990; Machina, 1989; Strack and Viefers, 2021) as well as a growing experimental (e.g. Heimer et al., 2023; Imas, 2016; Nielsen, 2019; Strack and Viefers, 2021) literature on behavioral stopping. On the one hand, we add to the theoretical literature by providing the first study of behavioral stopping in salience theory. On the other hand, we contribute to the experimental salience literature (for a survey, see Bordalo et al., 2022) by testing salience theory's predictions on behavioral stopping as well as by investigating whether it can coherently explain both static and dynamic choice under risk.

Most existing theoretical work on behavioral stopping deals with the implications of non-linear probability weighting for dynamic gambling, with a focus on the behavior predicted by cumulative prospect theory (henceforth: CPT, see Barberis, 2012; Ebert and Strack, 2015, 2018; He et al., 2019; Henderson et al., 2017, 2018; Karni and Safra, 1990; Machina, 1989; Xu and Zhou, 2013). This focus can be explained by the fact that non-linear probability weights imply (empirically relevant) time-inconsistent preferences (e.g., Machina, 1989). Predicted behavior depends, in particular, on whether or not the agent is *naïve* about his time-inconsistency. A naïve agent will revise his strategy throughout time, while a (fully) sophisticated agent foresees her intention to adjust certain strategies and chooses only strategies she will actually follow through with (e.g., Karni and Safra, 1990). With time-inconsistent preferences also the question of whether the agent can commit to a strategy becomes important. The literature has studied the stopping behavior of naïve agents without commitment (e.g., Barberis, 2012; Ebert and Strack, 2015) as well as with partial or full commitment (e.g., He et al., 2019; Henderson et al., 2017; Xu and Zhou, 2013).

For our purpose of testing salience theory's predictions, the setups with naïvete, but without commitment are relevant (Barberis, 2012; Ebert and Strack, 2015): naïvete is a more plausible assumption than sophistication (for instance, sophisticates should not start gambling with non-positive drifts), and only the absence of commitment allows to test whether predictions on time-inconsistent behavior hold true. In the seminal paper

by Barberis (2012), it is numerically shown that in finite discrete time setups naive CPT agents without commitment mostly choose loss-exit (as compared to gain-exit) strategies and start to gamble (at least for a wide range of parameters), but revise their strategies, so that ex post they exhibit gain-exit behavior. The reason is that close to the expiration date, agents cannot choose strongly skewed return distributions anymore, and therefore exit earlier than intended.[3] In continuous time setups, this mechanism is not at work: agents can always choose strongly skewed return distributions and, as a consequence, naive CPT-agents never stop (Ebert and Strack, 2015). This also holds true with an indefinite end date: for most empirically relevant cumulative prospect theory parameter values, a naive agent does not stop with probability one at any loss level (He et al., 2019). This never-stopping prediction can only be avoided by allowing for randomized stopping strategies and thereby some form of commitment (Henderson et al., 2017), or by imposing different functional forms for dynamic choices than those typically used for static choices (Duraj, 2020; Huang et al., 2020). We show in how far the predictions of salience theory differ from those by other models such as CPT, and show in particular that dynamic salience theory does not yield the (too) extreme never-stopping prediction that Ebert and Strack (2015) have derived for CPT.

We also contribute to the small, but growing experimental literature on behavioral stopping (e.g., Imas, 2016; Nielsen, 2019; Strack and Viefers, 2021). Unlike us, these papers focused on the question in how far stopping decisions are path-dependent; in particular, in how far the realization of previous gains and losses affects behavior. Closest related to us is the contemporaneous paper by Heimer et al. (2023), who study optimal stopping behavior using a process consisting of repeated (fair) coin tosses. Similarly to us, they focus on stop-loss and take-profit strategies, and they find, both in laboratory experiments as well as in observational brokerage data, that subjects ex-ante choose loss-exit strategies, but then deviate by revealing disposition-effect-like behavior. Both findings are also reflected in our data. In contrast to Heimer et al. (2023), our paper is focused on salience theory and establishes a novel link between static and dynamic skewness seeking.

[3]Heimer et al. (2023) provide direct experimental evidence for this prediction.

2.3 A Dynamic Version of Salience Theory of Choice under Risk

2.3.1 Static Model

Consider an agent who has to choose from some set $\mathcal{C}$ that contains exactly two non-negative random variables (or *lotteries*), X and Y, with a joint cumulative distribution function (CDF) $F : \mathbb{R}^2_{\geq 0} \to [0, 1]$. A state of the world here refers to a tuple of outcomes, $(x, y) \in \mathbb{R}^2_{\geq 0}$. We denote the state space by $\mathcal{S} \subseteq \mathbb{R}^2_{\geq 0}$. If a random variable is degenerate, we call it a *safe* option.

According to salience theory of choice under risk (Bordalo et al., 2012), the agent is a *salient thinker*, who evaluates a random variable by assigning a subjective probability to each state of the world $s \in \mathcal{S}$ that depends on the state's objective probability and on its salience. The salience of a state is assessed by a so-called *salience function*, which is defined as follows:

Definition 1 (Salience Function)**.** *We say that a symmetric, bounded, and twice differentiable function $\sigma : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{>0}$ is a* salience function *if and only if it satisfies the following two properties:*[4]

1. Ordering. *Let $x \geq y$. Then, for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$, we have*

$$\sigma(x + \epsilon, y - \epsilon') > \sigma(x, y).$$

2. Diminishing sensitivity. *For any $x > y$ and any $\epsilon > 0$, we have*

$$\sigma(x + \epsilon, y + \epsilon) < \sigma(x, y).$$

We say that a given state of the world $(x, y) \in \mathcal{S}$ is the more salient the larger its salience value is. The ordering property implies that a state of the world is the more

[4]Bordalo et al. (2012) also allow for random variables with negative outcomes and add a third property to ensure that diminishing sensitivity (with respect to zero) reflects to the negative domain: by the *reflection* property, for any $w, x, y, z \geq 0$, it holds that $\sigma(x, y) > \sigma(w, z)$ if and only if $\sigma(-x, -y) > \sigma(-w, -z)$.

salient the more the attainable outcomes in this state differ. In this sense, ordering captures the well-known *contrast effect* (e.g., Schkade and Kahneman, 1998), whereby large contrasts (in outcomes) attract a great deal of attention. Diminishing sensitivity reflects *Weber's law* of perception, and it implies that the salience of a state decreases if the outcomes in this state uniformly increase. Hence, diminishing sensitivity describes a *level effect*, according to which a given contrast in outcomes is less salient at a higher outcome level, thereby qualifying the contrast effect.

A salient thinker is intrinsically (weakly) risk-averse but may, depending on the salience of outcomes, sometimes behave as if being risk-seeking. He evaluates monetary outcomes via a strictly increasing, (weakly) concave, and twice differentiable value function $v : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, and forms an "expectation" by assigning each state of the world a (probability) weight that depends on how a given option compares to the alternative at hand in this state. More specifically, a salient thinker behaves as if maximizing a *salience-weighted utility*, which is defined as follows:

Definition 2. *The* salience-weighted utility *of a random variable X evaluated in $\mathcal{C} = \{X, Y\}$ equals*

$$U^s(X|\mathcal{C}) = \int_{\mathbb{R}^2_{\geq 0}} v(x) \cdot \frac{\sigma\Big(v(x), v(y)\Big)}{\int_{\mathbb{R}^2_{\geq 0}} \sigma\Big(v(s), v(t)\Big)\ dF(s,t)}\ dF(x,y),$$

where $\sigma : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{>0}$ is a salience function that is bounded away from zero.

Since the salience-weighted probabilities are normalized so that they sum up to one (e.g., Bordalo et al., 2012; Dertwinkel-Kalt and Köster, 2020), a salient thinker's valuation of a safe option $x \in \mathbb{R}_{\geq 0}$ is undistorted and given by $v(x)$, irrespective of the properties of the alternative option.

2.3.2 Dynamic Model

Stochastic process. Following Ebert and Strack (2015, 2018), we model an agent's wealth via a Markov diffusion. Specifically, we consider an *Arithmetic Brownian Motion* (ABM),

$$dX_t = \mu dt + \nu dW_t,$$

with an initial value $X_0 = x$, a constant drift $\mu \in \mathbb{R}$ and a constant volatility $\nu \in \mathbb{R}_{>0}$, as well as a standard Brownian Motion $(W_t)_{t\in\mathbb{R}_{\geq 0}}$.

To make the theory testable in the context of an incentivized lab experiment, we deviate from Ebert and Strack (2015, 2018) in two ways: First, we assume that the process is non-negative, and absorbing in zero. Second, we allow for a finite *expiration date* $T \in \mathbb{R}_{>0} \cup \{\infty\}$.

Stopping strategies. As in Ebert and Strack (2015), we represent the set of stopping strategies by the set of stopping times, where each stopping time τ refers to a deterministic plan of when to stop the process. The central feature of a stopping time is that it is based on past information only: that is, any τ is adapted to the natural filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_{\geq 0}}$ of the process $(X_t)_{t\in\mathbb{R}_{\geq 0}}$. For a fixed expiration date $T \in \mathbb{R}_{>0} \cup \{\infty\}$, choosing a stopping time $\tau \leq T$ (with probability one) implements a random wealth level X_τ with a cumulative distribution function denoted by F_τ.

For our first results, we do not impose any restrictions on the set of deterministic stopping times that the agent can choose from. But, to learn more about the strategies that are attractive to a salient thinker and the role of skewness for stopping behavior, we derive additional results under the assumption that the agent is restricted to choose a *threshold stopping time* $\tau_{a,b}$ — defined as the first leaving time of the interval (a, b) — that implements a random wealth level $X_{T\wedge\tau_{a,b}}$. The set of threshold stopping times represents the set of stop-loss and take-profit strategies, which are often proposed by retail banks to their customers (see., e.g., the brokerage data by Heimer et al., 2023) and which have attracted much attention in the behavioral and financial literature (e.g., Ebert and Strack, 2015; Fischbacher et al., 2017; Heimer et al., 2023; Xu and Zhou, 2013).

Solution concept under salience theory. Any form of non-linear probability weighting — whether it is salience-driven or mechanical — implies that an agent's optimal strategy at time t might no longer be optimal at some later point in time (e.g., Machina, 1989). Thus, optimal stopping behavior under salience theory depends on whether or not the salient thinker is aware of this time-inconsistency. We follow Ebert and Strack (2015) in assuming that the agent is naïve about his time-inconsistency. As we think of salience

effects as unconscious distortions of perception, we regard the assumption of naïvete as sensible. In Section 2.5 and Appendix B, we further discuss how to experimentally test this assumption within the salience framework.

As in Ebert and Strack (2015), we assume that "at every point in time the naïve [salient thinker] looks for a [...] strategy τ that brings him higher [salience-weighted utility] than stopping [...]. If such a strategy exists, he holds on to the investment — irrespective of his earlier plan." Assuming that the naïve salient thinker continues to gamble if and only if he strictly prefers to do so, the decision rule then reads as follows.

Definition 3 (Continuation Rule)**.** *Let $x_t \in \mathbb{R}_{\geq 0}$ be the current wealth level at time $t \in [0, T)$. A naïve salient thinker* continues at time t *if there exists a stopping time τ, such that $U(X_\tau | \{X_\tau, x_t\}) > u(x_t)$, that is, if the salient thinker finds a strategy that gives him a strictly higher salience-weighted utility than stopping at time t. Otherwise, the naïve salient thinker* stops at time t.

Our decision rule imposes the additional assumption that a naïve salient thinker evaluates each stopping strategy in isolation: at any point in time, the *consideration set* — that is, the set of strategies that the agent compares when making his stopping decision — includes a single strategy to continue with, X_τ, and the alternative to stop right now, x_t; the consideration set thus is assumed to be $\{X_\tau, x_t\}$. Since salience theory is a model of context-dependent behavior to derive testable predictions, it is necessary to impose some assumption on the consideration set. With infinitely many strategies to choose from, we regard the above specification as plausible. Moreover, our experimental design (see Section 4 for details) highlights a single strategy at a time, so that subjects likely evaluate this strategy in isolation. Still, one might argue that previously chosen strategies affect the perception of whatever strategy is considered next. Without any guidance on how the consideration set changes over time, however, it is impossible to provide a comprehensive analysis.[5] To tie our hands, we pre-registered our assumptions on the consideration set before running the experiment.

[5]When restricting attention to stop-loss and take-profit strategies, one could argue that the previously chosen lower bound a_p of a stop-loss and take-profit strategy provides a "reference point" for the newly selected lower bound a_n, and that the previously chosen upper bound b_p provides a "reference point" for the newly chosen upper bound b_n in the sense that the respective salience weights are $\sigma(v(a_n), v(a_p))$ and $\sigma(v(b_n), v(b_p))$. Then, conditional on not stopping the process, subjects would always adjust the upper

2.4 Stopping Behavior of a Naïve Salient Thinker

2.4.1 Motivating Example

To illustrate the salience mechanism, consider a salient thinker with a linear value function, $v(x) = x$, who decides when to stop a fair process with zero drift that does not expire ($T = \infty$). Suppose that the agent adopts a stop-loss and take-profit strategy, which can be represented by a threshold stopping time $\tau_{a,b}$ with a being the lower and b being the upper threshold. Such a strategy induces a binary lottery, $X_{\tau_{a,b}} = (a, p; b, 1-p)$, over wealth. Because the process has a drift of zero, at time t, the expected value of following this stop-loss and take-profit strategy is $\mathbb{E}_t[X_{\tau_{a,b}}] = x_t$. Does the salient thinker ever stop?

It is immediate to see that a salient thinker with a linear value function chooses a binary lottery with upside payoff b, downside payoff a and expected value x_t over the safe option paying x_t if and only if the lottery's upside b is assigned a larger salience weight than the lottery's downside a, that is, if and only if $\sigma(b, x_t) > \sigma(a, x_t)$. As a consequence, whenever $\sigma(b, x_t) > \sigma(a, x_t)$, following the stop-loss and take-profit strategy represented by $\tau_{a,b}$ is more attractive to the salient thinker than stopping at time t. Since $\sigma(b, x_t) > \sigma(x_t, x_t)$ due to ordering, and since the salience function is continuous, we can always find a stopping time $\tau_{a,b}$ — with a close enough to the current wealth level x_t — that the salient thinker prefers to stopping at time t. Hence, he never stops. It is easily verified that the result remains to hold for a finite expiration date. All missing proofs are provided in Appendix A.

Proposition 1. *Fix an initial wealth level $x \in \mathbb{R}_{>0}$ and expiration date $T \in \mathbb{R}_{>0} \cup \{\infty\}$. A naïve salient thinker with a linear value function does not stop a process with zero drift at any positive level of wealth.*

2.4.2 Main Theoretical Result

We are interested in how general the never-stopping result derived in the previous subsection is. By Definition 3, a salient thinker continues (or starts) to gamble if he can find

threshold by more than the lower threshold, as otherwise the lower threshold would be salient. While this prediction is inconsistent with the data that we present later on, there might be other specifications of the consideration that are consistent with our experimental findings.

a strategy that gives him strictly higher utility than not gambling. We will now show that there are two reasons why a salient thinker cannot find such a strategy and hence stops before the expiration date: either the drift of the process is sufficiently negative, or the salient thinker is intrinsically risk-averse. More precisely, while a naïve salient thinker with a linear value function also holds processes with a slightly negative drift until the expiration date, a salient thinker with a sufficiently concave value function never starts even a fair process, and this holds irrespective of his intrinsic risk-aversion (i.e., irrespective of how concave his value function is). This last prediction distinguishes salience theory from models like CPT (see Ebert and Strack, 2015), and it constitutes our main theoretical result.

Consider an ABM with an arbitrary drift $\mu \in \mathbb{R}$ and volatility $\nu \in \mathbb{R}_{>0}$. A naïve salient thinker never starts to gamble if and only if, for any stopping time $\tau \leq T$,

$$\int_{\mathbb{R}_{\geq 0}} \Big(v(z) - v(x)\Big)\sigma\Big(v(z), v(x)\Big)\ dF_\tau(z) \leq 0, \tag{2.1}$$

where F_τ denotes the CDF of the induced wealth level X_τ. Fixing the initial value $X_0 = x$, we define an *auxiliary utility function* $\tilde{u}(z) := \Big(v(z) - v(x)\Big)\sigma\Big(v(z), v(x)\Big)$, which is strictly increasing and differentiable in $z \in \mathbb{R}_{\geq 0}$. By construction, the condition derived in Eq. (2.1) is equivalent to

$$\int_{\mathbb{R}_{\geq 0}} \tilde{u}(z)\ dF_\tau(z) \leq \tilde{u}(x).$$

In words, for any fixed initial value $X_0 = x$, a naïve salient thinker never starts if and only if an EUT-agent with a utility function $\tilde{u}(\cdot)$ never starts. The main step in proving that a naïve salient thinker never starts any ABM with a sufficiently negative drift, is to derive a bound on how risk-seeking a salient thinker can ever be.

Our first result approximates a salient thinker's willingness to take risk by that of an EUT-agent with an exponential utility function. More precisely, there is an EUT-agent with exponential utility who takes up some risks that a salient thinker certainly avoids, thereby imposing a bound on the salient thinker's willingness to take risk. Given

that we can approximate a salient thinker's willingness to take risk by that of an EUT-agent with an exponential utility function, we can apply Proposition 1 in Ebert and Strack (2015) to show that a naïve salient thinker does not start any process with a sufficiently negative drift.

Theorem 1. *For any expiration date $T \in \mathbb{R}_{>0} \cup \{\infty\}$, any initial wealth level $x \in \mathbb{R}_{>0}$ and any volatility $\nu \in \mathbb{R}_{>0}$, there exists some $\tilde{\mu} \in \mathbb{R}$, such that a naïve salient thinker never starts any process with a drift $\mu < \tilde{\mu}$.*

Building on Theorem 1, we observe that an intrinsically risk-averse salient thinker may not even start a fair process; whether he does so depends on his intrinsic risk aversion (i.e., the concavity of his value function). We thus obtain the following corollary to the preceding theorem.

Corollary 1. *Depending on the concavity of their value function, salient thinkers may start or not start a process with zero drift.*

Theorem 1 and Corollary 1 allow us to distinguish between salience theory and its main alternative models (as discussed in detail in Section 2.7): EUT with a concave value function as well as reference-dependent preferences without probability weighting predict that a process with a non-positive drift is never started, while CPT and models of disappointment aversion predict that such a process is always started (and even never stopped before the expiration date). Salience theory permits for (an arguably more realistic) heterogeneity in gambling behavior.

That salience theory produces more realistic predictions than CPT is due to the boundedness of the salience function; if the salience function was unbounded, similar predictions as in CPT would prevail. While Bordalo et al. (2012) did not psychologically motivate the boundedness of the salience function they assumed, it is in line with the well-known fact that humans have difficulties interpreting numbers outside of the range they commonly experience, namely, very small and very large numbers (see Resnick et al., 2017, for a review). Salience distortions stem from large payoff contrasts that attract attention. But when humans cannot well understand the difference in magnitude between two large numbers, then it is natural to assume that increasing a large contrast even further does

not induce a (strong) behavioral reaction and, therefore, should also not distort salience weights much further (as salience theory wants to well-describe actual behavior). This is precisely the effect that the boundedness of the salience function produces.[6]

2.4.3 Gambling an (Un)Fair Process with Stop-Loss and Take-Profit Strategies

To learn more about the behavior of a naïve salient thinker, we restrict the choice set to all stop-loss and take-profit strategies, and consider only processes with a non-positive drift, $\mu \in \mathbb{R}_{\leq 0}$. First, we characterize the type of stop-loss and take-profit strategies that is attractive to a salient thinker. Second, using the additional structure, we derive a stronger result on the limits of naïve gambling. Third, we show that salience theory can rationalize the disposition effect; that is, the tendency to rather stop processes that have increased in value than those that have decreased in value (e.g., Imas, 2016; Odean, 1998; Shefrin and Statman, 1985; Weber and Camerer, 1998).

The role of skewness in naïve gambling. When referring to skewness, we use the most conventional definition of skewness, whereby skewness $S[X]$ of a lottery X is defined by the third standardized central moment

$$S[X] := \mathbb{E}\left[\left(\frac{X - \mathbb{E}[X]}{\sqrt{Var[X]}}\right)^3\right]. \tag{2.2}$$

We can then define the following notions.

Definition 4. *Lottery X is called* right-skewed *(or, equivalently,* positively skewed*) if $S(X) > 0$,* left-skewed *(or, equivalently,* negatively skewed*) if $S(X) < 0$, and* symmetric *otherwise.*

[6]Interestingly, a similar argument can also motivate the fact that probability distortions are unbounded in prospect theory. In standard prospect theory, agents mechanically overweight small probabilities, and in cumulative prospect theory, they overweight the probabilities of extreme events, which are also small for most probability distributions. When humans are bad at distinguishing between the magnitude of two small probabilities, it makes sense that shrinking an already small probability further does not affect the subjective probability by much. Hence, probabilities in prospect theory become more and more overweighted as they get smaller.

Having defined skewness, we can now define a subset of stop-loss and take-profit strategies — *loss-exit* strategies — that give rise to right-skewed return distributions. A loss-exit strategy derives its name from combining a relatively large upside with a moderate downside (so that the upper threshold is further away from the current wealth level than the lower threshold), which makes it likely to stop at a loss. To show that such a strategy induces a right-skewed return distribution, suppose the underlying process has a non-negative drift and no expiration date. Then, a loss-exit strategy induces a binary lottery that is positively skewed (as it is less likely that the upper than that the lower threshold is reached this directly follows from the fact that the upper threshold is reached with less than 50% probability, see Lemma 1 in Dertwinkel-Kalt and Köster, 2020). As a finite expiration date induces a very complicated CDF (see Lemma 3 in the Appendix), we cannot analytically prove that this holds true also with a finite expiration date; we can, however, back this claim with the help of numerical simulations, where we simulate repeatedly playing loss-exit strategies in our setup with a finite expiration date and calculate the skewness of the resulting empirical distribution (see Appendix H).

Definition 5 (Loss-Exit Strategy)**.** *Suppose current wealth level x_t. Then a* loss-exit strategy *is a stop-loss and take-profit strategy $\tau_{a,b}$ (with b denoting the upper and a the lower threshold) such that $b - x_t > x_t - a$.*

Contrast and level effect together imply that a salient thinker adopts a stop-loss and take-profit strategy only if it is a loss-exit strategy. To see why, consider again the case without an expiration date, and assume a drift of zero. In this case, any stop-loss and take-profit strategy is associated with a threshold stopping time $\tau_{a,b}$ and, because the process has zero drift, it induces a binary lottery $X_{\tau_{a,b}} = (a, p; b, 1-p)$ with an expected value of $\mathbb{E}[X_{\tau_{a,b}}] = x$. A salient thinker, thus, adopts a stop-loss and take-profit strategy only if the upside of the corresponding binary lottery is salient. If $b - x_t \leq x_t - a$, then not only is the payoff level lower in the state (a, x_t) compared to the state (b, x_t), but also the contrast in outcomes is weakly larger in the former state. Hence, since the value function is weakly concave, the downside a is more salient than the upside b, which makes this stop-loss and take-profit strategy unattractive to a salient thinker. Conversely, due

to the level effect and concavity of the value function, $b - x_t > x_t - a$ does not imply that the lottery's upside b is more salient than the downside a, so that a salient thinker does not find every loss-exit strategy attractive. All arguments extend to processes with a negative drift and to a setup with a finite expiration date. We obtain:

Proposition 2. *If a salient thinker does not stop a process, he always chooses a loss-exit strategy.*

As loss-exit strategies induce positively skewed return distributions (see our argumentation before Proposition 2), this proposition implies that a salient thinker is *skewness seeking*. Specifically, while she does not find every strategy leading to a positively skewed outcome distribution or even every loss exit strategy attractive, every strategy that she does find attractive is positively skewed.[7]

A stronger result on the limits of naïve gambling. Using Proposition 2, we can strengthen our result on the limits of naïve gambling: a salient thinker, who is restricted to choose a stop-loss and take-profit strategy, never starts if and *only if* the drift falls below some threshold.

To fix ideas, let us get back to the case of no expiration date, so that any stop-loss and take-profit strategy induces a binary lottery $X_{\tau_{a,b}} = (a, p; b, 1-p)$ over wealth. For any such strategy, the probability $p = p(a, b, \mu)$, with which the downside of the corresponding binary lottery is realized, monotonically decreases in the drift of the process. Hence, an increase in the drift μ improves the distribution induced by a stop-loss and take-profit strategy in terms of first-order stochastic dominance. By Proposition 1 in Dertwinkel-Kalt and Köster (2020), a salient thinker's certainty equivalent is monotonic with respect to first-order stochastic dominance shifts. This implies that, if a salient thinker is willing to gamble according to stopping time $\tau_{a,b}$ for a drift μ', then this stopping time is still more attractive than never starting for any drift $\mu > \mu'$. In sum, a naïve salient thinker never starts if and *only if* the drift falls below some threshold.

[7]Notably, in salience theory, it is not the case that when choosing among two positively skewed lotteries, the more skewed lottery is preferred (which has been formally proven in Corollary 2 in Dertwinkel-Kalt and Köster, 2020).

What happens if we allow for a finite expiration date instead? Because the drift of the process affects the probability of stopping before the expiration date, it is not clear, in general, whether the distribution of $X_{T\wedge\tau_{a,b}}$ improves in terms of first-order stochastic dominance as the drift increases. For loss-exit strategies, however, an increase in the drift does improve the distribution of $X_{T\wedge\tau_{a,b}}$ in terms of first-order stochastic dominance (Lemma 2 (d) in Appendix A). Hence, by Proposition 2, we can again invoke Proposition 1 in Dertwinkel-Kalt and Köster (2020) to establish that a salient thinker's gambling behavior is monotonic in the drift of the process.

Proposition 3. *For any expiration date $T \in \mathbb{R}_{>0}\cup\{\infty\}$, any initial wealth level $x \in \mathbb{R}_{>0}$ and any volatility $\nu \in \mathbb{R}_{>0}$, there exists some constant $\tilde{\mu} \in \mathbb{R}$, such that a naïve salient thinker — who is restricted to choose a stop-loss and take-profit strategy — never starts if and only if the drift of the process satisfies $\mu \leq \tilde{\mu}$.*

Proposition 3 differs from Theorem 1 in two aspects: on the one hand, the class of strategies that we consider is more restrictive as we only consider stop-loss and take-profit strategies here; but on the other hand, we obtain in Proposition 3 not just an "if", but an "if and only if" statement.

Salience theory and the disposition effect. Even if only stop-loss and take-profit strategies are available, so that *planned* behavior is path-independent, salience theory can explain *actual* behavior consistent with the disposition effect; i.e., the tendency to rather stop when the process has increased in value than decreased in value. Our salience-based explanation of the disposition effect is similar in spirit to the CPT-based one by Barberis (2012): it is not the exact path of the process, but the current wealth level that affects a salient thinker's disposition to stop.

To establish an intuition for when a salient thinker is likely to stop, let us again abstract from an expiration date. A naïve salient thinker stops at time t if and only if, for any $\epsilon, \epsilon' > 0$,

$$\frac{\sigma\Big(v(x_t-\epsilon), v(x_t)\Big)}{\sigma\Big(v(x_t+\epsilon'), v(x_t)\Big)} \times \frac{v(x_t)-v(x_t-\epsilon)}{v(x_t+\epsilon')-v(x_t)} \geq \frac{1-p}{p}, \tag{2.3}$$

where $p = p(\epsilon, \epsilon', \mu)$ denotes the probability of stopping at a loss relative to the current wealth level. Because of the constant drift, the right-hand side of (2.3) is independent of the current wealth level x_t (see Lemma 1 in Appendix A). If the left-hand side of (2.3) is increasing in x_t, the salience model, thus, predicts a *disposition effect*: in this case, stopping becomes more likely after the process has increased in value and less likely after it has decreased in value. If the left-hand side of (2.3) is decreasing in x_t, however, salience theory predicts the exact opposite behavior. In sum, salience theory can rationalize, but does not predict the disposition effect.

While in general we stay agnostic regarding the functional forms of salience and value functions, it could still be interesting to see whether common salience specifications could explain the disposition effect or not. So far, the salience literature has adopted the specification proposed in Bordalo et al. (2012), which uses a linear value function and salience function

$$\sigma(x, y) = \frac{|x - y|}{|x| + |y| + \theta}$$

with $\theta > 0$. In Appendix A.5 we examine whether this salience specification predicts the disposition effect and find that this is not the case. Hence, in order for salience theory to predict the disposition effect, other salience specifications are needed: either other salience functions or other value functions (e.g., piece-wise linear value function reflecting loss aversion, as used in the Online Appendix of Bordalo et al., 2012).

2.5 An Experiment on Dynamic Gambling Behavior

In this section, we present and discuss our experimental design. A translated version of the experiment is available at: `https://os-experiment-archive.herokuapp.com/demo`.[8]

2.5.1 Experimental Design

We conducted a pre-registered lab experiment in which subjects had to repeatedly decide at which price to sell different assets. Subjects made their selling decisions in (approx-

[8]The experimental design, including the fully specified salience model and its predictions, was pre-registered in the AEA RCT registry as trial AEARCTR-0005359.

imately) continuous time, and they could hold each asset for a maximum duration of 10 seconds. If a subject did not sell an asset within 10 seconds, it was automatically sold at the price reached at the expiration date. We set the initial price of each asset to $x = 100$ Taler, an experimental currency that was converted into € at a ratio of 10:1 at the end of the experiment.

The price of an asset followed an ABM with a drift $\mu \in \{0, -0.1, -0.3, -0.5, -1, -2\}$ and a volatility $\nu = 5$. The price was updated every tenth of a second (i.e. $T = 100$), with the price changes being drawn from a normal distribution with mean μ and variance ν^2.[9] Hence, although the implemented price paths are not truly continuous, the incentives provided to the subjects are exactly the same as in the continuous-time model introduced in Section 2.3. Moreover, while using a discrete number of time periods is necessary for implementation, the process looked smooth, and subjects could not know how many discrete steps it consisted of.

As it is illustrated in Figure 2.1, we restricted the choice set to all stop-loss and take-profit strategies: at every point in time, subjects could choose an upper and a lower stopping threshold. Once the price of the asset reached either threshold, subjects could decide whether to sell the asset at this price or to adjust the thresholds and continue the process (see the lower left panel); that is, the strategies were non-binding to rule out any form of commitment. Subjects could pause the process at any point in time to adjust the thresholds (see the upper right panel). But, importantly, subjects could set only one upper threshold and one lower threshold at a time, and thus observed a stopping strategy in isolation. At the beginning, the upper and lower threshold were centered symmetrically around the initial price (see the upper left panel). To start the process, subjects had to move each threshold at least once. Before starting the process, subjects could decide to sell the asset immediately (see also the upper left panel).

Overall, subjects made six selling decisions, one decision for each of the drift parameters. The order of the drifts was randomized at the subject level. It is important for our analysis to make sure that subjects understand before the start of the decision round that a non-positive drift of a process means that they will, on average, not gain money from

[9] Notice that the drift of an ABM is additive over time. To help subjects understand what the drift of a process is, we thus presented them aggregated drifts per second (i.e., 10 μ) in the experiment.

Figure 2.1: *Screenshots of the decision screen for the process with zero drift (in German). The text above the chart mentions the drift for this round ("The practice rounds are over now - now it's getting serious. Please make your selling decision. The drift in this round is 0."). The red lines indicate the upper and lower stopping thresholds. The blue button in the upper left panel says "Sell Immediately". The button in the upper right panel allows subjects to pause the process. The buttons in the lower left panel say "Sell" or "Adjust the bounds". The lower right panel shows the final selling price.*

gambling with this process. If they do not understand this, they might start gambling (even if they have EUT preferences), because they expect to earn money; then they would stop as soon as they have learned from observing the process during the decision round that they are losing money over time.

We have two approaches to rule out such learning about the drift during the decision round: First, at the beginning of each round, we inform the subjects about this round's drift. If subjects fully understand the implications of a given drift, this would be sufficient to rule out learning. However, subjects may not understand the meaning of a drift, despite our explanations in the instructions. Therefore, we also let subjects watch the development of three sample paths from the underlying process for 10 seconds each. Moreover, we show them an overview of ten additional sample paths of the process (see Figure 2.2). These thirteen sample paths that subjects see before making a decision

(which were randomly drawn at the subject level, meaning that different subjects saw different sample paths of the same underlying process) should give them a quite good understanding of the process and its drift. Thus, we feel confident that substantial learning during the decision round while watching the change of the process is not a driver of our results.

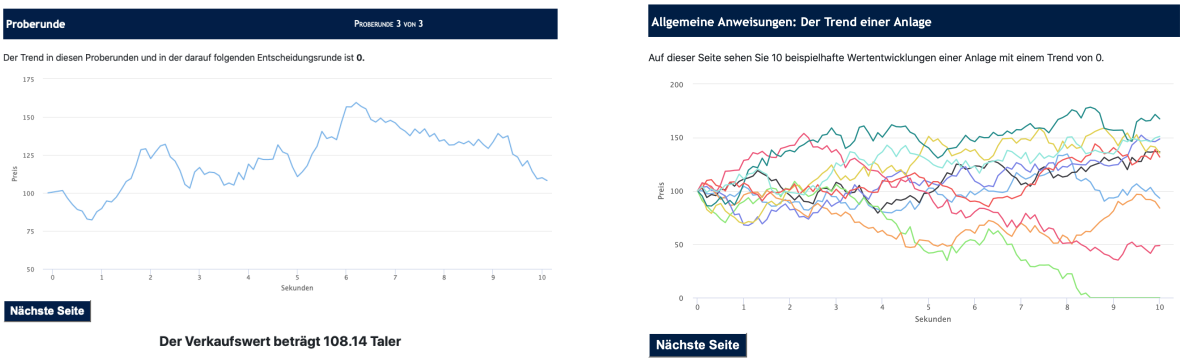


Figure 2.2: *Screenshots of the sampling screens for the process with zero drift (in German).*

After completing the six stopping problems, subjects faced a series of twelve (static) choices between a binary lottery and the safe option paying the lottery's expected value with certainty. We used two sets of lotteries with the exact same expected value (either €30 or €50) and the exact same variance, but different levels of skewness (see Table 2.2 in Appendix C for an overview). The order of the lotteries was randomized at the subject level. Finally, subjects answered five questions of a modified cognitive reflection test (CRT; closely aligned to Primi et al., 2016), and the five financial literacy questions proposed by Lusardi and Mitchell (2011). All additional questions are listed in Appendix C.

At the end of the experiment, for each subject, one of the six selling decisions was randomly drawn by the computer to be payoff-relevant. We randomly selected one subject in each session for whom, in addition, one of the twelve static choices was randomly chosen to be payoff-relevant. Subjects were further rewarded for correctly answering CRT and financial literacy questions (1 Taler per question), and they received an additional €4 for their participation.

We conducted 5 sessions with a total number of $n = 158$ subjects. The sessions took place in January 2020 in the experimental laboratory at the University of Cologne.

The experiment was conducted using the software oTree (Chen et al., 2016) and participants were invited via ORSEE (Greiner, 2015). The experiment lasted for around 45 minutes on average. Subjects earned on average slightly less than €15, with earnings ranging from €4 to €117.

2.5.2 Implementation and Discussion of the Design

In this subsection, we provide additional information on the implementation of the experiment, and we discuss in how far our design choices are essential given the objectives of our study.

Explanation of the process. To make the definition of the process easily accessible for subjects, we followed a mostly visual approach. In particular, we did not confront subjects with the differential equation that defines an ABM. Instead, we simply told subjects the following:

> "*In this experiment you will see assets of varying profitability. How profitable an asset is in the long run is described by the drift of the asset. The drift denotes the average change in the value of the process per second.*
>
> *A positive drift implies that the asset will increase in value in the long run, while a negative drift implies that the asset will decrease in value in the long run. Notice that the value of the asset varies. Hence, even an asset with a negative drift sometimes increases in value.*"

To get some understanding of the process and its drift, subjects were presented with exemplary sample paths from three processes with different drifts.[10] Subjects were told that the processes they would see in the experiment differ *only* in their drift. In particular, we told them that all processes have in common that they are non-negative and absorbing in zero.

[10]The sample paths we used in the instructions are exemplary for these processes in the sense that the final values after 10 seconds are 120 (for $\mu = 2$), 100 (for $\mu = 0$), and 80 (for $\mu = -2$), respectively. All subjects saw the exact same sample paths in the instructions.

Finally, to make sure that subjects really understood the stochasticity of an ABM (without confusing them by introducing a formal notion of variance), we told them that

> "*Independent of the drift, the value of the asset can, in principle, become arbitrarily large. The probability that the asset's value indeed becomes very large is the smaller the more negative the drift is. But even an asset with a very negative drift can attain a very large value.*"

This may raise the concern that subjects could think (at least if they did not carefully read the previous part of the instructions, stating that a negative drift gives, on average, a decrease in value) that even assets with negative drifts are on average a profitable investment. In this case, the total of thirteen sample paths that subjects see for each drift before making their selling decision should give subjects a rough understanding of the expected value of the process.

We regard this part of the instructions as particularly important since the predictions of salience theory rely on the assumption that subjects understand the potential skewness induced by stop-loss and take-profit strategies with a large upper stopping threshold. A translation of the full screen-by-screen instructions is provided in Appendix C.

Features of the process. To make our theory testable, we deviate from Ebert and Strack (2015) in two ways: First, since it is impossible to implement a process that can run forever with probability one, we implemented — similar as Heimer et al. (2023) — a finite expiration date. Alternatively, we could have implemented a random termination rule, according to which, at time t, the asset is automatically sold with probability $\omega_t \in [0, 1]$. A finite expiration date makes a theoretical analysis of stopping behavior feasible, while with a random termination rule the probability distribution associated with a given stop-loss and take-profit strategy would not be tractable anymore. A finite expiration date is also easier to explain to the subjects, which we regard — given the complexity of the experiment — as a major advantage. Second, to ensure incentive-compatability, we make

the process absorbing in zero.[11] We further restrict the drift of the process to be non-positive, because processes with a positive drift do not allow us to separate between the predictions of different models such as EUT, CPT, and salience theory.

Duration of the Process. A potential concern of our experimental design is that the process only runs for 10 seconds, which is significantly shorter than the time horizon of our motivating examples such as stock trading or job search. However, during the 10 seconds processes already change considerably, and this rather short time span makes it easy to visually follow the development of the process. Notably, skewness both intuitively and theoretically affects short- and long-run behavior likewise. Hence, abstracting from concerns about reaction times (which are discussed below), skewness effects can be studied with short and long processes alike. Notably, subjects can pause the process at any time and thus slow down their decision making. We view the short duration of the process even as an advantage as it allows subjects to stay focused on the task which would become increasingly difficult for longer horizons.

(Approximately) Continuous time. It is not feasible to implement a truly continuous process. Instead, we update the process every tenth of a second by drawing from a normal distribution with mean μ and variance ν^2. This way the problem of whether to stop after s seconds in our experiment is equivalent to that of stopping an ABM with drift μ and volatility ν at $10-s$ seconds before the expiration date.[12] Moreover, as the process looked smooth, subjects could never know whether and when only a few time periods were left. This design feature is crucial because it allows subjects to select a strongly skewed return distribution even near the expiration date.[13] This way our experimental implementation fits well to the continuous time process in our model. Our experimental

[11] In principle, we could have implemented losses up to the size of an endowment that subjects received at the beginning of the experiment. But even then we would have needed to bound the process from below.

[12] Subjects could really implement any combination of stop-loss and take-profit thresholds they like. For conciseness, assume that $X_t = 100$, and a subject would like to stop either at 110 or 99. If, in the next step, the process was updated to say 98, the subject is still paid according to his chosen stop-loss threshold of 99 (unless he revises his strategy to continue gambling). The same is true in case the process jumps above the take-profit threshold.

[13] This is not possible in the discrete-time setup in Barberis (2012) or the experimental setup in Heimer et al. (2023).

setup also closely approximates many real live situations where investor or gamblers can always choose strongly skewed return distributions.

Restriction of the choice set. Subjects could choose among all stop-loss and take-profit strategies.[14] This design choice was made based on both practical and theoretical considerations. First, we need an experimental design that allows us to learn something about the actual strategies that subjects choose. When simply providing subjects with a STOP-button, so that they could implement any strategy, we would not learn anything beyond realized stopping times. Stop-loss and take-profit strategies are not only easy to elicit but also enable subjects to choose highly skewed return distributions. This allows us to study the role of skewness in stopping problems. Second, stop-loss and take-profit strategies are highly relevant in practice, which is reflected in the large interest that this type of stopping strategy has attracted in the economics literature (e.g. Ebert and Strack, 2015; Fischbacher et al., 2017; Heimer et al., 2023; Xu and Zhou, 2013).

Non-binding strategies and costless adjustments. We allowed the subjects to costlessly adjust the stop-loss and take-profit thresholds over time: subjects could stop the process at any time, adjust one or both thresholds, and then continue the process. Moreover, the chosen strategies were non-binding in the sense that, once the price of the asset reached either threshold, subjects could decide whether to really sell the asset at this price or whether to adjust the thresholds and continue the process. Again we made both design choices for practical and theoretical reasons.

First, if either strategy adjustments were costly or if the strategies were binding, subjects could partially commit to a strategy. While the commitment effect of costly strategy adjustments is obvious, binding strategies introduce partial commitment when subjects anticipate that they will not be able to adjust their strategy fast enough; namely, before the process hits a threshold. In our main real-world examples, such as selling

[14]We intentionally designed the decision screen, where subjects set a single upper and a single lower threshold (see Figure 2.1), in a way that makes it hard for subjects to visualize a strategy that does not fall into the class of stop-loss and take-profit strategies. But, even if subjects adopted other strategies, the test of our main theoretical result — namely, Theorem 1 — would still be valid, as here we did not impose any restriction on the choice set.

an asset and gambling in a casino, investors or gamblers have (at best) very limited commitment power (as also demonstrated by Heimer et al., 2023, using brokerage data). Third, since subjects have a non-zero reaction time, non-binding strategies reduce noise in measuring the *intended* stopping time. Preventing this kind of noise seems particularly important, as it would be asymmetric — making stopping disproportionally more likely than non-stopping — and hard to model. This improves the fit between our experiment and our model, where stopping results from the unavailability of an attractive threshold stopping strategy. In the model, the agent chooses infinitely quick in continuous time, which is not feasible for the experimental subjects. However, upon the process hitting one of the thresholds subjects can take as much deliberation time as they need to figure out whether they want to continue gambling.

Importantly, even though strategy adjustments are costless, the exact thresholds are important and should be carefully set by the subject right from the beginning. The stop-loss threshold, for instance, gives a lower bound on the value that the process can reach and, therefore, should not be set below the level that the subject would not want to undercut. Likewise, the take-profit threshold should not be set too high, as otherwise moderate gains cannot be cashed in. Since the value of the process changes in (almost) continuous time, but subjects are not able to adjust the thresholds in continuous time, choosing the "right" thresholds to begin with is important. Subjects could, however, start with bounds that are tighter than the ones they actually want to play with, and plan to adjust them once one bound is hit. As this, however, involves an extra effort without any benefit, we would not think such behavior is a dominant force in our experiment.

Indicators of naïvete. When assuming a fixed expiration date and restricting the choice set to all stop-loss and take-profit strategies, we cannot interpret adjustments of the initial strategy as time-inconsistent behavior and thus as an indication of naïvete, since the remaining time until the expiration date conveys payoff-relevant information. Looking at processes with a non-positive drift, however, allows us to test the naïvete assumption within the salience framework.

A sophisticated salient thinker differs from her naïve counterpart in that she anticipates her future selves to act in a different way than her present self does, which she takes into account when making her stopping decision. A sophisticated salient thinker who lacks commitment then behaves as if she was playing a game with her future selves (e.g., Karni and Safra, 1990). To solve this game, we adopt the equilibrium concept of Ebert and Strack (2018), according to which a given stopping strategy constitutes an equilibrium if and only if at every point in time it is optimal to follow this strategy, taking as given that all future selves will do so.

As we show in Appendix B, a sophisticated salient thinker, who lacks commitment and chooses from the set of all stop-loss and take-profit strategies, never starts any process with a non-positive drift. Consequently, (partial) naïvete is a necessary assumption to rationalize gambling in the context of our experiment within the salience framework.

2.5.3 Experimental Predictions

We now state the precise predictions of salience theory that guided our experimental design. We slightly deviate from our pre-registration, which was based on the salience model with a linear value function: due to the weakly concave value function, Prediction 2 differs from what was pre-registered, while the remaining predictions are identical to the pre-registered ones.

At its core, our dynamic salience model is built to explain skewness-seeking behavior in dynamic choices under risk. Since subjects are restricted to play stop-loss and take-profit strategies in our experiment, we hypothesize (based on Proposition 2) that they play loss-exit strategies. This prediction is not shared by EUT, but by other models on skewness seeking such as CPT and models on disappointment aversion.

Prediction 1. *Conditional on not selling the asset, subjects choose a loss-exit strategy.*

The main theoretical contribution of our model, relative to existing models on skewness effects, is that it can rationalize gambling with moderately negative expected values while ruling out that agents accept any sufficiently skewed gamble regardless of how unfair it is. Therefore, our model also makes two predictions on the relation between the drift of the process and the subjects' decisions to start gambling.

Prediction 2. *If $\mu = 0$, subjects might start to gamble, and they might stop before the expiration date.*

Prediction 3. *The share of subjects selling the asset immediately monotonically decreases in the drift.*

By Corollary 2, salient thinkers may or may start to gamble. With the help of simulations, we also show that salient thinkers need not gamble until the expiration date. Due to the complexity of the Cumulative Distribution Function (CDF), we cannot formally prove this. However, there are two reasons why stopping before the expiration date is likely predicted by salience: First, as time passes, the current level of the process and, consequently, the salience utility change. Second, hitting fixed thresholds becomes less probable over time, so that they have less effect on the resulting return distribution when time passes. These two effects are also demonstrated in our simulations (see Appendix H), which gives Prediction 2. Prediction 3 directly follows from Proposition 3. These predictions distinguish our model from EUT with a concave utility function, which does not yield Prediction 2. It also distinguishes our model from CPT as modelled by Ebert and Strack (2015) as well as from models on disappointment aversion, both of which do neither yield Prediction 2 nor Prediction 3. Therefore, we regard it as important to test these predictions, even if they are very intuitive.

Lastly, since we extend a theory of "static" choice under risk to a dynamic setup, we are interested in the empirical relationship between a subject's attitude toward static and dynamic risks. If salience is indeed the psychological mechanism underlying our results, it should coherently explain behavior revealed in static and dynamic choices. As we show in Appendix D, a salient thinker chooses a binary lottery, with a fixed expected value and a fixed variance, over the safe option paying its expected value if and only if the lottery's skewness exceeds a certain threshold. By Proposition 2, this preference for positive skewness is also what drives a salient thinker's stopping behavior. We, therefore, classify both static and dynamic choices into being *skewness-seeking* or not. We say that a static choice is skewness-seeking if the subject chooses a right-skewed lottery over the safe option or the safe option over a left-skewed or symmetric lottery (see Table 2.2 in

Appendix C for the exact lotteries and classification). We further classify a stopping strategy as being skewness-seeking if it is a loss-exit strategy and, thus, induces a right-skewed return distribution. Based on Proposition 2, we expect a positive correlation between the share of skewness-seeking choices in static and dynamic decisions.

Prediction 4. *The share of skewness-seeking choices by a subject in the static decisions is positively correlated with the share of loss-exit strategies this subject chooses in the dynamic decisions.*

This prediction is not shared by any of the alternative models that we discuss throughout this paper, either because they do not predict skewness seeking (as EUT and reference-dependent preferences without probability weighting) or because they cannot explain the heterogeneity in gambling behavior that we observe (as it is the case for CPT and models of disappointment aversion).

2.6 Experimental Results on Dynamic Gambling Behavior

We first describe our data. Subsequently, we present our main experimental results, as well as evidence on how subjects revise their strategies over time, and we discuss in how far this speaks to the salience mechanism that drives our predictions on stopping behavior. Finally, we present exploratory results on disposition-effect-like behavior, and on the role of cognitive skills.

2.6.1 Data and Descriptive Statistics

For all subjects and all processes, our data includes the choice of whether to start the process as well as all chosen stop-loss and take-profit strategies (consisting of an upper bound and a lower bound). We also record the times when each strategy was chosen and the value of the process at each point in time. From these values, we can calculate at which time, if ever, subjects stopped a process, as well as the distance between the two bounds and the process at each point in time.

Table 4.2 shows descriptive statistics for the data from our experiment. We can see that the share of subjects who sell the asset immediately increases as the drift becomes more negative. The share of subjects who do not sell the asset before it expires and the average time the asset was held decreases as the drift becomes more negative. Moreover, the upper bounds are, on average, further away from the current value of the process than the lower bounds for all drifts. The termination value also decreases with the drift except for the comparison between the drifts -1 and -2. The higher termination value for a drift of -2 is likely driven by the higher fraction of subjects selling the asset immediately (i.e., at a value of 100).

Drift:	0.0	-0.1	-0.3	-0.5	-1.0	-2.0
% Sold Immediately	5.70%	9.49%	13.29%	22.15%	27.22%	41.14%
% Never Sold	18.99%	8.86%	9.49%	5.06%	3.80%	1.27%
Termination Value	95.45	92.34	87.98	81.89	73.97	82.74
	(40.19)	(38.99)	(33.63)	(32.59)	(34.03)	(26.70)
Stopping Time	6.31	5.03	3.77	2.99	2.04	0.74
	(3.72)	(3.69)	(3.72)	(3.59)	(3.16)	(1.85)
Upper Bound	136.92	133.57	124.97	124.93	122.64	122.14
	(27.99)	(28.47)	(29.51)	(25.22)	(26.63)	(26.80)
Lower Bound	71.82	75.88	68.40	71.62	66.42	77.75
	(31.23)	(33.82)	(31.38)	(33.15)	(32.24)	(30.04)
Distance Lower Bound to Value	19.02	16.80	19.38	18.52	19.99	14.94
	(18.85)	(17.80)	(20.61)	(20.33)	(22.52)	(21.07)
Distance Upper Bound to Value	33.76	28.86	26.74	24.55	25.47	24.99
	(27.74)	(27.02)	(22.40)	(25.98)	(27.02)	(31.30)

Table 2.1: *The table shows descriptive statistics for our experimental data. The values without parentheses are the means. The values in parentheses are the standard deviations. Each column shows data for one drift. "% Sold Immediately" is the percentage of subjects that sold the asset immediately and hence never started gambling. "% Never Sold" is the percentage of subjects who held the asset until the expiration date. The stopping time is either the time at which a subject sells the asset after it hits one of the bounds or it equals the value of 10 seconds, the time after which the asset process expires and the asset is sold automatically. The values for the upper and lower bounds include one data point for the initial bounds set by a subject before he can start the process and a data point for each time a bound was adjusted. Similarly, the variables "Distance Upper/Lower Bound to Value" include the (absolute) distance between a bound and the current value of the process every time a bound is adjusted.*

2.6.2 Main Test of our Salience Predictions

First, we show that, consistent with Prediction 1, a majority of subjects initially choose a loss-exit strategy. This result on initial strategies holds across all the different drifts that we considered (see Figure 2.10 in Appendix E).

Result 1 (a)**.** *Conditional on not selling immediately, 65% of initial strategies are loss-exit strategies.*

We also perform a t-test with standard errors clustered at the subject level and confirm that the share of subjects who initially choose loss exit strategies is significantly above 50%, implying that, on average, subjects are skewness seeking.

When aggregating all the strategies a subject has chosen throughout the experiment (including both initial and revised strategies), we observe that a majority of the subjects predominantly choose loss-exit strategies and that 17% of the subjects pick exclusively loss-exit strategies (see Figure 2.3 for the distribution across all subjects). This gives the second part of our result on chosen strategies:

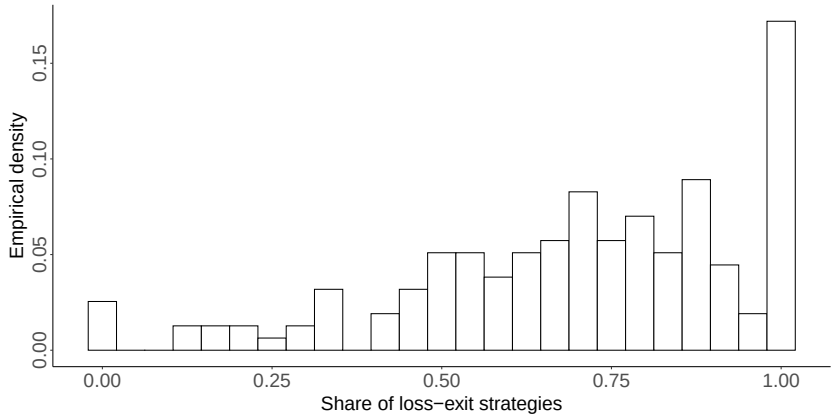


Figure 2.3: *The figure depicts the empirical distribution of the share of loss-exit strategies across subjects. The share is calculated on the subject level by taking all strategies, including both initial and revised strategies, aggregating across different drifts, and determining the percentage of those strategies that are loss-exit strategies.*

Result 1 (b)**.** *For the median subject, 73% of all strategies chosen throughout the experiment are loss-exit strategies.*

Overall, these results suggest that the majority of subjects are skewness seeking as predicted by our model. The fact that not all selected strategies are right-skewed can partially be explained by inherent noise in experimental data collection, and partially by subject heterogeneity (meaning that not all subjects are salient thinkers).

Our next result — as depicted in Figure 2.4 — is a monotonic relationship between the drift of the process and a subject's stopping behavior. Specifically, subjects do gamble (even if the drift is negative), but their behavior is sensitive to the drift of the process. At the hand of Figure 2.4, we will successively discuss the results corresponding to Predictions 2 and 3.

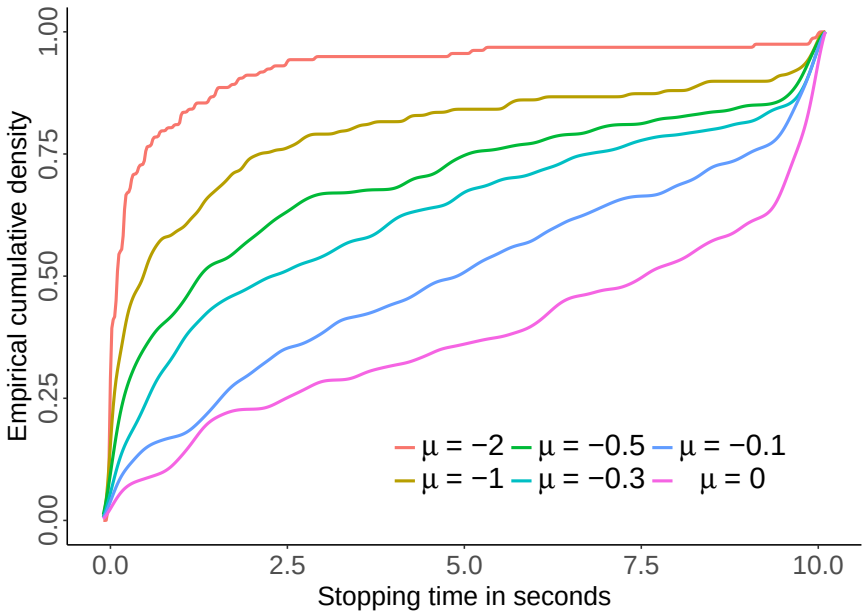


Figure 2.4: *The figure depicts the smoothed empirical cumulative distribution functions of stopping times, one for each of the different drifts.*

To address Prediction 2, we look into stopping behavior for the fair process with zero drift. Around a fifth of all subjects hold the asset with a drift of zero until the expiration date, while only about 5% of all subjects sell the asset with a drift of zero immediately. Moreover, 65% of the subjects hold this asset for more than 5 of the maximal 10 seconds. We summarize:

Result 2. *The median subject holds the fair asset with zero drift for 7.65 out of 10 seconds, and around 19% of the subjects hold the fair asset until the expiration date.*

While salience theory does not make a precise prediction on when subjects stop a fair process with zero drift (see Prediction 2), Result 2 is clearly inconsistent with both, EUT with a concave utility function — which predicts that subjects sell the asset immediately — as well as with CPT as modelled by Ebert (2015) — which predicts that subjects will hold the asset until the expiration date.

Next, we investigate how the drift affects a subject's decision whether to start a process and, after starting, when to stop it. The share of subjects stopping immediately monotonically decreases in the drift of the process (see the right panel of Figure 2.8 in Appendix E). Drift-sensitive stopping behavior is consistent with Prediction 3: estimating a linear probability model indicates that increasing the drift by one unit reduces the average probability of stopping immediately by 17.1 p.p. (p-value < 0.001, standard errors clustered at the subject level).

Result 3 (a)**.** *The share of subjects selling immediately monotonically decreases in the drift of the process.*

This result is also clearly inconsistent with EUT and CPT, both of which predict that the share of subjects who sell the asset immediately is constant in the drift, either because subjects should always stop immediately (EUT) or always gamble until the expiration date (CPT).

Figure 2.4 further shows that not only the share of subjects selling the asset immediately is monotonic in the drift, but that the whole distribution of stopping times shifts upward in the sense of first-order stochastic dominance as the drift increases.[15]

Result 3 (b)**.** *Subjects stop earlier for processes with more negative drifts.*

These results provide valuable insights into the strength of skewness effects. In our model, agents are intrinsically risk averse, so that they will not start a process with a non-positive drift unless they can implement a strategy that induces a skewed outcome distribution. When deciding whether to start a process with a drift of zero, subjects face a trade-off between the variance of the return distribution — which they dislike

[15]This is only violated for the processes with a drift of $\mu = -0.1$ and $\mu = -0.3$ in very few points, so that these violations are not even visible in the smoothed CDFs depicted in Figure 2.4.

— and the skewness of the return distribution — which they can select with the right bounds and which they like. We find that, interpreted through the lens of our model, for 95% of subjects skewness seeking is strong enough to make them gamble. But the more negative the drift of the process is, the stronger subjects' skewness seeking needs to be to render gambling attractive. Hence, our finding that the share of subjects who start to gamble monotonically decreases in the drift of the process, substantiates heterogeneity in the strength of skewness seeking.

We further look into whether subjects hold the process until it expires. Conditional on starting, most subjects do not hold the processes until the expiration date. Even for the fair process only 19% do so, which is consistent with salience theory but conflicts with models such as CPT that predict that agents will never stop gambling regardless how negative the drift of the process is.

Finally, we study whether skewness seeking in static and dynamic decisions is related. As depicted in Figure 2.5, subjects behave quite consistently in the static and the dynamic decision problems. To test for the link between static and dynamic skewness seeking, we regress the share of loss-exit strategies amongst all strategies chosen throughout the six dynamic problems on the share of skewness-seeking choices in the twelve static problems. We find a positive and statistically significant correlation, which gives our fourth result:

Result 4. *The share of skewness-seeking choices by a subject in the twelve static decisions is positively correlated with the share of loss-exit strategies this subject chooses in the six selling decisions.*

One might be concerned that Result 4 conflicts with the "discrepancy" between static and dynamic risk taking documented in Heimer et al. (2023). Heimer et al. (2023), however, compare the willingness to take risk when a fair coin is flipped once (i.e., in a static choice) — which induces a *symmetric* distribution of returns — and when a fair coin is flipped repeatedly (i.e., in a dynamic choice) — in which case the right stopping strategies allow to create very *right-skewed* distributions of returns. In other words, while we study the relationship between skewness seeking in static and dynamic decisions, Heimer et al. (2023) look at the difference in behavior between static and

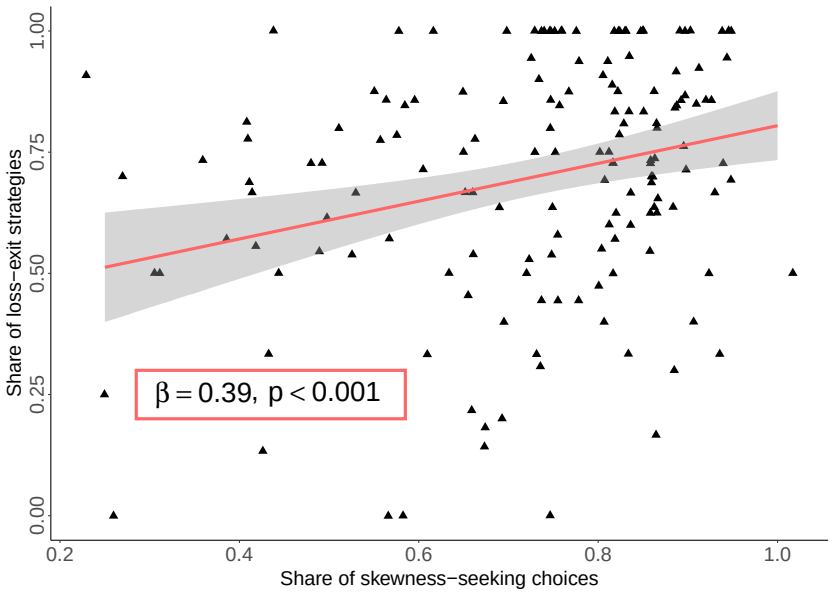


Figure 2.5: *The figure depicts the relationship between static and dynamic gambling behavior. We further provide the estimated slope-coefficient of the depicted linear regression, which is significantly larger than zero. The data points are scattered for illustrative purposes.*

dynamic problems that results from the fact that the latter enables subjects to choose a skewed distribution of returns.

2.6.3 On the Salience-Mechanism: Frequency and Direction of Strategy Adjustments

Consistent with our model, strategy revisions are ubiquitous and follow precise patterns. Altogether, (i) more than 93% of the subjects (147 out of 158) revised their initial strategy in at least one of the six selling tasks, (ii) conditional on starting, subjects adjust their strategies 1.6 times per task, and (iii) about 70% of the strategy adjustments happen in an attempt to prolong gambling after the process has hit one of the previously chosen thresholds. Moreover, if a subject chooses a loss-exit strategy and the process hits a threshold, the subject is — conditional on not stopping the process — more than six times as likely to again choose a loss-exit instead of a gain-exit strategy (see the left table of Figure 2.11 in Appendix E),[16] which is consistent with Prediction 1.

Conditional on not selling the asset immediately, around 45% of the processes are stopped "later" than when the subject initially planned to stop the process; that is, 45% of the processes pass (at least) one of the initial thresholds without being stopped. Notably, the share of processes being stopped later than initially planned monotonically increases in the drift of the process, from 20% (for the most negative drift) to around 54% (for zero drift). This suggests that subjects revise their strategies, as predicted by our model, and the fact that this behavior is more pronounced for processes with a less negative drift is again in line with our salience model's prediction that subjects are sensitive to the drift in a "reasonable" way.

We further observe that 35% of the processes fall below the initial stop-loss threshold, but only 12% of the processes rise above the initial take-profit threshold. Taken together these results indicate exactly the type of strategy revisions that our model (and also the model by Ebert and Strack, 2015, in an extreme form) builds on in order to explain excessive gambling: subjects choose loss-exit strategies and thereby positively skewed return distributions, and then adjust strategies as soon as the stop-loss threshold is hit in order to continue gambling with a newly chosen loss-exit strategy. In sum, the findings

[16]Strategy adjustments conditional on not hitting a threshold follow a similar pattern. Suppose that in the moment of pausing the process the currently played strategy is a loss-exit strategy. Then, all of our subjects have selected another loss-exit strategy and no one has switched to a gain-exit strategy (right table in Figure 2.11 in Appendix E).

on strategy adjustments indicate that our model gives a quite accurate description of the mechanism underlying our main experimental results.

2.6.4 On the Disposition to Stop and the Role of Cognitive Skills

Subjects reveal a disposition effect. As alluded to before, more processes fall below the initial stop-loss threshold (namely, 35%) than rise above the initial take-profit threshold (namely, 12%). Keeping the asset "too long" (compared to the subject's initial plan) when the process has decreased in value rather than increased in value is reminiscent of the disposition effect, whereby assets are rather sold in the gain domain and rather held in the loss domain.

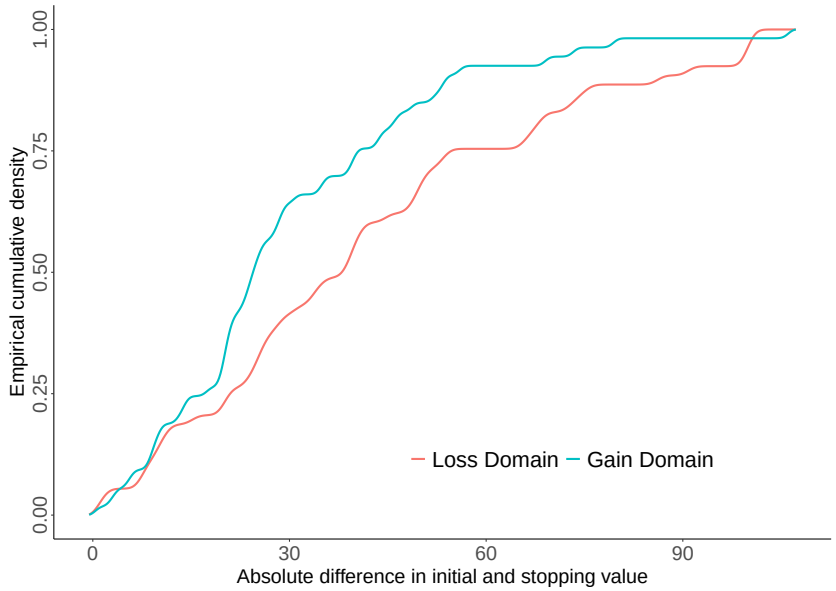


Figure 2.6: *The figure depicts the smoothed empirical CDF of stopping at a given distance to the initial value of the process, separately for processes that have gained and that have lost in value. We consider only fair processes with a drift of zero for which the initial strategy was adjusted at least once.*

Another test for the disposition effect is to compare the likelihood of selling assets that have gained a particular amount to that of selling assets that have lost exactly the same amount: by the disposition effect, the former assets should be more likely to be sold than the latter, which is precisely what we find. Those subjects, who have revised their initial strategy for a respective process at least once, are, on average, more likely

to sell a process at value $100 + x$ than one at value $100 - x$ (see Figure 2.6). To make selling decisions comparable, we here consider only processes with a drift of zero, for which gains and losses are equally likely.[17]

As we discussed in Section 2.4.3, the disposition effect is consistent with salience theory, but not with the standard specification of the salience model that the salience literature has adopted so far.

Cognitive skills matter. Below-median subjects in terms of cognitive skills — as measured by the sum of correct answers to the modified CRT and the financial literacy questions — are particularly likely to gamble in our experiment. For instance, for the process with a drift of zero, the share of below-median subjects holding the asset until the expiration date is twice as large as the share of above-median subjects doing so (see Figure 2.9 in Appendix E). Notably, both the below- and above-median subjects support Prediction 3: both are responsive to a change in the drift of the process.

2.6.5 Limitations of the model in explaining the data

While the majority of subjects choose mostly strategies inducing right-skewed return distributions, strategies inducing left-skewed return distributions are, unlike what our model predicts, also chosen. This can be explained by noise in the data and by heterogeneity in subjects' susceptibility to salience.

Another issue is that, despite the good fit of our model and our data, we cannot conclusively show that salience indeed is the mechanism that drives our results. Given the strong correlation between skewness effects revealed in static and in dynamic choices, there arguably is one cognitive mechanism that drives all of these skewness effects, but this doesn't have to be salience. In any case, we think of the model as a useful "as if" model.

[17] A similar picture also arises if all selling decisions including those for processes with a negative drift are taken into account; due to losses being more likely for negative drifts, however, the interpretation of the respective findings is less clear, which is why we focus on the fair processes here.

2.7 Discussion of Alternative Models

2.7.1 Expected Utility Theory

In order to explain basic findings in choice under risk — such as an aversion toward symmetric mean-preserving spreads — EUT needs to assume a strictly concave utility function (see, e.g., Bernoulli, 1738; Rothschild and Stiglitz, 1970). Under this assumption, however, EUT predicts that all assets with a non-positive drift will be immediately sold, and it, thus, cannot explain why subjects start to gamble in our experiment (see Result 2 in Section 2.6).

In order to rationalize Result 2 via EUT, we would need to assume that the utility function is convex over at least some range around the initial value of the asset. But, even if we would allow for a completely flexible utility function, which switches back-and-forth from being concave to being convex, EUT cannot explain the skewness-dependence of risk attitudes, as elicited in the static choices between a binary risk and its expected value: here, subjects seek, for *different* outcome levels, sufficiently right-skewed risks, but avoid left-skewed risks (see Figure 2.12 in Appendix E). While EUT could, in principle, rationalize this behavior for *one* outcome level via a utility function that is concave first and then becomes convex, it cannot do so for multiple outcome levels, as the inflection point from concave to convex would have to change with the outcome level. Salience theory, in contrast, predicts skewness-dependent risk attitudes for *any* outcome level (see Appendix D and Dertwinkel-Kalt and Köster, 2020), and is thus consistent with the data. Moreover, EUT — in contrast to salience theory — does in general not explain why subjects prefer loss-exit strategies over gain-exit strategies (Result 1 in Section 2.6). In sum, EUT cannot coherently explain our findings on static and dynamic risk attitudes.

2.7.2 Cumulative Prospect Theory

Abstracting from a finite expiration date, Ebert and Strack (2015) have shown that, under empirically weak assumptions on the probability weighting function, a CPT-agent will never stop an ABM, irrespective of how negative its drift is. This stark never-stopping result follows from the fact that the preference for positive skewness induced by common

CPT-specifications is so strong that the naïve CPT-agent can always find a stop-loss and take-profit strategy that is more attractive than never starting. As we numerically show in Appendix F, at the example of the representative CPT-agent proposed by Amos Tversky and Daniel Kahneman (1992),[18] the never-stopping result extends to processes with a finite expiration date. Consequently, common specifications of CPT can neither rationalize the fact that subjects stop a process with zero drift before the expiration date (i.e., Result 2) nor that stopping behavior is sensitive to the drift of the process (i.e., Result 3).[19] As a consequence of this never-stopping result, CPT is also inconsistent with the disposition effect in a setting like ours (or the one by Ebert and Strack, 2015, as they argue). CPT can, however, also account for Result 1: because a CPT-agent overweights the tails of a probability distribution, he likes the right-skewed distribution generated by loss-exit strategies (this has been also shown in Barberis, 2012; Ebert and Strack, 2015; Heimer et al., 2023).[20]

2.7.3 Reference-Dependent Preferences without Probability Weighting

Barberis and Xiong (2009, 2012) propose an explanation of the disposition effect based on a version of prospect theory *without* probability weighting, according to which gains and losses are experienced at the level of an individual asset in the moment of selling it.[21] Moreover, Barberis and Xiong (2012) derive results that are seemingly similar to the

[18]It is easily verified that the stark never-stopping result extends to finite expiration dates also for other common CPT-specifications. But, for expositional convenience and in line with the related literature (Barberis, 2012; Heimer et al., 2023), we focus on the representative CPT-agent based on the estimates by Amos Tversky and Daniel Kahneman (1992).

[19]CPT belongs to the class of rank-dependent utility models (see, e.g., Quiggin, 1982), which do not assume in general, however, that behavior is reference-dependent and affected by loss aversion. As the never-stopping result of CPT does not rely on either reference-dependence or loss aversion, it extends to a larger class of models within the RDU-family (as shown by Duraj, 2020). But, due to the flexibility of rank-dependent utility models, we do not obtain general predictions regarding the stopping behavior of an RDU-agent in our setup.

[20]The most striking difference between salience theory and alternative approaches — such as EUT and CPT — is that it predicts behavior to be context-dependent in the sense that the evaluation of a given option depends on the alternatives at hand. We ran an additional (pre-registered) experiment that documents context-dependent stopping behavior in line with salience theory. To focus on our main results, and not to disrupt the flow of the main text, we decided to relegate this additional experiment to Appendix G, however.

[21]Barberis and Xiong (2009) show that other, more common reference point specifications (such as annual gains and losses) do not allow CPT to explain the disposition effect.

drift-sensitivity of a naïve salient thinker that we establish in this paper. This apparent similarity, however, is driven by the different setup that they analyze: to establish their result, Barberis and Xiong assume, in particular, that (1) upon selling an asset the agent can immediately reinvest his wealth in another asset, (2) when selling an asset the agent pays positive transaction costs, and (3) the time-horizon is sufficiently long for discounting to play an important role. Our experimental design shares neither of these features, so that their results cannot be applied to our setting. Using a stylized version of the model by Barberis and Xiong (2012), we demonstrate in the following that their *realization-utility* approach, which has found some experimental support (e.g., Imas, 2016), cannot account for our experimental findings.

Without loss of generality, we abstract from a finite expiration date and from discounting. Adapting the model in Barberis and Xiong (2012) to our setup, we assume that the agent's utility is given by the sum of an asset's net present value and her realization utility from selling the asset, where the latter is given by a (piece-wise) linear function $u(\cdot)$ defined as follows: $u(x) = x - r$ if $x \geq r$ and $u(x) = \lambda(x - r)$ if $x < r$ for some loss-aversion parameter $\lambda \geq 1$ and a reference point $r = x_0$.[22] The agent's utility derived from selling the asset at time t is equal to

$$\underbrace{X_t}_{\text{net present value}} + \underbrace{u(X_t).}_{\text{realization utility}}$$

Now consider a threshold stopping time $\tau_{a,b}$ with $a < x_0 < b$, and denote by $p = p(a, b, x_0)$ the probability that the process is stopped at the stop-loss threshold a. The agent sells the asset immediately if and only if, for any such threshold stopping time, it holds that

$$\underbrace{pa + (1-p)b}_{\text{expected net present value}} + \underbrace{p\lambda(a - x_0) + (1-p)(b - x_0)}_{\text{expected realization utility}} \leq x_0,$$

[22] Precisely, the case of $\lambda = 1$ refers to Eq. (7) in Barberis and Xiong (2012), while $\lambda > 1$ corresponds to Eq. (18) in their paper.

or, equivalently,

$$2(1-p)(b-x_0) \leq (1+\lambda)p(x_0-a). \tag{2.4}$$

A sufficient condition for Eq. (2.4) to hold is that $(1-p)(b-x_0) \leq p(x_0-a)$ or, equivalently, $\mathbb{E}[X_{\tau_{a,b}}] \leq x_0$, which is satisfied for any process with a non-positive drift. We conclude that an agent with realization utility à la Barberis and Xiong (2012) would immediately sell any asset in our experiment; that is, their model can neither account for Result 2 nor Result 3.[23]

More generally, the preceding analysis highlights that some form of non-linear probability weighting is necessary to explain our results on skewness seeking, not only in the dynamic selling decisions but also in the static choices, which we analyze in Appendix E (see Figure 2.12). The former point is made in an informal way also in Heimer et al. (2023). Adding non-linear probability weighting to the model by Barberis and Xiong (2012) would yield a model that is essentially equivalent to the ones studied in Barberis (2012) or Ebert and Strack (2015), which we have already discussed in detail in the previous subsection.

2.7.4 Disappointment Aversion

Gul (1991) proposes a theory of disappointment aversion to explain the Allais paradox, in particular, the certainty effect.[24] The model can, in principle, rationalize skewness seeking and thereby gambling in the context of our experiment (Duraj, 2020, Proposition 4). But, as we will formally argue in the following, under the assumptions necessary to explain skewness seeking, it also predicts that subjects will not stop a process with zero drift before the expiration date, which is inconsistent with Result 2.

[23] As the setup in Barberis and Xiong (2012) shows substantial differences to our setup—for instance, asset selling goes along with substantial transaction costs—this is no contradiction to their Figure 1. This does also not change if we use a different variant of their model, namely one where we drop the "expected net present value" term. In that case, only gain-loss utility prevails, and as losses loom larger than gains a symmetric process (or one with a negative drift) cannot be attractive.

[24] Disappointment aversion is a special case of cautious expected utility (Cerreia-Vioglio et al., 2015), which is so flexible, however, that it can explain basically any kind of stopping behavior, including the stark never-stopping result predicted by CPT (see Proposition 6 in Duraj, 2020).

If we abstract from a finite expiration date (i.e., if $T = \infty$ holds), a disappointment-averse agent values the random variable induced by a threshold stopping time $\tau_{a,b}$ at

$$V(X_{\tau_{a,b}}) = \frac{p(1+\beta)}{1-p+p\beta}u(a) + \frac{1-p}{1-p+p\beta}u(b),$$

where u is a classical utility function and $\beta > -1$ captures the agent's disappointment aversion.

As illustrated in Gul (1991), we need to assume $\beta > 0$ in order to rationalize puzzling behavior like the Allais paradox. But, given that $\beta > 0$, the only way to rationalize a preference for sufficiently right-skewed risks is to assume a convex utility function $u(\cdot)$. Precisely, with a concave utility function, the disappointment-averse agent would reject any fairly priced risk, and he would thus sell any asset with a non-positive drift immediately, which contradicts both our results on dynamic (i.e., Result 2) and static choices (see Figure 2.12 in Appendix E).

So, let us assume not only that $\beta > 0$, but also that $u(\cdot)$ is convex. As in our experiment, we assume that the agent can only choose stop-loss and take-profit strategies. A disappointment-averse agent stops a process with zero drift at time t, if and only if, for any stopping time $\tau_{a,b}$,

$$\frac{\frac{u(b)-u(x_t)}{b-x_t}}{\frac{u(x_t)-u(a)}{x_t-a}} \leq 1+\beta.$$

Since $u(\cdot)$ is convex by assumption, the left-hand side of the preceding inequality is strictly increasing in b (and strictly decreasing in a). Again since $u(\cdot)$ is convex, for any fixed $a \geq 0$, the left-hand side approaches infinity, as b approaches infinity. But this implies that, for any fixed $\beta > 0$, we can find a finite b, such that the above inequality is violated. Consequently, a disappointment-averse agent with a convex utility function never stops a process with zero drift, which contradicts the fact that a large majority of subjects stop the process with zero drift before the expiration date (i.e., Result 2). All the preceding arguments carry over to the case of a finite expiration date. In sum, we conclude that a model of disappointment aversion cannot coherently explain the findings on skewness seeking in static and dynamic settings.

2.8 Conclusion

While we find that people take up symmetric gambles if they can obtain skewed return distributions through the choice of their stopping strategies, theoretical considerations suggest similar behavior in case the underlying process is skewed itself. On the one hand, even if the underlying process is negatively skewed, the return distribution associated with the "right" loss-exit strategy is again positively skewed (Ebert, 2020). And if the process itself is positively skewed — which is indeed the case in many real-world applications — our results are likely to be amplified.

A first example refers to processes underlying many casino gambles (as discussed in Ebert and Strack, 2015, Section V) and many asset values, which are not symmetric, but positively skewed. Skewness seeking, as modelled by salience theory, then suggest that consumers gamble or over-invest all the more, as the skewness created with their stopping strategies is exacerbated by the skewness of the process. As an alternative example, we could think of teenagers or young adults who decide whether to pursue the career of a professional athlete, actor, or musician. While the probability of actually making it to the professional level is small, it requires substantial investments of time and other resources to take the shot at becoming a superstar. A teenager who practices excessively for a particular sport, for instance, might as a result neglect school or studies, thereby lowering the attainable wage in the likely case that he fails to become a professional athlete. Now suppose that, as suggested by our model, this teenager adopts the following strategy: each year, he hopes for a breakthrough, but plans to quit sports and instead study otherwise. This strategy generates a positively skewed return distribution, which can be particularly appealing due to the skewness that is inherent to the process of becoming a superstar. After each failure, however, the teenager revises his plans and decides to try it for *one more year*, as this way he can again experience a right-skewed distribution of returns. This idea of excessively pursuing a career is not only consistent with our model but it is also supported by empirical studies (e.g., Choi et al., 2022; Grove et al., 2021). A similar type of argument applies to the classical problem of searching for a job, one of our introductory examples from the classical stopping literature. Here, skewness seeking can explain why people pass on too many mediocre jobs, thereby forgoing pay over a longer

time horizon, in the hope of finding one of very few outstanding jobs with excellent pay. Also in this example, the skewness of the return distribution that results from the chosen stopping strategy is complemented by the skewness of the process itself. In sum, skewness seeking can explain time-inconsistent behavior in trying to reach an elusive goal.

References

Barberis, Nicholas (2012). "A Model of Casino Gambling". In: *Management Science* 58.1, pp. 35–51.

Barberis, Nicholas and Wei Xiong (2009). "What drives the disposition effect? An analysis of a long-standing preference-based explanation". In: *Journal of Finance* 64.2, pp. 751–784.

— (2012). "Realization utility". In: *Journal of Financial Economics* 104.2, pp. 251–271.

Bernoulli, Daniel (1738). "Specimen Theoriae Novae de Mensura Sortis". In: *Comentarii Academiae Scientiarum Imperialis Petropolitanae* 5.2, 175–92 [translated by L. Sommer (1954) in *Econometrica*, 22, 23–36].

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2022). "Salience". In: *Annual Review of Economics* 14.1, pp. 521–544.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). "Salience theory of choice under risk". In: *Quarterly Journal of Economics* 127.3, pp. 1243–1285.

Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva (2015). "Cautious expected utility and the certainty effect". In: *Econometrica* 83.2, pp. 693–728.

Chen, Daniel, Martin Schonger, and Chris Wickens (2016). "oTree - An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

Choi, Darwin, Dong Lou, and Abhiroop Mukherjee (2022). *Superstar Firms and College Major Choice*. Unpublished.

Dertwinkel-Kalt, Markus and Mats Köster (2020). "Salience and Skewness Preferences". In: *Journal of the European Economic Association* 18.5, pp. 2057–2107.

Duraj, Jetlir (2020). *Optimal stopping with general risk preferences*. Unpublished.

Ebert, Sebastian (2015). "On skewed risks in economic models and experiments". In: *Journal of Economic Behavior & Organization* 112, pp. 85–97.

— (2020). *On Taking a Skewed Risk More Than Once*. unpublished.

Ebert, Sebastian and Philipp Strack (2015). "Until the bitter end: On prospect theory in a dynamic context". In: *American Economic Review* 105.4, pp. 1618–1633.

— (2018). *Never, Ever Getting Started: On Prospect Theory Without Commitment*. Unpublished.

Fischbacher, Urs, Gerson Hoffmann, and Simeon Schudy (2017). "The causal effect of stop-loss and take-gain orders on the disposition effect". In: *Review of Financial Studies* 30.6, pp. 2110–2129.

Greiner, Ben (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". In: *Journal of the Economic Science Association* 1.1, pp. 114–125.

Grove, Wayne A, Michael Jetter, and Kerry L Papps (2021). "Career lotto? Labor supply in a superstar market". In: *Journal of Economic Behavior & Organization* 183, pp. 362–376.

Gul, Faruk (1991). "A theory of disappointment aversion". In: *Econometrica* 59, pp. 667–686.

He, X, S Hu, J Obloj, and X Zhou (2019). "Optimal exit time from casino gambling: Strategies of pre-committed and naive gamblers". In: *SIAM Journal on Control and Optimization* 57, pp. 1854–1868.

Heimer, Rawley, Zwetelina Iliewa, Alex Imas, and Martin Weber (2023). *Dynamic Inconsistency in Risky Choice: Evidence from the Lab and Field.*

Henderson, Vicky, David Hobson, and SL Alex (2017). "Randomized strategies and prospect theory in a dynamic context". In: *Journal of Economic Theory* 168, pp. 287–300.

— (2018). "Probability weighting, stop-loss and the disposition effect". In: *Journal of Economic Theory* 178, pp. 360–397.

Huang, Yu-Jui, Adrien Nguyen-Huu, and Xun Yu Zhou (2020). "General Stopping Behaviors of Naïve and Noncommitted Sophisticated Agents, with Application to Probability Distortion". In: *Mathematical Finance* 30.1, pp. 310–340. ISSN: 1467-9965. DOI: 10.1111/mafi.12224.

Imas, Alex (2016). "The realization effect: Risk-taking after realized versus paper losses". In: *American Economic Review* 106.8, pp. 2086–2109.

Kahneman, D and A Tversky (1979). "Prospect theory: An analysis of decision under risk". In: *Econometrica* 47.2, pp. 263–291.

Karni, Edi and Zvi Safra (1990). "Behaviorally consistent optimal stopping rules". In: *Journal of Economic Theory* 51.2, pp. 391–402.

Lusardi, Annamaria and Olivia Mitchell (2011). *Financial Literacy and Planning: Implications for Retirement Wellbeing.* Oxford University Press, pp. 17–39.

Machina, Mark J (1989). "Dynamic consistency and non-expected utility models of choice under uncertainty". In: *Journal of Economic Literature* 27.4, pp. 1622–1668.

Nielsen, Kirby (2019). "Dynamic risk preferences under realized and paper outcomes". In: *Journal of Economic Behavior & Organization* 161, pp. 68–78.

Odean, Terrance (1998). "Are investors reluctant to realize their losses?" In: *Journal of Finance* 53.5, pp. 1775–1798.

Primi, Caterina, Kinga Morsanyi, Francesca Chiesi, Maria Anna Donati, and Jayne Hamilton (2016). "The development and testing of a new version of the cognitive reflection test applying item response theory (IRT)". In: *Journal of Behavioral Decision Making* 29.5, pp. 453–469.

Quiggin, J (1982). "A Theory of Anticipated Utility". In: *Journal of Economic Behavior & Organization* 3, pp. 323–343.

Resnick, Ilyse, Nora S Newcombe, and Thomas F Shipley (2017). "Dealing with big numbers: Representation and understanding of magnitudes outside of human experience". In: *Cognitive science* 41.4, pp. 1020–1041.

Rothschild, Michael and Joseph E Stiglitz (1970). "Increasing risk: I. A definition". In: *Journal of Economic Theory* 2.3, pp. 225–243.

Schkade, D A and D Kahneman (1998). "Does living in California make people happy? A focusing illusion in judgments of life satisfaction". In: *Psychological Science* 9.5, pp. 340–346.

Shefrin, Hersh and Meir Statman (1985). "The disposition to sell winners too early and ride losers too long: Theory and evidence". In: *Journal of Finance* 40.3, pp. 777–790.

Strack, Philipp and Paul Viefers (2021). "Too proud to stop: Regret in dynamic decisions". In: *Journal of the European Economic Association* 19.1, pp. 165–199.

Tversky, Amos and Daniel Kahneman (1992). "Advances in prospect theory: Cumulative representation of uncertainty". In: *Journal of Risk and Uncertainty* 5.4, pp. 297–323.

Weber, Martin and Colin F Camerer (1998). "The disposition effect in securities trading: An experimental analysis". In: *Journal of Economic Behavior & Organization* 33.2, pp. 167–184.

Xu, Z and X Zhou (2013). "Optimal stopping under probability distortion". In: *Annals of Applied Probability* 23.1, pp. 251–282.

2.A Proofs

2.A.1 Preliminary Results on Arithmetic Brownian Motions

Fix an initial wealth level $x \in \mathbb{R}_{>0}$ and an expiration date $T \in \mathbb{R}_{>0}$. Throughout this section, we take the perspective of period $t = 0$ and consider a threshold stopping time $\tau_{a,b}$ with $a < x < b$. Our first result describes the distribution $X_{T\wedge\tau_{a,b}}$ conditional on stopping before expiration.

Lemma 1. *If $\mu \neq 0$, then, for any threshold stopping time $\tau_{a,b}$ with $a < x < b$, we have*

$$\mathbb{P}_0[X_{\tau_{a,b}} = a] = \frac{\exp(-(2\mu/\nu^2)b) - \exp(-(2\mu/\nu^2)x)}{\exp(-(2\mu/\nu^2)b) - \exp(-(2\mu/\nu^2)a)}. \tag{2.5}$$

If $\mu = 0$, then $\mathbb{P}_0[X_{\tau_{a,b}} = a] = \frac{b-x}{b-a}$. In particular, an increase in the drift of the process improves the distribution of $X_{\tau_{a,b}}$ in terms of first-order stochastic dominance.

Proof. Fix some $a, b \in \mathbb{R}_{\geq 0}$ with $a < x < b$. For any threshold stopping time $\tau_{a,b}$, we have

$$\mathbb{P}_0[X_{\tau_{a,b}} = a] = \frac{\Psi(b) - \Psi(x)}{\Psi(b) - \Psi(a)},$$

where $\Psi : \mathbb{R} \to \mathbb{R},\ z \mapsto \Psi(z) = \int_0^z \exp\left(-\int_0^y 2\frac{\mu}{\nu^2}\, dv\right) dy = \int_0^z \exp\left(-2\frac{\mu}{\nu^2}y\right) dy$ is a strictly increasing *scale function* (e.g., Revuz and Yor, 1999, pp. 302). For any $\mu \neq 0$, we obtain

$$\Psi(z) = \int_0^z \exp\left(-2\frac{\mu}{\nu^2}y\right)\ dy = \frac{\nu^2}{2\mu}\Big[1 - \exp(-(2\mu/\nu^2)z)\Big],$$

while for $\mu = 0$, we have $\Psi(z) = \int_0^z 1\ dy = z$, which yields the claim. The last part of the lemma follows from taking the partial derivative of the right-hand side of Eq. (2.5) with respect to μ. □

Our second result derives the probability of reaching the expiration date, and describes several properties of the distribution of $X_{T\wedge\tau_{a,b}}$ conditional on stopping at the expiration date.

Lemma 2. *(a) The probability of stopping at the expiration date equals*

$$\mathbb{P}_0[\tau_{a,b} \geq T | X_0 = x] = \int_a^b q(y, T | X_0 = x)\ dy,$$

where the integrand is given by

$$q(y,T|X_0=x)=\frac{2\exp\left(\frac{\mu(y-x)}{\nu^2}-\frac{T}{2}\frac{\mu^2}{\nu^2}\right)}{(b-a)}\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}.$$

(b) $\lim_{a\to x}\mathbb{P}_0[\tau_{a,b}\geq T|X_0=x]=0$ *and* $\lim_{x\to b}\mathbb{P}_0[\tau_{a,b}\geq T|X_0=x]=0$.

(c) For any stopping time $\tau_{a,b}$ *with* $a<x<b$*, the CDF of* X_T *conditional on* $\tau_{a,b}\geq T$ *equals*

$$\mathbb{P}_0[X_T\leq z|X_0=x,\tau_{a,b}\geq T]=\frac{\int_a^z\exp\left(\frac{\mu(y-x)}{\nu^2}\right)\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}\,dy}{\int_a^b\exp\left(\frac{\mu(y-x)}{\nu^2}\right)\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}\,dy}.$$

Suppose that $a=x-\epsilon$ *and* $b=x+\epsilon'$ *for some* $\epsilon'>\epsilon>0$.

(d) If $\mu<0$*, then* $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_T\leq z|X_0=x,\tau_{a,b}\geq T]<0$ *for any* $z\in[a,b)$.

(e) If $\mu=0$*, then* $\lim_{\epsilon\to0}\mathbb{P}_0[X_T\leq x|X_0=x,\tau_{a,b}\geq T]=0$.

Suppose that $a=x-\epsilon-\epsilon'$ *and* $b=x+\epsilon$ *for some* $\epsilon>0$ *and* $\epsilon'\geq0$*. In addition, let* $\alpha\in(0,\epsilon)$.

(f) If $\mu\leq0$*, then* $\mathbb{P}_0[X_T\leq x-\alpha|X_0=x,\tau_{a,b}\geq T]\geq\mathbb{P}_0[X_T>x+\alpha|X_0=x,\tau_{a,b}\geq T]$*, holding with a strict inequality whenever* $\mu<0$.

Proof. PART (a). Example 5.1 in Cox and Miller (1977).

PART (b). We prove only the first part here, as the proof of the second part is analogous. To establish the first part, it is sufficient to show that

$$\lim_{a\to x}\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}=0. \tag{2.6}$$

As $|\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)|\leq1$ and as $\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\leq\exp\left(-\frac{T}{2}\frac{n\pi\nu^2}{(b-a)^2}\right)$ and as

$$\sum_{n=1}^{\infty}\exp\left(-\frac{t}{2}\frac{n\pi\nu^2}{(b-a)^2}\right)=\frac{1}{\left(\exp\left(\frac{t}{2}\frac{\pi\nu^2}{(b-a)^2}\right)-1\right)}<\infty,$$

we can take the limit in (2.6) inside the summation. The claim follows from the fact that $\sin(0)=0$.

PART (c). Follows immediately from Part (a).

PART (d). Consider a threshold stopping time $\tau_{a,b}$ with $a = x - \epsilon$ and $b = x + \epsilon'$ for some $\epsilon' > \epsilon > 0$. The CDF of the corresponding random variable $X_{\tau_{a,b} \wedge T}$ is given by

$$\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z] = \begin{cases} \mathbb{P}_0[\tau_{a,b} < T] \cdot \mathbb{P}_0[X_{\tau_{a,b}} = a] & \text{if } z = a, \\ \mathbb{P}_0[\tau_{a,b} < T] \cdot \mathbb{P}_0[X_{\tau_{a,b}} = a] + \int_a^z q(y, T|X_0 = x)\ dy & \text{if } a < z < b, \\ 1 & \text{if } z = b. \end{cases}$$

Taking the partial derivative of the CDF at $z \in [a, b)$ with respect to the drift of the process yields

$$\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z] = \mathbb{P}_0[\tau_{a,b} < T] \cdot \frac{\partial}{\partial\mu}\mathbb{P}_0[X_{\tau_{a,b}} = a] + \mathbb{P}_0[X_{\tau_{a,b}} = a] \cdot \frac{\partial}{\partial\mu}\mathbb{P}_0[\tau_{a,b} < T] + \int_a^z \frac{\partial}{\partial\mu} q(y, T|X_0 = x)\ dy.$$

The first two terms in $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z]$ are negative or, at least, non-positive: First, by Lemma 1, $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{\tau_{a,b}} = a] < 0$. Second, because (i) the drift of the process is negative and (ii) $x-a < b-x$ by assumption, a marginal increase in the drift shifts the distribution of $\tau_{a,b}$ upward, so also $\frac{\partial}{\partial\mu}\mathbb{P}_0[\tau_{a,b} < T] \leq 0$. Together these two observations imply that $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq a] < 0$.

The remainder of the proof proceeds in two steps: First, we will show that *there exists some* $\hat{z} \in (a, b]$ *such that* $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z] < 0$ *if and only if* $z < \hat{z}$. Second, we will show that $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} = b] > 0$ *and, thus,* $\frac{\partial}{\partial\mu}\Big\{1 - \mathbb{P}_0[X_{T\wedge\tau_{a,b}} = b]\Big\} < 0$*, which in turn implies that* $\hat{z} = b$.

1. Step: Using the definition of $q(y, T|X_0 = x)$, as provided in Part (a), we obtain

$$\begin{aligned}\int_a^z \frac{\partial}{\partial\mu} q(y, T|X_0 = x)\ dy &= \frac{1}{\nu^2}\left[\int_a^z yq(y, T|X_0 = x)\ dy - (x + T\mu)\int_a^z q(y, T|X_0 = x)\ dy\right] \\ &= \frac{\mathbb{P}_0\Big[\tau_{a,b} \geq T, X_T \leq z\Big]}{\nu^2}\left[\mathbb{E}_0\Big[X_T|\tau_{a,b} \geq T, X_T \leq z\Big] - \mathbb{E}_0\Big[X_T\Big]\right].\end{aligned}$$

We distinguish two cases, depending on whether the difference in brackets is positive or not. If $\mathbb{E}_0\Big[X_T|\tau_{a,b} \geq T, X_T \leq z\Big] \leq \mathbb{E}_0\Big[X_T\Big]$, then $\int_a^z \frac{\partial}{\partial\mu} q(y, T|X_0 = x)\ dy \leq 0$, which together with our observations on the first two terms of $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z]$ implies that $\frac{\partial}{\partial\mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z] < 0$. Otherwise, because $\mathbb{E}_0\Big[X_T|\tau_{a,b} \geq T, X_T \leq z\Big]$ and $\mathbb{P}_0\Big[\tau_{a,b} \geq$

$T, X_T \leq z\Big]$ are increasing in z, it follows that $\int_a^z \frac{\partial}{\partial \mu} q(y, T|X_0 = x)\, dy$ increases in z whenever it is positive. In sum, there exists some $\hat{z} \in (a, b]$ such that $\frac{\partial}{\partial \mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} \leq z] < 0$ if and only if $z < \hat{z}$, which was to be proven.

2. Step: Now, it is sufficient to show that $\frac{\partial}{\partial \mu}\mathbb{P}_0[X_{T\wedge\tau_{a,b}} = b] > 0$. For the sake of a contradiction, we assume that $\mathbb{P}_0[X_{T\wedge\tau_{a,b}} = b|\mu = \mu'] \geq \mathbb{P}_0[X_{T\wedge\tau_{a,b}} = b|\mu = \mu'']$ for some $\mu' < \mu'' < 0$.

Denote as $M_t := \max_{0\leq s\leq t} X_s$ the running maximum of the process. By our assumption toward a contradiction, it has to be true that, for some $t \in (0, T)$,

$$\mathbb{P}_0[M_t \geq b|X_s > a, \forall s \leq t, \mu = \mu'] \geq \mathbb{P}_0[M_t \geq b|X_s > a, \forall s \leq t, \mu = \mu''];$$

otherwise, reaching the upper threshold b would be strictly more likely under the process with a less negative drift. Recall that, because both $\mathbb{P}_0[X_{\tau_{a,b}} = a]$ and $\mathbb{P}_0[\tau_{a,b} < T]$ (weakly) decrease with μ, increasing the drift of the process makes it less likely to reach the lower threshold a. Hence, a necessary condition for our assumption toward a contradiction to hold is

$$\mathbb{P}_0[M_t \geq b|\mu = \mu'] \geq \mathbb{P}_0[M_t \geq b|\mu = \mu''];$$

that is, at any time $t \in (0, T)$, conditioning on not having reached the lower threshold is more restrictive for the process with a more negative drift. By the same argument, we can ignore the fact that the process is absorbing in zero. In this case, by the Reflection Principle and Girsanov's Theorem, for an arbitrary drift $\mu < 0$, the distribution of the running maximum, M_t, satisfies

$$\mathbb{P}_0[M_t \geq b] = \exp\left(\frac{2\epsilon'\mu}{\nu^2}\right)\left[1 - \Phi\left(\frac{\epsilon' + \mu t}{\sqrt{t}\nu}\right)\right] + \left[1 - \Phi\left(\frac{\epsilon' - \mu t}{\sqrt{t}\nu}\right)\right],$$

with $b = x + \epsilon'$ for some $\epsilon' > 0$ and with Φ being the standard normal CDF (see Example

5.1 in Cox and Miller, 1977, or Corollary 7.2.2 in Shreve, 2004). Hence, we obtain

$$\begin{aligned}\frac{\partial}{\partial\mu}\mathbb{P}_0[M_t \geq b] &= \frac{2\epsilon'}{\nu^2}\exp\left(\frac{2\epsilon'\mu}{\nu^2}\right)\left[1-\Phi\left(\frac{\epsilon'+\mu t}{\sqrt{t}\nu}\right)\right] - \exp\left(\frac{2\epsilon'\mu}{\nu^2}\right)\phi\left(\frac{\epsilon'+\mu t}{\sqrt{t}\nu}\right) + \phi\left(\frac{\epsilon'-\mu t}{\sqrt{t}\nu}\right) \\ &= \frac{2\epsilon'}{\nu^2}\exp\left(\frac{2\epsilon'\mu}{\nu^2}\right)\left[1-\Phi\left(\frac{\epsilon'+\mu t}{\sqrt{t}\nu}\right)\right] + \frac{\sqrt{t}}{\sqrt{2\pi}\nu}\left[\exp\left(\frac{(\epsilon'-\mu t)^2}{2t\nu^2}\right) - \exp\left(\frac{(\epsilon'-\mu t)^2}{2t\nu^2}\right)\right] \\ &= \frac{2\epsilon'}{\nu^2}\exp\left(\frac{2\epsilon'\mu}{\nu^2}\right)\left[1-\Phi\left(\frac{\epsilon'+\mu t}{\sqrt{t}\nu}\right)\right] \\ &> 0,\end{aligned}$$

where the second equality follows from plugging in the density ϕ of a standard normal distribution. But this implies that $\mathbb{P}_0[M_t \geq b|\mu=\mu'] \geq \mathbb{P}_0[M_t \geq b|\mu=\mu'']$ cannot hold; a contradiction.

PART (e). By Part (b), we have $\lim_{\epsilon\to 0}\int_{x-\epsilon}^{x+\epsilon'} q(y,T|X_0=x)\,dy = 0$ and, as a consequence, also $\lim_{\epsilon\to 0}\int_{x-\epsilon}^{x} q(y,T|X_0=x)\,dy = 0$. Now, to determine $\lim_{\epsilon\to 0}\ \mathbb{P}_0[X_T \leq x|X_0=x,\tau_{a,b}\geq T]$, we will apply L'Hospital's rule. For that, we have to make a few preliminary observations.

First, if the partial derivative $\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)$ exists, then it is given by

$$\begin{aligned}\frac{\partial}{\partial\epsilon}q(y,T|X_0=x) = &\frac{\epsilon'}{(\epsilon+\epsilon')^2}\sum_{n=1}^{\infty}\left\{\pi n\cos\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\sin\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\} \\ &+\frac{\epsilon'-(y-x)}{(\epsilon+\epsilon')^2}\sum_{n=1}^{\infty}\left\{\pi n\sin\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\cos\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\} \\ &+\frac{T\nu^2}{(\epsilon+\epsilon')^3}\sum_{n=1}^{\infty}\left\{\pi^2n^2\sin\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\sin\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\},\end{aligned}$$

and since we have

$$\begin{aligned}\left|\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)\right| \leq &\left[\frac{1}{\epsilon'}+\frac{\epsilon'-2(y-x)}{\epsilon'^2}\right]\frac{\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)\pi}{\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{4\epsilon'^2}\right)-1\right)^2} \\ &+\frac{T\nu^2}{\epsilon'^3}\frac{\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)+1\right)\pi^2}{\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{4\epsilon'^2}\right)-1\right)^3} < \infty,\end{aligned} \tag{2.7}$$

it indeed exists. To apply L'Hospital's rule, we need to compute the limit for ϵ approach-

ing zero:

$$\begin{aligned}\lim_{\epsilon\to 0}\ \frac{\partial}{\partial\epsilon}q(y,T|X_0=x) &= \frac{1}{\epsilon'}\sum_{n=1}^{\infty}\lim_{\epsilon\to 0}\left\{\pi n\cos\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\sin\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\}\\ &+\frac{\epsilon'-(y-x)}{\epsilon'^2}\sum_{n=1}^{\infty}\lim_{\epsilon\to 0}\left\{\pi n\sin\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\cos\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\}\\ &+\frac{T\nu^2}{\epsilon'^3}\sum_{n=1}^{\infty}\lim_{\epsilon\to 0}\left\{\pi^2n^2\sin\left(\frac{\pi n\epsilon}{\epsilon+\epsilon'}\right)\sin\left(\frac{\pi n(y-x+\epsilon)}{\epsilon+\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(\epsilon+\epsilon')^2}\right)\right\}\\ &=\frac{1}{\epsilon'}\sum_{n=1}^{\infty}\left\{\pi n\sin\left(\frac{\pi n(y-x)}{\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{\epsilon'^2}\right)\right\}\geq 0,\end{aligned} \tag{2.8}$$

where the first equality follows from the fact that, by the considerations in Eq. (2.7), we can take the limits into the summations; the second equality holds as $\sin(0)=0$ and $\cos(0)=1$; and the inequality follows from the fact that $q(y,T|X_0=x)\geq 0$ and $\lim_{\epsilon\to 0}q(y,T|X_0=x)=0$, because otherwise $q(y,T|X_0=x)$ would be negative for ϵ sufficiently close to zero.

Second, we observe that

$$\begin{aligned}&\left|\int_{x-\epsilon}^{x}\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)\ dy\right|\leq\int_{x-\epsilon}^{x}\left|\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)\right|\ dy\\ &\quad\leq\epsilon\left[\frac{1}{\epsilon'}+\frac{\epsilon'-2(y-x)}{\epsilon'^2}\right]\frac{\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)\pi}{\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{4\epsilon'^2}\right)-1\right)^2}+\epsilon\frac{T\nu^2}{\epsilon'^3}\frac{\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{\epsilon'^2}\right)+1\right)\pi^2}{\left(\exp\left(\frac{T}{2}\frac{\pi\nu^2}{4\epsilon'^2}\right)-1\right)^3}\ \xrightarrow{\epsilon\to 0}\ 0,\end{aligned} \tag{2.9}$$

where the first inequality follows by the triangle inequality, and where the second inequality follows from Eq. (2.7). Taking the limit of the final expression is straightforward.

Third, we conclude that

$$\begin{aligned}\lim_{\epsilon\to 0}\int_{x}^{x+\epsilon'}\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)\ dy &= \int_{x}^{x+\epsilon'}\lim_{\epsilon\to 0}\frac{\partial}{\partial\epsilon}q(y,T|X_0=x)\ dy\\ &=\frac{1}{\epsilon'}\int_{x}^{x+\epsilon'}\sum_{n=1}^{\infty}\left\{\pi n\sin\left(\frac{\pi n(y-x)}{\epsilon'}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{\epsilon'^2}\right)\right\}\ dy\\ &=\frac{1}{\epsilon'}\int_{0}^{1}\sum_{n=1}^{\infty}\left\{\sin(\pi nz)\,\pi n\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{\epsilon'^2}\right)\right\}\ dz,\end{aligned}$$

where the first equality holds by the Theorem of Dominated Convergence, the second one holds by the second to last line in (2.8), and the third equality follows by substitution.

Recall that

$$\sum_{n=1}^{\infty}\left\{\sin(\pi n z)\,\pi n \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{\epsilon'^2}\right)\right\} \geq 0 \tag{2.10}$$

for any $z \in (0,1)$, and notice that this inequality is strict for any $z = \frac{1}{k}$ with $k \in \mathbb{N}_{\geq 2}$. The latter follows from the fact that $\sin(\pi n \frac{i}{k}) = -\sin(\pi n \frac{k+i}{k})$ for any $i \leq k$, and $\sin(\pi \frac{i}{k}) \geq 0$ for any $i \leq k$ with a strict inequality for any $i \notin \{0,k\}$, and $\pi n \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{\epsilon'^2}\right)$ being strictly decreasing in n. Since (2.10) is continuous in z, we conclude that it is strictly positive on a dense interval around any $z = \frac{1}{k}$ with $k \in \mathbb{N}_{\geq 2}$. This, in turn, implies that $\lim_{\epsilon\to 0}\int_x^{x+\epsilon'} \frac{\partial}{\partial\epsilon} q(y,T|X_0 = x)\,dy > 0$.

Combining all the considerations above, we finally conclude that

$$\begin{aligned}
\lim_{\epsilon\to 0}\mathbb{P}_0[X_T \leq x | X_0 = x, \tau_{a,b} \geq T] &= \lim_{\epsilon\to 0}\frac{\frac{\partial}{\partial\epsilon}\int_{x-\epsilon}^{x} q(y,T|X_0=x)\,dy}{\frac{\partial}{\partial\epsilon}\int_{x-\epsilon}^{x+\epsilon'} q(y,T|X_0=x)\,dy} \\
&= \lim_{\epsilon\to 0}\frac{\int_{x-\epsilon}^{x}\frac{\partial}{\partial\epsilon} q(y,T|X_0=x)\,dy}{\int_{x-\epsilon}^{x+\epsilon'}\frac{\partial}{\partial\epsilon} q(y,T|X_0=x)\,dy} \\
&= \frac{\lim_{\epsilon\to 0}\int_{x-\epsilon}^{x}\frac{\partial}{\partial\epsilon} q(y,T|X_0=x)\,dy}{\lim_{\epsilon\to 0}\int_{x-\epsilon}^{x+\epsilon'}\frac{\partial}{\partial\epsilon} q(y,T|X_0=x)\,dy} \\
&= 0,
\end{aligned}$$

where the first equality follows by L'Hospital's rule (given that the limit on the right-hand side exists), the second equality follows by the Theorem of Dominated Convergence and by Leibniz's integral rule, the third equality follows from the fact that the limit of the numerator and the limit of the denominator exist, and the last equality holds by (2.9) and by the fact that $\lim_{\epsilon\to 0}\int_x^{x+\epsilon'}\frac{\partial}{\partial\epsilon} q(y,T|X_0 = x)\,dy > 0$ and therefore, by $\lim_{\epsilon\to 0}\frac{\partial}{\partial\epsilon} q(y,T|X_0 = x) \geq 0$, also $\lim_{\epsilon\to 0}\int_{x-\epsilon}^{x+\epsilon'}\frac{\partial}{\partial\epsilon} q(y,T|X_0 = x)\,dy > 0$. This completes the proof.

PART (f). To begin with, let $\epsilon' = 0$. By Part (c), we have to show that

$$\begin{aligned}
&\int_a^{x-\alpha}\exp\left(\frac{\mu(y-x)}{\nu^2}\right)\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}\,dy \\
&\quad\geq \int_{x+\alpha}^{b}\exp\left(\frac{\mu(y-x)}{\nu^2}\right)\sum_{n=1}^{\infty}\left\{\sin\left(\frac{\pi n(x-a)}{b-a}\right)\sin\left(\frac{\pi n(y-a)}{b-a}\right)\exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\}\,dy
\end{aligned}$$

for any $\alpha \in (0, \epsilon)$, with a strict inequality if $\mu < 0$. For any $\mu \leq 0$, we have $\exp\left(\frac{\mu(y-x)}{\nu^2}\right) \geq 1$ if and only if $y \leq x$, holding with a strict inequality whenever $y < x$ and $\mu < 0$. This implies that

$$\begin{aligned}
&\int_a^{x-\alpha} \exp\left(\frac{\mu(y-x)}{\nu^2}\right) \sum_{n=1}^{\infty} \left\{ \sin\left(\frac{\pi n(x-a)}{b-a}\right) \sin\left(\frac{\pi n(y-a)}{b-a}\right) \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\} \, dy \\
&\quad \geq \int_{\frac{\pi n}{2}}^{\frac{\pi n}{2}\frac{\alpha}{\epsilon}} \sum_{n\in\mathbb{N}, n \text{ odd}} \sin\left(\frac{\pi n}{2}\right) \sin\left(\frac{\pi n}{2} - z\right) \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{4\epsilon^2}\right) \, dz \\
&\quad = \int_{\frac{\pi n}{2}\frac{\alpha}{\epsilon}}^{\frac{\pi n}{2}} \sum_{n\in\mathbb{N}, n \text{ odd}} \sin\left(\frac{\pi n}{2}\right) \sin\left(\frac{\pi n}{2} + z\right) \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{4\epsilon^2}\right) \, dz \\
&\quad \geq \int_{x+\alpha}^{b} \exp\left(\frac{\mu(y-x)}{\nu^2}\right) \sum_{n=1}^{\infty} \left\{ \sin\left(\frac{\pi n(x-a)}{b-a}\right) \sin\left(\frac{\pi n(y-a)}{b-a}\right) \exp\left(-\frac{T}{2}\frac{n^2\pi^2\nu^2}{(b-a)^2}\right)\right\} \, dy,
\end{aligned}$$

where the two inequalities follow from the fact that $\frac{x-a}{b-a} = \frac{1}{2}$ and $\sin\left(\frac{\pi n}{2}\right) = 0$ for any even $n \in \mathbb{N}$, while the equality holds since $\sin\left(\frac{\pi n}{2} - z\right) = \sin\left(\frac{\pi n}{2} + z\right)$ for any odd $n \in \mathbb{N}$ and any $z \in (0, \frac{\pi n}{2})$. The claim follows from the fact that the inequalities are strict whenever $\mu < 0$.

Fix some $\epsilon > 0$ and $\mu \leq 0$. Now, if $\epsilon' > 0$, the probability that X_T is weakly below x, conditional on reaching the expiration date when playing according to the stopping time $\tau_{a,b}$, $\mathbb{P}_0[X_T \leq x | X_0 = x, \tau_{a,b} \geq T]$, increases compared to the case with $\epsilon' = 0$. This follows from the fact that due to $\epsilon' > 0$ there is now more room below x than above x to reach the expiration date T and from the continuity of the sample paths. □

2.A.2 Motivating Example

Proof of Proposition 1. We have to show that, when the value function is linear, *then for any point in time $t < T$ with $X_t = x_t$ there exists a stopping time $\tau_{a,b}$ such that $U^s\left(X_{T\wedge\tau_{a,b}} | \mathcal{C}\right) > x_t$ or, equivalently,*

$$\begin{aligned}
\mathbb{P}_0[\tau_{a,b} < T - t] \cdot \underbrace{\left[p(a - x_t)\sigma(a, x_t) + (1-p)(b - x_t)\sigma(b, x_t)\right]}_{(\star)} \\
+ \mathbb{P}_0[\tau_{a,b} \geq T - t] \cdot \underbrace{\int_{(a,b)} (z - x_t)\sigma(z, x_t) \, d\Phi_\mu(z)}_{(\star\star)} > 0,
\end{aligned}$$

where the probability $p = p(a, b, \mu)$ *is defined in Eq.* (2.5) *and where* $\mathbb{P}_0[\tau_{a,b} < T - t]$ *as well as the conditional CDF* $\Phi_\mu(z) := \mathbb{P}_0[X_T \leq z | X_0 = x_0, \tau_{a,b} \geq T - t]$ *are described in Lemma 2.*

Consider a threshold stopping time $\tau_{a,b}$ with $a = x_t - \epsilon$ and $b = x_t + \epsilon'$ for some $\epsilon' > \epsilon > 0$. First, we show that *there exists some threshold* $\hat{\epsilon} > 0$ *such that for any* $\epsilon < \hat{\epsilon}$, *it holds that* $(\star) > 0$. Since $p = \frac{b - x_t}{b - a} = \frac{\epsilon'}{\epsilon + \epsilon'}$, it follows that $(\star) > 0$ holds if and only if

$$\frac{\epsilon\epsilon'}{\epsilon + \epsilon'}\Big[\sigma(x_t + \epsilon', x_t) - \sigma(x_t - \epsilon, x_t)\Big] > 0.$$

The claim then follows from the fact that — due to ordering — the salience weight $\sigma(x_t - \epsilon, x_t)$ monotonically increases in ϵ, and $\sigma(x_t + \epsilon', x_t) > \sigma(x_t, x_t)$ holds.

Second, we show that *there exists some* $\check{\epsilon} > 0$, *such that for any* $\epsilon < \check{\epsilon}$, $(\star\star) > 0$. We have

$$\begin{aligned}
\int_{(a,b)} (z - x_t)\sigma(z, x_t)\, d\Phi_\mu(z) &\geq \int_{(x-\epsilon,x)} (z - x_t)\sigma(z, x_t)\, d\Phi_\mu(z) + \int_{(x+\epsilon,x+\epsilon')} (z - x_t)\sigma(z, x_t)\, d\Phi_\mu(z) \\
&> -\epsilon\overline{\sigma} \int_{(x-\epsilon,x)} d\Phi_\mu(z) + \epsilon\underline{\sigma} \int_{(x+\epsilon,x+\epsilon')} d\Phi_\mu(z) \\
&= \epsilon\Big[\Big(1 - \Phi_\mu(x + \epsilon)\Big)\underline{\sigma} - \Phi_\mu(x)\overline{\sigma}\Big],
\end{aligned}$$

where the weak inequality holds as $\epsilon > 0$ and the strict inequality follows by the definition of $\overline{\sigma} := \sup_{(x,y)\in\mathbb{R}^2_{\geq 0}} \sigma(x, y)$ and $\underline{\sigma} := \inf_{(x,y)\in\mathbb{R}^2_{\geq 0}} \sigma(x, y)$. Now recall that $\overline{\sigma} < \infty$ and $\underline{\sigma} > 0$ by assumption. By Lemma 2 (e), we have $\lim_{\epsilon\to 0} \Phi_\mu(x) = 0$, which yields the claim. □

2.A.3 Main Theoretical Result

Proof of Lemma **??**. We have to *find some* $\alpha, \beta \in \mathbb{R}_{>0}$, *so that* $\Big[\tilde{u}(z) + \beta\Big] \leq \Big[\tilde{u}(x) + \beta\Big] \exp\Big(\alpha(z - x)\Big)$ *holds for any* $z \geq 0$. Set $\beta = \frac{\sigma(v(x),v(x))v'(x)}{\alpha}$ for some $\alpha > 0$. We need to find some $\alpha > 0$ so that

$$\Big(v(z) - v(x)\Big)\sigma\Big(v(z), v(x)\Big) + \beta \leq \Big[\Big(v(x) - v(x)\Big)\sigma\Big(v(x), v(x)\Big) + \beta\Big] \exp\Big(\alpha(z - x)\Big)$$

holds for all $z \geq 0$. This condition is indeed satisfied if and only if

$$\alpha \frac{v(z)-v(x)}{v'(x)} \frac{\sigma\Big(v(z),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} + 1 \leq \exp\Big(\alpha(z-x)\Big)$$

or, equivalently,

$$\frac{v(z)-v(x)}{v'(x)} \frac{\sigma\Big(v(z),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} \leq \frac{\exp\Big(\alpha(z-x)\Big)-1}{\alpha} \tag{2.11}$$

holds for all $z \geq 0$. By construction, (2.11) holds at $z = x$. We distinguish two cases:

1. CASE: Let $z < x$. Divide both sides of (2.11) by $z - x < 0$, which gives

$$\frac{\frac{v(x)-v(z)}{x-z}}{v'(x)} \frac{\sigma\Big(v(z),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} \geq \frac{\frac{\exp\Big(\alpha(x-x)\Big)-\exp\Big(\alpha(z-x)\Big)}{x-z}}{\alpha}. \tag{2.12}$$

Since the exponential function is strictly convex, such that, for any $z < x$, we have

$$\frac{\exp\Big(\alpha(x-x)\Big)-\exp\Big(\alpha(z-x)\Big)}{x-z} < \alpha \exp\Big(\alpha(x-x)\Big) = \alpha,$$

the right-hand side of (2.12) is strictly is less than 1. Since the value function is (weakly) concave, which implies that, for any $z < x$, we have

$$\frac{v(x)-v(z)}{x-z} \geq v'(x),$$

and since $\sigma\Big(v(z),v(x)\Big) > \sigma\Big(v(x),v(x)\Big)$ holds by ordering, the left-hand side of (2.12) is strictly larger than 1. In sum, we conclude that, for any $z < x$, Condition (2.11) is satisfied for any $\alpha > 0$.

2. CASE: Let $z > x$. Since both sides of (2.11) are zero at $z = x$, we can re-write (2.11) as follows

$$\int_x^z \frac{\partial}{\partial w}\left[\frac{\exp\Big(\alpha(w-x)\Big)-1}{\alpha}\right] - \frac{\partial}{\partial w}\left[\frac{v(w)-v(x)}{v'(x)} \frac{\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}\right] dw \geq 0,$$

which holds if and only if

$$\int_x^z \exp\Big(\alpha(w-x)\Big) - \left[\frac{v'(w)}{v'(x)}\frac{\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} + \Big[v(w)-v(x)\Big]\frac{v'(w)}{v'(x)}\frac{\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}\right] dw \geq 0. \tag{2.13}$$

A sufficient condition for (2.13) to hold is that

$$\exp\Big(\alpha(w-x)\Big) \geq \frac{v'(w)}{v'(x)}\frac{\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} + \Big[v(w)-v(x)\Big]\frac{v'(w)}{v'(x)}\frac{\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}$$

for any $w \geq x$. When evaluated at $w = x$, this inequality is tight, since the salience function is differentiable and thus $\lim_{w\to x}\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big) = 0$. Also, if the right-hand side of this inequality is non-positive, the condition is certainly met. So, from now on, consider only $w > x$ for which the right-hand side is positive. Then, we can re-state the condition as follows

$$\alpha \geq \frac{\ln\left(\frac{v'(w)}{v'(x)}\frac{\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)} + \Big[v(w)-v(x)\Big]\frac{v'(w)}{v'(x)}\frac{\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}\right)}{w-x}, \tag{2.14}$$

which has to hold for all relevant $w > x$. The right-hand side of (2.14) is bounded from above by

$$h(w) := \frac{\ln\left(\frac{\sigma\Big(v(w),v(x)\Big)+\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}\right)}{w-x},$$

since v is (weakly) concave. Hence, a sufficient condition for (2.14) to hold is given by $\alpha \geq \max_{w\in(x,\infty)} h(w)$. Since we are free to choose any $\alpha > 0$, it is thus sufficient to show that $\max_{w\in(x,\infty)} h(w) < \infty$. First, since $h(w) \geq 0$ for any $w \geq x$, we know that, if the limit $\lim_{w\to\infty} h(w)$ does not exist, then it has to be positive infinity. Then, by

L'Hospital's rule, we conclude that

$$
\begin{aligned}
0 \leq \lim_{w\to x} h(w) &= \lim_{w\to x} \frac{\frac{2v'(w)\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)+v'(w)\Big[v(w)-v(x)\Big]\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}}{\frac{\sigma\Big(v(w),v(x)\Big)+\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(x),v(x)\Big)}} \\
&= \frac{\lim_{w\to x} 2v'(w)\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)+v'(w)\Big[v(w)-v(x)\Big]\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big)}{\lim_{w\to x}\sigma\Big(v(w),v(x)\Big)+\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)} \\
&= \frac{2v'(x)}{\sigma\Big(v(x),v(x)\Big)}\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)\Big|_{w=x} < \infty.
\end{aligned}
$$

where the third equality as well as the inequality follow from the fact that the salience function is twice differentiable, which implies, in particular, $\lim_{w\to x}\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big) = 0$ as well as $\lim_{w\to x}\Big[v(w)-v(x)\Big]\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big) = 0$. The fact that the limit exists further justifies the application of L'Hospital's rule. Second, again by L'Hospital's rule, we have

$$
\begin{aligned}
0 \leq \lim_{w\to\infty} h(w) &= \lim_{w\to\infty} \frac{2v'(w)\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)+v'(w)\Big[v(w)-v(x)\Big]\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big)}{\sigma\Big(v(w),v(x)\Big)+\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)} \\
&\leq \lim_{w\to\infty} \frac{2v'(w)\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)+v'(w)\Big[v(w)-v(x)\Big]\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big)}{\Big[v(w)-v(x)\Big]\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)} \\
&= \lim_{w\to\infty}\left\{\frac{2v'(w)}{v(w)-v(x)}+v'(w)\frac{\frac{\partial^2}{\partial v(w)^2}\sigma\Big(v(w),v(x)\Big)}{\frac{\partial}{\partial v(w)}\sigma\Big(v(w),v(x)\Big)}\right\} \\
&\leq \lim_{w\to\infty}\frac{2v'(w)}{v(w)-v(x)} = 0,
\end{aligned}
$$

where the second inequality follows from the fact that $\lim_{\Delta\to\infty}\frac{\partial^2}{\partial\Delta^2}\sigma(x+\Delta,x) \leq 0$, as otherwise $\sigma(x+\Delta,x)$ could not be strictly increasing in Δ on $(0,\infty)$ and bounded from above.

Since $h(z)$ is continuous on (x,∞), it follows from $\lim_{w\to x} h(w) < \infty$ and $\lim_{w\to\infty} h(w) < \infty$ that $\max_{w\in(x,\infty)} h(w)$ exists. This, in turn, implies that there exists a constant $\alpha > 0$, such that Condition (2.13) is satisfied for any $z > x$, which was to be proven. □

Proof of Theorem 1. The statement follows from Proposition 1 in Ebert and Strack (2015). We re-state their argument here in terms of our notation. By Lemma **??**, the auxiliary utility function is of exponential growth at $z = x$, so that we can find $\alpha, \beta \in \mathbb{R}_{>0}$ such that, for any $z \geq 0$, we have

$$\Big[\tilde{u}(z) + \beta\Big] \leq \Big[\tilde{u}(x) + \beta\Big] \exp\Big(\alpha(z - x)\Big). \tag{2.15}$$

Recall that the preferences of an EUT agent are invariant under positive affine transformations, which implies that the utility function $\hat{u}(z) := \tilde{u}(z) + \beta$ represents the exact same preferences. We should also keep in mind that $\hat{u}(x) = \beta > 0$.

Consider an EUT agent with a utility function $\hat{u}$, and an ABM $X_t = x + \mu t + \nu W_t$ with an initial value x and a drift $\mu < -\frac{1}{2}\alpha\nu^2 =: \tilde{\mu}$. For any stopping time τ with $\mathbb{P}_0[\tau > 0] > 0$, we have

$$\begin{aligned}
\mathbb{E}[\hat{u}(X_\tau)] &\leq \hat{u}(x)\mathbb{E}\Big[\exp\Big(\alpha(X_\tau - x)\Big)\Big] \\
&= \hat{u}(x)\mathbb{E}\Big[\exp\Big(\alpha\mu\tau + \alpha\nu W_\tau\Big)\Big] \\
&= \hat{u}(x)\mathbb{E}\left[1 + \int_0^\tau \left(\alpha\mu + \frac{1}{2}\alpha^2\nu^2\right)\exp\Big(\alpha\mu s + \alpha\nu W_s\Big)\, ds + \int_0^\tau \alpha\nu \exp\Big(\alpha\mu s + \alpha\nu W_s\Big)\, dW_s\right] \\
&= \hat{u}(x)\mathbb{E}\left[1 + \int_0^\tau \left(\alpha\mu + \frac{1}{2}\alpha^2\nu^2\right)\exp\Big(\alpha\mu s + \alpha\nu W_s\Big)\, ds\right] \\
&< \hat{u}(x),
\end{aligned}$$

where the first inequality holds by (2.15), the second equality holds by Itô's Lemma, the third equality holds by Doob's Optional Sampling Theorem, and the second inequality holds by $\hat{u}(x) > 0$ and the assumption that $\mu < \tilde{\mu}$ (so that the expectation is less than one). Hence, an EUT agent with a utility function $\hat{u}$ and, thus, the naïve salient thinker stop immediately. □

Proof of Corollary 1. Consider an Arithmetic Brownian Motion $X_t = x + \mu t + \nu W_t$ with a negative drift, which a naïve salient thinker with a linear value function would stop immediately. We transform the process using the strictly increasing *scale function* (Revuz

and Yor, 1999, pp. 302)

$$\Psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, z \mapsto \int_0^z \exp\left(-2\frac{\mu}{\nu^2}y\right) dy,$$

which yields a scaled process $(\Psi(X_t))_{t\in\mathbb{R}_{\geq 0}}$ with zero drift and an initial value $\Psi(x)$.

Now consider a salient thinker with the exact same salience function, but a value function $v(z) = \frac{\nu}{2|\mu|} \ln\left(1 + \frac{2|\mu|}{\nu} z\right)$, which is strictly increasing and concave. Since $v(z) = \Psi^{-1}(z)$, we conclude that, for any stopping time τ, it has to hold that

$$\mathbb{E}\Big[\Big(v\Big(\Psi(X_\tau)\Big) - v\Big(\Psi(x)\Big)\Big)\sigma\Big(v\Big(\Psi(X_\tau)\Big), v\Big(\Psi(x)\Big)\Big] = \mathbb{E}\Big[\Big(X_\tau - x\Big)\sigma(X_\tau, x)\Big] \leq 0,$$

where the inequality follows from the assumption that the naïve salient thinker with a linear value function immediately stops the process $(X_t)_{t\in\mathbb{R}_{\geq 0}}$. Consequently, the naïve salient thinker with a value function $v(\cdot)$ immediately stops the scaled process $(\Psi(X_t))_{t\in\mathbb{R}_{\geq 0}}$ with zero drift. □

2.A.4 Additional Results on Stop-Loss and Take-Profit Strategies

Proof of Proposition 2. Consider a stop-loss and take-profit strategy that corresponds to the threshold stopping time $\tau_{a,b}$ with $a := x_t - \epsilon - \epsilon'$ and $b := x_t + \epsilon$ for some $\epsilon > 0$, $\epsilon' \geq 0$, and $\epsilon + \epsilon' \leq x_t$, and that is therefore not a loss-exit strategy. Again, we denote as

$$\Phi_\mu(z) := \mathbb{P}_t[X_T \leq z | X_t = x_t, \tau_{a,b} \geq T]$$

the CDF of X_T conditional on reaching the expiration date. Then, it follows that

$$\begin{aligned}
U^s\Big(X_{T\wedge\tau_{a,b}}|\mathcal{C}\Big) - v(x_t) \propto \mathbb{P}_t[\tau_{a,b} < T] \times \Big[& p\Big(v(x_t-\epsilon-\epsilon') - v(x_t)\Big)\sigma\Big(v(x_t-\epsilon-\epsilon'), v(x_t)\Big) \\
& + (1-p)\Big(v(x_t+\epsilon) - v(x_t)\Big)\sigma\Big(v(x_t+\epsilon), v(x_t)\Big)\Big] \\
& + \mathbb{P}_t[\tau_{a,b} \geq T] \times \int_{(a,b)} \Big(v(z) - v(x_t)\Big)\sigma\Big(v(z), v(x_t)\Big)\; d\Phi_\mu(z) \\
< \mathbb{P}_t[\tau_{a,b} < T] \times \sigma\Big(v(x_t+\epsilon), v(x_t)\Big) & \times \Big[pv(x_t-\epsilon-\epsilon') + (1-p)v(x_t+\epsilon) - v(x_t)\Big] \\
& + \mathbb{P}_t[\tau_{a,b} \geq T] \times \int_{(-\epsilon,\epsilon)} \Big(v(x_t+z) - v(x_t)\Big)\sigma\Big(v(x_t+z), v(x_t)\Big)\; d\tilde{\Phi}_\mu(z) \\
< \mathbb{P}_t[\tau_{a,b} \geq T] \times & \int_{(-\epsilon,\epsilon)} \Big(v(x_t+z) - v(x_t)\Big)\sigma\Big(v(x_t+|z|), v(x_t)\Big)\; d\tilde{\Phi}_\mu(z) \\
\leq \mathbb{P}_t[\tau_{a,b} \geq T] \times & \int_{(0,\epsilon)} \Big(\Big[v(x_t+z) - v(x_t)\Big] - \Big[v(x_t) - v(x_t-z)\Big]\Big) \\
& \times \sigma\Big(v(x_t+z), v(x_t)\Big)\; d\tilde{\Phi}_\mu(z) \leq 0,
\end{aligned}$$

with $p = p(a,b,\mu)$ defined in Eq. (2.5). The first inequality follows from ordering, diminishing sensitivity, and the fact that $v(x_t+\epsilon) - v(x_t) \leq v(x_t) - v(x_t-\epsilon-\epsilon')$ due to the concavity of the value function as well as the construction of $\tilde{\Phi}_\mu$, which is defined as $\tilde{\Phi}_\mu(z) := \Phi_\mu(x_t+z)$ for any $z \geq -\epsilon$ and $\tilde{\Phi}_\mu(z) := 0$ for any $z < -\epsilon$. The second one follows from the drift being non-positive in combination with Jensen's Inequality, and diminishing sensitivity of the salience function. The weak inequality holds by Lemma 2 (f), and the last inequality holds by concavity of the value function, which implies $v(x_t+z) - v(x_t) \leq v(x_t) - v(x_t-z)$ for any $z > 0$. □

Proof of Proposition 3. Let $\mu' < 0$. If the naïve salient thinker does not stop immediately, there exists some threshold stopping time $\tau_{a,b}$ such that $U^s\Big(X_{T\wedge\tau_{a,b}}|\mathcal{C}\Big) > v(x)$. By Proposition 2, the stopping time $\tau_{a,b}$ represents a loss-exit strategy; that is, the stopping thresholds satisfy $x - a < b - x$. By Lemma 2 (d), as the drift increases to $\mu'' > \mu'$, the distribution of $X_{T\wedge\tau_{a,b}}$ improves in terms of first-order stochastic dominance. Hence, by Proposition 1 in Dertwinkel-Kalt and Köster (2020), also $U^s\Big(X_{T\wedge\tau_{a,b}}\mathcal{C}\Big)$ increases as we move from a drift μ' to a drift μ''.

In sum, if the naïve salient thinker does not stop immediately a process with drift $\mu' < 0$, he does not stop immediately any process with drift $\mu'' > \mu'$. Likewise, if the

naïve salient thinker immediately stops a process with $\mu' \leq 0$, then he does so for any process with $\mu'' < \mu'$. □

2.A.5 Salience and the Disposition Effect

Salience theory predicts the disposition effect if

$$\frac{\sigma\Big(v(x_t-\epsilon), v(x_t)\Big)}{\sigma\Big(v(x_t+\epsilon'), v(x_t)\Big)} \times \frac{v(x_t)-v(x_t-\epsilon)}{v(x_t+\epsilon')-v(x_t)} \tag{2.16}$$

is increasing in x_t. With a linear value function, this term simplifies to

$$\frac{\sigma(x_t-\epsilon,\ x_t)}{\sigma(x_t+\epsilon',\ x_t)} \times \frac{\epsilon}{\epsilon'},$$

and substituting Bordalo et al. (2012)'s salience function into this term gives

$$\frac{\frac{|x_t-(x_t-\epsilon)|}{|x_t|+|x_t-\epsilon|+\theta}}{\frac{|x_t-(x_t+\epsilon')|}{|x_t|+|x_t+\epsilon'|+\theta}} \times \frac{\epsilon}{\epsilon'} = \frac{|x_t|+|x_t+\epsilon'|+\theta}{|x_t|+|x_t-\epsilon|+\theta} \times \frac{|\epsilon|}{|-\epsilon'|} \times \frac{\epsilon}{\epsilon'}.$$

Since we exclude negative outcomes, x_t, $x_t+\epsilon'$, and $x_t-\epsilon$ must be positive, and we can remove the absolutes. Similarly, because ϵ and ϵ' are positive by definition, we have $|\epsilon| = \epsilon$ and $|-\epsilon'| = \epsilon'$, which gives

$$\frac{\epsilon^2(2x_t+\epsilon'+\theta)}{(\epsilon')^2(2x_t-\epsilon+\theta)}.$$

Taking the first derivative with respect to x_t yields

$$-\frac{2\epsilon^2(\epsilon'+\epsilon)}{(\epsilon')^2(2x_t+\theta-\epsilon)^2}.$$

Since ϵ and ϵ' are positive, the fraction's numerator and denominator are always positive. Hence, the whole term is always negative, and the original term (2.16) with the standard salience function and a linear value function is decreasing in x_t. Hence, salience theory's standard specification cannot explain the disposition effect.

2.B Sophisticated Stopping Behavior Without Commitment

2.B.1 Statement of the Result

To solve for a sophisticate's stopping behavior, we adopt the equilibrium concept of Ebert and Strack (2018), which says that a stopping time τ constitutes an equilibrium if and only if at any point in time it is optimal to follow the strategy, taking as given that all future selves will do so.

Definition 6 (Equilibrium)**.** *A stopping time τ constitutes an* equilibrium *if and only if at every point in time t it is optimal to take the prescribed decision, given that all future selves will follow this strategy.*

We find that a sophisticated salient thinker, who is restricted to choose stop-loss and take-profit strategies, never gambles any process with a non-positive drift, which implies that naïvete is a necessary assumption to explain (unfair) casino gambling within the salience framework.

Proposition 4. *Suppose that the agent can only choose stop-loss and take-profit strategies. Fix an initial wealth level $x \in \mathbb{R}_{>0}$, and consider only processes with a non-positive drift. Then, in any equilibrium, the sophisticated salient thinker stops immediately.*

To fix ideas, let us assume that $T = \infty$. For any threshold stopping time $\tau_{a,b}$, there exists some wealth level $y' \in (a, b)$ such that the downside of the binary lottery $X_{\tau_{a,b}}$ is salient when evaluated in the choice set $\mathcal{C} = \{X_{\tau_{a,b}}, y'\}$. Moreover, if the process has a non-positive drift, then, at any wealth level y, we have $\mathbb{E}[X_{\tau_{a,b}}] \leq y$. Since a salient thinker, with a weakly concave value function, values a binary lottery with a salient downside strictly less than its expected value, the sophisticated agent anticipates to stop no later than at wealth level y'. Thus, by Definition 6, the threshold stopping time $\tau_{a,b}$ cannot constitute an equilibrium. In contrast, at any initial wealth level $x \in \mathbb{R}_{>0}$, stopping immediately can be supported as an equilibrium outcome: given that all future selves will stop immediately, the current self is indifferent between stopping immediately

and continuing the process, so that it is indeed optimal to stop at every point in time. As we prove in the following, the argument extends to processes with a finite expiration date.

2.B.2 Proof of Proposition 4 for a Finite Expiration Date

Fix an initial wealth level x and a non-positive drift $\mu \leq 0$. It remains to be shown that the arguments presented above extend to processes with a finite expiration date $T \in \mathbb{R}_{>0}$.

Consider a stop-loss and take-profit strategy, which can be represented by a threshold stopping time $\tau_{a,b}$. We now argue that it cannot be an equilibrium to play according to stopping time $\tau_{a,b}$ with $b \in (x, \infty)$. At any time t with a wealth level $X_t = y \in (a, b)$, a salient thinker follows the stop-loss and take-profit strategy that is represented by $\tau_{a,b}$ if and only if

$$\mathbb{P}_t[\tau_{a,b} < T] \times \Big[p\Big(v(a) - v(y)\Big)\sigma\Big(v(a), v(y)\Big) + (1-p)\Big(v(b) - v(y)\Big)\sigma\Big(v(b), v(y)\Big)\Big] \\ + \mathbb{P}_t[\tau_{a,b} \geq T] \times \int_{(a,b)} \Big(v(z) - v(y)\Big)\sigma\Big(v(z), v(y)\Big)\, d\Phi_\mu(z) \geq 0$$

holds, where the probability $p = (a, b, \mu)$ is defined as in Eq. (2.5) and where the conditional CDF $\Phi_\mu(z) := \mathbb{P}_0[X_T \leq z | X_0 = x_0, \tau_{a,b} \geq T - t]$ is described in Lemma 2.

Notice that $\sigma(v(a), v(y)) > \sigma(v(b), v(y))$ holds for any wealth level y sufficiently close to b. Also, we have $\mathbb{E}_t[X_{\tau_{a,b}} | X_t = y] \leq y$ due to the non-positive drift. This implies, together with the concave value function, that, for any wealth level y sufficiently close to b, it holds that

$$p\Big(v(a) - v(y)\Big)\sigma\Big(v(a), v(y)\Big) + (1-p)\Big(v(b) - v(y)\Big)\sigma\Big(v(b), v(y)\Big) < 0.$$

Since, for any fixed t, we have $\lim_{y \to b} \mathbb{P}_t[\tau_{a,b} < T] = 1$ by Lemma 2 (b), we thus conclude that for any $\tau_{a,b}$ there is some $y' \in (a, b)$ such that

$$\mathbb{P}_t[\tau_{a,b} < T] \times \Big[p\Big(v(a) - v(y)\Big)\sigma\Big(v(a), v(y)\Big) + (1-p)\Big(v(b) - v(y)\Big)\sigma\Big(v(b), v(y)\Big)\Big] \\ + \mathbb{P}_t[\tau_{a,b} \geq T] \times \int_{(a,b)} \Big(v(z) - v(y)\Big)\sigma\Big(v(z), v(y)\Big)\, d\Phi_\mu(z) < 0.$$

Hence, if the agent is restricted to choose from the set of all stop-loss and take-profit strategies, there exists no equilibrium in which a sophisticated salient thinker does not stop immediately. By the same arguments as for the case of $T = \infty$, stopping immediately can be supported as an equilibrium outcome, which proves the claim. □

2.C Additional Details on the Main Experiment

This appendix contains supplementary material to the experiment that we conducted.

2.C.1 Parameters and Layout of the Static Choices

After making the six selling decisions, subjects had to chooses twelve times between a binary lottery and the safe option payings its expected value. The parameters of the lotteries as well as the classification of skewness-seeking choices are depicted in Table 2.2. Figure 2.7 further illustrates the layout that we used for these static choices in the experiment.

Lottery	Safe Option	Skewness	Skewness-Seeking Choice
(37.5, 80%; 0, 20%)	30	-1.5	Safe
(41.25, 64%; 10, 36%)	30	-0.6	Safe
(45, 50%; 15, 50%)	30	0	Safe
(60, 20%; 22.5, 80%)	30	1.5	Lottery
(75, 10%; 25, 90%)	30	2.7	Lottery
(135, 2%; 27.85, 98%)	30	6.9	Lottery
(57.5, 80%; 20, 20%)	50	-1.5	Safe
(61.25, 64%; 30, 36%)	50	-0.6	Safe
(65, 50%; 35, 50%)	50	0	Safe
(80, 20%; 42.5, 80%)	50	1.5	Lottery
(95, 10%; 45, 90%)	50	2.7	Lottery
(155, 2%; 47.85, 98%)	50	6.9	Lottery

Table 2.2: *Lotteries used to elicit skewness preferences in static choices. The safe option is equal to the lottery's expected value. In addition, all lotteries have the same variance, so that the first and the second set of lotteries, respectively, differ only in terms of skewness.*

2.C.2 Experimental Instructions

Screen 1 — Instructions: Overview of the Experiment

Please note that you are not allowed to use your mobile phone or talk to other participants during the experiment. After you have finished the experiment and your payment appears on the screen, please stay seated and wait for the other participants to finish. At this point you are allowed to use your phone again. If you have a question, please raise your

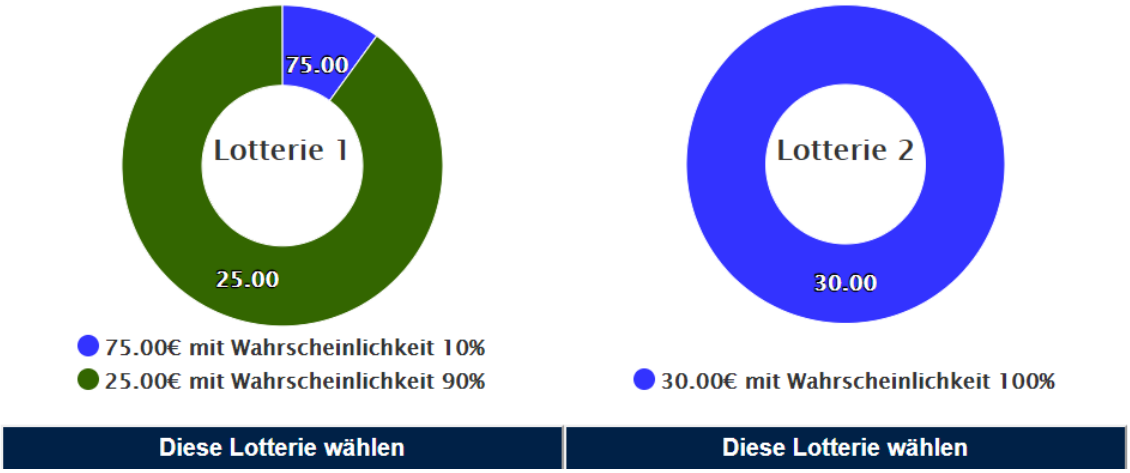


Figure 2.7: *The figure illustrates the layout of the static choices in the experiment (in German).*

hand and a lab assistant will come to you.

In this experiment you will make investment decisions. More precisely, you will have to decide at what time and price you want to sell an asset. The price at which you can sell the asset will change over time.

In total you will make 6 such investment decisions. At the end of the experiment, we will choose one of your decisions at random and pay you the price at which you sold this asset. Irrespective of this, you will receive a show up fee of 4 Euro. During the experiment, we will denote all monetary values in the currency Taler, which will be converted to Euro at an exchange rate of 1 Euro = 10 Taler.

The only thing that changes between the different decisions is the long-term profitability of the asset. The maximum time for which you can hold the asset will be 10 seconds in all decisions. If you do not sell the asset in the first 10 seconds, it will automatically be sold at its price after 10 seconds. The initial value of the asset will always be 100 Taler.

In the following, we will explain to you the development of the asset step by step. In particular, we will show you how the long-term profitability varies across the different assets. Moreover, we will explain in detail which selling strategies you will be able to

choose.

Screen 2 — Instructions: Development of the Asset Price

Below you can see a graph, which depicts the development of the price of an asset. As soon as you press "Start", a line which represents the value of the asset will appear.

Please press "Start" now.

[Subjects are shown a graph of an exemplary price path with a final price of 100 Taler.]

As mentioned previously, you cannot hold the asset for longer than 10 seconds. The final asset price is 100 Taler.

Screen 3 — Instructions: Different Drifts

In this experiment, you will see assets of varying profitability. How profitable an asset is in the long run is described by the drift of the asset. The drift denotes the average change in the value of the process per second.

A positive drift implies that the asset will increase in value in the long run, while a negative drift implies that the asset will decrease in value in the long run. Notice that the value of the asset varies over time. Hence, even an asset with a negative drift sometimes increases in value.

In order for you to get a feeling for how the value of an asset changes with the drift, we will show you a few examples of different drifts on the next screens.

Screens 4-6 — Instructions: The Drift of an Asset

The drift of this asset is 0 [or 2 or -2]. Please press "Start" and watch the development of the asset's price.

[Subjects are shown a graph of an exemplary price path with a final price of 100 Taler for the process with drift 0, 120 Taler for the process with drift 2, and 80 for the process with drift -2.]

Screen 7 — Instructions: No Negative Prices

The asset does never take a negative value. Once the asset's value reaches zero, it does not rise again, but will stay at zero permanently. Please press "Start" and watch the development of the asset's price.

[Subjects are shown a graph illustrating that the process is absorbing in zero.]

Screen 8 — Instructions: The Process is not Bounded from Above

Independent of the drift, the value of the asset can, in principle, become arbitrarily large. The probability that the asset's value indeed becomes very large is the smaller the more negative the drift is. But even an asset with a very negative drift can attain a very large value.

Screen 9 — Instructions: Strategies with an Upper and a Lower Bound

In each decision, you will set an upper and a lower bound at which you are willing to sell the asset. If the price reaches the upper bound, the process will stop and you will be able to sell the asset. If you sell the asset, you will receive the price that you have set as the upper bound. If the price reaches the lower bound, you can also sell the asset. In this case you will receive the price that you have set as the lower bound.

The upper bound must always be above the current value of the process. The lower bound must always be below the current value of the process. You can adjust the bounds by clicking on the red lines and moving them around. Important: throughout the experiment, you will have to move the upper and the lower bound at least once, before you can start the process. Please move the bounds now and then click "Start".

[Subjects are shown a graph similar to the ones depicted in Figure 2.1. After moving the bounds and starting the process, subjects cannot pause the process or adjust the bounds anymore.]

Screen 10 — Instructions: Pausing the Process

After you have started the process, you can pause it at any time. While the process is paused, you can move the upper and the lower bound. While the process is moving, you cannot move the bounds.

Now you have to complete the following steps in the order listed below:

1. Move the upper and the lower bound.
2. Start the process.
3. Pause the process.
4. Move the upper and the lower bound again.
5. Start the process again.

[Subjects are shown a graph similar to the ones depicted in Figure 2.1, but without the opportunity to sell the asset immediately.]

Screen 11 — Instructions: Sell Immediately

Before you start the process, you can instead sell the asset immediately by clicking on "Sell Immediately". You can only do this before you start the process for the first time. After you have started the process, the "Sell Immediately" button will disappear. Afterwards the process will only stop prematurely if it either hits the upper or the lower bound. You can now either "Sell Immediately" or — after moving each bound at least once — start the process.

[Subjects are shown a graph similar to the ones depicted in Figure 2.1.]

Screen 12 — Instructions: Change Bounds Before Starting the Process

In the first 10 seconds on each decision screen, you will only be able to move the bounds. After that you can "Sell Immediately" or start the process.

If the process reaches either bound, it stops and you can sell the asset at the price at which this bound is set. Alternatively, you can move the bounds and restart the process. Corresponding buttons for both options will appear once the process reached a bound.

[Subjects are shown a graph similar to the ones depicted in Figure 2.1.]

Screen 13 — The Task is About to Start

You will now participate in three practice rounds. Afterwards you will play the decision round. The drift in the practice rounds will be identical to the one in the decision round. The practice rounds will give you the opportunity to get an intuition for how the process will develop during the decision round.

The drift in the practice rounds and the subsequent decision round will be 0 [or -1 or -3 or -5 or -10 or -20].

Screen 14 — Practice Round

The drift in the practice rounds and the subsequent decision round will be 0 [or -1 or -3 or -5 or -10 or -20].

[Subjects are shown a graph as depicted in the left panel of Figure 2.2.]

Screen 15 — Instructions: The drift of an asset

On this page you see 10 exemplary paths of an asset with a drift of 0 [or -1 or -3 or -5 or -10 or -20].

[Subjects are shown a graph as depicted in the right panel of Figure 2.2.]

Screen 16 — Decision

The practice rounds are over — now it is getting serious! Please make your selling decision. The drift in this round is 0 [or -1 or -3 or -5 or -10 or -20].

[Subjects are shown a graph as depicted in Figure 2.1.]

Screen 17 — Additional Questions I: Instructions

On the next pages you will make 12 choices between a lottery and a safe payoff. From now on all outcomes will be displayed in Euro.

At the end of the experiment, we will select one participant of this session at random. For this participant, we will randomly select one of the 12 decisions and determine the outcome of the chosen lottery. This participant will receive the corresponding payoff from the chosen lottery.

Example

If you select Lottery 1 in the example below, you will receive either 135 Euro or 27.85 Euro. The probability that you receive 135 Euro is 2% and the probability that you receive

27.85 Euro is 98%. Alternatively, if you select Lottery 2, you will receive 30 Euro for sure.

[Subjects are shown the graph depicted in Figure 2.7.]

Screen 18: Additional Questions I - Decision 1

Please choose a lottery. As soon as you have chosen a lottery, a button labelled "Next Page" will appear.

[Subjects are shown the graph as depicted in Figure 2.7.]

Screen 19 — Additional Questions II

Please answer the following questions. For every correct answer, you will receive one Taler.

If 10 machines take 10 minutes to make 10 nails, how many minutes do 100 machines need to make 100 nails?

A part of a pond is covered with water lilies. Every day the area covered with water lilies doubles. If it takes 24 days until the whole pond is covered with water lilies, how many days does it take until half of the pond is covered with water lilies?

If three elves can wrap three presents in one hour, how many elves does it take to wrap six presents in two hours?

Jerry has both the 15th best and the 15th worst grade in his class. How many students are in the class?

In a sports team tall members are three times as likely to win medals as short members. This year the team won 60 medals in total. How many medals were won by short team members?

Screen 20 — Additional Questions III

Please answer the following questions. For every correct answer, you will receive one Taler.

Suppose you had 100 Euro in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

[Options: "More than 102 Euro", "Exactly 102 Euro", "Less than 102 Euro".]

Suppose you had 100 Euro in a savings account and the interest rate was 20% per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total?

[Options: "More than 200 Euro", "Exactly 200 Euro", "Less than 200 Euro".]

Imagine that the interest rate on your savings account was 1% per year and the inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

[Options: "More than today", "As much as today", "Less than today".]

Assume a friend inherits 10.000 Euro today and his brother inherits 10.000 Euro three years from now. There is a positive interest rate. Who is richer because of the inheritance?

[Options: "My friend", "Her brother", "Both are equally rich".]

Suppose that your income and all prices double in the next year. How much will you be able to buy with your income?

[Options: "More than today", "As much as today", "Less than today".]

Screen 21 — General Information About You

Please enter your age:

Please choose your gender:

Screen 22 — Enter Station Number

Please enter your station number:

Screen 23 — Payment

Your decision from round 2 will be paid.

You sold the asset for 100.00 Taler.

You received 1 Taler from answering the additional questions.

You are the participant whose lottery choice is paid. You receive an additional 80.00 Euro from the lottery.

Your payment including the show up fee of 4 Euro is 94.10 Euro.

2.D Salience Predictions on Static Skewness Preferences

In this section, we extend a result from Dertwinkel-Kalt and Köster (2020) on a salient thinker's skewness preferences in static settings from the case of a linear value function to the case of a weakly concave value function. Assuming a linear value function, Dertwinkel-Kalt and Köster (2020) study, in particular, a salient thinker's choice between a binary lottery with an expected value E, a variance V, and a skewness S, which we denote by $L(E, V, S)$, and the safe option paying the lottery's expected value E with certainty, and they show that:

Proposition 5 (Dertwinkel-Kalt and Köster, 2020)**.** *There exists some $\hat{S} = \hat{S}(E, V) \in \mathbb{R}$, such that a salient thinker with a linear value function chooses $L(E, V, S)$ over E if and only if $S > \hat{S}$.*

The proposition says that a salient thinker with a linear value function chooses a binary lottery over its expected value if and only if this lottery is sufficiently skewed. In the following, we will show that the same comparative static holds when assuming a weakly concave value function. This provides a theoretical foundation for why we look at the empirical relationship between a subject's share of skewness-seeking choices in the static choices and the share of loss-exit strategies this subject has chosen in the stopping problems (see Result 44 and Figure 2.5). A positive correlation between the two measures indicates that a subject revealing skewness preferences consistent with salience theory in static choices does so also in dynamic choices.

To begin with, recall that the parameters of the binary lottery $L(E, V, S)$ — i.e. the outcomes x_1 and x_2 as well as the probability p that x_1 is realized — are uniquely defined by (Ebert, 2015):

$$x_1 = E - \sqrt{\frac{V(1-p)}{p}},\ x_2 = E + \sqrt{\frac{Vp}{1-p}},\ \text{ and } p = \frac{1}{2} + \frac{S}{2\sqrt{4+S^2}}.$$

Now consider a salient thinker with a weakly concave value function $v(\cdot)$, who faces the choice between the lottery $L(E, V, S)$ and the safe option paying its expected value E with

certainty. The salient thinker chooses lottery $L(E,V,S)$ over its expected value E if and only if

$$p\left[v\left(E-\sqrt{\frac{V(1-p)}{p}}\right)-v(E)\right]\sigma\left(v\left(E-\sqrt{\frac{V(1-p)}{p}}\right),v(E)\right)$$
$$+(1-p)\left[v\left(E+\sqrt{\frac{Vp}{1-p}}\right)-v(E)\right]\sigma\left(v\left(E+\sqrt{\frac{Vp}{1-p}}\right),v(E)\right)>0,$$

or, equivalently,

$$\pi\frac{v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)}{v(E)-v\left(E-\sqrt{V\pi}\right)}>\frac{\sigma\left(v\left(E-\sqrt{V\pi}\right),v(E)\right)}{\sigma\left(v\left(E+\sqrt{\frac{V}{\pi}}\right),v(E)\right)},\tag{2.17}$$

where $\pi := \frac{1-p}{p}$ denotes the relative likelihood of the lottery's upside. To establish that a salient thinker chooses the lottery if and only if it is sufficiently skewed, we will show that both the left-hand side and the right-hand side of (2.17) are monotonic in the likelihood ratio π; namely, that the left-hand side decreases in π, while the right-hand side increases in π.

First, by the ordering property, the right-hand side of (2.17) monotonically increases in π. Second, we observe that the left-hand side monotonically decreases in the likelihood ratio π:

$$\begin{aligned}
\frac{\partial}{\partial\pi}\left(\pi\frac{v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)}{v(E)-v\left(E-\sqrt{V\pi}\right)}\right) &= \frac{v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)}{v(E)-v\left(E-\sqrt{V\pi}\right)} \\
&\quad-\frac{\sqrt{\frac{V}{\pi}}v'\left(E+\sqrt{\frac{V}{\pi}}\right)\left[v(E)-v\left(E-\sqrt{V\pi}\right)\right]+\sqrt{V\pi}v'\left(E-\sqrt{V\pi}\right)\left[v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)\right]}{\left[v(E)-v\left(E-\sqrt{V\pi}\right)\right]^2} \\
&\propto \left[v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)-\sqrt{\frac{V}{\pi}}v'\left(E+\sqrt{\frac{V}{\pi}}\right)\right]\left[v(E)-v\left(E-\sqrt{V\pi}\right)\right] \\
&\qquad-\sqrt{V\pi}v'\left(E-\sqrt{V\pi}\right)\left[v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)\right] \\
&\propto 1-\underbrace{\frac{v'\left(E+\sqrt{\frac{V}{\pi}}\right)}{\frac{v\left(E+\sqrt{\frac{V}{\pi}}\right)-v(E)}{\sqrt{\frac{V}{\pi}}}}}_{\geq 0 \text{ since } v(\cdot) \text{ is increasing}}-\underbrace{\frac{v'\left(E-\sqrt{V\pi}\right)}{\frac{v(E)-v\left(E-\sqrt{V\pi}\right)}{\sqrt{V\pi}}}}_{\geq 1 \text{ since } v(\cdot) \text{ is concave}} \\
&\leq 0,
\end{aligned}$$

where, after taking the derivative, we first multiply by $\left[v(E) - v\left(E - \sqrt{V\pi}\right)\right]^2$ and rearrange, and then divide by $v(E) - v\left(E - \sqrt{V\pi}\right)$ and $v\left(E + \sqrt{\frac{V}{\pi}}\right) - v(E)$ to arrive at the final expression. Combining these two observations, we conclude that there exists some $\hat{\pi} \geq 0$, such that (2.17) is satisfied if and only if $\pi < \hat{\pi}$. Since π monotonically decreases in the probability p, and since the probability p monotonically increases in the skewness S, we arrive at the following result:

Proposition 6. *There exists some $\hat{S} = \hat{S}(E, V) \in \mathbb{R} \cup \{\infty\}$, such that a salient thinker with a weakly concave value function chooses $L(E, V, S)$ over E if and only if $S > \hat{S}$.*

This proposition confirms that the comparative static on the lottery's skewness derived in Dertwinkel-Kalt and Köster (2020), under the assumption of a linear value function, is robust to allowing for a weakly concave value function. The only difference compared to the result in Dertwinkel-Kalt and Köster (2020) is that a salient thinker with a sufficiently concave value function will not choose the binary lottery, irrespective of how skewed it is. Formally, it follows that the threshold value $\hat{S}$ in Proposition 6 satisfies $\hat{S} < \infty$ if and only if

$$\lim_{\pi \to 0} \frac{\partial}{\partial \pi} \left(\pi \frac{v\left(E + \sqrt{\frac{V}{\pi}}\right) - v(E)}{v(E) - v\left(E - \sqrt{V\pi}\right)} - \frac{\sigma\left(v\left(E - \sqrt{V\pi}\right), v(E)\right)}{\sigma\left(v\left(E + \sqrt{\frac{V}{\pi}}\right), v(E)\right)} \right) < 0,$$

which depends both on the curvature of the value and on the curvature of the salience function. But, as illustrated in Proposition 5, the above inequality is certainly satisfied for a linear value function and, by continuity, it will hold for mildly concave value functions as well.

2.E Additional Figures

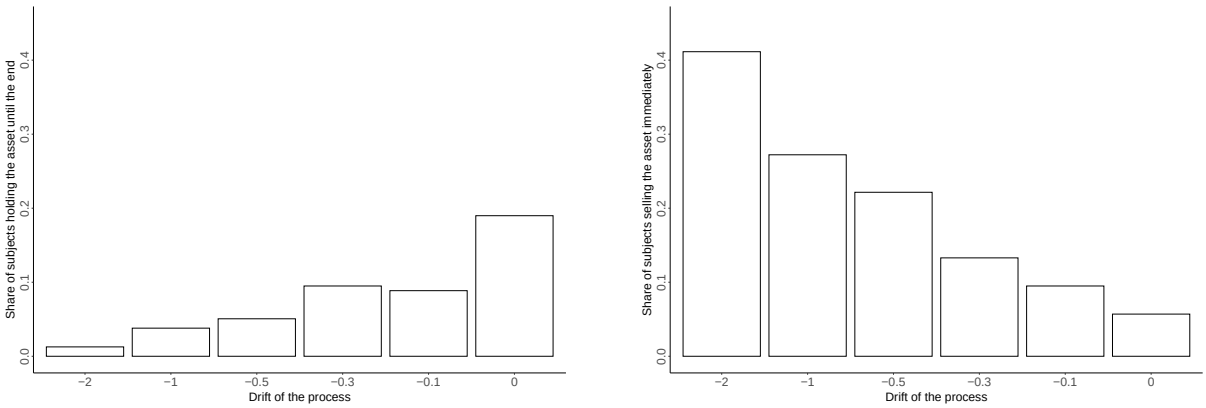


Figure 2.8: *The left panel depicts the share of subjects holding the asset until the expiration date, separately for the different drifts. The right panel depicts the share of subjects selling the asset immediately.*

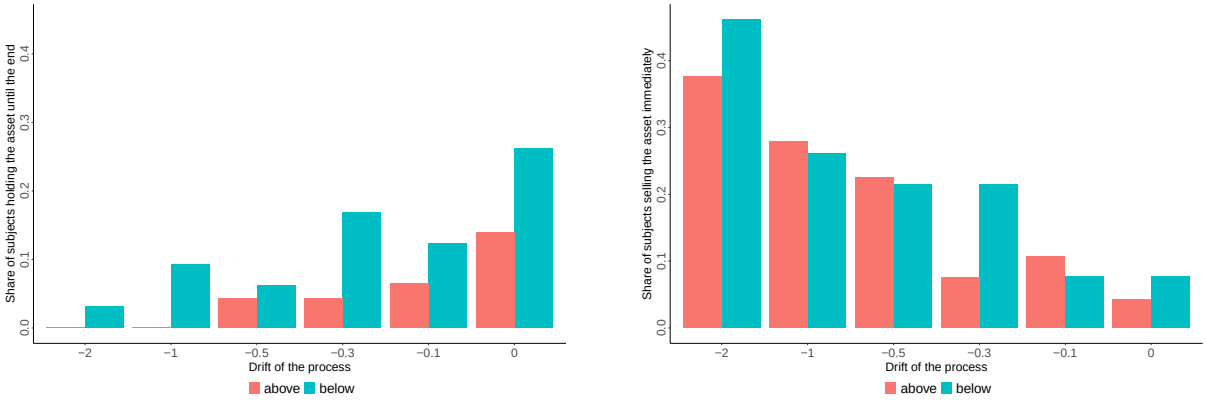


Figure 2.9: *The left panel depicts the share of subjects holding the asset until the expiration date, separately for the different drifts and below- and above-median subjects in terms of cognitive skills. The right panel depicts the share of below- and above-median subjects selling the asset immediately.*

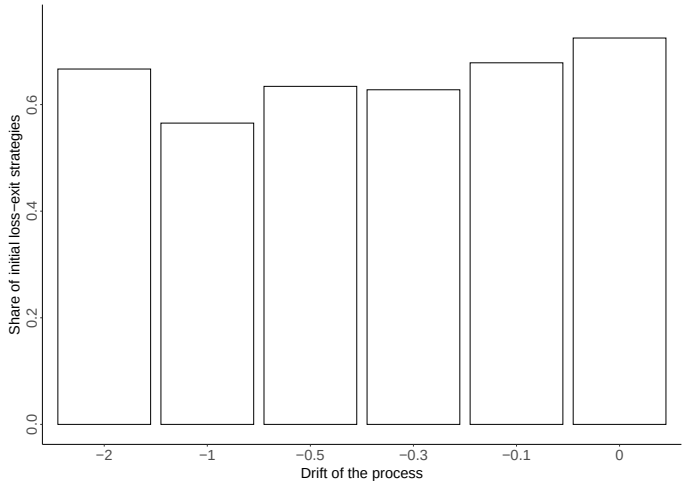


Figure 2.10: *The figure depicts the share of initial loss-exit strategies chosen for the different drifts.*

		After	
		Loss-Exit	**Gain-Exit**
Before	**Loss-Exit**	63.31%	10.31%
	Gain-Exit	12.01%	14.37%

	After	
	Loss-Exit	**Gain-Exit**
Loss-Exit	72.65%	0%
Gain-Exit	3.48%	23.87%

Figure 2.11: *The left (right) table gives a categorization of all strategy adjustments that we observe throughout the experiment when a bound (no bound) is hit. "Before" indicates, in the left table, which type of strategy the subject has chosen last, and, in the right table, the type of strategy that is played in the moment in which the process is paused.*

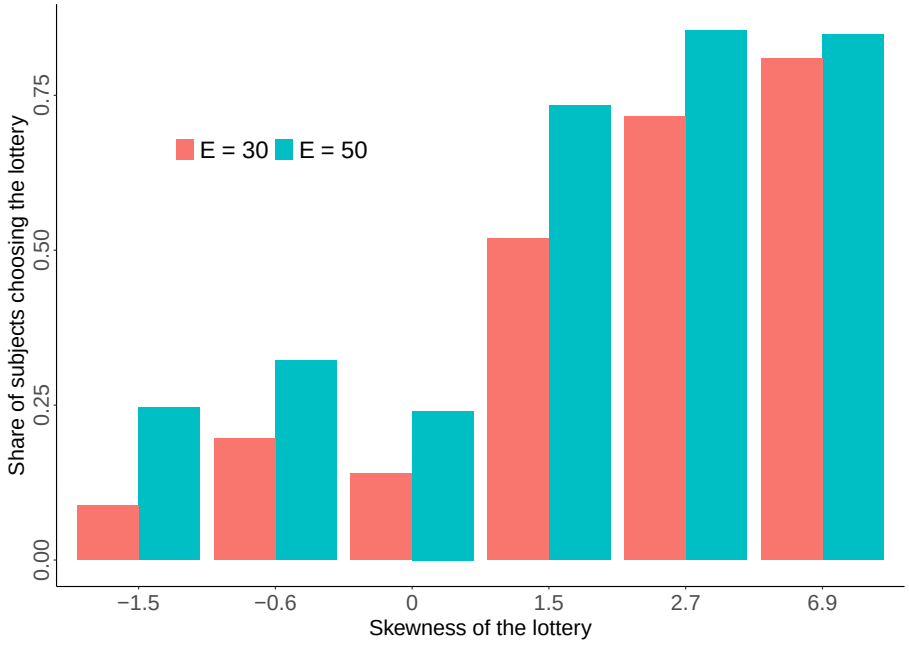


Figure 2.12: *The figure depicts the share of subjects choosing each of the lotteries depicted in Table 2.2 over its expected value.*

2.F Stopping Behavior under Cumulative Prospect Theory

In this section, we analyze the stopping behavior of a naïve CPT-agent under the assumption of a finite expiration date. Ebert and Strack (2015) study the case without an expiration date and show that, under mild regularity assumptions on the probability weighting function,[25] a naïve CPT-agent never stops an ABM irrespective of its drift. In what follows, we will show numerically that this strong result still holds for a finite expiration date.

CPT preferences. Let X be a real-valued random variable. A CPT-agent evaluates each outcome of this random variable relative to a *reference point* $r \in \mathbb{R}$ via a strictly increasing *value function* $U : \mathbb{R} \to \mathbb{R}$. All outcomes larger than the reference point are classified as *gains*, while outcomes smaller than the reference point are classified as *losses*. Throughout this section, we assume a (weakly) *S-shaped* value function (Ebert and Strack, 2015, Online Appendix W.2),

$$U(x) = \begin{cases} (x-r)^\alpha & \text{if } x \geq r, \\ -\lambda \cdot (r-x)^\alpha & \text{if } x < r, \end{cases} \tag{2.18}$$

with parameters $\alpha \in (0,1]$ and $\lambda > 1$.[26] According to Tversky and Kahneman (1992), cumulative probabilities are distorted by a *weighting function*. More precisely, there are (potentially different) non-decreasing weighting functions $w^-, w^+ : [0,1] \to [0,1]$ for gains and losses with $w^-(0) = w^+(0) = 0$ and $w^-(1) = w^+(1) = 1$. Throughout this section, we use the following weighting functions, which have been proposed by Tversky and Kahneman (1992):[27]

$$w^-(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{1/\delta}} \quad \text{and} \quad w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \qquad \text{for } 0.279 < \delta, \gamma < 1.$$

[25] In their Online Appendix W.1, Ebert and Strack (2015) verify that common CPT specifications satisfy the sufficient conditions that they impose on the probability weighting function to derive their main result.

[26] As argued in Wakker (2010, p. 270), the model is ill-specified when taking different α for gains and losses.

[27] The bounds on the parameters are taken from Dhami (2016, p. 122).

The CPT value of the random variable X can be defined as (see Kothiyal et al., 2011)

$$\begin{aligned} CPT(X) &:= \int_{\mathbb{R}_+} w^+\Big(\mathbb{P}[U(X) > y]\Big)\, dy - \int_{\mathbb{R}_-} w^-\Big(\mathbb{P}[U(X) < y]\Big)\, dy \\ &= \int_{\mathbb{R}_+} w^+\Big(\mathbb{P}[X > r + y^{1/\alpha}]\Big)\, dy - \int_{\mathbb{R}_-} w^-\Big(\mathbb{P}[X < r - (-y)^{1/\alpha}/\lambda]\Big)\, dy, \end{aligned} \tag{2.19}$$

where the second equality holds due to a (weakly) S-shaped value function in Eq. (2.18).

Stopping strategies. Consider a threshold stopping time $\tau_{a,b}$ and therefore induces a random wealth level $X_{T\wedge\tau_{a,b}}$. If the reference point $r \in \mathbb{R}$ satisfies $r \in [a, b]$, then the CPT value associated with this random variable is given by

$$\begin{aligned} CPT(X_{T\wedge\tau_{a,b}}) = \int_{(0,(b-r)^\alpha)} & w^+\Big(\mathbb{P}_t[X_{T\wedge\tau_{a,b}} > r + y^{1/\alpha}]\Big)\, dy \\ & - \int_{(-\lambda(r-a)^\alpha,0)} w^-\Big(\mathbb{P}_t[X_{T\wedge\tau_{a,b}} < r - (-y)^{1/\alpha}/\lambda]\Big)\, dy. \end{aligned} \tag{2.20}$$

For $a \geq r$, in contrast, the CPT value of the random variable $X_{T\wedge\tau_{a,b}}$ equals

$$CPT(X_{T\wedge\tau_{a,b}}) = \int_{((a-r)^\alpha,(b-r)^\alpha)} w^+\Big(\mathbb{P}_t[X > r + y^{1/\alpha}/\lambda]\Big)\, dy + (a-r)^\alpha, \tag{2.21}$$

while for $b \leq r$ it is given by

$$CPT(X_{T\wedge\tau_{a,b}}) = -\int_{(-\lambda(r-a)^\alpha,-\lambda(r-b)^\alpha)} w^-\Big(\mathbb{P}_t[X < r - (-y)^{1/\alpha}/\lambda]\Big)\, dy - \lambda(r-b)^\alpha. \tag{2.22}$$

At time $t < T$ with a current wealth level $x_t \in \mathbb{R}_{>0}$, we consider the following class of threshold stopping times:[28] for $k \in \mathbb{R}_{>0}$ and $p \in (0, \frac{1}{2})$, define $a_{t,k} = x_t - k \cdot p$ and $b_{t,k} = x_t - k \cdot (1-p)$. Notice that, for any drift $\mu \leq 0$, these threshold stopping times are not only loss-exit strategies, but also induce a right-skewed distribution of returns.

Numerical analysis of stopping behavior. To ease the illustration of the results, we assume that the reference point constantly adjusts to the current wealth level (i.e., $r_t = x_t$ for any t). This implies, in particular, that the wealth level itself does not matter for a CPT-agent's stopping behavior, which makes the numerical analysis much more

[28] These strategies are similar to those used in the proof of Theorem 1 in Ebert and Strack (2015).

convenient. Based on the estimates in Tversky and Kahneman (1992), we set $\alpha = 0.88$ and $\lambda = 2.25$ as well as $\delta = 0.69$ and $\gamma = 0.61$.

Assuming a drift of $\mu = -2$ and a volatility of $\nu = 5$, Figure 2.13 depicts, for a given point in time t, the CPT value of the random variable $X_{T \wedge \tau_{a_{t,k}, b_{t,k}}}$ as a function of the remaining time, $T - t$, until the expiration date for the (naïvely) right-skewed strategies with $k \in \{2, 4, 6, 8, 10\}$ and $p = 0.01$. Since we have $r_t = x_t$ by assumption, a CPT-agent does not stop at time t as long as there exists a stopping strategy that yields a strictly positive CPT value. We observe from Figure 2.13 that for any remaining time until the expiration date, there indeed exists a threshold stopping time that yields a strictly positive CPT value. When shifting the stopping thresholds closer to the current wealth level (by shifting the parameter k closer to zero), we obtain a similar picture for any arbitrarily negative drift. Hence, at least for the chosen parameter values, a naïve CPT-agent does not stop before the expiration date or, in other words, the stark never-stopping result derived by Ebert and Strack (2015) still holds for a finite expiration date.

Figure 2.13 highlights a couple of numerical regularities that are suggestive for the result not to hinge on the exact parameters chosen here: First, the CPT value derived from the depicted stopping strategies becomes flat in the remaining time until the expiration date relatively quickly and the earlier so the closer the stopping thresholds are to the current wealth level (i.e., the closer is k to zero). This suggests that the result derived by Ebert and Strack (2015) — which is proven by the explicit use of strategies with thresholds close to the current wealth level — does not rely on T being infinity, but should hold already for rather short expiration dates. Second, as the remaining time until the expiration date gets smaller, the CPT value of the depicted loss-exit strategies increases (before it eventually falls to zero). Both patterns are robust to different specifications (e.g., a piece-wise linear value function with $\alpha = 1$ or a reference point of $r_t = 0$). This suggests that the never-stopping result derived by Ebert and Strack (2015) is indeed robust to allowing for a finite expiration date. All numerical results are available upon request.

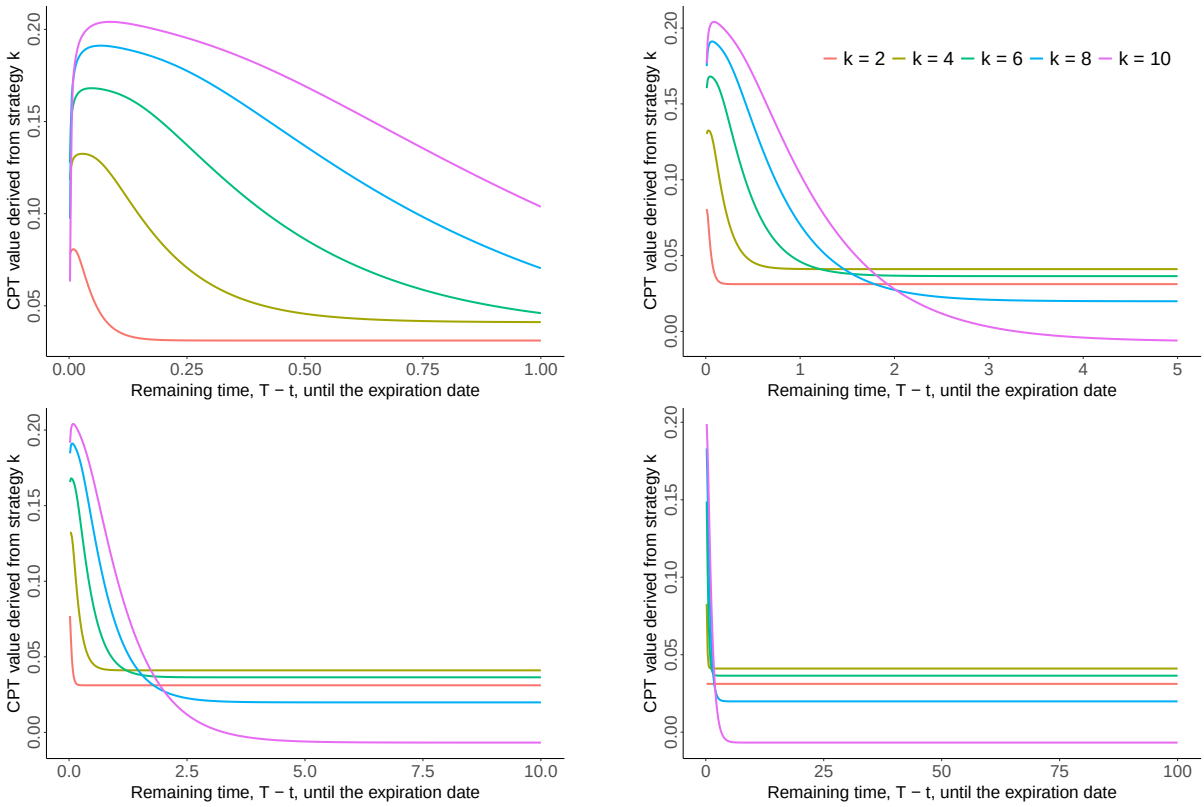


Figure 2.13: *The figure depicts* $CPT(X_{T\wedge\tau_{a_{t,k},b_{t,k}}})$ *as a function of the remaining time,* $T-t$*, until the expiration date for time invariant strategies with* $k \in \{2,4,6,8,10\}$ *and* $p = 0.01$ *as described above. We assume a drift parameter of* $\mu = -2$ *and a volatility parameter of* $\nu = 5$*. The preference parameters are set to* $\alpha = 0.88$ *and* $\lambda = 2.25$*, and the parameters of the weighting function are* $\delta = 0.69$ *and* $\gamma = 0.61$*.*

2.G Salience-Driven Decoy Effects

2.G.1 Overview of the Experimental Design and Results

The most striking difference compared to the alternative models discussed above is that salience theory is first and foremost a model of context-dependent behavior: the evaluation of an option depends on the alternatives at hand. To conclusively rule out models of context-independent behavior, we next study the role of (dominated) *decoys* in stopping problems. Such decoy effects are not only inconsistent with the models discussed in the preceding section, but also with other prominent models such as expectation-based loss aversion (Kszegi and Rabin, 2007).

The modified stopping problem. Suppose there is no longer just one asset, but two assets. We refer to these assets as $G(reen)$ and $B(lue)$, and assume that their prices follow the ABMs

$$dX_t = \mu_G dt + \nu dW_t \quad \text{and} \quad dY_t = \mu_B dt + \nu dU_t, \quad \text{respectively.}$$

Both assets share the same initial value $X_0 = z = Y_0$ and the same volatility $\nu \in \mathbb{R}_{>0}$, but Asset G has a larger drift than Asset B, $\mu_G > \mu_B$. We further assume that the standard Brownian Motions $(W_t)_{t\in\mathbb{R}_{\geq 0}}$ and $(U_t)_{t\in\mathbb{R}_{\geq 0}}$ are independent of each other. There is no expiration date.

	$p_G p_B$	$p_G(1-p_B)$	$(1-p_G)p_B$	$(1-p_G)(1-p_B)$
Invest in Asset G	a	a	b	b
Invest in Asset B	a	b	a	b
No Investment	z	z	z	z

Table 2.3: *Joint distribution of the different options.*

If the agent invests in either asset, he is restricted to choose a loss-exit strategy, represented by the threshold stopping time $\tau_{a,b}$ with $a < z < b$, which he *cannot* revise over time. Since there is no expiration date, the asset is sold at one of these thresholds, and since the agent cannot revise his strategy, this happens in finite time with probability one. Denote as $p_G := p(a, b, \mu_G)$ the probability of Asset G being sold at the lower price $a < z$, and as $p_B := p(a, b, \mu_B)$ the corresponding probability for Asset B. Since $\mu_G > \mu_B$ and, therefore, $p_G < p_B$, investing in Asset B is dominated (in the sense of first-order stochastic dominance) by investing in Asset G. We, therefore, refer to Asset B as a *(dominated) decoy* (e.g., Huber et al., 1982). We compare two scenarios: the choice set is either $\{X_{\tau_{a,b}}, z\}$ (*no decoy*) or $\{X_{\tau_{a,b}}, Y_{\tau_{a,b}}, z\}$ (*decoy*).

Salience-driven decoy effects. According to salience theory, investing in Asset G can become more attractive in the presence of the dominated Asset B. More precisely, a salient thinker compares the selling price of Asset G state-by-state to some *reference point* given by a convex combination of the other options' outcomes (see Table 2.3 for the joint

distribution of the different options).[29] As a consequence, when adding the dominated decoy to the choice set, the salience of Asset G's upside — i.e. selling it at a price b — changes: by the contrast effect, it becomes more salient whenever the dominated asset is sold at a price $a < z$ and less salient when it is sold at $b > z$. Likewise, the downside of investing in Asset G — i.e. selling it at a price a — becomes more salient when the dominated asset is sold at a price $b > z$ and less salient when it is sold at a price $a < z$ instead. Since the dominated asset is the more likely to be sold at the lower price the more negative its drift μ_B is, the upside of Asset G becomes relatively more salient as μ_B decreases. In other words, according to salience theory, (i) the presence of the dominated asset can boost demand for Asset G, and (ii) this decoy effect is (weakly) stronger for dominated assets with a more negative drift (see Appendix G.3 for a formal derivation).

Experimental design and implementation. We conducted an online experiment to test for such decoy effects on stopping behavior.[30] As in the main experiment, we set the initial price to $z = 100$ Taler (this time converted to £ at a ratio of 60:1) and the volatility to $\nu = 5$. Throughout, Asset G has zero drift, while the dominated asset's drift is either $\mu_B = -10$ or $\mu_B = -20$. Subjects could always choose the outside option of *No Investment*. Figure 2.14 illustrates the decision screens with (right panel) and without (left panel) a dominated decoy.

Figure 2.14: *Screenshots of the decision screens with and without a decoy.*

[29] While Bordalo et al. (2012) assume that the reference point is given by the state-wise average over the alternative options, we allow for a much more flexible functional form (see Appendix G.2 for details).

[30] The experimental design, including the fully specified salience model and its predictions, was pre-registered in the AEA RCT registry as trial AEARCTR-0007398. Moreover, the experiment is available via the following link: `https://os-experiment-archive.herokuapp.com/demo`.

If a subject decides to invest in an asset, he can sell it only at pre-specified prices 90 and 190. More precisely, if a subject invests, the price of the asset will change until it reaches either 90 or 190. Recall that such a loss-exit strategy is potentially attractive to a salient thinker (see Proposition 2). If a subjects does not invest, he receives an asset's initial price with certainty.

Each subject made three investment decisions: one decision with a binary choice set (*Asset Green* vs. *No Investment*) and two with a larger choice set (*Asset Green* vs. *Asset Blue* vs. *No Investment*) for $\mu_B \in \{-10, -20\}$. The order of decisions was randomized at the subject level.

As in the main experiment, to explain the drift of an ABM to the subjects, they had to successively draw three sample paths from the underlying process, and then saw an overview of five additional sample paths of this process before making a decision (see Figure 2.15 for examples of the latter with and without a decoy). The sample paths were randomly drawn at the subject level; that is, different subjects saw different sample paths of the same underlying process.

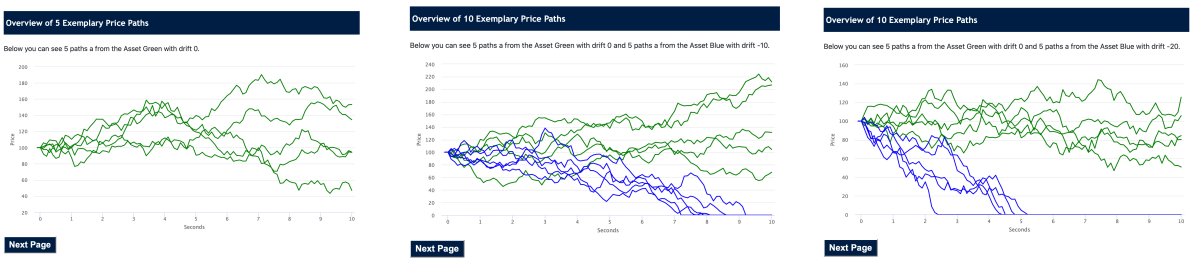


Figure 2.15: *Screenshots of the sampling screens with and without a decoy.*

At the end of the experiment, for each subject, one of the three investment decisions was randomly drawn by the computer to be payoff-relevant. All subjects received an additional £3 for their participation in the experiment.

We conducted 7 sessions in March 2021 via the Oxford Nuffield CESS lab, using the software oTree (D. Chen et al., 2016). A total number of $n = 247$ subjects completed the experiment.[31] The experiment lasted for around 11 minutes on average. Subjects earned £4.68 on average, with earnings ranging from £4.50 to £6.17.

[31]All of these subjects passed an attention check that tested for comprehension of the experimental instructions.

Experimental results. The share of subjects investing in the viable Asset *Green* increases by roughly 10 p.p. when a dominated decoy is available (see Figure 2.16). As pre-registered, when a decoy was available, we drop the subjects who chose the dominated option — roughly, 4% of the subjects for $\mu_B = -10$ and 2% of the subjects for $\mu_B = -20$ — before calculating the share of subjects investing in the viable asset. This entails the implicit assumption that choosing the decoy is a mistake that is independent of whether a subject actually wants to invest or not. It is important to note, however, that the observed decoy effect is robust (but slightly smaller) to dealing with subjects who chose the decoy in a different way (e.g. assuming that all of them actually prefer not to invest).[32] This finding is consistent with salience theory, but it is clearly inconsistent with models of context-independent behavior such as EUT or CPT.

We do not find a significant effect of the dominated decoy's drift on how attractive Asset *Green* appears to be. At the same time, we cannot reject the null-hypothesis that the share of subjects choosing Asset *Green* is (weakly) larger for $\mu_B = -20$ than for $\mu_B = -10$. Hence, the observed behavior is indeed consistent with salience theory.

2.G.2 A More General Salience Model

Consider a choice set $\mathcal{C} = \{X_i\}_{i=1}^n$. The random variables (or *lotteries*) X_1 to X_n are non-negative with a joint cumulative distribution function $F : \mathbb{R}^n_{\geq 0} \to [0, 1]$. A state of the world refers to a tuple of outcomes, $(x_1, \ldots, x_n) \in \mathbb{R}^n_{\geq 0}$. A salient thinker compares the value of a given lottery, $v(X_i)$, to a *reference point* $R_i = \phi(v(X_1), \ldots, v(X_{i-1}), v(X_{i+1}), \ldots, v(X_n))$. Bordalo et al. (2012) assume that the reference point is given by the state-wise average over all alternative options: $R_i = \frac{1}{n-1} \sum_{j \neq i} v(X_j)$. We, in contrast, allow for a more general reference point $\phi : \mathbb{R}^{n-1} \to \mathbb{R}$ that (i) strictly increases in each of its arguments and (ii) satisfies $\phi(z, \ldots, z) = z$.

[32] The most conservative way to test for decoy effects is to assume that all subjects who chose the decoy actually prefer not to invest. When making this assumption, the decoy effect ranges from 7.7 p.p. (if $\mu_B = -10$) to 8.5 p.p. (if $\mu_B = -20$). Using t-tests with standard errors being clustered at the subject level, we find that the latter effect is statistically significant at the 5%-level (p-value = 0.027), while the former effect almost is (p-value = 0.053).

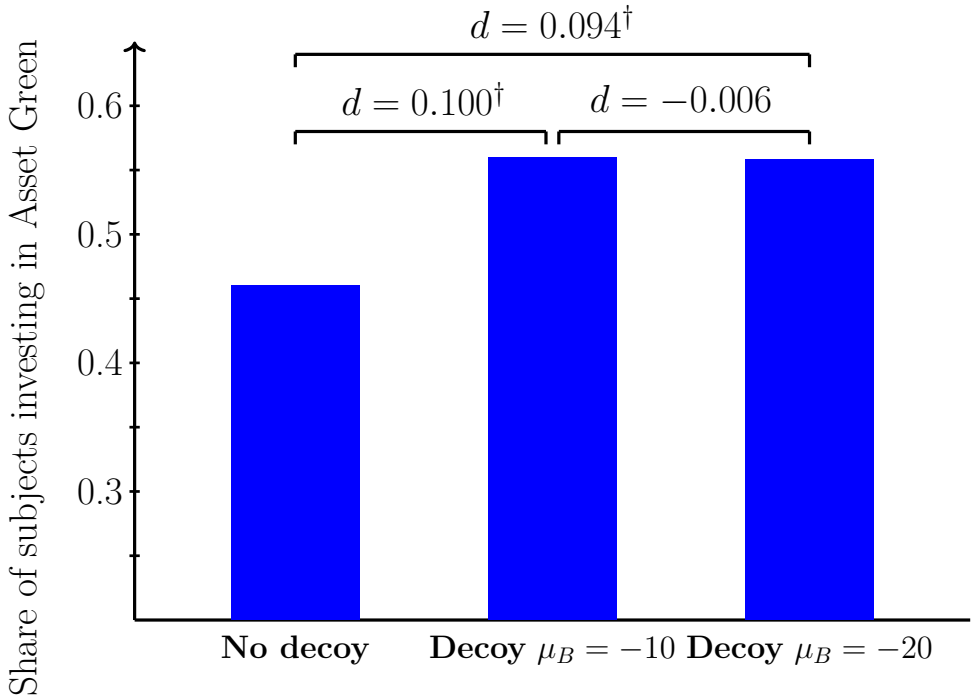


Figure 2.16: *The figure depicts the share of subjects that invested in Asset Green with and without a decoy. As pre-registered, when a decoy was available, we drop the subjects who chose the dominated option (roughly, 4% for $\mu_B = -10$ and 2% for $\mu_B = -20$) before calculating these shares. We further present the results of t-tests with standard errors being clustered at the subject level. †: p-value = 0.013.*

Definition 7. *The* salience-weighted utility *of lottery X_i evaluated in $\mathcal{C} = \{X_j\}_{j=1}^n$ equals*

$$U^s(X|\mathcal{C}) = \frac{1}{\int_{\mathbb{R}^2_{\geq 0}} \sigma\Big(v(x_i), r_i\Big)\ dF(x_1, \ldots, x_n)} \int_{\mathbb{R}^2_{\geq 0}} v(x_i) \cdot \sigma\Big(v(x_i), r_i\Big)\ dF(x_1, \ldots, x_n),$$

where $r_i = \phi(v(x_1), \ldots, v(x_{i-1}), v(x_{i+1}), \ldots, v(x_n))$, and where $\sigma : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{>0}$ is a salience function that is bounded away from zero.

2.G.3 Salience Predictions on Decoy Effects

Consider the choice set $\mathcal{C} = \{X_{\tau_{a,b}}, Y_{\tau_{a,b}}, z\}$. Here, the reference point R_k relative to which Asset $k \in \{G, B\}$ is evaluated has the following distribution:

If the choice set includes also the dominated Asset B and if this dominated asset has a sufficiently negative drift, then — compared to the case with a binary choice set — Asset G becomes more attractive to a salient thinker. Hence, a sufficiently "bad"

	$p_G p_B$	$p_G(1-p_B)$	$(1-p_G)p_B$	$(1-p_G)(1-p_B)$
R_G	$\phi\Big(v(z),v(a)\Big)$	$\phi\Big(v(z),v(b)\Big)$	$\phi\Big(v(z),v(a)\Big)$	$\phi\Big(v(z),v(b)\Big)$
R_B	$\phi\Big(v(z),v(a)\Big)$	$\phi\Big(v(z),v(a)\Big)$	$\phi\Big(v(z),v(b)\Big)$	$\phi\Big(v(z),v(b)\Big)$

Table 2.4: *Distribution of the reference points in the larger choice set.*

Asset B serves as a decoy that boosts demand for Asset G. More formally, we obtain the following proposition:

Proposition 7 (Salience-Driven Decoy Effect)**.**

(a) The salient thinker will never invest in the dominated Asset B.

(b) The salience-weighted utility derived from investing in Asset G monotonically increases in p_B.

(c) There is some $\hat{\mu} \in \mathbb{R} \cup \{-\infty\}$, so that a salient thinker invests in Asset G if and only if $\mu_B < \hat{\mu}$.

(d) If the salient thinker invests in Asset G when facing the binary choice set $\{X_{\tau_{a,b}}, z\}$, then $\hat{\mu} \in \mathbb{R}$.

Proof. Part (b). The salience-weighted utility from investing in Asset G is

$$U^s(X_{\tau_{a,b}}|\mathcal{C}) = v(a)\pi(p_G,p_B) + v(b)\Big[1-\pi(p_G,p_B)\Big],$$

where

$$\pi(p_G,p_B) := \frac{p_G s_a(p_B)}{p_G s_a(p_B) + (1-p_G)s_b(p_B)},$$

with $s_a(p_B) := p_B\sigma\Big(v(a),\phi(v(z),v(a))\Big) + (1-p_B)\sigma\Big(v(a),\phi(v(z),v(b))\Big)$ being the average salience of a and $s_b(p_B) := p_B\sigma\Big(v(b),\phi(v(z),v(a))\Big) + (1-p_B)\sigma\Big(v(b),\phi(v(z),v(b))\Big)$ being that of b.

Since $\phi(v(z),v(b)) > \phi(v(z),v(a))$ and, thus, $\sigma\Big(v(a),\phi(v(z),v(b))\Big) > \sigma\Big(v(a),\phi(v(z),v(a))\Big)$ by the ordering property, $s_a(p_B)$ is strictly decreasing in p_B. Analogously, ordering implies that $\sigma\Big(v(b),\phi(v(z),v(a))\Big) > \sigma\Big(v(b),\phi(v(z),v(b))\Big)$, so that $s_b(p_B)$ is strictly increasing

in p_B. It follows that $\pi(p_G, p_B)$ is strictly decreasing and, thus, $U^s(X_{\tau_{a,b}}|\mathcal{C})$ is strictly increasing in p_B.

Part (a). Next, we observe that

$$\begin{aligned}\frac{\partial}{\partial p_G}\pi(p_G,p_B) &= \frac{s_a(p_B)\Big[p_G s_a(p_B)+(1-p_G)s_b(p_B)\Big]-p_G s_a(p_B)\Big[s_a(p_B)-s_b(p_B)\Big]}{\Big(p_G s_a(p_B)+(1-p_G)s_b(p_B)\Big)^2}\\ &= \frac{s_a(p_B)s_b(p_B)}{\Big(p_G s_a(p_B)+(1-p_G)s_b(p_B)\Big)^2} > 0,\end{aligned}$$

which, in turn, implies that $U^s(X_{\tau_{a,b}}|\mathcal{C})$ is strictly decreasing in p_G.

Combining this with Part (b) and the fact that $p_B > p_G$, we conclude:

$$\begin{aligned}U^s(X_{\tau_{a,b}}|\mathcal{C}) &= v(a)\pi(p_G,p_B)+v(b)\Big[1-\pi(p_G,p_B)\Big]\\ &> v(a)\pi(p_G,p_G)+v(b)\Big[1-\pi(p_G,p_G)\Big]\\ &> v(a)\pi(p_B,p_G)+v(b)\Big[1-\pi(p_B,p_G)\Big] = U^s(Y_{\tau_{a,b}}|\mathcal{C}).\end{aligned}$$

Part (c). Follows immediately from the fact that $U^s(X_{\tau_{a,b}}|\mathcal{C})$ is increasing in p_B.

Part (d). Given the binary choice set $\mathcal{C}' = \{X_{\tau_{a,b}}, z\}$, the salience-weighted utility from investing in Asset G is given by $U^s(X_{\tau_{a,b}}|\mathcal{C}') = v(a)\tilde{\pi} + v(b)\Big[1-\tilde{\pi}\Big]$, where

$$\tilde{\pi} := \frac{p_G\sigma\Big(v(a),v(z)\Big)}{p_G\sigma\Big(v(a),v(z)\Big)+(1-p_G)\sigma\Big(v(b),v(z)\Big)}.$$

Notice that

$$\lim_{p_B\to 1}\pi(p_G,p_B) = \frac{p_G\sigma\Big(v(a),\phi(v(z),v(a)\Big)}{p_G\sigma\Big(v(a),\phi(v(z),v(a)\Big)+(1-p_G)\sigma\Big(v(b),\phi(v(z),v(a)\Big)} < \tilde{\pi},$$

since, by ordering, $\sigma\Big(v(a),\phi(v(z),v(a)\Big) < \sigma\Big(v(a),v(z)\Big)$ and $\sigma\Big(v(b),\phi(v(z),v(a)\Big) > \sigma\Big(v(b),v(z)\Big)$. Hence, if $U^s(X_{\tau_{a,b}}|\mathcal{C}') > 0$, then also $\lim_{p_B\to 1} U^s(X_{\tau_{a,b}}|\mathcal{C}) > 0$. □

2.G.4 Experimental Instructions

Screen 1 — Instructions: Overview of the Experiment

In this experiment you will make investment decisions. More precisely, in each decision, you will have to decide whether to invest in an asset and, if so, in which one. In total you will make 3 investment decisions.

At the end of the experiment, we will choose one of your 3 decisions at random to be payoff-relevant. Your payment depends on whether you invested in an asset and, if so, at which price the asset is sold. During the experiment, we will denote all monetary values in the currency Taler which will be converted to č at an exchange rate of č1 = 60 Taler.

You will receive a show up fee of č3 for your participation in the experiment.

Screen 2 — Instructions: Development of the Asset Price
Below you can see a graph, which depicts the development of the price of an asset, for a period of 10 seconds. As soon as you press "Start", a line which represents the value of the asset will appear. Please press "Start" now.
[Subjects are shown a graph of an exemplary price path with a final price of 100 Taler]

Screen 3 — Instructions: Different Drifts
The assets that you can choose in this experiment differ only in their long-term profitability. How profitable an asset is in the long run is described by the drift of the asset. The drift denotes the average change in the assets value per second.

A positive drift implies that the asset will increase in value in the long run, while a negative drift implies that the asset will decrease in value in the long run. A drift of zero implies that the assets value will neither in- nor decrease in the long run.

Notice, though, that even an asset with a negative drift sometimes increases in value, and that independent of its drift the value of an asset can, in principle, become arbitrarily large. The probability that an asset's value indeed becomes very large is the smaller the more negative the drift is.

To give you a feeling of how the long-run value of an asset changes with its drift we will show you now a few examples of assets with different drifts.

Screens 4-6 — Instructions: The Drift of an Asset
The drift of this asset is 0 [or 2 or -2]. Please press "Start" and watch the development of the asset's price.

[Subjects are shown a graph of an exemplary price path with a final price of 100 Taler for the process with drift 0, 120 Taler for the process with drift 2, and 80 for the process with drift -2.]

Screen 7 — Instructions: Upper and Lower Bounds on the Assets Price
If you decide to invest in an asset, it will be sold at pre-specified prices: either at 90 (lower bound) or at 190 (upper bound). All available assets will be automatically sold as soon as their prices hit one of these bounds. If the price of an asset reaches the upper bound of 190, the asset will be sold and you will receive 190 Taler. If the price reaches the lower bound of 90, the asset will be sold as well and you will receive 90 Taler.

Importantly, there is no expiration time. The price of an asset will change until it hits one of the two bounds.

[Subjects are shown a graph with a button labelled "Asset Green". When they click the button, a price path appears]

Screen 8 — Instructions: Upper and Lower Bounds on the Assets Price
You have to decide repeatedly whether to invest in one of at most two different assets. You can always opt for the alternative of no investment in which case you will receive 100 Taler (i.e. the starting value of each asset) with certainty.

In the example below you can choose among the following three options:

- No investment
- Invest in Asset Green (which has drift 0)
- Invest in Asset Blue (which has drift -2)

Please make a decision now.

[Subjects are shown a graph with three buttons labelled "Asset Green", "Asset Blue" and "Sell Immediately"]

Screen 9 — Questions about the Instructions

[Subjects see a graph with an upper bound of 190, a lower bound of 90 and a starting value of 100]

[Subjects need to answer the following questions]

- Suppose you choose "No Investment". How many Taler do you get?
- What is the starting value of the asset?
- Suppose you have invested in the asset and it hits the lower bound. How many Taler do you get?
- If an asset has a drift of 2 it will:

[In the last question they choose between the following options]

- increase by an amount of 2 every second
- increase by an amount of 2 per second on average over a long time period
- vary by an amount of 2 every second

Screen 10 — Decision round x of 3

In the upcoming decision you will be able to choose between the following options:

- No investment. (100 Taler with certainty)
- An asset with a drift of 0. (Green Asset)
- An asset with a drift of [-10 / -20] (Blue Asset)

Before you make your decision we will show you a few exemplary price paths of the asset.

Screen 11 — Sampling page x of 3

Below you can see exemplary price paths of Asset Green (with drift 0), for a period of ten seconds.

[Subjects see one individual price path from Asset Green. If there is also an Asset Blue, they see a price path from Asset Blue as well. This screen is shown 3 times in a row]

Screen 12 — Overview of 5 Exemplary Price Paths

Below you can see 5 paths from the Asset Green with drift 0.

[Subjects see 5 price paths from Asset Green. If there is also an Asset Blue, they see 5 price paths from Asset Blue as well.]

Screen 13 — Decision x of 3

Please choose between Asset Green (with drift 0) [and Asset Green (with drift -10/-20)] and "No Investment".

[Subjects are shown a graph with three buttons labelled "Asset Green", "Asset Blue" and "Sell Immediately"]

Screen 14 — Payoff

Your decision from round [a] will be paid.

You sold the asset for [b] Taler.

You earned a bonus payment of č[c] on top of your show up fee of č3. Your total payment is č[d].

2.H Stopping a Process Prematurely

In this section, we present simulations that suggest that (1) loss-exit strategies induce right-skewed return distributions, and (2) salience theory could predict, after having started, stopping before the expiration date.

To do this, we simulate the development of an ABM with starting value $x_t = 100$, drift $\mu = 0$, standard deviation $\sigma = 5$ and expiration date $T = 100$, for a given loss-exit strategy $\tau_{60,180}$ with lower threshold 60 and upper threshold 180 exactly 500 million times at each $t \in \{0, \ldots, 100\}$. If the ABM hits the bounds at any time, we terminate the process prematurely and register the boundary value as the outcome. This approach provides us with an approximation of the outcome distribution for the truncated ABM, enabling us to compute its skewness.

Figure 2.17 shows that loss-exit strategies indeed induce right-skewed return distributions, but close to the expiration date, the return distribution is approximately symmetric as the selected thresholds are unlikely to be met before the expiration date.

For (2), we show that gambling becomes less attractive as time passes (i.e., t increases and the remaining time $T - t$ decreases). Specifically, we show that the salience-weighted utility of gambling (Definition 2) is decreasing in t for a range of points in time. For that, we need to assume a specific functional form of the salient thinker's value and salience functions. We use a linear value function and as a salience function we use an adaptation of the salience function $|x - y|/(|x| + |y| + 0.1)$ that Bordalo et al. (2012) use in their rank-based salience model, namely

$$\sigma(x, y) = \delta^{\frac{|x-y|}{|x|+|y|+0.1}}.$$

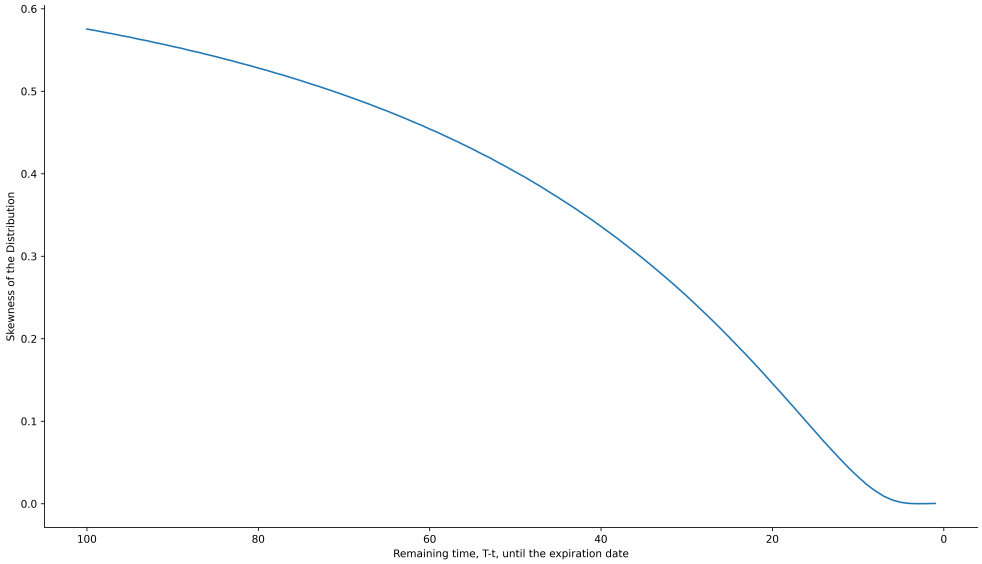


Figure 2.17: *Skewness of the outcome distribution induced by* $\tau_{60,180}$

Here, $\delta > 1$ controls the strengths of the salience distortion, and as there are no previous calibrations of the continuous salience model that could guide our choice of δ, we set $\delta = 2$.[33] With these functional-form assumptions and the simulated outcome distribution, we can calculate the salience-weighted utility of gambling at each $t \in \{0, \ldots, 100\}$. However, to make this computationally feasible, we need to reduce the number of states by assigning the outcomes of our simulated distributions into bins of size 0.1.[34]

Figure 2.18 shows that, like the skewness of the outcome distribution, the salience-weighted utility of gambling is decreasing over time for a considerable range of points in time. Initially, the salience-weighted utility exceeds 100, so that gambling is attractive, but falls below this value as the expiration date approaches. This demonstrates that a salient thinker might start to gamble and stop before the expiration date. Notably, the fact that the salience-weighted utility is increasing close to the expiration date follows from decreasing variance of the return distribution together with

[33]Because we use the continuous salience model, we cannot rely on the calibration of Bordalo et al. (2012) and others that relied on the rank-based salience model. Assuming $\delta = 2$ gives us only quite mild salience distortion, which however suffices for our purpose of demonstrating in-between stopping.

[34]This is necessary because each simulation could lead to a different final value of the ABM, leaving us with up to 500 million states.

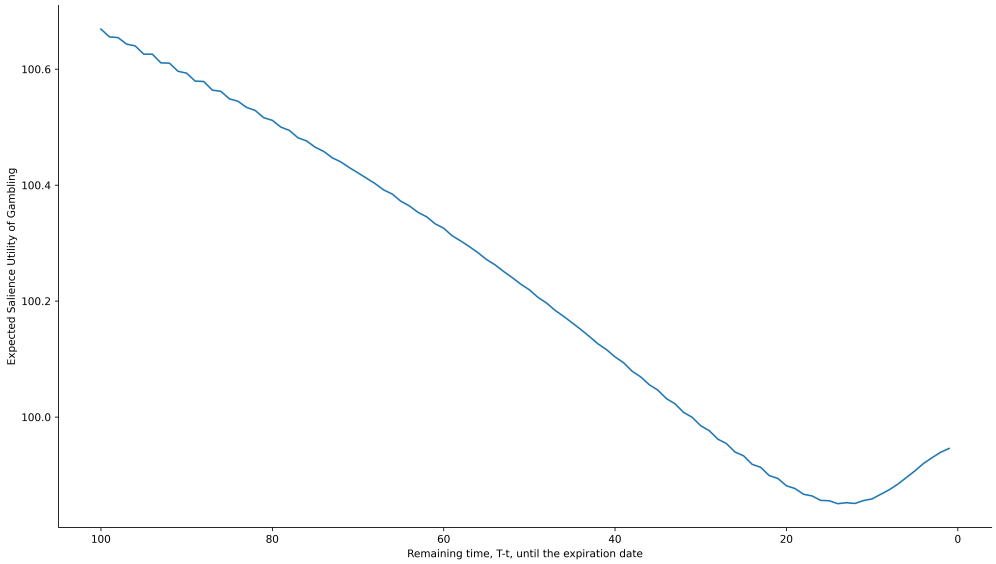


Figure 2.18: Salience-Weighted Utility of Gambling over Time

diminishing sensitivity.[35]

[35]The intuition behind this result is the following: The return distribution close to the expiration date is quite symmetric as the bounds are unlikely to be hit, and by diminishing sensitivity the lower payoffs are overweighted relative to the higher payoffs. This effect, however, decreases when the variance of the process decreases.

Chapter 3

Do contrast effects influence earnings expectations?

Abstract

I study how professional analysts' earnings forecasts made shortly after a firm's earnings announcement are affected by news about other firms. I find that positive earnings surprises of other firms are positively linked to both forecast revisions and forecast errors, suggesting that the other firms' surprises are relevant news to which analysts underreact. A psychological "contrast effect" as in Hartzmark and Shue (2018) can explain the result. I study if the revisions of expectations that are explained by other firms' surprises are linked to abnormal returns and find mixed evidence.

3.1 Introduction

Expectations are an integral part of any economic or financial decisions involving an uncertain future outcome. For instance, asset prices are driven by expectations about future cash flows, firms' investment decisions depend on expectations about the macroeconomic conditions and consumption decisions depend on expectations about future income. Therefore, it is pivotal to understand how economic agents form expectations. Since Lucas and Prescott (1971) rational expectations has been the workhorse assumption of economists. While it provides a useful benchmark a vast amount of research documents systematic deviations from rational expectations. Evidence comes both from laboratory experiments (e.g. Afrouzi et al. (2021) and Glaser et al. (2019)) and from survey data of professional forecasters and households (e.g. Coibion and Gorodnichenko (2012, 2015) and Greenwood and Shleifer (2014)).

The existing research has focussed on documenting under and overreaction to news about the forecasted variable and examined if the way that information is presented or the properties of the forecasted process can explain deviations from rationality. In this paper, I take a different approach and examine how forecasts depend on news about a different but related variable. Specifically, I examine how earnings forecasts for a firm depend on earnings surprises of other firms that announced their earnings shortly prior to the forecasted firm. To do this, I use earnings forecasts by professional analysts as well as other data described in Section 3.2.

In Section 3.3, I formulate my hypotheses. The other firms' earnings surprises can influence the analysts' forecasts through two possible channels. First, they may contain relevant news about the future earnings of the forecasted firm. In this case, they should influence the analysts' forecast revisions. The link can be either positive or negative, depending on whether a positive surprise of the contrast firm is good or bad news for the forecasted firm. If analysts process this news without bias, the other firm's surprise should have no explanatory power for forecast errors. Second, the surprises of other firms may cause a psychological contrast effect. In the context of earnings forecasts, contrast effects imply that a given earnings surprise of a firm will look more positive (negative) the worse (better) it compares to recent earnings surprises of other firms. This implies

that forecast errors should be predictable from the contrast surprises. In the absence of a news channel, it also implies that contrast surprises should be negatively linked to forecast revisions. Hartzmark and Shue (2018) (henceforth HS) find evidence of a contrast effect in the market reaction to earnings announcements and the phenomenon has also been documented across a wide range of other domains.

Section 3.4 contains my main results. I find that analysts revise their forecast of a firm's earnings upwards if other firms have positive earnings surprises. However, the other firms' surprises are positively linked to forecast errors, suggesting that analysts' forecasts become more negative relative to the actually realised future earnings. This finding is consistent with a positive news effect of other firms' surprises together with a psychological contrast effect, whereby the news effect dominates quantitatively.

Subsequent forecast revisions occurring between 15 and 45 days after the announcement are positively linked to contrasts suggesting that the influence of the psychological contrast effect gets corrected over time. However, I also find evidence that forecast errors remain positively linked to contrast surprises up to 75 days after the announcement, implying that this correction occurs slowly. Sticky expectations (i.e. the anchoring of new forecasts in previous forecasts) (Coibion and Gorodnichenko, 2015) could explain the persistence of the contrast effect.

Moreover, I study the impact of contrast effects on absolute forecast error and find evidence that some types of contrast surprises are linked to higher absolute forecast errors and, thus, overall less accurate forecasts. The fact that some contrast surprises are not linked to larger forecast errors can be explained by the fact that on average forecasters are too optimistic. Therefore, a contrast effect that makes them more pessimistic can increase their accuracy by canceling out their existing over-optimism.

To shed further light on the relation between contrast surprises, subsequent revisions and resulting forecast errors I use a two-stage regression procedure. In the first step, I regress forecast revisions on contrast surprises, controlling for the firm's own surprise. I then regress forecast errors on the part of the revision that is explained by contrast surprises. Coibion and Gorodnichenko (2015) show that a positive coefficient in a regression of forecast error on forecast revision implies underreaction while a negative coefficient

implies overreaction. I find that the revisions explained by the contrast surprise are positively linked to forecast errors. This is consistent with positive contrast surprises being positive news to the firm, to which forecasters underreact due to contrast effects.

The magnitude of the contrast effect on forecasts is small. However, because analysts revise their forecasts infrequently, I have to include forecasts made several days after the announcement to achieve a sufficient sample size. Thus, I likely underestimate the true effect of contrasts on the expectations of market participants. Moreover, given that HS find a substantial effect of other firms' surprises on returns, understanding the mechanism behind their result is important, even if the effect sizes I can document for forecasts are small.

In Section 3.5, I study if changes in expectations can explain the contrast effect in return reactions to earnings announcements discovered by Hartzmark and Shue (2018). Moreover, I test if expectation data can be used to predict returns later in the quarter. I find no consistent link between the contrast effect in expectations and subsequent returns. I discuss potential explanations for the lack of a link and conclude that methodological challenges in linking forecast and return data are a likely explanation. In Appendix 3.B I discuss in more detail how my findings can be reconciled with the results of HS.

My research is related to the existing literature that studies deviations from rationality in expectation formation. Over the last decade, a vast amount of research using survey data has documented systematic errors in expectations across a wide range of domains such as inflation forecasts and forecasts of other macroeconomic variables (Bordalo et al., 2020; Coibion and Gorodnichenko, 2012, 2015), firm earnings (Bordalo et al., 2019; La Porta, 1996), bond yields (d'Arienzo, 2020) and stock returns (Adam et al., 2017; Amromin and Sharpe, 2013; Greenwood and Shleifer, 2014).

Survey forecasts have been repeatedly found to be linked to actual behaviour or market price movements (Bacchetta et al., 2009; Gennaioli et al., 2016; Greenwood and Shleifer, 2014; Ma et al., n.d.) and models with biased expectations that are calibrated on survey data can explain patterns in the stock market (Bouchaud et al., 2019; De La O and Myers, 2021; Engelberg et al., 2018; Greenwood and Shleifer, 2014). Moreover, in machine learning models that use both fundamental information and analyst forecasts,

the forecasts have been found to be a significant component of the optimal forecast (van Binsbergen et al., 2022; de Silva and Thesmar, 2021). These results alleviate potential concerns that unincentivised survey measures do not capture the true expectations of economic agents. Moreover, similar deviations violations of rational expectations have also been found in incentivised laboratory experiments (see Afrouzi et al. (2021) for a recent large-scale experiment and an overview of the recent literature and Assenza et al. (2014) for a review of the older literature).

Overall there is evidence for both overreaction and underreaction to news about the forecasted variable. More recently there have also been efforts to determine under which conditions each of these two occur. Kwon and Tang (2020) suggest that the extremeness of the fundamentals associated with the news can explain this.

I add to the literature by studying the impact of contrast effects on expectations, which is crucial because they are orthogonal to any form of bias in processing news directly related to the forecasted variable. Hence, they are not captured by the commonly used existing behavioural models of expectation formation. Moreover, readily available contrasts exist for virtually any economic variable about which agents need to form expectations. Firm-related news can be contrasted to news of other firms, country-level news such as inflation or GDP growth data can be contrasted to news about other countries and product-related news e.g. price developments can be contrasted to news about similar products. Given the widespread availability of contrast and the significant evidence that contrast effects are prevalent across a wide range of domains including low-level information processing, it is likely that they have an effect on many economic decisions that has not yet been studied systematically.

Moreover, I add to the literature on contrast effects. A large literature in psychology has found evidence for contrast effects across a wide range of domains such as scholarship candidate evaluations (Radbruch and Schiprowski, 2020), judgements of crime severity (Pepitone and DiNubile, 1976), dating decisions (Bhargava and Fisman, 2014), perception of beauty (Cash et al., 1983; Kenrick and Gutierres, 1980) and life satisfaction (Schkade and Kahneman, 1998). Contrast effects have also been found in decisions made by non-human animals (Waite, 2001), as well as the firing rate of (animals') neurons (Wallis

and Kennerley, 2011). The evidence, therefore, suggests that contrast effects are hard-wired into the perception and decision processes of the brain. More recently studies have started to examine the impact of contrast effects on economic and financial decision-making and have found that they affect the market response to earnings announcements (Hartzmark and Shue, 2018) and stock investment decisions (Antoniou et al., 2021; Kim and Hoffman, 2020). Moreover, they are a fundamental component of the salience theory of choice under risk (Bordalo et al., 2012), which is a widely applied theory that can explain a multitude of puzzles in decision-making under risk.

In many situations in which previous studies find evidence of contrast effects, expectations likely play a critical role. For instance, expectations about future earnings are a key driver of stock returns and changed expectations should explain a large share of the stock market reaction to earnings announcements. However, so far there is no direct evidence of how contrast effects influence expectations. Therefore, I add to the literature by documenting a contrast effect in the forecasts of professional earnings analysts.

Finally, because of the key role earnings announcements play in the corporate news cycle, determining how investors react to them is an important topic that has received significant attention. A central finding in this literature is the post-earnings announcement drift (Ball and Brown, 1968) which is generally seen as evidence that the market underreacts to the earnings news and that this underreaction is slowly corrected over the weeks after the announcement (Bernard and Thomas, 1989). Another important phenomenon is that firms tend to earn abnormal positive returns after announcements (Chari et al., 1988; Cohen et al., 2007).

Beyond uncovering return patterns around earnings announcements, researchers have studied which conditions affect the market's response to announcements. Attention is an important driver. For instance, On Fridays, when investors are distracted by the start of the weekend the market reaction to announcements is less pronounced (Dellavigna and Pollet, 2009). Furthermore, investors react less to a given announcement on days during which they have to process more concurrent announcements (Hirshleifer et al., 2009) and if the announcement happens during trading hours (Michaely et al., 2014). Retail (but not professional) investors' attention also seems to be affected by the amount of non-stock

related news (daily news pressure) on the announcement day (Israeli et al., 2021). There is also evidence that firms strategically time announcements and report more negative earnings surprises on low attention days (deHaan et al., 2015).

Most closely related to this paper are studies which examine the impact that a firm's announcement has on other firms. As discussed above, Hartzmark and Shue (2018) show that contrast effects can help explain the market reaction to earnings announcements. Savor and Wilson (2016) highlight that "earnings reports provide valuable information about the prospects of not only the issuing firms but also their peers and more generally the entire economy" and show that this can help explain the earnings announcement premium. I find evidence both for the contrast effect and for learning about other firms' prospects from a firm's announcement in the exaptations of analysts. This adds to existing results (Bordalo et al., 2019) showing that expectations data can be used to understand the market reaction to announcements.

3.2 Data

3.2.1 Sources and Inclusion Criteria

I use data on analyst forecasts from the IBES detail history file. I obtain individual-level forecasts of earnings per share (EPS) for between one and eight quarters ahead. In addition, I gather data on announcement dates and actual EPS from IBES.

Data on prices comes from the CRSP daily stock database. I gather data on the daily closing price and the shares outstanding which I use to calculate returns and market capitalisations.

Accounting data comes from the CRSP/Compustat merged database. I obtain additional EPS data (basic, excluding extraordinary earnings), as the IBES dataset only contains EPS for announcements for which there is at least one forecast available. I use the actual EPS data from IBES to calculate forecast errors and the EPS data from CRSP/Compustat for everything else. Moreover, I collect data on book equity at the end of the year. I collect data on the daily level of the S&P500 from CRSP. In addition,

I use Industry classifications from the Kenneth French data library. My data covers the years from 1980 to 2022.

In the key parts of the analysis, I examine the impact of news about other firms on expectations measured by analyst forecasts and study how these expectations are linked to returns. Therefore, I require that the earnings surprise, as well as the forecast revision around the announcement, can be calculated in order for a data point to be included in my dataset. Details on how these variables are calculated can be found in section 3.2.2.

3.2.2 Calculation of Derived Variables

In this section, I describe the calculation of important variables used throughout the analysis section.

Earnings Surprise

Following HS I define earnings surprise as the difference between the announced earnings and the market's expectation of earnings before the announcement. Since the market's expectations are unobservable I use the median (consensus) analyst forecast prior to the announcement as a proxy. To avoid using stale information I include only forecasts made within 15 days before the announcement in the calculation of the consensus forecasts. Moreover, I only include forecasts from at least two days before the announcement to avoid confounds between earnings surprise and other variables of interest measured at time t-1 or t. If an analyst makes more than one forecast in the 15 days time window I use the most recent one. Since the level of a company's EPS and therefore also the difference between expected and actual EPS depends on the share price I normalise by price to make different companies comparable. Therefore, the earnings surprise of firm i on day t is

$$\frac{Actual\ earnings_{i,t} - Consensus\ forecast_{i,[t-15,t-2]}}{Price_{i,t-3}}$$

The earnings surprise is winsorized at the 1% level to deal with outlier issues.

Contrast Surprises

The contrast surprises for a firm are the earnings surprises of other firms that announced their earnings at a similar time as the firm. I define three different contrast surprises. $Contrast_{t-1}$ is the value-weighted earnings surprise of all large firms (defined as being above the 90th percentile of NYSE market capitalisation) that announced on the day prior to the firm's own announcement. This definition follows HS. $Contrast_t$ is defined like $Contrast_{t-1}$ except that it captures the surprise of other firms that announced on the same day but at an earlier time as the firm. In addition to these two firm-level contrasts, I define an analyst-level contrast variable. $Contrast_{SA}$ is the value-weighted surprise of all firms which announced up to 15 days before the firm and which were covered by the same analyst as the firm. The contrast surprises are formed from the winsorized earnings surprises of the firms and are then winsorized again at the 1% level.

Forecast Revision

A forecast revision is defined as the forecasted value for a certain quarter minus the forecasted value for the same quarter from a forecast made at an earlier time. I include forecasts for between one and eight quarters ahead. My analysis focuses on forecast revisions around earnings announcements. For these revisions, the forecast after the announcement needs to be made within 15 days to make sure it is mostly driven by the news coming from the announcement of the firm and the comparison firms. The forecast before the announcement can be made at any point during the preceding quarter. I also construct an additional dataset in which the forecast before the announcement also has to be made within 15 days of the announcement. This should reduce noise by further limiting the amount of unrelated information that becomes available during the revision window but it also reduces the sample size by roughly 75% of the main dataset since it is relatively uncommon for analysts to update their more long-term forecasts shortly before an announcement.

Characteristics adjusted return

Unless stated otherwise I use characteristics adjusted returns instead of raw returns throughout the paper. To calculate characteristics adjusted returns I subtract returns of a portfolio matched on size, book value of equity divided by market value of equity and momentum from the raw returns. In calculating these portfolios I follow HS who in turn follow Daniel et al. (1997). I sort companies into quintiles based on market capitalisation, book value of equity divided by market value of equity and momentum return calculated as the return from 252 up to 20 trading days before the date t.[1] While I use data from all available US stocks in the calculation, I calculate the market capitalisation cut-offs for the quintiles using only firms listed at the NYSE to avoid using too many small companies in the calculation. The book-to-market ratio is calculated yearly using the book value at the end of the firm's fiscal year during the preceding calendar year and the market value at the end of December of the preceding year.

3.2.3 Comparing Earnings per Share Across Companies

A company's earnings per share can only be meaningfully interpreted in conjunction with its share price. EPS crucially depend on the number of shares outstanding and can easily be changed by (reverse) stock splits. Because of this two otherwise identical companies can have vastly different earnings per share if they have a different amount of shares outstanding. This scaling difference is an issue for any analysis that examines a link between EPS of two companies. In particular, the magnitude of regression coefficients depends on the scaling of the dependent relative to the independent variables and differences in scaling within the dependent or independent variables can influence the weight of an individual observation on the regression and therefore affect not only the magnitude but also the sign of coefficients. The same issues arise in analysis that does not use EPS directly but includes variables related to EPS, such as forecasted EPS, forecast revisions or forecast errors.

[1]There is a possibility that drifts around earnings announcements may influence the momentum return of the matched portfolio. However, HS show that excluding each company and potential contrast companies from the calculation of the momentum return yields similar results.

To reduce the difference in magnitude between EPS of different companies, I scale all EPS-related variables by price and hence transform all measures from earnings per share to earnings per $ of market capitalisation. This variable is unaffected by the number of shares outstanding and is only driven by measures which are economically meaningful. While this scaling reduces the difference in magnitude of EPS between companies it does not eliminate it. Since the market capitalisation of a company does not only depend on its current earnings but also other factors, most notably expected growth, earnings per $ of market capitalisation are still different between companies. These differences are large enough that asset managers and researchers use them to categorise stocks into value stocks, which have high earnings-to-market capitalisation ratio and growth stocks, which have a low earnings-to-market capitalisation ratio Fama and French (1998). Therefore, I take additional steps to reduce the impact of scaling on my analysis most notably winsorization and the use of outlier robust estimation methods which are described in detail ins Section 3.3.

3.2.4 Descriptive Statistics

Table 3.1 describes the data. The forecast error in the table refers to the error of the forecast made after the announcement. Both forecast revision and forecast error are negative on average. A negative forecast error implies that analysts are too optimistic after the announcements even after the downward revision they make on average. Interestingly, both the own surprise and all contrast surprises are positive on average and even the 25th percentile is positive for the three contrast surprises. As these surprises are based on the consensus forecast including all forecasts made within 15 days of the announcement, this suggests that in the very short term analysts' forecasts are too pessimistic rather than too optimistic. This may be the result of earnings management, where firms explicitly try to report earnings that exceed analyst expectations. In line with the on-average positive earnings surprise, announcement returns are also positive on average. A striking feature of the data is the extremely large kurtosis. This illustrates the issues of EPS scaling described in Section 3.2.3.

Table 3.1: Descriptive Statistics

	N	Mean	SD	Kurtosis	P25	P50	P75
$forecast\ revision$	1978737	-0.0009	0.0065	13.39	-0.0016	-0.0002	0.0009
$forecast\ error$	1978737	-0.0020	0.0181	10.59	-0.0043	0.0000	0.0025
$own\ earnings\ surprise$	1978737	0.0012	0.0092	16.00	-0.0002	0.0006	0.0024
$contrast_{t-1}$	1829441	0.0013	0.0034	30.11	0.0003	0.0009	0.0017
$contrast_t$	1648382	0.0011	0.0022	9.41	0.0003	0.0008	0.0016
$contrast_{SA}$	905115	0.0013	0.0050	10.44	0.0000	0.0007	0.0021
$return_{t-1,t+1}$	1616976	0.0006	0.0777	8.68	-0.0361	0.0006	0.0380
$return_{t+2,t+25}$	1616751	-0.0002	0.1061	36.57	-0.0508	-0.0029	0.0449
$return_{t+26,t+50}$	1614787	-0.0021	0.1096	17.31	-0.0535	-0.0032	0.0466

The table shows descriptive statistics for the main variables used in the analysis from 1980 to 2022. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecast Error and Forecast Revision are pooled across forecast horizons. Forecast revision, Forecast Error, as well as own surprise, are winsorised at the 1% level. The contrast surprises are formed using the winsorised individual surprises and are again winsorised at the 1% level. Forecast error, forecast revision as well as own and contrast earnings surprises are scaled by the firm's price 3 days prior to the corresponding earnings announcement.

Table 3.2 shows descriptive statistics for forecast errors and forecast revisions separated by forecast horizon. One notable finding is that data availability decreases with horizons. Forecast errors are negative across all horizons and their magnitude monotonically increases in the horizon which is likely driven by the increased difficulty of long-term forecasts due to higher uncertainty. Forecast revisions are also negative across horizons and seem to get smaller as the horizon increases. This could be driven by the fact that the current announcement is less informative for the long-term prospects of the firm relative to other available information.

Table 3.2: Descriptive Statistics by Horizon

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
forecast revision							
N	507973	421336	347575	242402	182993	123974	77926
$Mean$	-0.0015	-0.0010	-0.0007	-0.0003	-0.0007	-0.0004	-0.0002
SD	0.0075	0.0067	0.0059	0.0058	0.0056	0.0050	0.0047
$Kurtosis$	9.8976	13.1105	12.7092	10.6184	12.0472	11.3984	9.4993
forecast error							
N	507973	421336	347575	242402	182993	123974	77926
$Mean$	-0.0003	-0.0018	-0.0026	-0.0028	-0.0029	-0.0034	-0.0038
SD	0.0131	0.0171	0.0186	0.0190	0.0201	0.0207	0.0212
$Kurtosis$	14.2332	12.3860	10.0336	8.8692	7.4681	6.6584	6.1872

The table shows descriptive statistics for forecast error and forecast revision separated by forecast horizon using data from 1980 to 2022. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecast revision and Forecast Error are winsorised at the 1% level and scaled by the firm's price 3 days prior to the corresponding earnings announcement

3.3 Methodology

3.3.1 Questions & Hypothesis

In my main analysis, I study how news about other firms are related to firm-specific expectation variables. This news could influence expectations via two channels: an information channel and a psychological channel. News about other firms' earnings should contain information that is relevant to the firm. In particular, it is relevant to assess the general economic condition and market environment. If other firms did surprisingly well this indicates generally favourable economic conditions and since most firms' earnings are cyclical this is generally positive news for the firm. Thus. news about other firms' earnings should be positively linked to changes in expected future EPS in most cases. In addition, the other firm's surprise may also contain other information, if the two firms are related e.g. by being competitors, customers or suppliers. However, due to my limited sample size, I cannot use a granular industry classification to study these effects in detail. I provide some results on industry effects using the 30 industry classification of Fama and French in Section 3.4.4.

In addition to the information channel, I expect to observe a contrast effect like HS found in the context of the reaction of returns to earnings news about other firms. This effect suggests that the firm's own earnings surprise looks worse the better the surprise of contrast firms is, which implies a negative link between contrast surprises and expectations. I have no direct way of disentangling the channels so I cannot make a clear prediction for the direction of the effect that contrast surprises have on the revision of analysts' expectations. However, it is an empirically interesting exercise to study the direction of the link which leads us to my first open question:

Question 1. *Are contrast surprises positively or negatively linked to forecast revisions?*

Unlike for forecast revisions, the existing evidence suggests a clear hypothesis for the link between contrast surprises and forecast errors. If analysts react optimally to the information in the contrast surprise they should be unrelated to forecast errors. However, there is ample evidence of both under and overreaction of expectations to news. If analysts underreact this means that after a positive (negative) contrast surprise they are too pessimistic (optimistic) which means they have a positive (negative) forecast error. Hence, if analysts underreact I expect a positive link between forecast errors and contrast surprises. Conversely, if they overreact I expect a negative link. While evidence for both under- and overreaction exists, in the context of earnings forecasts both Bordalo et al. (2019) and Bouchaud et al. (2019) find underreaction to news after the announcement for earnings per share in the short horizon that I study. While they do not study the reaction to news about other firms it is likely that the reaction is similar. In addition, the contrast effect suggests that analysts are too pessimistic (optimistic) if the contrast firms did well (poorly) which also suggests a positive link between forecast errors and other firms' surprises. This leads to my first hypothesis.

Hypothesis 1. *Forecast errors are positively linked to contrast surprises*

If Hypothesis 1 holds, an interesting question is how persistent the effects are. If analysts make forecasts at a later point after the announcement they should no longer think about the contrast earnings surprises, so the contrast effect should no longer influence their forecasts. If the initial contrast effect and the underreaction to the news

component in the other firms' surprises caused them to become overly too pessimistic (optimistic) after positive (negative) contrast surprises, there should be a positive link between contrast surprises and subsequent revisions and the effect on forecast errors should disappear. However, there is also evidence that forecasts are sticky in the sense that new forecasts are anchored in old forecasts (Bouchaud et al., 2019; Coibion and Gorodnichenko, 2015). This anchoring in old forecasts predicts a persistence of the contrast effect in later forecasts even if there is no direct contrast effect in these forecasts. This leads us to my next hypotheses:

Hypothesis 2. *A positive (negative) contrast surprise predicts subsequent upwards (downwards) revisions after the initial revision around the announcement.*

Hypothesis 3. *A positive (negative) contrast surprise predicts a positive (negative) forecast error for forecasts made more than 15 days after the announcement, which is weaker than the initial effect.*

As a next step, I decompose the changes in expectation (forecast revisions) into the components that are explained by the firm's own surprise and the other firms' surprises. I then link forecast errors to these decomposed revisions. Coibion and Gorodnichenko (2015) showed that the link between forecast revisions and forecast errors can be used to determine if analysts under or overreact. If the link is positive this implies underreaction and if it is negative this implies overreaction. Since I established that analysts are likely going to underreact to the news in the other firms' surprises and that the contrast effect also works in the same direction as underreaction to this news my fourth hypothesis is

Hypothesis 4. *Forecast errors are positively linked to the component of forecast revision that is explained by contrast surprises*

In addition to studying the link between contrast and expectations, I also analyse how these changes in expectations are linked to returns. Since expected future earnings are a key determinant of firm value, there should be a clear link. To test this, I regress returns on revisions explained by own and contrast surprise. In general, an upwards change in expectations should lead to positive returns, so I expect a positive link between forecast

revisions and returns. If I decompose the revisions into the part explained by the own surprise and the part explained by the contrast surprise, the own surprise component should clearly also be positively linked to returns. The link to the revisions explained by contrast surprises is less clear. If the market reacts to contrasts in the same way that analyst expectation does, I should see a positive link. This should be the case if either the analysts respond to the news about other firms in a rational way or if the market is affected by the same biases in information processing as the analyst. If the reaction of analysts to the contrast surprise is purely behavioural whereas the market is rational and does not react to contrast surprise, I should see no link. However, if the market and the analysts respond to the contrast surprise in the opposite direction, e.g. because for one of them, the information channel dominates, while for the other the contrast effect dominates I should see a negative link. A reason why expectation and returns could move in a different direction is that I use characteristics-adjusted returns. Due to this, any news component of the contrast surprise that does not only affect the firm but also affects all other firms in the market equally is filtered out by the characteristics-adjustment. This leads to the next question:

Question 2. *Is there a link between the revision that is explained by contrast surprises and the return around the announcement?*

In addition to looking at the return around the announcement, I also study the link between revisions explained by own and contrast surprise and subsequent returns. For this horizon, it is even more difficult to predict the direction of a possible link. In addition to the potential difference in the relation of returns and expectations to contrast this link also depends on the initial market reaction. However, it is empirically interesting to study. Since my revisions include forecasts made up to 15 days after the announcement any link between expectations and the announcement return is purely correlational. However, if I find a link between forecasts and later returns this would imply return predictability. Because HS find evidence for a significant link between contrasts and returns from t+26 to t+50 I focus on this time period and the question:

Question 3. *Do revisions that are explained by contrast surprises predict returns between t+26 and t+50?*

3.3.2 Regression Specification

To test my hypotheses and answer the questions I mainly use two regression specifications. To link expectations to contrasts I estimate the following regression model:

$$dependent\ variable = \beta_1 + \beta_2 Own\ Surprise_i + \beta_3 contrast_i + \epsilon_i \tag{3.1}$$

Unless otherwise stated, I follow HS and do not use the firm's own surprise directly in the regression. Instead, I sort own surprises into 20 bins and create an additional bin for surprises that are exactly zero. This allows me to capture non-linear effects of the firm's own surprise.

To estimate the impact of revisions explained by own surprise and contrast surprise on variables, I use a two-stage regression procedure. In the first stage, I estimate Equation 3.1 using forecast revisions as the dependent variable. I then use the coefficients from the first stage to get the components of the revisions that are predicted by the own surprise and the contrast surprise

$$revPredOwn_{i,a} = \beta_1 + \beta_2 ownSurprise_i \tag{3.2}$$

$$revPredContrast_{i,a} = \beta_1 + \beta_3 contrast_i \tag{3.3}$$

In the second stage, I regress the dependent variable on these predicted revisions:

$$dependent\ variable = \beta_4 + \beta_5 revPredOwn_{i,a} + \beta_6 revPredContrast_{i,a} + e_i \tag{3.4}$$

To deal with outliers caused by the scaling issue described in Section 3.2.3 I use a robust estimation method that decreases the weight of outliers for my main specifications.[2] Unless specified otherwise, I pool forecast revisions and forecast errors for different horizons in all specifications.

[2]Precisely, I estimate a model using a HuberT norm via iteratively reweighed least squares as implemented by the Python statsmodels module.

3.4 Expectations

3.4.1 Main Result

To examine the impact of contrast on forecasts I first study if they affect forecast revisions, which are made around the announcement (Question 1).

Table 3.3: Revision on Contrast

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	0.0218 (SE=0.0051) (p<0.001)	0.0123 (SE=0.0015) (p<0.001)				
$Contrast_t$			0.0527 (SE=0.0101) (p<0.001)	0.0242 (SE=0.0027) (p<0.001)		
$Contrast_{SA}$					0.0052 (SE=0.0030) (p=0.086)	0.0016 (SE=0.0008) (p=0.052)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0464	0.0651	0.0428	0.0619	0.0444	0.0603
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast revision on contrast. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Table 3.3 shows the results of a regression of forecast revision on contrast. For $contrast_{t-1}$ and $contrast_t$ I find a positive link that is significant at the 0.1% level. The effect of $contrast_{SA}$ goes in the same direction but is only significant at the 10% level and the magnitude is also smaller. This is to be expected given that this contrast includes surprises of small firms which likely contain less relevant information about the general economic state.

Overall, these results show that a positive contrast surprise makes analysts more optimistic about the firm that they forecast. A possible concern is that the other firm's surprise has no intrinsic information and just serves as a proxy for already available news about the general state of the economy that has not yet been included in the forecast. This is of particular concern because the forecast before the announcement can be from any time in the preceding quarter so there is significant potential that there is positive economic news that was not available at the time the forecast was made but did affect

the contrast companies' earnings. I show in Appendix 3.A that the result is robust to controlling for several indicators of the economic condition including the S&P 500 return between the date the forecasts before and after the announcement were made, which should capture any news about the economy at large. As discussed in 3.3.1 this result suggests that the contrast surprises contain information and the information channel dominates a potential contrast effect channel.

To examine if there is evidence of a contrast effect and test Hypothesis 1, I regress the forecast error of the forecasts made within 15 days after the announcement on the contrast surprises.

Table 3.4: Error on Contrast

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	0.0224 (SE=0.0147) (p=0.127)	0.0085 (SE=0.0049) (p=0.084)				
$Contrast_t$			0.1238 (SE=0.0300) (p<0.001)	0.0554 (SE=0.0077) (p<0.001)		
$Contrast_{SA}$					0.0765 (SE=0.0123) (p<0.001)	0.0313 (SE=0.0033) (p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0168	0.0280	0.0155	0.0266	0.0167	0.0275
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast error on contrast. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

In line with Hypothesis 1, Table 3.4 shows that all three types of contrast surprises are positively linked to forecast errors. The coefficients of $contrast_t$ and $contrast_{SA}$ are also highly significant both using OLS and using outlier robust regression, whereas the coefficient of $contrast_{t-1}$ is insignificant with OLS and only marginally significant with robust regression. Overall the results support Hypothesis 1 that more positive contrast surprises are linked to more pessimistic forecasts for a firm, in line with the return result of HS.

The effects of the contrast surprises are relatively small. Looking at revisions, a change of 1 standard deviation in $contrast_t$ changes forecasted earnings per $ of market capitalisation by $0.00027 based on the OLS regression coefficient. The 90th percentile of

NYSE market capitalisation which is used as a threshold for a large company was around \$24 billion on December 31st 2022. So, an increase by 1 standard deviation in $contrast_t$ increases analyst expected earnings of a large firm by around \$2.8 million. For forecast errors an increase of 1 standard deviation of $contrast_t$ leads to a more positive forecast error of \$0.00027 based on the OLS regression or around \$6.5 million.

I do not see the small effect size as particularly concerning. Analyst forecasts are available infrequently so I have to include forecasts made up to 15 days after the announcement. The contrast effect is likely weaker the more time passes between the announcement of the contrast surprise and the time a forecast was made. Hence I likely underestimate the contrast effect on the market's earnings expectations. Moreover, despite the small effect size, I can establish a contrast effect in earnings expectations and thus provide a mechanism for the contrast effect that HS find for returns, which does have a sizeable effect size.

In the main specifications, I pool forecast errors and forecast revisions across horizons. To see if there is a differential effect for different horizons I rerun the regressions using robust regression on the data separated by horizon. For revisions, the results look very similar for each horizon. However, as Figure 3.1 shows the impact of the contrast surprises tends to increase as the forecast horizon increases from 1 quarter to 4 quarters ahead. For longer horizons, the result is less clear. The effect continues to grow for $Contrast_{SA}$, plateaus for $Contrast_t$ and even decreases for $Contrast_{t-1}$. Tables with the exact coefficients for the regressions of both forecast revision and forecast error on the contrast surprises, separated by forecast horizon can be found in Appendix 3.A.

Having shown that forecast errors are positively linked to contrast surprises, I next study the link between contrast surprises and absolute forecast errors. The existing link between contrasts and forecast errors does not necessarily imply that the contrast effect makes forecasts less accurate. In particular, my descriptive statistics (Table 3.1) show that forecasters are on average overly optimistic (the mean forecast error is negative). Therefore, a bias induced by contrast effects could potentially even increase forecast accuracy if it counteracts an existing overoptimism bias. Table 3.5 shows the result of a regression of absolute forecast error on contrast surprises. For $contrast_{t-1}$ the

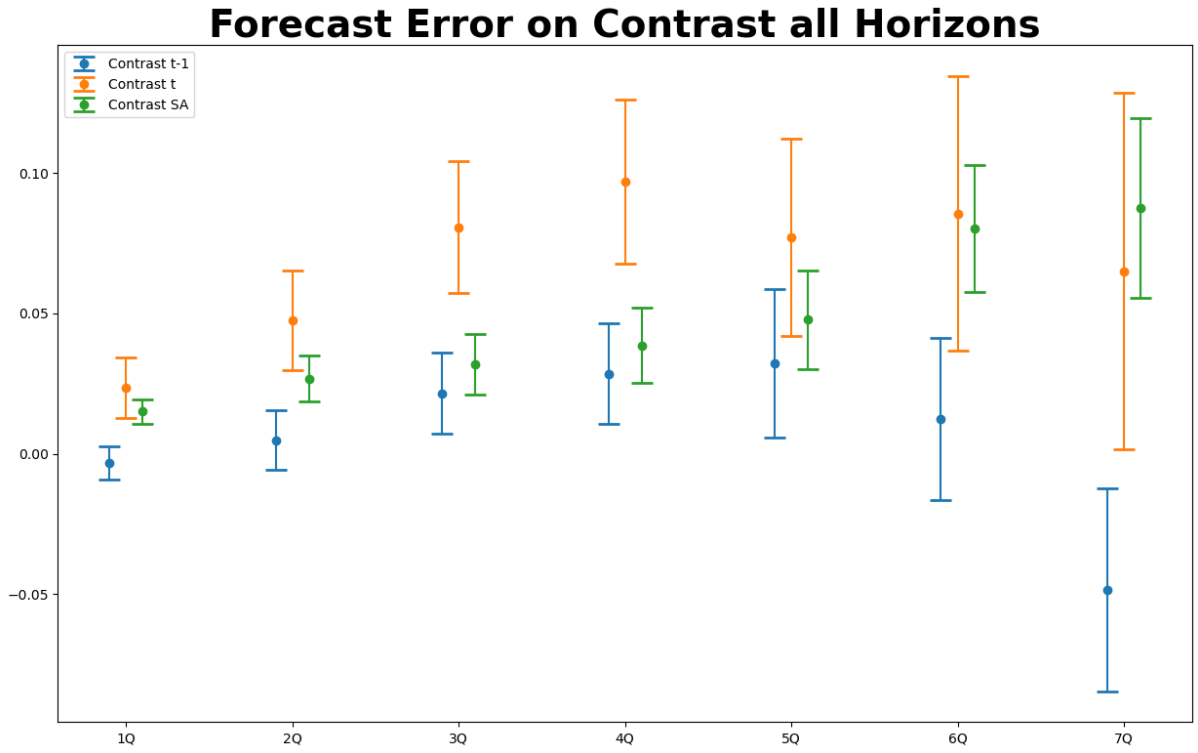


Figure 3.1: The figure shows the coefficients from a regression of forecast error on contrasts separated by forecast horizon. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $Contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $Contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $Contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. The regression is run using an outlier robust estimation method and with standard errors clustered at the firm and the analyst level.

coefficient is negative insignificant using OLS regression (1) and robust regression (2). For $contrast_t$ the coefficient is positive insignificant using OLS (3) and positive and marginally significant using robust regression (4). For $contrast_{SA}$ the coefficient is positive significant both using OLS (5) and robust regression (6).

Overall, the results suggest no clear effect of the two surprises based on large firms on forecast accuracy. However, the contrast based on firms covered by the same analyst is associated with less accurate forecasts. An explanation for this is that those forecasts contain less relevant information so that the psychological contrast channel dominates.

Table 3.5: Absolute Error on Contrast

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	-0.0025 (SE=0.0155) (p=0.874)	-0.0030 (SE=0.0039) (p=0.446)				
$Contrast_t$			0.0256 (SE=0.0285) (p=0.368)	0.0127 (SE=0.0069) (p=0.064)		
$Contrast_{SA}$					0.0530 (SE=0.0138) (p<0.001)	0.0203 (SE=0.0029) (p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.1950	0.2174	0.1923	0.2174	0.1878	0.2095
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of absolute forecast error on contrast. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.4.2 Persistence of the Effect

Next, I study how long the contrast effects persist. I first examine the link to subsequent upwards revisions (Hypothesis 2) In Table 3.6 I find that errors introduced by the contrast surprises get corrected over time. Both $contrast_t$ (2) $contrast_{SA}$ (3) are linked to upwards revisions for forecasts made between 16 and 45 days after the announcement, which contracts the excessive possimism initially induced by these contrasts. Somewhat surprisingly, the effect of $contrast_{t-1}$ (1) negative and very small. This can partially be explained by the fact that the link between $contrast_{t-1}$ and the forecast error of forecasts made immediately after the announcement was less clear in the first place.

Turning to the link between contrasts and errors of forecasts made later during the quarter (Hypothesis 3), Columns 4 to 9 show these forecasts remain positively linked to the contrast surprises. The link is highly significant except for the effect of $contrast_{t-1}$ on forecasts made between 16 and 45 days after the announcement (4). This suggests that, as hypothesised, the upwards revision in forecasts after the announcement is not sufficient to fully correct the contrast effect, which can be explained by sticky forecasts.

Table 3.6: Persistence of Contrast Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$revision_{[t+1,t+15],\,[t+16t+45]}$			$Error_{t+16\,t+45}$			$Error_{t+46\,t+75}$		
$Contrast_{t-1}$	-0.0007			0.0129			0.0192		
	(SE=0.0032)			(SE=0.0128)			(SE=0.0076)		
	(p=0.814)			(p=0.313)			(p=0.012)		
$Contrast_t$		0.0123			0.0493			0.0453	
		(SE=0.0060)			(SE=0.0190)			(SE=0.0128)	
		(p=0.041)			(p=0.010)			(p<0.001)	
$Contrast_{SA}$			0.0059			0.0403			0.0396
			(SE=0.0021)			(SE=0.0073)			(SE=0.0056)
			(p=0.004)			(p<0.001)			(p<0.001)
Robust Regression	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R^2	0.0061	0.0051	0.0050	0.0161	0.0144	0.0151	0.0194	0.0181	0.0187
N	220471	197672	125587	220449	197654	125574	445521	398839	254135

The table shows the result of a regression of forecast revisions from forecasts made between t+1 and t+15 to forecasts made between t+16 and t+45 (columns 1-3), forecast errors for forecasts made between t+16 and t+45 (columns 4-6) and forecast errors for forecasts made between t+46 and t+75 (columns 7-9) on contrast. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.4.3 The Link Between Errors and Revisions

Next, I examine the link between error and revision. Coibion and Gorodnichenko (2015) established that a regression of forecast error on forecast revision allows to detect under and overreaction. If forecast revisions are positively linked to errors in the updated forecast, this means that if the revision was positive the forecaster is too pessimistic and if it was negative she is too optimistic, which implies underreaction. Similarly, a negative link between errors and revisions implies overreaction.

I first examine if there is evidence for under or overreaction in forecast revision (without looking at the effect of contrast surprises). Table 3.7 shows a positive link between error and revision for both specifications which implies underreaction.

This result is in line with previous research that also finds underreaction of analysts for horizons of up to 3 years using consensus level and yearly data (Bordalo et al., 2019; Bouchaud et al., 2019). To establish if analysts under or overreact to contrasts I use the 2-stage regression procedure described in Section 3.3.2. Table 3.8 shows that in line with Hypothesis 4 the revisions explained by the three types of contrast surprise are positively linked to forecast errors. The link is highly significant for $contrast_t$ and $contrast_{SA}$. For $contrast_{t-1}$ it is only marginally significant and only with robust regression. Overall the results is in line with the idea that a positive contrast surprise contains positive news but at the same time makes the firm's own surprise look worse.

Table 3.7: Error on Revision

	(1)	(2)
Forecast Revision	0.1037	0.0460
	(SE=0.0153)	(SE=0.0044)
	(p<0.001)	(p<0.001)
Robust Regression	no	yes
Adjusted R^2	0.0014	0.0009
N	1904179	1904179

The table shows the result of a regression of forecast error on forecast revision. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. Standard errors are clustered at the firm and at the analyst level.

Table 3.8: Error on Revision Predicted by Contrast

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	1.0285	0.6870				
	(SE=0.6746)	(SE=0.3979)				
	(p=0.127)	(p=0.084)				
$Contrast_t$			2.3494	2.2893		
			(SE=0.5684)	(SE=0.3185)		
			(p<0.001)	(p<0.001)		
$Contrast_{SA}$					14.6672	19.4585
					(SE=2.3628)	(SE=2.0466)
					(p<0.001)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0168	0.0280	0.0155	0.0266	0.0167	0.0275
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast error on revision predicted by contrast. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.4.4 Heterogeneity in the Contrast Effect

In this Section, I examine heterogeneity in the effect of the other firms' earnings surprises on expectations.

Industry of the Contrast Firm

One possible source of heterogeneity in the effect of the contrast surprise is whether the contrast firm comes from the same or from a different industry as the announcing firm. A positive surprise of a firm from a different industry is likely linked to the

expectations about the firm because it contains new information about the economy as a whole. However, the surprise of a competitor from the same industry could either have a stronger positive link, if it is driven by the growth of the entire industry or a more negative link if it is driven by rising market share at the expense of the forecasted firm.

Empirically, I find in Table 3.9 that the surprises by firms from a different industry generally are more positively linked to forecast errors for $contrast_{t-1}$ and $contrast_t$. The fact that contrast surprises of firms from different industries lead to more pessimistic expectations can either be explained by a stronger contrast effect or by more underreaction to the information contained in the contrast surprise. A stronger contrast effect would be surprising, given that if anything analysts should spend more attention on firms from the same industry, which should lead to a stronger contrast effect. For $contrast_{SA}$ the effects of the same and different industry surprises are very similar. This supports the idea that the stronger effect of different industry contrasts for $contrast_{t-1}$ and $contrast_t$ is driven by the information channel rather than the contrast effect channel as it likely plays a smaller role for $contrast_{SA}$.

Table 3.9: Error on Contrast by Industry FF30

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1,SI}$	-0.0091	0.0010				
	(SE=0.0205)	(SE=0.0057)				
	(p=0.658)	(p=0.863)				
$Contrast_{t-1,DI}$	0.0382	0.0090				
	(SE=0.0267)	(SE=0.0080)				
	(p=0.152)	(p=0.261)				
$Contrast_{t,SI}$			0.0211	0.0209		
			(SE=0.0320)	(SE=0.0099)		
			(p=0.510)	(p=0.034)		
$Contrast_{t,DI}$			0.1333	0.0595		
			(SE=0.0475)	(SE=0.0136)		
			(p=0.005)	(p<0.001)		
$Contrast_{SA,SI}$					0.0963	0.0238
					(SE=0.0327)	(SE=0.0082)
					(p=0.003)	(p=0.004)
$Contrast_{SA,DI}$					0.0988	0.0328
					(SE=0.0322)	(SE=0.0087)
					(p=0.002)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0187	0.0271	0.0152	0.0250	0.0152	0.0286
N	831046	831046	629451	629451	100329	100329

The table shows the result of a regression of forecast error on contrast. The contrast surprises are separated by industry. If a firm is in the same industry as the announcing firm its surprise is part of the same industry contrast. Otherwise, it is p of the different industry contrast. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. Industries are defined using the French 30 industry classification. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Size of the Firm

Another possible source of heterogeneity in the effect of the other firms' surprises is the size of the firm for which the forecasts are made. There are various potential reasons for this. For instance, $contrast_{t-1}$ and $contrast_t$ only use large firms and there may be a difference in how a surprise of large firms affects other large firms compared to smaller firms. On the one hand, the announcement of a large company may be more informative of the earnings of another large company than of those of a small company. On the other hand, the larger similarity between the large contrast company and another large company may make the contrast between them more salient and lead to a larger contrast effect. However, analysts may spend more effort on the forecasts they make for large important firms which could make them less prone to the contrast effect bias. Finally, the degree to which analysts use news included in the announcement of other firms depends on how precise their prior beliefs are. This in turn depends on the amount

of available information. Small firms likely have less available information on average so any information contained in the announcement of other firms may have a larger effect. Overall, there are various possible reasons why firm size may moderate the effect of other firms' surprises which make different predictions on the direction of the effect.

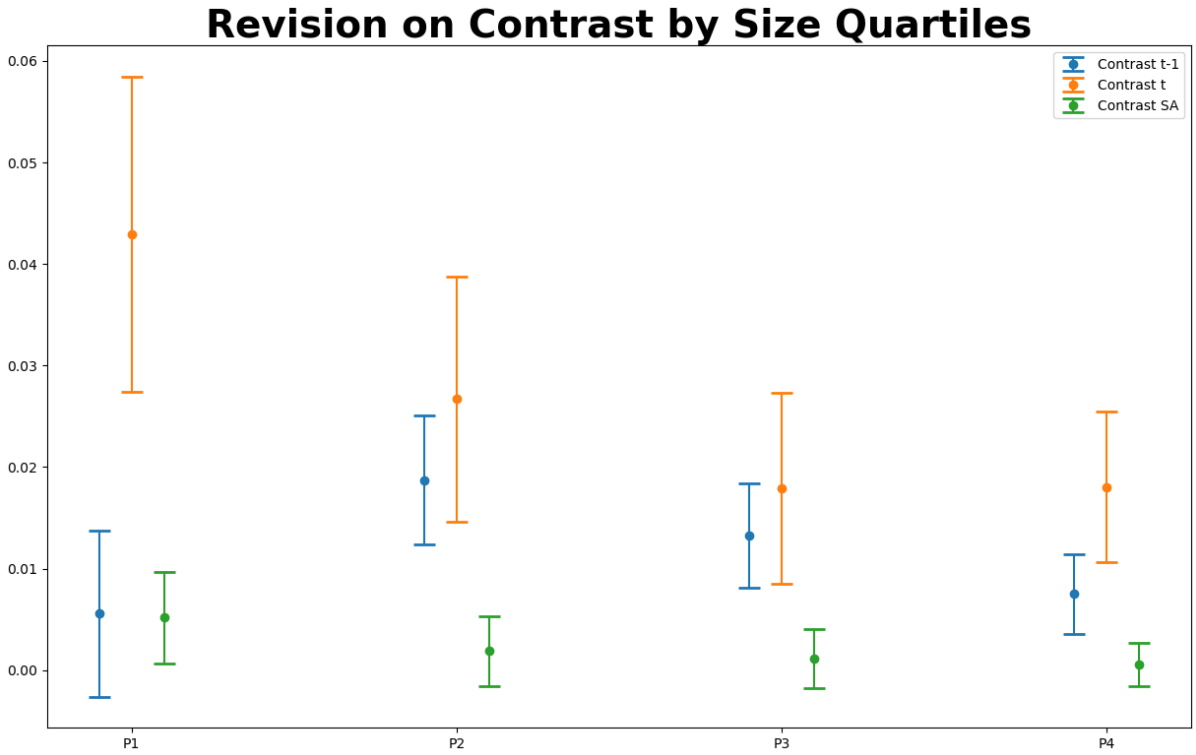


Figure 3.2: The figure shows the coefficients from a regression of forecast revision on contrasts with the data separated into quartiles based on the size of the firm. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement.

Figure 3.2 shows the results of a regression of forecast revision on contrasts with the firms split into quartiles based on size. The effect of $contrast_t$ seems to somewhat decrease in the firm's size but all confidence intervals overlap so no strong conclusion can be drawn. For the other two contrast surprises there is no clear effect.

Figure 3.3 shows the results of a regression of forecast error on contrasts with the firms split into quartiles based on size. There is no clear effect as all confidence intervals overlap and there is no clear trend in the coefficients either.

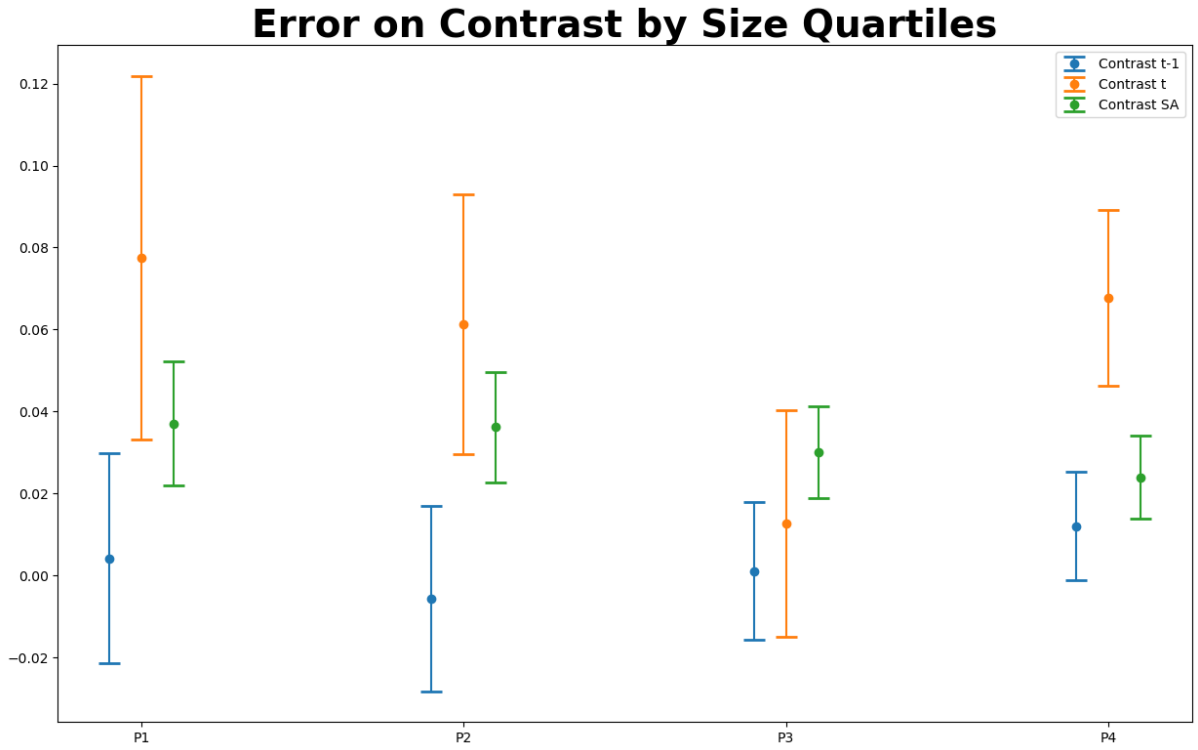


Figure 3.3: The figure shows the coefficients from a regression of forecast error on contrasts with the data separated into quartiles based on the size of the firm. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $Contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $Contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $Contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement.

3.5 Returns

HS show that $contrast_{t-1}$ negatively predicts returns after the announcement. They interpret their results as an error in perception rather than in expectation. They note: *"Under a perceptual error such as contrast effects, agents hold a biased quality assessment about the next case only after seeing the next case. Under an expectational error, seeing one case causes agents to hold mistaken beliefs about the quality of future cases, before the future cases are directly observed"*. However, while the timing of their abnormal returns rules out an impact of changed expectations for the upcoming announcement, it does not rule out a change in expectations about future announcements after the firm's own announcement is observed. Given the importance of cash flow expectations for returns, it would be surprising to observe the result of HS in the absence of a contrast effect

on expectations which I document in Section 3.4. In this Section, I link the results on contrast effects in expectations to the return result of HS.

As a first step, I aim to reproduce their result. Table 3.10 reports the results of a regression of the return around an earnings announcement on contrast surprises. Columns 2, 4, 6 and 8 use a dataset that follows the specification of HS. The remaining columns use my main dataset, which is restricted to data points with an available forecast revision. For comparability with HS all specifications use standard OLS. However, I cluster forecasts on the firm level to be consistent with my analysis using forecasts, while they cluster on the announcement level. I first examine the return around the earnings announcement (from t-1 to t+1). For $contrast_{t-1}$ the most direct comparison is my column 2 and column 6 of Table II of HS. I find a coefficient that goes in the same direction (negative) but is more than an order of magnitude smaller and only marginally significant. For $contrast_t$ all coefficients are far from significant. For this contrast, HS do not have a directly comparable result, but they perform a similar exercise in their Table X and find suggestive but weak evidence of a same day contrast effect.

Next, I analyse if there is reversal in the abnormal returns. HS find a significant reversal of returns between 26 and 50 days after the announcement (Their Table IV). I do not find this positive effect.

There are two possible explanations for the differences between my result and that of HS. First, I use an additional 9 years of data, 7 of which are from after HS published their working paper [3] and two of which are post publication. It may be that the effect has disappeared or diminished in that time, potentially due to traders becoming aware of it. The other potential explanation is that both HS and I use forecasts from IBES to calculate earnings surprises. The IBES database is not persistent and historical data can change when brokers add their forecasts to or pull them from the database so that two datasets of IBES forecasts covering the same time period can be different if they are downloaded at different times.[4] In Appendix 3.B I explore the difference between

[3] using the date it was posted on SSRN

[4] HS also use a slightly different characteristics adjusted return. They exclude firms that are part of the contrast surprises from the characteristics portfolios, however, they find that this has no relevant effect so it is unlikely to explain the difference between our results.

my results and those of HS in more detail.

Table 3.10: Return on Contrast

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$return_{t-1,t+1}$				$return_{t+26,t+50}$			
$Contrast_{t-1}$	-0.0987 (SE=0.0543) (p=0.069)	-0.0719 (SE=0.0372) (p=0.053)			-0.0544 (SE=0.0877) (p=0.535)	-0.0482 (SE=0.0644) (p=0.455)		
$Contrast_t$			-0.0928 (SE=0.1336) (p=0.487)	-0.1147 (SE=0.1263) (p=0.364)			0.2347 (SE=0.2215) (p=0.289)	0.4263 (SE=0.2290) (p=0.063)
Robust Regression	no	no	no	no	no	no	no	no
Adjusted R^2	0.0732	0.0632	0.0771	0.0734	0.0009	0.0011	0.0008	0.0010
N	90536	114608	71924	81301	90396	112899	71795	79836

The table shows the result of a regression of returns contrast. Columns 2,4,6 and 8 use a dataset that does not require an available forecast around an announcement for a data point to be included. All returns are characteristics adjusted by subtracting the return of a portfolio matched on size, book-to-market value and momentum from the raw return. The reference date t for the return windows is the day the firm announced its earnings. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm level.

Moving beyond the re-examination of the contrast effect on returns found by HS, I examine if this effect is driven by an expectations channel (Question 2). To do that I use the two-stage regression procedure explained in Section 3.3.2. Table 3.11 reports the results of a regression of the return around the announcement (t-1 to t+1) on the revisions that are explained by contrast surprises. This result is correlational as the return window and the time during which the expectations are recorded overlap. All coefficients are negative and all except the one for $contrast_t$ using OLS (3) are significant. This suggests a discrepancy between the reaction of analysts and that of the market to the contrast surprise. A possible explanation for this result is that following HS, I adjust returns for momentum, size and book-to-market ratio, by subtracting the returns of a portfolio matched on these characteristics from the firm's own return. As a side effect, this filters out all news that affects firms equally. The results discussed in Section 3.4, showed that analysts revise expectations upwards after a positive contrast surprise (likely because it implies good news for the firm) but are nevertheless more pessimistic relative to the true value. However, for returns, the news channel is filtered out and the reaction to a good surprise is negative as shown in Table 3.10. Thus the link between the component of forecast revisions that is explained by contrast surprises and returns is negative.

Moving to returns later during the quarter (Question 3), Table 3.12 shows the results a regression of characteristics adjusted returns between t+26 and t+50 on revisions

Table 3.11: Return [t-1, t+1] on Revision Predicted by Surprises

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	-11.5013	-19.1224				
	(SE=4.5674)	(SE=5.2887)				
	(p=0.012)	(p<0.001)				
$Contrast_t$			-1.0504	-9.7538		
			(SE=3.1194)	(SE=4.7803)		
			(p=0.736)	(p=0.041)		
$Contrast_{SA}$					-16.6182	-45.6477
					(SE=7.5198)	(SE=11.0762)
					(p=0.027)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0820	0.1027	0.0822	0.1033	0.0818	0.1024
N	1440937	1440937	1303981	1303981	708180	708180

The table shows the result of a regression of Return [t-1, t+1] on revision predicted by contrast. All returns are characteristics adjusted by subtracting the return of a portfolio matched on size, book-to-market value and momentum from the raw return. The reference date t for the return windows is the day the firm announced its earnings. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

predicted by contrast. Since the last date for which forecasts are included is t+15 I can interpret this regression as revisions predicting returns. I focus on the time period between t+25 and t+50 since HS find evidence for significant return reversals here. In Appendix 3.D I also look at returns starting from t+16 which is the earliest date that allows us to interpret the regression results as predictive. All coefficients are positive, but only those for $contrast_t$ are significant and only at the 10% level. Thus, no conclusions can be drawn. However, if taken at face value, these results would be consistent with the idea that expectations initially underreact to the news contained in the contrast surprise and the later correction of this underreaction is associated with higher returns.

Table 3.12: Return [t+26, t+50] on Revision Predicted by Surprises

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	1.9971 (SD=6.7449) (p=0.767)	8.2349 (SD=7.2673) (p=0.257)				
$Contrast_t$			7.7172 (SD=4.6382) (p=0.096)	12.8181 (SD=6.7811) (p=0.059)		
$Contrast_{SA}$					12.9258 (SD=12.0871) (p=0.285)	16.9548 (SD=17.3429) (p=0.328)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0018	0.0015	0.0019	0.0017	0.0023	0.0019
N	1438906	1438906	1301974	1301974	707070	707070

The table shows the result of a regression of Return [t+26, t+50] on revision predicted by contrast. All returns are characteristics adjusted by subtracting the return of a portfolio matched on size, book-to-market value and momentum from the raw return. The reference date t for the return windows is the day the firm announced its earnings. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.6 Conclusion

This paper shows that changes in expected future earnings after earnings announcements depend not only on a firm's own news but also on the earnings surprise of contrasting firms. The evidence is consistent with an effect via two channels a news channel and a contrast effect channel. This complements Hartzmark and Shue (2018) in showing that it is crucial to study firms' earnings announcements not in isolation but in the context of other firms' announcements. My results suggest that analysts revise their expectations upwards if contrasting firms had positive earnings surprises suggesting that good news for one firm are also good news for other firms.

Moreover, I show that contrast surprises are positively linked to forecast errors, suggesting that earnings expectations become more negative relative to the actual future earnings, consistent with a contrast effect in expectation formation. This effect has not been documented previously and has potentially wide-ranging implications for three reasons. First, any decision that is influenced by unknown future states of the world requires economic agents to form expectations. Second, contrast effects have been found across a wide range of domains and the evidence suggests that they are a fundamental

component in the reaction to information down to the level of the firing rate of neurons. Finally, contrasts are arguably available for almost every scenario in which expectations are formed and updated. Behavioural explanations of economic phenomena have often (rightfully) been criticised as being ad-hoc. Given the large number of established cognitive biases, there is a risk that one of them predicts behaviour that fits data ex-post by coincidence. However, this risk is greatly reduced if the same bias can explain data across a wide range of contexts. A recent example of this is the diagnostic expectations model by Bordalo et al. (2018) which has been shown to explain phenomena in a wide range of domains.[5] Contrast effects have the potential to be another portable mechanism that are at work across different setups. Moreover, because they are unrelated to any distortion in processing news directly related to the forecasted variable, they complement existing models. Therefore, this paper is just a first step in a wider research agenda and future work should examine context effects in other setups with available expectations data.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017). "Stock Price Booms and Expected Capital Gains". In: *American Economic Review* 107.8, pp. 2352–2408. ISSN: 0002-8282. DOI: 10.1257/aer.20140205.

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma, David Thesmar, and MIT Sloan (2021). *Overreaction in Expectations: Evidence and Theory*. preprint.

Amromin, Gene and Steven A. Sharpe (2013). "From the Horse's Mouth: Economic Conditions and Investor Expectations of Risk and Return". In: *Management Science* 60.4, pp. 845–866. ISSN: 0025-1909. DOI: 10.1287/mnsc.2013.1806.

Antoniou, Constantinos, Junyang Guo, and Neil Stewart (2021). "It Depends Who You Ask: Context Effects in the Perception of Stock Returns". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3777778.

D'Arienzo, Daniele (2020). "Maturity Increasing Overreaction and Bond Market Puzzles". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3733056.

Assenza, Tiziana, Te Bao, Cars Hommes, and Domenico Massaro (2014). "Experiments on Expectations in Macroeconomics and Finance". In: *Research in Experimental Economics* 17, pp. 11–70. DOI: 10.1108/S0193-230620140000017002.

[5] Examples are earnings growth, macroeconomic expectations and interest rate volatility (as discussed in Section 3.1) as well as credit cycles, (Bordalo et al., 2018) and asset price bubbles (Bordalo et al., 2021).

Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop (2009). "Predictability in Financial Markets: What Do Survey Expectations Tell Us?" In: *Journal of International Money and Finance* 28.3, pp. 406–426. ISSN: 0261-5606. DOI: 10.1016/j.jimonfin.2008.09.001.

Ball, Ray and Philip Brown (1968). "An Empirical Evaluation of Accounting Income Numbers". In: *Journal of Accounting Research* 6.2, p. 159. ISSN: 00218456. DOI: 10.2307/2490232.

Bernard, Victor L. and Jacob K. Thomas (1989). "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" In: *Journal of Accounting Research* 27, p. 1. ISSN: 00218456. DOI: 10.2307/2491062.

Bhargava, Saurabh and Ray Fisman (2014). "Contrast Effects in Sequential Decisions: Evidence from Speed Dating". In: *The Review of Economics and Statistics* 96.3, pp. 444–457. ISSN: 0034-6535. DOI: 10.1162/REST_a_00416.

Van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira (2022). *Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases*. DOI: 10.2139/ssrn.3625279. preprint.

Bordalo, Pedro, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer (2021). "Diagnostic Bubbles". In: *Journal of Financial Economics* 141.3, pp. 1060–1077. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2020.06.019.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic Expectations and Stock Returns". In: *The Journal of Finance* 74.6, pp. 2839–2874. ISSN: 1540-6261. DOI: 10.1111/jofi.12833.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020). "Overreaction in Macroeconomic Expectations". In: *American Economic Review* 110.9, pp. 2748–2782. ISSN: 0002-8282. DOI: 10.1257/aer.20181219.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). "Salience Theory of Choice Under Risk". In: *The Quarterly Journal of Economics* 127.3, pp. 1243–1285. ISSN: 0033-5533. DOI: 10.1093/qje/qjs018.

— (2018). "Diagnostic Expectations and Credit Cycles". In: *The Journal of Finance* 73.1, pp. 199–227. ISSN: 1540-6261. DOI: 10.1111/jofi.12586.

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar (2019). "Sticky Expectations and the Profitability Anomaly". In: *The Journal of Finance* 74.2, pp. 639–674. ISSN: 1540-6261. DOI: 10.1111/jofi.12734.

Cash, Thomas F., Diane Walker Cash, and Jonathan W. Butters (1983). ""Mirror, Mirror, on the Wall...?": Contrast Effects and Self-Evaluations of Physical Attractiveness". In: *Personality and Social Psychology Bulletin* 9.3, pp. 351–358. ISSN: 0146-1672. DOI: 10.1177/0146167283093004.

Chari, V. V, Ravi Jagannathan, and Aharon R Ofer (1988). "Seasonalities in Security Returns: The Case of Earnings Announcements". In: *Journal of Financial Economics* 21.1, pp. 101–121. ISSN: 0304-405X. DOI: 10.1016/0304-405X(88)90033-5.

Cohen, Daniel A., Aiyesha Dey, Thomas Z. Lys, and Shyam V. Sunder (2007). "Earnings Announcement Premia and the Limits to Arbitrage". In: *Journal of Accounting and Economics* 43.2, pp. 153–180. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2007.01.008.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What Can Survey Forecasts Tell Us about Information Rigidities?" In: *Journal of Political Economy* 120.1, pp. 116–159. ISSN: 0022-3808. DOI: 10.1086/665662.

— (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". In: *American Economic Review* 105.8, pp. 2644–2678. ISSN: 0002-8282. DOI: 10.1257/aer.20110306.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997). "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks". In: *The Journal of Finance* 52.3, pp. 1035–1058. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1997.tb02724.x.

De La O, Ricardo and Sean Myers (2021). "Subjective Cash Flow and Discount Rate Expectations". In: *The Journal of Finance* 76.3, pp. 1339–1387. ISSN: 1540-6261. DOI: 10.1111/jofi.13016.

deHaan, Ed, Terry Shevlin, and Jacob Thornock (2015). "Market (in)Attention and the Strategic Scheduling and Timing of Earnings Announcements". In: *Journal of Accounting and Economics* 60.1, pp. 36–55. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2015.03.003.

Dellavigna, Stefano and Joshua M. Pollet (2009). "Investor Inattention and Friday Earnings Announcements". In: *The Journal of Finance* 64.2, pp. 709–749. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2009.01447.x.

Engelberg, Joseph, R. David Mclean, and Jeffrey Pontiff (2018). "Anomalies and News". In: *The Journal of Finance* 73.5, pp. 1971–2001. ISSN: 1540-6261. DOI: 10.1111/jofi.12718.

Fama, Eugene F. and Kenneth R. French (1998). "Value versus Growth: The International Evidence". In: *The Journal of Finance* 53.6, pp. 1975–1999. ISSN: 1540-6261. DOI: 10.1111/0022-1082.00080.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer (2016). "Expectations and Investment". In: *NBER Macroeconomics Annual* 30, pp. 379–431. ISSN: 0889-3365. DOI: 10.1086/685965.

Glaser, Markus, Zwetelina Iliewa, and Martin Weber (2019). "Thinking about Prices versus Thinking about Returns in Financial Markets". In: *The Journal of Finance* 74.6, pp. 2997–3039. ISSN: 1540-6261. DOI: 10.1111/jofi.12835.

Greenwood, Robin Marc and Andrei Shleifer (2014). "Expectations of Returns and Expected Returns". In: *Review of Financial Studies* 27.3, pp. 714–746. ISSN: 0893-9454. DOI: 10.1093/rfs/hht082.

Hartzmark, Samuel M. and Kelly Shue (2018). "A Tough Act to Follow: Contrast Effects in Financial Markets". In: *The Journal of Finance* 73.4, pp. 1567–1613. ISSN: 1540-6261. DOI: 10.1111/jofi.12685.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh (2009). "Driven to Distraction: Extraneous Events and Underreaction to Earnings News". In: *The Journal of Finance* 64.5, pp. 2289–2325. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2009.01501.x.

Israeli, Doron, Ron Kasznik, and Suhas A. Sridharan (2021). "Unexpected Distractions and Investor Attention to Corporate Announcements". In: *Review of Accounting Studies*. ISSN: 1573-7136. DOI: 10.1007/s11142-021-09618-4.

Kenrick, Douglas T. and Sara E. Gutierres (1980). "Contrast Effects and Judgments of Physical Attractiveness: When Beauty Becomes a Social Problem". In: *Journal of Personality and Social Psychology* 38.1, pp. 131–140. ISSN: 1939-1315(Electronic),0022-3514(Print). DOI: 10.1037/0022-3514.38.1.131.

Kim, Jae Hyoung and Elizabeth Hoffman (2020). "Contrast Effects in Investment and Financing Decisions". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3256087.

Kwon, Spencer Yongwook and Johnny Tang (2020). *Extreme Events and Overreaction to News*. SSRN Scholarly Paper ID 3724420. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3724420.

La Porta, Rafael (1996). "Expectations and the Cross-Section of Stock Returns". In: *The Journal of Finance* 51.5, pp. 1715–1742. ISSN: 0022-1082. DOI: 10.2307/2329535.

Lucas, Robert E and Edward C. Prescott (1971). "Investment Under Uncertainty". In: *Econometrica* 39.5, pp. 659–681. ISSN: 0012-9682. DOI: 10.2307/1909571.

Ma, Yueran, Teodora Paligorova, and Jose-Luis Peydro (n.d.). "Expectations and Bank Lending". In: (), p. 42.

Michaely, Roni, Amir Rubin, and Alexander Vedrashko (2014). "Corporate Governance and the Timing of Earnings Announcements*". In: *Review of Finance* 18.6, pp. 2003–2044. ISSN: 1572-3097. DOI: 10.1093/rof/rft054.

Pepitone, Albert and Mark DiNubile (1976). "Contrast Effects in Judgments of Crime Severity and the Punishment of Criminal Violators". In: *Journal of Personality* 33.4, pp. 448–459. ISSN: 0022-3514.

Radbruch, Jonas and Amelie Schiprowski (2020). *Interview Sequences and the Formation of Subjective Assessments*. 045. University of Bonn and University of Cologne, Germany.

Savor, Pavel and Mungo Wilson (2016). "Earnings Announcements and Systematic Risk". In: *The Journal of Finance* 71.1, pp. 83–138. ISSN: 1540-6261. DOI: 10.1111/jofi.12361.

Schkade, David A. and Daniel Kahneman (1998). "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction". In: *Psychological Science* 9.5, pp. 340–346. ISSN: 0956-7976. DOI: 10.1111/1467-9280.00066.

De Silva, Tim and David Thesmar (2021). *Noise in Expectations: Evidence from Analyst Forecasts*. Working Paper 28963. National Bureau of Economic Research. DOI: 10.3386/w28963.

Waite, Thomas A. (2001). "Background Context and Decision Making in Hoarding Gray Jays". In: *Behavioral Ecology* 12.3, pp. 318–324. ISSN: 1045-2249. DOI: 10.1093/beheco/12.3.318.

Wallis, Jonathan D. and Steven W. Kennerley (2011). "Contrasting Reward Signals in the Orbitofrontal Cortex and Anterior Cingulate Cortex". In: *Annals of the New York Academy of Sciences* 1239.1, pp. 33–42. ISSN: 1749-6632. DOI: 10.1111/j.1749-6632.2011.06277.x.

3.A Robustness of the Effect of Contrast on Forecast Revisions and Errors

A potential concern in studying the link between forecast revisions and other firms' surprises is that forecasts are made at infrequent intervals. The forecast before the announcement can come from any time in the preceding quarter, making it potentially up to three months old. Because of this, a positive contrast surprise may not be news in itself, but just capturing publicly available good news that positively affects the forecasted firm as well as the contrast firm and that was not yet included in the forecast. To deal with this issue I rerun the regression while controlling for the S&P500 return between the forecast before and after the announcement, which should capture news that affect the whole stock market. Moreover, I directly control for several macroeconomic variables namely GDP growth, investment growth and unemployment changes over the preceding quarter. Table 3.13 shows that the results hold in the regression with controls. Compared to the results without controls in Table 3.3, the coefficients get slightly smaller in all specifications and are no longer significant for $contrast_{SA}$, but overall the results are very similar.

Another potential concern is that I pool data across forecast horizons, which may mask heterogeneity. In Figure 3.1 in Section 3.4 I have shown graphically that the contrast effect on forecast error does seem to increase with the forecast horizon. Tables 3.14 the results of a regression of forecast revision on the contrast surprises separated by horizon using robust regression. The coefficients for $contrast_{t-1}$ and $contrast_t$ are

Table 3.13: Revision on Contrast and Economic Controls

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	0.0117	0.0079				
	(SD=0.0051)	(SD=0.0015)				
	(p=0.022)	(p<0.001)				
$Contrast_t$			0.0334	0.0166		
			(SD=0.0099)	(SD=0.0027)		
			(p<0.001)	(p<0.001)		
$Contrast_{SA}$					0.0041	0.0011
					(SD=0.0030)	(SD=0.0008)
					(p=0.174)	(p=0.184)
$S\&P500_{between\ forecasts}$	0.0071	0.0032	0.0073	0.0034	0.0074	0.0032
	(SD=0.0004)	(SD=0.0001)	(SD=0.0004)	(SD=0.0001)	(SD=0.0005)	(SD=0.0001)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
$\Delta\ Gdp_{q-1}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)
	(p=0.177)	(p=0.002)	(p=0.105)	(p=0.003)	(p=0.525)	(p=0.002)
$\Delta\ Investment_{q-1}$	0.0001	0.0000	0.0001	0.0000	0.0001	0.0000
	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
$\Delta\ unemployment_{m-1}$	-0.0001	-0.0000	-0.0001	-0.0000	-0.0001	-0.0000
	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)
	(p=0.007)	(p<0.001)	(p=0.017)	(p<0.001)	(p=0.029)	(p<0.001)
$\Delta\ unemployment_{m-2}$	-0.0005	-0.0002	-0.0006	-0.0002	-0.0005	-0.0002
	(SD=0.0001)	(SD=0.0000)	(SD=0.0001)	(SD=0.0000)	(SD=0.0001)	(SD=0.0000)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
$\Delta\ unemployment_{m-3}$	0.0001	0.0000	0.0001	0.0000	0.0001	0.0000
	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)	(SD=0.0000)
	(p=0.002)	(p<0.001)	(p=0.002)	(p<0.001)	(p=0.004)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0524	0.0700	0.0496	0.0675	0.0505	0.0649
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast revision on contrast. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins, the S&P500 return between the forecast before the announcement and the forecast after the announcement, the GDP growth over the quarter preceding the announcement, as well as monthly changes in unemployment for the preceding 3 months. Standard errors are clustered at the firm and at the analyst level.

positive significant across all horizons. For $contrast_{SA}$ the picture is less clear. While all horizons except 1 quarter ahead have a positive coefficient not all of them are significant. Table 3.15 shows the results of a regression of forecast error on the contrast surprises separated by horizon using robust regression. $contrast_t$ and $contrast_{SA}$ have positive and significant coefficients across horizons, while the result for $contrast_{t-1}$ is less clear.

Overall. the results suggest that the effect of the contrast surprises on expectations is relatively similar across horizons.

Finally, I use an alternative approach to the inclusion of controls to explore the issue that the link between forecast revision and the contrast surprise may be driven

Table 3.14: Revision on Contrast by Horizon

	1Q	2Q	3Q	4Q	5Q	6Q	7Q
$Contrast_{t-1}$	0.0130	0.0133	0.0105	0.0134	0.0160	0.0119	0.0069
	(SE=0.0019)	(SE=0.0016)	(SE=0.0016)	(SE=0.0019)	(SE=0.0021)	(SE=0.0022)	(SE=0.0023)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p=0.003)
	N=468540	N=389448	N=320728	N=222635	N=170344	N=115487	N=72544
$Contrast_t$	0.0300	0.0227	0.0189	0.0232	0.0295	0.0241	0.0166
	(SE=0.0037)	(SE=0.0029)	(SE=0.0028)	(SE=0.0034)	(SE=0.0037)	(SE=0.0038)	(SE=0.0043)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
	N=414352	N=348217	N=288402	N=202836	N=156763	N=106571	N=66890
$Contrast_{SA}$	-0.0004	0.0022	0.0002	0.0036	0.0044	0.0027	0.0050
	(SE=0.0015)	(SE=0.0012)	(SE=0.0012)	(SE=0.0015)	(SE=0.0018)	(SE=0.0018)	(SE=0.0020)
	(p=0.765)	(p=0.071)	(p=0.853)	(p=0.016)	(p=0.013)	(p=0.127)	(p=0.015)
	N=228608	N=190501	N=157936	N=110267	N=84180	N=57970	N=37300

This table shows the result of a regression of forecast revision on contrast surprise separated by forecast horizon, from 1 quarter ahead to 7 quarters ahead (for the forecasts after the announcement). Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level. Robust regression is used for all horizons.

Table 3.15: Error on Contrast by Horizon

	1Q	2Q	3Q	4Q	5Q	6Q	7Q
$Contrast_{t-1}$	-0.0033	0.0049	0.0216	0.0285	0.0322	0.0123	-0.0484
	(SE=0.0030)	(SE=0.0055)	(SE=0.0074)	(SE=0.0092)	(SE=0.0134)	(SE=0.0148)	(SE=0.0185)
	(p=0.269)	(p=0.375)	(p=0.004)	(p=0.002)	(p=0.016)	(p=0.406)	(p=0.009)
	N=468540	N=389448	N=320728	N=222635	N=170344	N=115487	N=72544
$Contrast_t$	0.0234	0.0474	0.0807	0.0968	0.0772	0.0856	0.0650
	(SE=0.0055)	(SE=0.0091)	(SE=0.0120)	(SE=0.0149)	(SE=0.0179)	(SE=0.0250)	(SE=0.0324)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p=0.045)
	N=414352	N=348217	N=288402	N=202836	N=156763	N=106571	N=66890
$Contrast_{SA}$	0.0151	0.0267	0.0318	0.0386	0.0478	0.0802	0.0875
	(SE=0.0022)	(SE=0.0041)	(SE=0.0055)	(SE=0.0069)	(SE=0.0090)	(SE=0.0115)	(SE=0.0163)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
	N=228608	N=190501	N=157936	N=110267	N=84180	N=57970	N=37300

This table shows the result of a regression of forecast error on contrast surprise separated by forecast horizon, from 1 quarter ahead to 7 quarters ahead (for the forecasts after the announcement). Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level. Robust regression is used for all horizons

by economic news that because available after the forecast before the announcement was made and that influences both the forecasted firm and the contrast firm. I limit my sample to revisions with a forecast made within 15 days before the announcement is available, which greatly reduces the potential for such news to appear. Table 3.16 shows that the link between forecast revisions and contrast surprises holds in the restricted sample and Table 3.17 shows that this is also the case for forecast errors.

Table 3.16: Revision on Contrast - Restrictive Dataset

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	0.0136	0.0112				
	(SD=0.0058)	(SD=0.0020)				
	(p=0.019)	(p<0.001)				
$Contrast_t$			0.0298	0.0174		
			(SD=0.0146)	(SD=0.0035)		
			(p=0.042)	(p<0.001)		
$Contrast_{SA}$					0.0163	0.0086
					(SD=0.0062)	(SD=0.0020)
					(p=0.009)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0522	0.0793	0.0478	0.0751	0.0455	0.0702
N	483021	483021	439126	439126	291645	291645

The table shows the result of a regression of forecast revision on contrast. For a datapoint to be included it must have an available forecast revision, where the forecast before the announcement is made within 15 days of the announcement instead of at any time during the preceding quarter as in my main specifications. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Table 3.17: Error on Contrast - Restrictive Dataset

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	0.0409	0.0166				
	(SD=0.0238)	(SD=0.0077)				
	(p=0.086)	(p=0.031)				
$Contrast_t$			0.1154	0.0495		
			(SD=0.0436)	(SD=0.0115)		
			(p=0.008)	(p<0.001)		
$Contrast_{SA}$					0.1031	0.0478
					(SD=0.0267)	(SD=0.0072)
					(p<0.001)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0159	0.0293	0.0144	0.0272	0.0178	0.0313
N	483021	483021	439126	439126	291645	291645

The table shows the result of a regression of forecast error on contrast. For a datapoint to be included it must have an available forecast revision, where the forecast before the announcement is made within 15 days of the announcement instead of at any time during the preceding quarter as in my main specifications. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.B More Detailed Comparison with Hartzmark and Shue

In Table 3.10 I reported that the contrast effect I find has less than a tenth of the magnitude of the one reported by HS. I proposed two possible explanations: differences in my dataset caused by retractions from or additions to the IBES database and the additional data beyond 2013 that is available now. To distinguish between these explanations I split my dataset into data up to 2013 and data from after 2013. Table 3.18 reports the results

of a regression of characteristics adjusted return on $contrast_{t-1}$ using the dataset that does not require an available forecast revision around the announcement. In the absence of changes to the underlying data and inconsistencies that can arise in matching different databases this dataset should be completely equivalent to the HS dataset if I restrict it to data up to 2013. Column (1) which should be equivalent to column 6 of HS's Table 2 has 81694 data points compared to HS who have 75897. Overall, I find that the coefficient for the regression of returns around the announcement on the contrast surprise is very similar for the data before and after 2013. The coefficients differ for returns between 2 and 25 as well as 26 and 50 days after the announcement but all of these coefficients are far from being statistically significant so no conclusions can be drawn.

Table 3.18: HS Results Before After 2013

	(1)	(2)	(3)	(4)	(5)	(6)
	Up to 2013			After 2013		
	[t-1, t+1]	[t+2, t+25]	[t+26, t+50]	[t-1, t+1]	[t+2, t+25]	[t+26, t+50]
$Contrast_{t-1}$	-0.0749	-0.0773	-0.0621	-0.0876	0.0926	-0.0034
	(SE=0.0397)	(SE=0.0677)	(SE=0.0733)	(SE=0.0986)	(SE=0.1412)	(SE=0.1278)
	(p=0.059)	(p=0.253)	(p=0.397)	(p=0.375)	(p=0.512)	(p=0.979)
Robust Regression	no	no	no	no	no	no
Adjusted R^2	0.0594	0.0016	0.0015	0.0745	0.0006	0.0001
N	81694	81519	81192	32914	32827	31707

"The table shows the result of a regression of return [t-1, t+1] contrast. The dataset used in this table is designed to mimic that used by Hartzmark and Shue (2018). The main difference to the default dataset is that datapoints do not need to have an available forecast revision to be included. Columns 1-3 only use data up to 2013 to match the sample period in Hartzmark and Shue (2018). Columns 4-6 use data up to 2022. All returns are characteristics adjusted by subtracting the return of a portfolio matched on size, book-to-market value and momentum from the raw return. The reference date t for the return windows is the day the firm announced its earnings. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm.

3.C Effect of the Own Surprise

Most specifications in the analysis control for the firm's own surprise using 20 bins. This follows HS and is done to allow for a non-linear effect of the own surprise. The downside of this is that it makes it impossible to report coefficients of the own surprise. In this Section, I repeat some of the analysis using the own surprise directly. Table 3.19 shows the result of a regression of forecast error and forecast revision on own surprise. Forecasts revisions are positively linked to earnings surprises. This is unsurprising: if the current earnings were better than expected, analysts should revise their expectations for future earnings

upwards. If analysts processed earnings information in an unbiased manner earnings surprise should not be linked to forecast errors, instead, I find a strong positive link which indicates that a more positive earnings surprise causes forecasts to be more pessimistic relative to the true value. This indicates that analysts underreact to news at the short horizon which is in line with existing findings (Bordalo et al., 2019; Bouchaud et al., 2019).

Table 3.19: Error on Own Surprise

	(1)	(2)	(3)	(4)
	Forecast Revision		*Forecast Error*	
Own Surprise	0.0949	0.1086	0.2112	0.2057
	(SE=0.0060)	(SE=0.0022)	(SE=0.0150)	(SE=0.0045)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
Robust Regression	no	yes	no	yes
Adjusted R^2	0.0178	0.0557	0.0119	0.0282
N	1904179	1904179	1904179	1904179

The table shows the result of a regression of forecast error on own surprise. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Table 3.20 shows the result of a regression of forecast revision on the firm's own surprise and contrast surprises. The results for contrast surprises are qualitatively the same and also of a similar magnitude as in Table 3.3 which controls for own surprise using bins. The coefficient on the own surprise is positive and significant in all specifications. For contrast errors (Table 3.21) we see a similar picture. All coefficients are similar in magnitude as in Table 3.4. The coefficient on the own surprise is also positive significant in all specifications, implying that the underreaction to the news in the own surprise discussed above does not disappear when the contrast surprises are included.

Table 3.20: Revision on Own and Contrast Surprise

	(1)	(2)	(3)	(4)	(5)	(6)
Own Surprise	0.0951	0.1079	0.0905	0.1034	0.0917	0.0996
	(SE=0.0064)	(SE=0.0024)	(SE=0.0071)	(SE=0.0025)	(SE=0.0062)	(SE=0.0024)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
$Contrast_{t-1}$	0.0209	0.0130				
	(SE=0.0055)	(SE=0.0015)				
	(p<0.001)	(p<0.001)				
$Contrast_t$			0.0468	0.0235		
			(SE=0.0104)	(SE=0.0028)		
			(p<0.001)	(p<0.001)		
$Contrast_{SA}$					0.0006	0.0014
					(SE=0.0032)	(SE=0.0008)
					(p=0.847)	(p=0.098)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0182	0.0557	0.0168	0.0525	0.0173	0.0496
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast revision on own surprise and contrast. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Table 3.21: Error on Own and Contrast Surprise

	(1)	(2)	(3)	(4)	(5)	(6)
Own Surprise	0.2108	0.2056	0.2034	0.2008	0.2094	0.2066
	(SE=0.0158)	(SE=0.0047)	(SE=0.0167)	(SE=0.0049)	(SE=0.0176)	(SE=0.0053)
	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)	(p<0.001)
$Contrast_{t-1}$	0.0212	0.0097				
	(SE=0.0148)	(SE=0.0049)				
	(p=0.152)	(p=0.049)				
$Contrast_t$			0.1189	0.0567		
			(SE=0.0305)	(SE=0.0079)		
			(p<0.001)	(p<0.001)		
$Contrast_{SA}$					0.0721	0.0316
					(SE=0.0125)	(SE=0.0034)
					(p<0.001)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0119	0.0281	0.0113	0.0273	0.0118	0.0276
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of forecast error on own surprise and contrast. Forecast error is defined as the announced EPS minus the forecasted EPS for a given announcement. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. Forecasts are pooled across forecast horizons. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

3.D Additional Return Results

In my main specifications, I follow HS and use characteristics-adjusted returns. That is, I define returns as raw returns minus the returns of a portfolio matched on size, book-to-

market value and momentum. This has the potential issue, that it filters out any effect of contrast surprises on returns that influences the whole market. In particular, this likely filters out the news effect of the contrast firm since this will influence all firms in the market if it is related to the general economic condition. Because I cannot do a similar characteristics adjustment for my forecast revision, any attempt to link forecast revisions to characteristics-adjusted returns is potentially confounded. I explore if this affects my results by repeating my return analysis using raw returns.

Table 3.22 is a direct match to table 3.10. It shows that the contrast effect on the announcement return disappears if I use raw returns. The effect becomes positive albeit insignificant. This is in line with the idea that positive contrast surprises are a sign of favourable economic conditions and that this effect gets filtered out in the characteristics-adjustment. For returns between 26 and 50 days after the announcement, the coefficients on the contrast surprises are positive or negative depending on the specification so there is no clear picture.

Table 3.23 is a direct match to Table 3.11. The results for raw returns are less clear than those for characteristic adjusted returns. The coefficients for $contrast_{t-1}$ are positive but insignificant. The coefficients for contrast $contrast_t$ and $contrast_{SA}$ are negative and with robust regression, they remain highly significant. This suggests that analysts and the market still react to these contrasts in the opposite direction. For the characteristics adjusted returns, this could be explained by the fact that the news channel of the contrast surprise gets filtered out by the adjustment, but for raw returns it is a puzzling result.

Table 3.24 is a direct match to table 3.12. Like for the characteristics adjusted returns, all coefficients are positive and statistical significance is somewhat higher with raw returns. This supports the idea that contrast surprises cause underreaction of expectations which leads to a subsequent correction and higher returns.

Finally, Table 3.25 shows the results of a regression of returns between 16 and 50 days after the announcement on forecast revisions predicted by the contrast surprises. Since I include forecasts made within 15 days after my announcement in the forecast revisions, returns starting 16 days after the announcement for which the return window does not overlap with the expectations and the results can be interpreted as expectations

Table 3.22: Raw return on Contrast

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$raw\ return_{t-1,t+1}$				$raw\ return_{t+26,t+50}$			
$Contrast_{t-1}$	0.0926	0.0027			-0.2234	-0.3231		
	(SE=0.0551)	(SE=0.0408)			(SE=0.1079)	(SE=0.0824)		
	(p=0.093)	(p=0.946)			(p=0.038)	(p<0.001)		
$Contrast_t$			0.1567	0.1149			1.2148	0.0899
			(SE=0.1321)	(SE=0.1297)			(SE=0.3705)	(SE=1.0087)
			(p=0.235)	(p=0.376)			(p=0.001)	(p=0.929)
Robust Regression	no	no	no	no	no	no	no	no
Adjusted R^2	0.0631	0.0438	0.0654	0.0473	0.0019	0.0007	0.0018	0.0007
N	114068	145132	91136	104452	114014	143048	91088	102665

The table shows the result of a regression of raw return on contrast. The reference date t for the return windows is the day the firm announced its earnings. Columns 1,3,4 and 7 use the standard dataset. Columns 2,4,6 and 8 use a dataset that does not require an available forecast around an announcement for a datapoint to be included. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm level.

Table 3.23: Raw return [t-1, t+1] on Revision Predicted by Surprises

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	2.6330	1.8159				
	(SE=4.5910)	(SE=5.5948)				
	(p=0.566)	(p=0.746)				
$Contrast_t$			-0.8780	-12.5454		
			(SE=3.7352)	(SE=5.3218)		
			(p=0.814)	(p=0.018)		
$Contrast_{SA}$					-17.1594	-65.3500
					(SE=9.6487)	(SE=19.5710)
					(p=0.075)	(p<0.001)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0712	0.0905	0.0709	0.0910	0.0710	0.0904
N	1759726	1759726	1584031	1584031	866762	866762

The table shows the result of a regression of raw return [t-1, t+1] on revision predicted by contrast. The reference date t for the return windows is the day the firm announced its earnings. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

predicting returns. Overall, the results are very similar to those for returns between 26 and 50 days after the announcement reported in Table 3.12, except that the coefficients for $contrast_{t-1}$ are negative and insignificant instead of positive and insignificant.

Table 3.24: Raw return [t+26, t+50] on Revision Predicted by Surprises

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	1.4267	13.7187				
	(SE=7.3453)	(SE=8.4193)				
	(p=0.846)	(p=0.103)				
$Contrast_t$			19.7965	26.7832		
			(SE=6.0509)	(SE=7.9792)		
			(p=0.001)	(p<0.001)		
$Contrast_{SA}$					48.6288	70.2443
					(SE=18.1321)	(SE=32.3683)
					(p=0.007)	(p=0.030)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0033	0.0027	0.0036	0.0032	0.0032	0.0031
N	1759319	1759319	1583678	1583678	866566	866566

The table shows the result of a regression of raw return [t+26, t+50] on revision predicted by contrast The reference date t for the return windows is the day the firm announced its earnings. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Table 3.25: Return [t+16, t+50] on Revision Predicted by Surprises

	(1)	(2)	(3)	(4)	(5)	(6)
$Contrast_{t-1}$	-1.7739	-1.1212				
	(SE=8.3707)	(SE=8.5489)				
	(p=0.832)	(p=0.896)				
$Contrast_t$			5.9512	8.7304		
			(SE=5.2892)	(SE=7.8304)		
			(p=0.261)	(p=0.265)		
$Contrast_{SA}$					16.7037	34.3953
					(SE=13.3381)	(SE=19.6067)
					(p=0.210)	(p=0.079)
Robust Regression	no	yes	no	yes	no	yes
Adjusted R^2	0.0029	0.0019	0.0030	0.0021	0.0032	0.0023
N	1438806	1438806	1301887	1301887	707018	707018

The table shows the result of a regression of Return [t+16, t+50] on revision predicted by contrast All returns are characteristics adjusted by subtracting the return of a portfolio matched on size, book-to-market value and momentum from the raw return. The reference date t for the return windows is the day the firm announced its earnings. Forecast revision is defined as the most recent forecast for a target date made after the announcement minus the most recent forecast for the same date made before the announcement. Forecast revision predicted by contrast is obtained by first regressing forecast revision on the contrast surprise and controls and then using the coefficients on the contrast surprise to predict revisions. $contrast_{t-1}$ is the value-weighted surprise of all firms announcing on the day prior to the firm, $contrast_t$ is the value-weighted surprise of all firms announcing on the same day at an earlier time as the firm, $contrast_{SA}$ is the value-weighted surprise of firms covered by the same analyst that announced within 15 days prior to the firm's own announcement. All specifications include controls for 20 own surprise bins. Standard errors are clustered at the firm and at the analyst level.

Chapter 4

Which stock return predictors reflect mispricing?

Abstract

A large and quickly growing set of stock characteristics has been found to predict the cross-section of returns. Return predictability can be driven by risk or mispricing and the nature of most return predictors remains an open question. I propose a new test that uses data on analysts' earnings forecasts to determine if a return predictor reflects mispricing. I apply this test to a dataset of 173 significant return predictors and find that at least 40% of them represent mispricing. This challenges the efficient market hypothesis. I further study whether the mispricing predictors' abnormal returns capture the divergence of prices from fundamental value (build-up predictors) or their convergence back to fundamental value (resolution predictors). Build-up predictors are less common than resolution predictors but do exist, implying that trading on certain return predictors can exacerbate rather than eliminate mispricing.

4.1 Introduction

Over the past decades, hundreds of cross-sectional stock return predictors have been discovered. For instance, small-cap stocks offer higher returns than large-cap stocks and stocks with a high book-to-market ratio outperform those with a low book-to-market ratio (Fama and French, 1992). The vast majority of these predictors were discovered by studying patterns in the data rather than being derived from theory. Therefore, for most predictors, there is no consensus on the channel through which they drive returns.

Given that the fundamental value of a stock is its discounted stream of future cash flows to the investor, stock returns over a given period can be driven by two possible components: the discount factor applied to future cash flows, and changes in dividend expectations Campbell (1991)[1].

In this paper, I develop a method to test if a given return predictor represents mispricing driven by biased expectations rather than risk. If a return predictor is driven by risk, it is linked to returns through the discount factor. Specifically, investors receive compensation for exposure to priced risk, that is risk which cannot be eliminated through diversification. The more priced risk a stock is exposed to, the more investors will discount its future cash flows, leading to a higher return.

In contrast, changes in dividend expectations are unrelated to risk. Moreover, under (full information) rational expectations they should be unpredictable because all available information is already incorporated optimally into the current forecast and can thus not explain future changes in this forecast. Since dividend expectations drive prices, biased expectations imply mispricing. If a return predictor is associated with risk, it should therefore be linked to the discount factor but not to changes in dividend expectations, whereas a predictor that can predict changes in earnings expectations in addition to returns is linked to mispricing.

[1]While early studies suggested that returns are almost exclusively driven by changes in required returns (i.e., discount factors) more recent studies suggest that changes in dividend expectations play a significant role (L. Chen et al., 2013; De La O and Myers, 2021; Pruitt, 2023).

For my test, I use the earnings forecast of professional analysts as a proxy for the market's dividend expectations. I can then say that a return predictor is a mispricing predictor (henceforth also called a "mispricing") if it also predicts more positive earnings forecast revisions for the stocks for which it predicts more positive returns[2]. This is a sufficient but not a necessary condition because mispricing can also be driven by biased beliefs about future required returns, as I discuss in Section 4.2.

The nature of return predictors is directly related to several central questions of finance. For instance, mispricing predictors pose a challenge to the efficient market hypothesis, whereas risk factors do not. Similarly, trading on mispricing predictors promises abnormal risk-adjusted returns, whereas trading on risk factors does not. Therefore, it is crucial to understand which predictors reflect mispricing. However, it is also challenging because predictor discovery is mainly driven by empirical research, and it is difficult to establish a predictor's nature from the return distribution without an economic model (Kozak et al., 2018). Therefore, considerable disagreement about the nature of most predictors remains in the literature (see Holcblat et al., 2022, for a review) and arguably the CAPM market factor (Lintner, 1965; Sharpe, 1964) is the only predictor that is universally accepted to represent risk and not mispricing. Recent studies further complicate the problem by raising the concern that many published predictors may result from multiple hypothesis testing issues and publication bias (Chordia et al., 2020; Harvey et al., 2016; Hou et al., 2020) and hence offer no out of sample return predictability.

In Section 4.2, I describe my data and methodology. I use a dataset created by A. Y. Chen and Zimmermann (2021) that contains monthly stock returns as well as data on 207 characteristics, which showed significant evidence of return predictability in the paper that introduced them (predictors) as well as 113 characteristics that lacked such evidence (placebos). I merge this data with analyst forecasts from IBES.

In Section 4.3, I first show that 173 out of the 207 predictors offer significant return

[2]I require that the predictability of returns and expectations goes in the same direction because only positive expectation changes can explain positive returns. However, risk factors should not be able to predict changes in expectations of any kind.

predictability if they are extended beyond their original sample period using data up to the end of 2021. Applying my mispricing test to these 173 predictors, I find that around 40% of them also predict earnings forecast revisions and are hence classified as mispricing. Papers that discover a new predictor generally offer guidance on whether the authors believe it represents risk or mispricing, even though they often do not have a formal test or model. However, recent evidence suggests that the interpretation in the original paper can be unreliable (A. Y. Chen et al., 2022). In line with that, a comparison between my classification and the interpretation of the original authors suggests that at least one-third of supposed risk factors are actually mispricings. Therefore, either structural models or empirical tests, like the one I propose in this paper, are required to interpret the nature of return predictors.

A potential concern for the validity of my results is that some of the return predictors in my sample may be spurious. My test could potentially classify such predictors as mispricing, because stocks that are ex-post selected to have high (low) returns plausibly also have more positive (negative) news which in turn implies more positive (negative) earnings forecast revisions. This can lead to spurious predictors being classified as mispricing (Engelberg et al., 2018). I do two analyses to assess how likely this is to occur. First, I apply my test to the 66 predictors that had insignificant evidence of return predictability both in the original paper that discovered them and in my sample. I find that in this dataset, around 20% of the predictors are classified as mispricing, which is around half as much as in the sample of significant predictors. While some insignificant predictors are classified as mispricing it is also plausible that they are true return predictors that just lack statistical significance in this sample. Therefore, I perform a second test in which I generate simulated spurious predictor portfolios by randomly sorting stocks into long and short portfolios and retaining those portfolios that offer significant and sizeable long-minus-short returns by chance. Then, I apply my classification procedure to these spurious predictors and find that my preferred specification does not misclassify a single one as mispricing. This suggests that falsely

classifying spurious predictors as mispricing is unlikely to be relevant in practice.

The abnormal returns associated with each mispricing can either reflect the convergences of prices back to fundamental values (resolution of mispricing) or their divergence from fundamental values (build-up of mispricing). In Section 4.4, I study which mispricings capture the build-up and the resolution of mispricing respectively. In a recent paper, van Binsbergen et al. (2023) perform this classification using a method that relies on calculating the fundamental value of stocks based on realised dividends and an assumption about the correct stochastic discount factor (SDF). They find that although most predictors capture the resolution of mispricing, a relevant share also captures its build-up. They also highlight the practical importance of this distinction: traders who exploit build-up predictors exacerbate mispricing rather than reduce it.

My classification method is based on the predictability of forecast errors, an orthogonal approach that does not require a stance on the correct SDF. For resolution predictors, stocks that predictably earn higher returns are initially underpriced relative to those earning lower returns. The subsequent higher returns then correct this mispricing. Hence, earnings expectations should also initially be significantly more pessimistic for stocks with predicted high returns relative to those with low returns. I find that around 45% of the mispricings can be classified as resolution. For build-up predictors, all stocks are initially priced correctly, and the subsequent higher returns of some stocks cause them to be overpriced relative to those that earned lower returns[3]. This implies that, after the period of excess returns, earnings expectations are too positive for stocks that earned higher returns relative to those with lower returns. I find that around 10% of mispricing predictors can be classified as build-up.

I am not the first to provide a method to test if a return predictor represents mispricing or risk. Pukthuanthong et al. (2019) develop a set of conditions that categorise a predictor as a risk factor if it is related to the covariance matrix of returns, priced and has a

[3]It is also possible that the stocks earning higher returns are already overpriced initially, and become more overpriced over time

reasonable Sharpe ratio[4]. Gafka et al. (2021) compare predictors' ability to forecast returns on days when important information is released onto the market (announcement days) and the remaining days. They argue that if the return predictability of a given predictor is concentrated on announcement days it is likely associated with risk. Holcblat et al. (2022) develop a test that classifies a predictor as mispricing if the hypothesis that every risk-averse individual would prefer to invest in its long portfolio rather than its short portfolio cannot be rejected.

My approach has two advantages over previous methods. First, it offers a clean statistical test that classifies a predictor as mispricing if its null hypothesis is rejected. In contrast, the test by Holcblat et al. (2022) classifies a predictor as mispricing if the null hypothesis is not rejected[5]. Gafka et al. (2021) make the intuitive argument that the return of risk predictors should be stronger on macroeconomic information announcement days, but some mispricing predictors likely also have significant returns on these days given that there is evidence that expectations do not respond optimally to new information. Finally, the protocol suggested by Pukthuanthong et al. (2019) requires an assumption about the "reasonable" upper bound for the Sharpe ratio of a risk factor and the authors note that it classifies mispricing that is correlated with common factors as discussed in Kozak et al. (2018) as risk.

The second advantage of my test is that, unlike previous approaches, it does not rely on the same return data used to discover the predictors. Therefore, it is very unlikely to falsely classify spurious predictors, which just offer high in-sample returns by chance, as mispricing.

Besides issues with interpretability, the sheer number of return predictors poses a challenge to asset pricing researchers, as it is not tractable to work with models that include hundreds of predictors. An important first step to reduce the number of predictors

[4]An older paper by Charoenrook and Conrad (2005) follows a similar approach but only derives a necessary but not sufficient condition.

[5]They show that their test asymptotically never fails to reject a false null hypothesis, but in finite samples, a lack of statistical power can nevertheless lead to a false classification of a predictor as mispricing, despite the tests good finite sample properties.

is to remove those that are spurious (i.e., offer no out-of-sample return predictability). Several recent papers have addressed this issue (Chordia et al., 2020; Giglio et al., 2021; Harvey et al., 2016; Hou et al., 2020). Going beyond validity, it has been shown that several predictors capture the same source of underlying return predictability making it unnecessary to include them all in a model. As a response, researchers have tried to test which predictors add additional information relative to other predictors (Feng et al., 2020; Harvey and Liu, 2021). Furthermore, Jensen et al. (2023) find that a large set of predictors can be grouped into 13 themes. Moreover, He et al. (2022) suggest shrinking the predictor dimension by forming linear combinations out of a set of existing predictors.

I add to this literature in two ways. First, I show that a predictor that is classified as mispricing by my test is unlikely to be spurious implying that the test not only provides evidence on the nature but also the validity of a predictor. Second, I show that the return predictability of around 40% of the predictors is associated with biassed expectations. This suggests that a single expectation bias return predictor could potentially subsume the return predictability of a large set of predictors. However, developing such a predictor is beyond the scope of this paper[6].

Finally, my paper contributes to the literature that shows links between expectation data and asset returns (see Adam and Nagel (2023) for a review) by showing that this data can not only be used directly to predict returns but also to understand the nature of return predictor characteristics. Some previous work in this literature has linked individual return predictors (Bouchaud et al., 2019; Bradshaw et al., 2006; Jackson and Johnson, 2006; La Porta et al., 1997) or small sets of predictors (Ben-Rephael et al., 2021; Bordalo et al., 2022; Grinblatt et al., 2018) to biassed expectations[7]. Closest to my paper is a strand of the literature showing that analysts are pessimistic about stocks in predictor short portfolios. (van Binsbergen et al., 2022; Engelberg et al., 2018; Kozak

[6]Developing an expectation bias return predictor is challenging because it needs to predict returns and, hence, cannot be based on ex-post forecast errors. Moreover, build-up and resolution predictors predict different levels of bias at different points in the return cycle

[7]A tangentially related literature also shows that analysts' stock return predictions (rather than firm earnings predictions) are related to return predictors (Engelberg et al., 2020; Guo et al., 2020).

et al., 2018). Notably, this approach implicitly assumes that all mispricings are resolution predictors since the opposite should be the case for build-up predictors[8]. I add to this literature by providing a test based on forecast revisions over the return period that is valid for both build-up and resolution predictors. Moreover, I provide results on the share of mispricings among a large set of predictor variables and an individual classification for each predictor. In contrast, previous research either studied only a few predictors or aggregated a large set of predictors into a single score which does not allow conclusions about the nature of individual predictors.

4.2 Data & Methodology

4.2.1 Motivation of the Empirical Specification

In this section, I motivate the empirical specification used to detect mispricings among a set of predictors potentially containing mispricings and risk factors. Derivations of the equations used in this Section can be found in Appendix 4.A.

The most common way to establish that a characteristic can be used to predict stock returns is the following procedure: First, the cross-section of stocks is assigned to long and short portfolios at regular intervals, based on the level of the characteristic for each stock at the time of portfolio formation and the hypothesised relation between the characteristic and returns. For example, for the size predictor, the 10% of stocks with the lowest market capitalisation are assigned to the long portfolio and the 10% of stocks with the highest market capitalisation are assigned to the short portfolio. In the second step, the long-minus-short return is calculated for each month by subtracting the (potentially weighted) average return of stocks in the short portfolio from that of stocks in the long portfolio. A characteristic is then considered to be a return predictor if the average long-minus-short

[8]For a build-up predictor, there is no difference in expectations at portfolio formation. Depending on the source of the mispricing, analysts should be either too optimistic about stocks in the long portfolio or too pessimistic about stocks in the short portfolio at the end of the return period. In either case, average expectations should be more positive for stocks in the long portfolio than those in the short one.

return across time is significantly larger than zero.

To better understand the source of return predictability, I decompose the long-minus-short return of a predictor characteristic into several components. I start with a stock level decomposition of the return, first proposed by Campbell (1991) based on the Campbell-Shiller decomposition (Campbell and Shiller, 1988a,b).

$$r_{i,t+1} = \mathbb{E}_t(r_{i,t+1}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} \rho_i^s g_{i,t+1+s} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} \rho_i^s r_{i,t+1+s} \tag{4.1}$$

In the equation above, $r_{i,t+1}$ is the log return of stock i between time t and time $t+1$, $g_{i,t+1+s}$ is its log dividend growth rate ($d_{t+1+s} - d_{t+s}$ where d_t is the log dividend paid at time t) and the parameter ρ_i is related to the average price dividend ratio and is close to but smaller than 1. While Campbell (1991) uses Equation 4.1 to study a market portfolio, it also holds for other portfolios and individual stocks since it follows directly from a dividend discount model.

Therefore, under rational expectations, three components drive a stock's realised return between time t and $t+1$: the expected return for this period at time t, and changes in dividend growth expectations, as well as changes in expected future return between time t and $t+1$. All else equal, an increase in expected future dividend growth will cause the realised return to be higher than expected, and an increase in future expected/required returns will cause it to be lower than expected. Importantly, under rational expectations, changes in expected dividend growth and expected future required returns can only be caused by new information arriving at time $t+1$ and cannot be predicted at time t. Hence, any potential return predictability must stem from differences in $\mathbb{E}_t(r_{i,t+1})$ (Bordalo et al., 2022).

Next, I transform the stock-level decomposition into a portfolio-level decomposition. Equation 4.1 can not only be applied to individual stocks but also to portfolios, by first taking the (potentially weighted) averages of the relevant variables across the stocks in the portfolios and then applying the log linearisation.

By applying Equation 4.1 separately to the aggregated long and short portfolios and taking the difference we can write:

$$r_{t+1}^{LS} \approx \mathbb{E}_t(r_{t+1}^{LS}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^{LS})^s g_{t+1+s}^{LS} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^{LS})^s r_{t+1+s}^{LS} \tag{4.2}$$

In the equation above, r_{t+1}^{LS} is the long-minus-short return of a predictor characteristic between time t and $t+1$ and g_{t+1}^{LS} the average log dividend growth rate between time t and $t+1$ in the long portfolio minus that of the short portfolio. Moreover, ρ^{LS} is a constant that is close to but smaller than one. Technically, different values of ρ need to be applied to the returns and dividend growth rates in the long portfolio than to those in the short portfolio, making it impossible to aggregate across portfolios. However, since ρ is always close to one by the construction of the Campbell-Shiller log linearisation, it is similar in both portfolios and we can approximate the ρ of each portfolio by ρ^{LS}. The resulting small approximation error is of no concern because I only use Equation 4.2 to qualitatively discuss the impact of the individual return components in order to give an intuition for my empirical specification.

In Equation 4.2, $\mathbb{E}_t\left(r_{t+1}^{LS}\right)$ captures the difference in expected/required returns between the long and the short portfolio, that is, the differences in compensation for risk. Under rational expectations, this is the only predictable component of Equation 4.2. Therefore, if a predictor is a risk factor and not associated with biassed expectations, its ability to predict returns stems exclusively from its ability to predict $\mathbb{E}_t\left(r_{t+1}^{LS}\right)$. However, if we allow for biased expectations (denoted by $\tilde{\mathbb{E}}$), then both $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ and $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ are potentially predictable. If a characteristic can indeed predict either of these terms, this implies that its ability to predict returns is not (fully) driven by its association with risk ($\mathbb{E}_t\left(r_{t+1}^{LS}\right)$) but is at least partially explained by predictably biased beliefs. Since bias in the dividend growth expectations implies that $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s}) \neq [\mathbb{E}_{t+1} - \mathbb{E}_t](g_{t+1+s})$, realised returns will differ from their counterparts under rational

expectations, indicating that the long-minus-short portfolio is mispriced. By a similar argument, bias in $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ also implies mispricing. This gives Proposition 1:

Proposition 1. *A return predictor characteristic is associated with mispricing if and only if it can predict either* $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ *or* $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ *for any* $s \in (0, \infty)$*, in addition to returns.*

Notably, the ability to predict changes in future expected dividend growth or required returns does not mean that a characteristic cannot also predict $\mathbb{E}_t(r^{LS}_{t+1})$ and is therefore associated with risk. However, since it is universally accepted that risk affects stock returns, but contested if predictable mispricing exists in the stock market, I will consider a return predictor to be a mispricing if its return predictability is at least partially driven by mispricing.

Neither $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ nor $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ are observable. However, for the former, forecasts by professional stock analysts can be used as a proxy. Therefore, I focus on the predictability of changes in dividend growth for the rest of the paper. Finding such predictability is a sufficient but not a necessary condition to identify that a predictor's returns are (partially) driven by mispricing because mispricing may also stem from the predictability of changes in future required returns.

4.2.2 Dataset and Variable Construction

This section describes the data I use to classify return predictors into risk factors and mispricings. My primary dataset consists of 207 predictor portfolios from Chen and Zimmermann (2021). I use their 2022.3 data release, which includes data up to the end of 2021. All included predictors showed evidence of return predictability in their original paper. I provide detailed descriptions of all predictors mentioned by name in this paper in Appendix 4.D. The portfolio data includes monthly stock returns from CRSP as well as data on the predictor characteristic. I merge the portfolio data with earnings forecasts from the IBES unadjusted detailed history file, adjusted for stock splits using CRSP

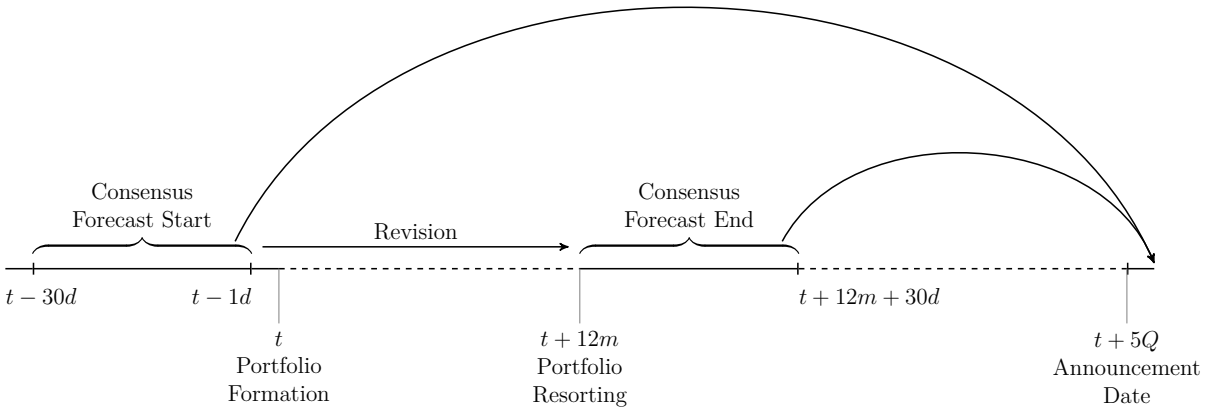
data. I use IBES data up to the end of 2022.

I use earnings forecasts instead of dividend forecasts due to their better data availability. These variables are closely related, as all earnings must eventually be paid out over a firm's complete life cycle. In fact, earnings forecasts may be a better proxy of the market's dividend growth expectations since I focus on forecasts for a single horizon for practical reasons. Arguably, a single earnings forecast is more informative about the firm's future dividend growth than a single dividend forecast, as earnings reflect the ability to pay dividends and are unaffected by pay-out policies such as dividend smoothing. Focussing on a single horizon is also sufficient to establish if a return predictor is associated with mispricing by Proposition 1. Finally, I use levels instead of growth forecasts. This can be done because growth rates are derived from levels, and under rational expectations, forecast revisions of any kind are unpredictable.

Figure 4.1 illustrates how I construct the forecast revision variable, which I will use as a proxy for the market's changes in dividend growth expectations. First, I calculate the median (consensus) five quarters ahead forecasts at the time of portfolio formation t using all forecasts made between 30 days ($t - 30d$) and 1 day ($t - 1d$) before portfolio formation. Then, I subtract this from the consensus forecasts for the same earnings announcement, including forecasts made within 30 days after the next portfolio resorting. Stocks are resorted into portfolios every one to twelve months in my sample of predictors. Hence, between zero and four announcements can happen between two sortings. Thus, the forecast made after resorting is for between one and five quarters ahead

The methodology above has the downside of only yielding one data point for each firm per resorting. Thus, I would only obtain one data point per year for predictors with yearly resorting. To increase my number of observations, I slightly adjust the procedure to generate data on a monthly frequency. Since predictors only have (confirmed) predictive power between the time of portfolio formation and the next resorting, I keep the date for the second forecast fixed at the time of resorting. The date of the first forecast is changed on a monthly frequency. Both the initial forecast and the one at the time of the next

Figure 4.1: Timing of the forecast revisions



The Figure illustrates the timing of the forecasts involved in the forecast revisions, at the example of a predictor with yearly portfolio resorting. t is the date of portfolio formation. d stands for days, m stands for months and Q stands for quarters.

Table 4.1: Overview of Forecast Matching

Date initial forecast made	Horizon initial forecast	Target date (both forecasts)	Date forecast after resorting	Horizon forecast after resorting
1st June year t	5Q	1st July Year t+1	1st June year t+1	1Q
1st July year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st August year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st September year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st October year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st November year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st December year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st January year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st February year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st March year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st April year t+1	5Q	1st July Year t+2	1st June year t+1	5Q
1st May year t+1	5Q	1st July Year t+2	1st June year t+1	5Q

The table illustrates the matching of forecasts used to compare forecast revisions in the predictor long and short portfolios at the example of a predictor with yearly resorting on the first of June and for a firm with earnings announcements happening on the first of January, April, July and October, respectively.

resorting are for the same target date. The initial forecast is always for five quarters ahead. Since the date at which the initial forecast is made moves every month, whereas the date of the second forecast only changes when a resorting occurs, this methodology implies that the difference in forecast horizon between the initial forecast and the forecast at the end of the resorting period changes over time.

To illustrate this, Table 4.1 shows an example of a predictor with yearly resorting on the first of June and for a firm with earnings announcements on the first of January, April, July and October, respectively.

I choose five quarters ahead forecasts since this is the shortest horizon for which I can have a forecast for the same announcement date at portfolio formation and after the next resorting for resorting periods of 12 months which is the longest period I study[9]. The choice of a short horizon has the upside of better data availability and the downside that long-term growth expectations (which are also available from IBES) are theoretically the more important driver of stock returns. I may therefore fail to capture some mispricing predictors by focusing on less important shorter horizons. In Appendix 4.B, I repeat my main analysis with long-term growth (LTG) expectation data.

I calculate the forecast revision of the consensus forecast for stock i at month t with x month until the next resorting for the earnings announced at date d as:

$$forecast\,revision_{i,d,t,x} = \frac{forecast_{i,d,t+x} - forecast_{i,d,t}}{price_{i,t}}$$

I follow the standard practice and scale forecast revisions (and any other EPS-related variables) by price[10]. Moreover, I winsorize all variables derived from EPS forecasts. Winsorization is done based on the interquartile range (IQR) method. Let $Q1$ be the value of the first quartile and $Q3$ the value of the third quartile. I winsorize all data below $Q1 - 1.5(Q3 - Q1)$ or above $Q3 + 1.5(Q3 - Q1)$. I do this because even after scaling by the price, some firms still have significantly larger EPS, which can cause them to dominate my analysis quantitatively.

In addition to forecast revisions, parts of my analysis also use forecast errors defined as:

$$forecast\,error_{i,d,t} = \frac{actual\,value_d - forecast_{i,d,t}}{price_{i,t}}$$

[9]There is a single predictor with a 36-month resorting period in my sample which I exclude

[10]The reason for this scaling is that a companys earnings per share can only be meaningfully interpreted in conjunction with its share price because EPS crucially depend on the number of shares outstanding and can easily be changed by (reverse) stock splits. This can cause two otherwise identical companies to have vastly different earnings per share if they have different amounts of outstanding shares. Scaling EPS-related variables by price transforms them to earnings per $ of market capitalisation, which is unaffected by the number of shares outstanding and is only driven by economically meaningful measures.

4.2.3 Multiple Hypothesis Testing

Since I am testing significance for a large set of predictors, multiple-hypothesis testing is a concern. There are two common approaches to address this issue: controlling the false discovery rate (FDR) and controlling the family-wise error rate (FWER). Controlling the FDR limits the share of false positive results to α (i.e. 5% for the most common value of α). Controlling the FWER limits the probability of a single false positive among all hypothesis α. Controlling the FWER is stricter than controlling the FDR, particularly for a large number of tests, since for the FDR, the number of acceptable false positives scales with the number of tests, whereas controlling the FWER tries to rule out any false positives regardless of the number of tests. To achieve this, controlling the FWER requires higher and higher critical t-values as the number of tests increases, leading to a large fraction of false negatives.

In my main analysis, I report results without MHT controls and results that control the FDR using the method by Benjamini and Hochberg (1995). I chose to control the FDR and not the FWER because I am trying to provide an accurate picture of the fraction of mispricings among the total predictors. Therefore, I am not only concerned about classifying a predictor as mispricing by mistake (type one error) but also about failing to classify a predictor as mispricing (type two error). Since controlling the FDR strikes a balance between limiting both error types while controlling the FWER tries to rule out type one errors at the expense of allowing more type two errors, controlling the FDR is more suited to my research question.

4.2.4 Descriptive Statistics

Table 4.2 shows descriptive statistics for my main variables, aggregated across the 207 predictors, that showed significant evidence of return predictability in the original sample. To calculate the aggregated descriptive statistics, I first calculate each variable's mean and standard deviation separately for each predictor and then take the mean of those means

and standard deviations across the 207 predictors. Only the values for N, the number of data points with available forecast revisions per predictor, are calculated differently. This variable is already defined on the predictor level. Hence, the mean and standard deviation across predictors can be calculated directly. As expected, the return is substantially larger in the long portfolios (5.26%) compared to the short portfolios (4.39%). Forecasts are also more positive in the long portfolio, both at portfolio formation and after the next resorting. Forecast revisions are negative in both portfolios but more so in the short portfolio, suggesting that there are at least some mispricing predictors in my sample. As discussed above, forecast data availability is an issue. On average, I have forecast revisions for slightly more than 10% of the data points for which I have return observations. This will reduce the statistical power of the forecast analysis relative to the return analysis and can lead to some mispricings not being detected.

Table 4.2: Descriptive Statistics

	long mean	long sd	short mean	short sd
Return	5.26	31.72	4.39	30.41
Forecast Start	0.0174	0.0191	0.0161	0.0192
Forecast End	0.0166	0.0190	0.0151	0.0189
Forecast Revision	-0.000953	0.005835	-0.001152	0.005758
Forecast Error Start	-0.003416	0.013100	-0.003903	0.012972
Forecast Error End	-0.002105	0.010326	-0.002366	0.010178
% Forecast Available	0.137937	0.058887	0.118441	0.054135
N	56945	57271	53561	66081

This table shows descriptive statistics for the main variables using return data from 1984 to 2021 and earnings (forecast) data from 1984 to 2022. It includes all data points with available forecasts. Forecast-related variables are scaled by price and winsorized. The data is separated into long and short portfolios. For all variables except N, I first calculate the mean and the standard deviation of the variable for each portfolio of each predictor and then aggregate across predictors by taking the mean of the individual means and standard deviations. The variable N is already defined on the predictor level, so the mean and standard deviation across predictors can be calculated directly. *Return* is the monthly stock return. *Forecast Start* and *Forecast End* are the consensus forecast values at portfolio formation and after the next resorting, respectively. *Forecast revision* is defined as *Forecast end* minus *Forecast Start*. *Forecast Error Start* and *Forecast Error End* are the actually announced value minus the consensus forecast values at portfolio formation and after the next resorting, respectively. *% Forecast Available* is the number of return observations with available forecast revisions divided by the total number of return observations. A more detailed description of the variable construction can be found in Section 4.2.2.

4.3 Mispricings

4.3.1 Main Results

As a first step, I test which predictors offer a significantly positive long-minus-short return using the data up to the end of 2021 (i.e. including the post-publication period). To do this, I calculate the monthly portfolio return using equal or value weighting depending on which the original paper used. Next, I do t-tests to determine if the long minus short return is significantly larger than zero with a t-statistic above 1.96 and find that this is the case for 173 predictors. Since significant return predictability is a prerequisite for

a predictor, I only use these predictors in the subsequent analysis. I do not control for multiple hypothesis testing in this first step because doing so would cause me to lose predictors even if their level of return predictability is the same as in the original paper since the original papers generally do not use MHT controls. This goes against my goal of classifying as many valid predictors as possible. If this approach causes a few false predictors to pass the initial test, this will likely cause me to underestimate the share of mispricings among the true predictors in my later analysis, as spurious predictors are unlikely to be classified as mispricing by my test. In Table A2 in Appendix 4.C, I show that my results are similar using only predictors that have significantly positive long minus short returns after controlling the false discovery rate.

By Proposition 1, a predictor is unable to predict revisions in dividend (or earnings) expectations if it is a risk factor but may do so if it is a mispricing because predictable changes in earnings expectations are one of the two channels through which biased beliefs may drive return predictability. I, therefore, say that a predictor represents mispricing if I find evidence that it predicts forecast revisions which I test by running the following regression using monthly-level data:

$$forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x} \tag{4.3}$$

In the regression, i indicates a firm, d indicates an announcement date, t is a month, and x is the number of months from t to the following portfolio resorting. $long_{i,t}$ is a dummy that is one if stock i is sorted into the long portfolio at time t and zero if it is sorted into the short portfolio. If a stock is neither in the short nor the long portfolio, it is excluded from the regression. Finally, $X_{i,t}$ is a vector of potential control variables. A positive and significant $long_{i,t}$ indicates that the predictor reflects mispricing since, for risk factors, there should be no difference in expectations between the long and the short portfolio. If $long_{i,t}$ is not positively significant, this suggests that the predictor is a risk factor or a mispricing related to biased beliefs about future required

returns. It could also mean that the analysis lacks statistical power to uncover changes in expectations in the noisy analyst forecast data or that earnings expectations, for a horizon I do not study, drive the predictor's abnormal returns. Corollary 1 then follows from Proposition 1 and Equation 4.3:

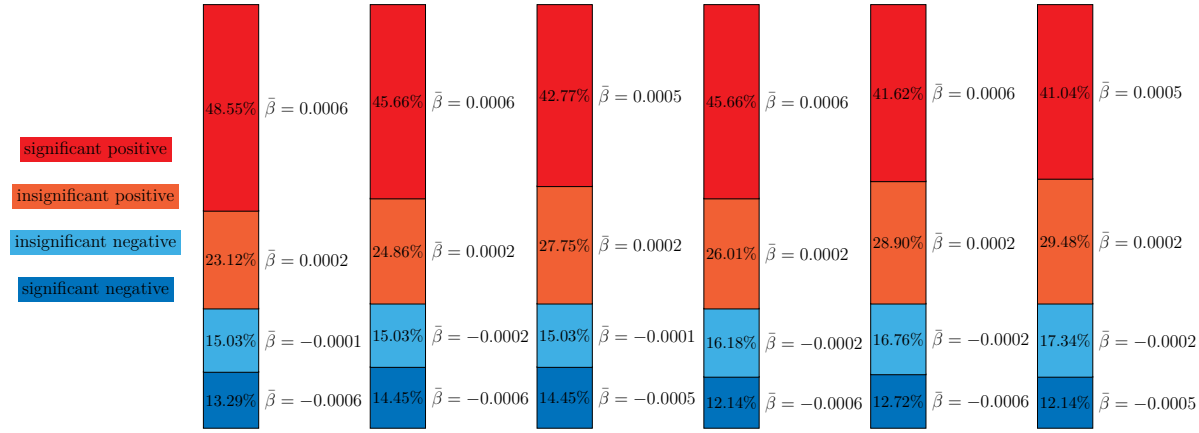
Corollary 1. *A return predictor can be classified as mispricing if β_1 is significantly larger than zero in the regression described in Equation 4.3.*

A crucial assumption behind Corollary 1 is that forecasts by stock market analysts are a valid proxy of the expectation of market participants. Previous research has found that data on analyst expectations can be used to explain patterns in stock returns and asset prices (see Adam and Nagel (2023) for a review), suggesting that they can indeed be used to proxy for the market's expectations. Another potential concern is that analysts (partially) extrapolate from returns when revising their forecasts (as suggested by, e.g. Ben-Rephael et al. (2021)). This could cause forecasts to be different in the long and the short portfolio, even for risk factors. I can address this issue by controlling for the return during the time between forecast revisions. However, since I form a consensus forecast by aggregating individual forecasts made at different points in time, I cannot perfectly control for the return information that was available to each individual analyst when they made their forecast, which means that there is some concern left that by results are biased upwards by analysts who extrapolate from returns. However, controlling for returns also introduces a downward bias in my result as it filters out firm-relevant news incorporated into both forecasts and prices. Nevertheless, existing evidence suggests that analysts add private information and do not just extrapolate returns van Binsbergen et al. (2022), which suggests that substantial variation will be left in the analyst forecast even when controlling for returns.

I run separate regression for each predictor with significantly positive long minus short returns in my data. The aggregated results are shown in Table 4.3.

The share of predictors that have a significant positive long-dummy and are classified

Table 4.3: Share of Mispricings among the Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
significant positive	48.55% $\bar{\beta} = 0.0006$	45.66% $\bar{\beta} = 0.0006$	42.77% $\bar{\beta} = 0.0005$	45.66% $\bar{\beta} = 0.0006$	41.62% $\bar{\beta} = 0.0006$	41.04% $\bar{\beta} = 0.0005$
insignificant positive	23.12% $\bar{\beta} = 0.0002$	24.86% $\bar{\beta} = 0.0002$	27.75% $\bar{\beta} = 0.0002$	26.01% $\bar{\beta} = 0.0002$	28.90% $\bar{\beta} = 0.0002$	29.48% $\bar{\beta} = 0.0002$
insignificant negative	15.03% $\bar{\beta} = -0.0001$	15.03% $\bar{\beta} = -0.0002$	15.03% $\bar{\beta} = -0.0001$	16.18% $\bar{\beta} = -0.0002$	16.76% $\bar{\beta} = -0.0002$	17.34% $\bar{\beta} = -0.0002$
significant negative	13.29% $\bar{\beta} = -0.0006$	14.45% $\bar{\beta} = -0.0006$	14.45% $\bar{\beta} = -0.0005$	12.14% $\bar{\beta} = -0.0006$	12.72% $\bar{\beta} = -0.0006$	12.14% $\bar{\beta} = -0.0005$
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	53322	52867	52867	53322	52867	52867
Number of Regressions	173	173	173	173	173	173

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\,revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

as mispricing varies from around 41% to around 49%. It is highest in the baseline specification (1), with no controls or fixed effects and no adjustment for multiple hypothesis testing. Interestingly, controlling for the return does not matter much: between specifications (1) and (2), the share of significantly positive long-dummies only drops by 2.89pp. This suggests that analysts learning from prices is not a significant concern. Column (3) adds year/month fixed effects relative to column (2), which is useful for two reasons: First, due to limited data availability, the number of firms in the long and short portfolio varies between months which could bias my result upwards if I have more (less) stocks in a month with generally good (bad) news in the long portfolio relative to the short portfolio. Second, due to the way the forecast revision is constructed (described in Section 4.2.2), the difference in forecast horizons between the initial forecast and the

revised forecast varies between months and the year/month will absorb any potential effect of this difference. Including fixed effects decreases the share of significant strategies by 2.89pp. Columns 4,5 & 6 are identical to columns 1,2 & 3, respectively, except that they control for multiple hypothesis testing (MHT) by keeping the false discovery rate at 5%. Naturally, this leads to a lower share of significant strategies. However, the effect is less than 5pp for all specifications, suggesting that the significant long-dummies generally have t statistics comfortably above the standard threshold of 1.96.

The coefficients are remarkably stable across specifications. Among the strategies with a significant positive long-dummy, the average coefficient is always either 0.0005 or 0.0006. This is approximately one-tenth of a standard deviation of the forecast revision and hence quite small. However, this is not particularly concerning: since excessive long-minus-short returns associated with mispricing are driven by changes in expectations for all future earnings, it is not unexpected that there is only a small effect for a single future quarterly earnings. Moreover, analyst forecasts are a noisy measure of the market's expectations. Finally, the small effect size also does not matter for the economic significance of my result: the predictability of forecast revisions has no direct economic relevance and cannot be traded upon. It just serves as a tool to classify predictors as mispricing. More relevant are the effect sizes of the mispricings' long-minus-short returns, which are large. Thus, the analysis has significant economic implications.

Interestingly, around 12% of predictors also have significant negative long dummies, which is predicted for neither mispricings nor risk factors. Since my controls for multiple hypothesis testing limit the rate of false positives to 5%, this is unlikely to be purely a statistical artefact. While this result is not predicted for mispricings, it is also no contradiction. Since returns are driven by changes in expectations for all future time periods, it is possible that expectations for a single period change in the opposite direction. In contrast, a risk factor should not be able to predict forecast revisions for any time period and in any direction. Hence, one potential interpretation for the result is that the associated predictors predict lower short-term but higher long-term earnings. Another

potential explanation is that the predictor is associated with both risk and mispricing but in the opposite direction. If the predictor's long portfolios are riskier but also overpriced at portfolio formation and the compensation for risk quantitatively dominates the mispricing effect, returns would be higher in the long portfolio relative to the short portfolio while expectation revisions in the long portfolio are more negative than those in the short portfolio at the same time.

In Appendix 4.B, I repeat my analysis using long-term growth forecasts, which capture the analysts' growth expectations over a firm's full business cycle. I find that the fraction of predictors that are classified as mispricing is comparable to that using quarterly forecasts. However, there is some difference in which predictors are classified as mispricing. Focussing on the specification with return controls, fixed effects and MHT controls, I find that 38 predictors are classified as mispricing by both methods, 33 are only classified as mispricing using quarterly forecasts and 23 are only classified as mispricing using LTG data. Since mispricing can stem from biased expectations at any horizon whereas risk factors should be unrelated to expectations at any horizon, this result suggests that mispricing might be even more widespread than the results focusing on quarterly forecasts suggest.

My sample contains predictors with different resorting periods. For roughly half of the predictors, the stocks are resorted into portfolios every month. The other predictors are almost all resorted yearly, while a few predictors also have 3-monthly and six-monthly resorting periods. In Figure A1 in Appendix 4.C, I examine if the share of detected mispricings among the predictors varies with the resorting frequency and find that it is broadly similar for predictors with monthly and yearly resorting. Surprisingly, all predictors with three-monthly and six-monthly resorting are mispricings, but there are not enough data points to draw a firm conclusion from this result.

The return predictability of each predictor in my dataset was tested separately. Many predictors are similar to each other, suggesting that some predictors are likely redundant. Feng et al., 2020 address the issue of which predictors to select from a large set of

candidates using a stepwise procedure. They start with the four predictors from the Carhart (1997) four-factor model: Market Beta, Size, Book-to-Market and Momentum. Then, they test for each predictor in their sample of 150 predictors how much it adds to the four-factor model. In the next step, they add the predictor that added most to the existing predictors to the model and test all remaining predictors against the new model with five factors. They repeat this procedure until no more predictor adds significantly to the model. Using this procedure, they select 26 predictors, including the original four that were used as a starting point. Out of those, I have 18 in my data, 14 of which offer significant long minus short returns.

Table 4.4 reports the results. Out of the 14 predictors, five have a t-statistic above 1.96 and are, thus, classified as mispricing by my method.

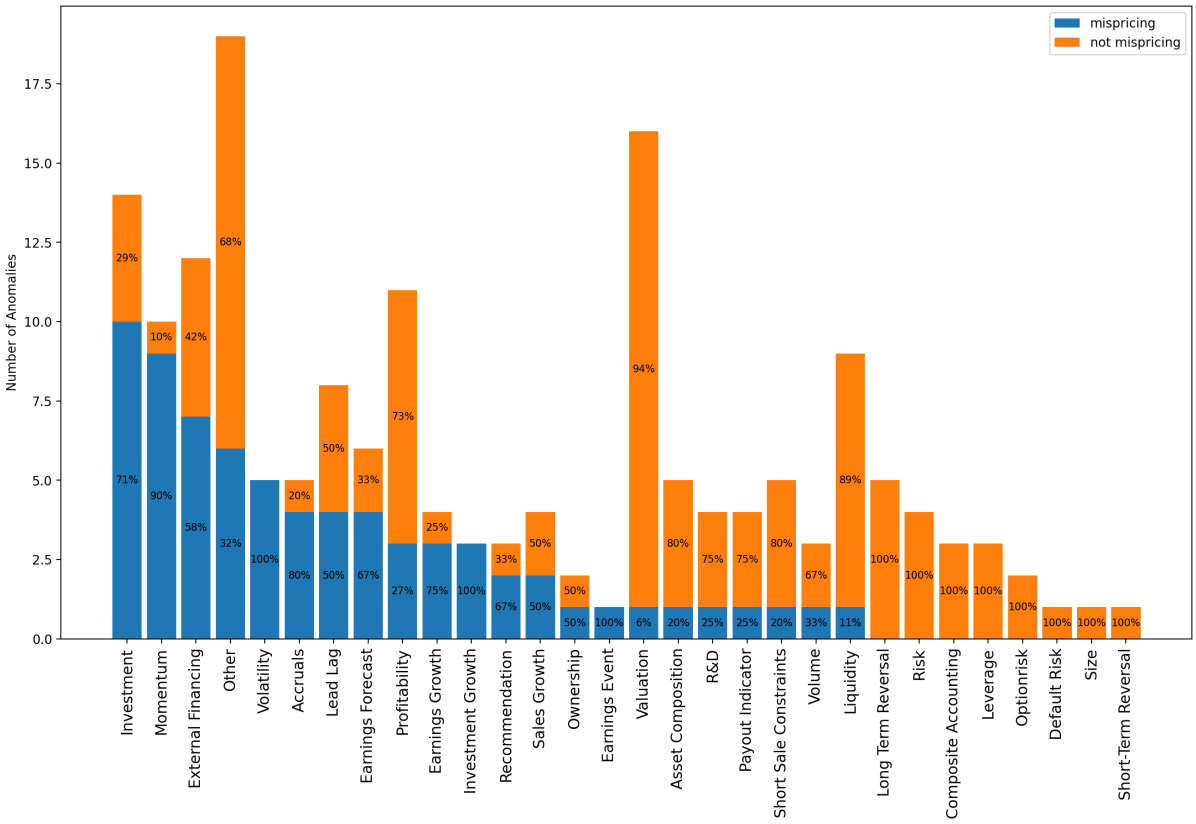
Table 4.4: Mispricing test for the stepwise selected predictors of Feng et al. (2020)

name	t value	parameter	standart error	Description
Accruals	2.43	0.0004	0.000157	Accruals
AdExp	0.94	0.0002	0.000193	Advertising Expense
BetaLiquidityPS	-3.09	-0.0004	0.000144	Pastor-Stambaugh liquidity beta
BMdec	-0.62	-0.0001	0.000204	Book to market using December ME
ChInvIA	3.11	0.0004	0.000116	Change in capital inv (ind adj)
DelFINL	3.57	0.0003	0.000096	Change in financial liabilities
Illiquidity	-2.35	-0.0004	0.000186	Amihud's illiquidity
Mom12m	8.60	0.0010	0.000119	Momentum (12 month)
NOA	3.81	0.0002	0.000055	Net Operating Assets
sinAlgo	-0.72	-0.0001	0.000109	Sin Stock (selection criteria)
Size	-6.31	-0.0007	0.000107	Size
SP	-2.04	-0.0005	0.000229	Sales-to-price
Tax	0.63	0.0001	0.000113	Taxable income to income
VolSD	-0.59	-0.0000	0.000083	Volume Variance

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\,revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor that survived the stepwise selection procedure from Feng et al. (2020), is included in my dataset and has significant long minus short returns in my data. The regression controls for the return over the period between the two forecasts included in the revision and uses Year/Month fixed effects.

4.3.2 Which Predictors are Mispricings?

Figure 4.2: Share of Mispricings by Predictor Category



The figure shows the share of predictors that are classified as mispricing by predictor category. The categories are taken from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification as mispricing is based on the regression of the forecast revision on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Specification (6) from Table 4.3).

Figure 4.2 separates the predictors into categories, following the classifications in the "Open Source Asset Pricing" dataset (A. Y. Chen and Zimmermann, 2021), and shows the share of mispricing predictors in each category using the regression specification that controls for return, includes fixed effects and controls the false discovery rate (specification 6 in Table 4.3). An individual-level overview of all mispricing predictors can be found in Table A8 in Appendix 4.C. *Momentum* is among the categories with the highest share of mispricing predictors (nine out of ten), which is reassuring since momentum is widely considered to be caused by under or overreaction to news. Related

to that, all three *investment growth*, three of four *earnings growth* and two out of four *sales growth* predictors are classified as mispricings. Overall this suggests that the vast majority of predictors based on past trends are mispricings. In addition, four out of five *volatility* predictors are classified as mispricing. Three of these come from Ang et al. (2006), who interpret their results as risk. Among these are two idiosyncratic volatility measures, namely the standard deviation of CAPM and FF3 Factor model residuals, which my approach classifies as mispricing. The third predictor is a measure of exposure to aggregate volatility as measured by the VIX, which has a t statistic of -0.96 in my regression and hence is not classified as mispricing. The other volatility measures are the earnings forecast dispersion from Diether et al. (2002) and the maximum return over the last month from Bali et al. (2011) who both interpret their results as mispricing.

Furthermore, four out of five predictors related to *accruals* are classified as mispricing, which is in line with a common interpretation of accrual predictors, for example by Sloan (1996): The accrual component of earnings is less persistent than the cash flow component, but investors do not make this distinction and focus on earnings. Hence, they overestimate the future earnings of companies with high accruals growth.

Finally, another class of predictors with many mispricings are those related to investment activities: 58% of the *external financing* predictors and 71% of the *investment* predictors are mispricings. This could be driven by biased reactions to the news related to investment decisions.

Among the categories with a low share of mispricings is *risk*, which contains predictors related to tails risk and skewness with 0% mispricing, *short sale constraints* with 0% mispricing and *valuation* related predictors with 6% mispricing. Somewhat surprisingly, 0% of the *long-term reversal* predictors are classified as mispricing. A common explanation for momentum and reversal is that momentum comes from overreaction to good news, which is then reversed later on. This would make both momentum and reversal related to mispricing. However, as I will show in Section 4.4, my results suggest that momentum captures the resolution of existing mispricing rather than the build-up of mispricing. This

goes against the interpretation above and suggests that there should not be reversal. A potential explanation that would reconcile these findings is that the reversal predictors stem from bias in earnings expectation for a more long-term horizon.

4.3.3 Detecting Spurious Predictors

So far, I have assumed that all return predictors in my dataset are either risk factors or mispricings. However, It is also possible that some are not valid return predictors at all but instead the result of data mining or multiple hypothesis testing. This is a potential concern for my methodology. Recall that by Equation 4.1, a stock has a high return between t and $t+1$ for three possible reasons: either its required return for this period is high, or the market revised its expectation for the discounted sum of future dividend growths upwards or because the market revised its future required returns downwards. If stocks in the long portfolio of a spurious predictor have higher in-sample returns than those in the short portfolio, this will likely be driven by a mixture of these three reasons. Therefore, these stocks likely have more positive dividend growth expectation revisions on average. Importantly, these more positive revisions do not imply mispricing, as they are not predictable. Instead, stocks with good (bad) firm-related news are selected ex-post to be in the long (short) portfolio, as also discussed by (Engelberg et al., 2018).

In theory, if returns and revisions of earnings forecasts are correlated ex-post for at least some stocks in a spurious predictor's long or short portfolio, this will cause my empirical specification to misclassify such a spurious predictor as mispricing. However, in practice, two factors make such a misclassification less likely. First, I control for the return in my main specification, which filters out a mechanical correlation between returns and analyst forecast revisions. Second, changes in dividend expectations are only one of the three reasons a stock could have higher returns in a given period. Therefore, parts of the return spread between a spurious predictor's long and short portfolio will also be driven by higher required returns or changes in future required returns. In contrast, a

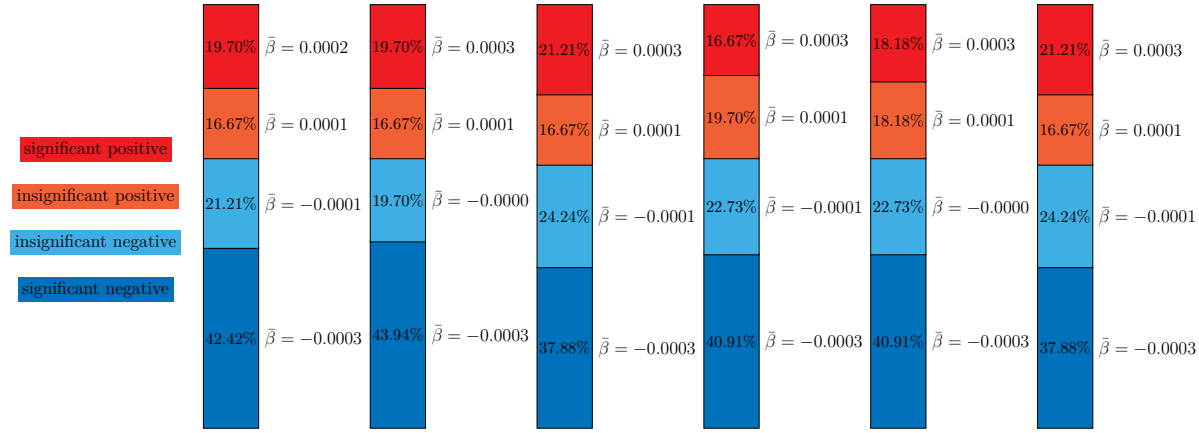
valid mispricing predictor can only predict changes in dividend expectations but not the other factors. Therefore, dividend expectations will drive most of its long-minus-short return. With limited statistical power, it is thus much less likely that the condition I use to classify a predictor as mispricing (Corollary 1) holds for a spurious predictor than for a mispricing predictor.

I do two robustness tests, to study how likely it is in practice that a spurious predictor gets classified as mispricing. First, I apply my test to a set of 66 predictors from A. Y. Chen and Zimmermann (2021) that had no significant evidence of return predictability both in the original paper that tested them and in the extended dataset using data up to the end of 2021. In this sample, around 20% of predictors are classified as mispricing. While this is only half of the share in the sample of significant return predictors, it highlights that some potentially spurious predictors are classified as mispricing. However, each of these predictors was initially tested because the researchers had the hypothesis that it might predict returns, so there are likely some true return predictors that just do not pass the statistical significance hurdle in this sample.

To assess how my test performs with a sample of predictors that are unambiguously spurious, I apply my methodology to a set of simulated spurious predictors. I generate these predictors by randomly sorting stocks into portfolios each month. I exclude the stocks with the worst 2% returns from the long portfolio and those with the best 2% of returns from the short portfolio to ensure I can find portfolios with sizeable long-minus-short returns in a reasonable time. I retain a simulated predictor if its long minus-short return is above the 90th percentile of long-minus-short returns for actual predictors.

To ensure that my simulated predictors are representative of the real predictors, they are designed to match them on the rebalancing frequency, the share of stocks assigned to the long/short portfolio and whether returns are equal or value-weighted. For example, after removing some predictors with more complex assignment rules, which I cannot represent in my simulated data, 38% of the predictor portfolios in my main analysis have monthly rebalancing, are equal-weighted and assign the most extreme deciles into the

Table 4.5: Detected Mispricings among Insignificant Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	63624	63460	63460	63624	63460	63460
Number of Regressions	66	66	66	66	66	66

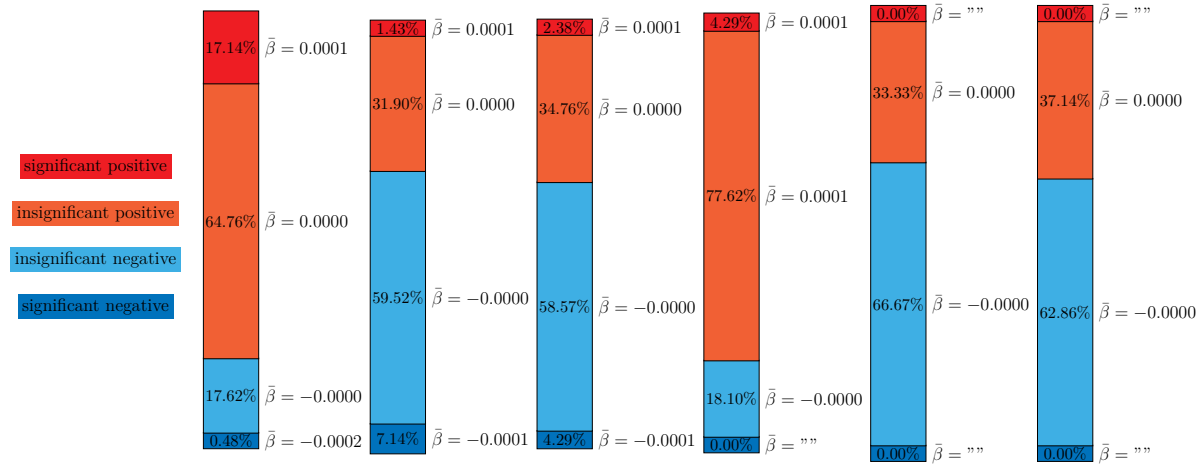
The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecastrevision$. $forecastrevision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in more detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for 210 simulated predictor portfolios that are designed to be spurious, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

long and short portfolio respectively. Consequently, also 38% of my simulated portfolios have these properties. I generate a number of simulated portfolios that is as close as possible to the number of real predictor portfolios (173) while maintaining the correct share of different types of predictors, giving me a target number of 210 portfolios.

I then run the regression specified in Equation 4.3 for each simulated portfolio. Table 4.6 shows the results, and its columns are directly comparable to those of Table 4.3, which uses the actual data. In the baseline specification, 17.14% of the simulated predictors are classified as mispricing. However, adding return controls reduces this share to 1.43% and even without return controls, only 4.29% of simulated predictors are classified as mispricing once multiple hypothesis testing is controlled for. In my main specification, with fixed effects, return controls and MHT controls, no simulated predictor is classified as

mispricing, suggesting that my methodology is extremely unlikely to misclassify spurious predictors as mispricing.

Table 4.6: Detected Mispricings among Simulated Spurious Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	62476	62476	62476	62476	62476	62476
Number of Regressions	210	210	210	210	210	210

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecastrevision$. $forecastrevision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in more detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for 210 simulated predictor portfolios that are designed to be spurious, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

4.4 Build-up or Resolution?

4.4.1 Main Results

In a recent paper, van Binsbergen et al. (2023) raise the question of whether mispricing predictors capture the build-up or the resolution of mispricing. They calculate the fundamental value of each firm in a predictor portfolio by discounting realised future dividends and capital gains with an estimated SDF. They then compare fundamental

values to market prices to calculate price wedges, i.e. deviations of the stock price from the fundamental value. By examining how the price wedges line up with the abnormal returns of the predictor portfolios they can determine if a predictor is build-up or resolution. They find that around a third of predictors are build-up. In Section 4.3, I identified 73 out of my total 173 predictors that capture mispricing driven by biased earnings expectations (based on specification 3 of Table 4.3). In this section, I classify these predictors into build-up and resolution predictors using an approach that is based on forecast errors which is complementary to the price-wedge-based approach.

For build-up predictors, prices are either correct at portfolio formation or the long portfolio is already overpriced relative to the short portfolio. The higher returns of the long portfolio lead to a (further) divergence of prices from fundamental value over the return period. This divergence can capture excessive optimism about stocks in the long portfolio, excessive pessimism about stocks in the short portfolio or both. In all of these cases, stocks in the long portfolio should be overpriced relative to those in the short portfolio after the return period. This should also be reflected in more optimistic earnings expectations for stocks in the long portfolio relative to stocks in the short portfolio at the end of the return period. By similar reasoning, for resolution predictors, the long portfolio is underpriced relative to the short portfolio at the time of portfolio formation. This implies that earnings expectations are more negative for stocks in the long portfolio than stocks in the short portfolio.

Finding evidence of biassed expectations at the time of portfolio formation (at the end of the return period) is insufficient to classify a variable as build-up (resolution). Another requirement is that changes in expectation during the return period have led to a divergence (convergence) of expectations. However, since I only classify predictors as mispricing if they had more positive forecast revisions for stocks in the long portfolio than for stocks in the short portfolio, this is the case for all mispricings. Consequently, it is sufficient to focus on the forecast errors in this section.

To determine if a variable reflects the resolution of mispricing I, therefore, run the fol-

lowing regression:

$$forecast\,error_{i,d,t} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t} \tag{4.4}$$

To determine if a variable reflects the build-up of mispricing, I run the following regression:

$$forecast\,error_{i,d,t+x} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t} \tag{4.5}$$

In the regression formulas, i indicates a firm, d indicates an announcement date, t is the month of a portfolio formation, and x is the number of months from t to the following portfolio resorting. $long_{i,t}$ is a dummy that is one if stock i is sorted into the long portfolio at time t and zero if it is sorted into the short portfolio. Unlike in the analysis in Section 4.3, which used monthly level data, I only use one data point per portfolio formation in this analysis. This implies that the data frequency depends on the resorting period and varies between monthly and yearly depending on the predictor. I do not use monthly data because, for resolution predictors, the difference in expectations between the long and the short portfolio diminishes over the return period. Hence, the further a monthly data point is away from the portfolio formation, the smaller the difference in expectations should be, and at the time of the next rebalance, the mispricing may be fully resolved. By a similar logic, there is initially no difference between expectations for the long and the short portfolio for build-up predictors, and the difference has only fully materialised at the end of the period for which the predictor predicts abnormal returns. Thus, I only have precise predictions for differences in forecast errors between the long and the short portfolios at the time of portfolio formation and after the next resorting. Hence, I cannot use monthly-level data.

As with the mispricing analysis, a potential issue with my methodology is that analysts may learn from returns. This issue is less severe here because I only use predictors classified as mispricing. Hence, provided that this classification is valid, there should be no

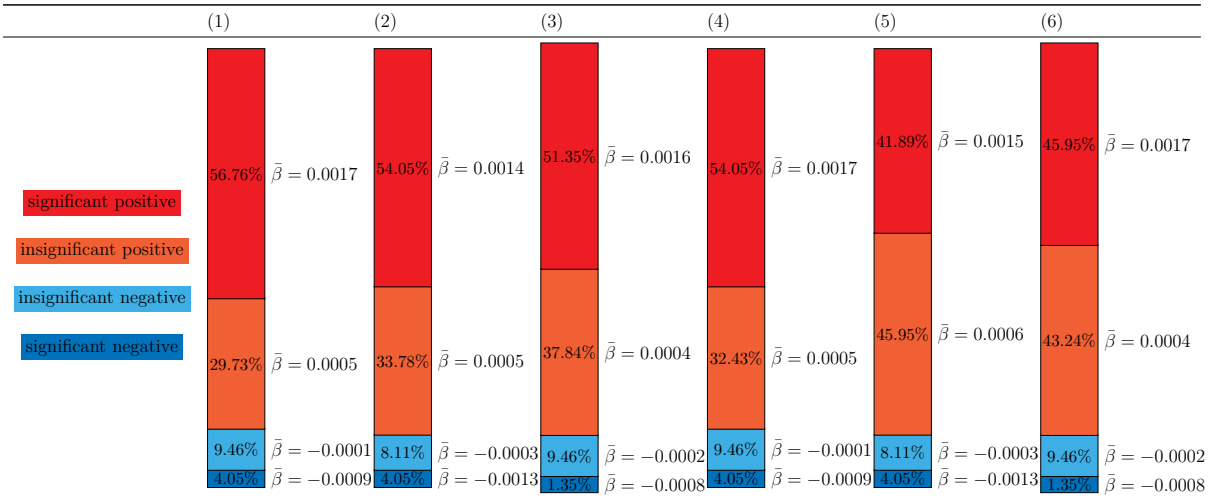
predictors where the link between the predictor sorts and the analysts' forecasts is purely driven by return extrapolation in this analysis. Nevertheless, extrapolation from returns would still affect the results. Therefore, I control for the return. For resolution predictors, I ideally want to control for the return over the horizon during which the mispricing that is being resolved has built up. Since this is unknown, I choose to control for the monthly returns of the 12 months preceding the forecast. For build-up predictors, mispricing is associated with the predictors' abnormal returns. Consequently, I control for the return between the formation of the predictor portfolio and the time the forecast was made.

I use the predictors that were classified as mispricing using the specification with return controls and year/month fixed effects but without MHT control (Column 3 of Table 4.3). I do not control for multiple hypothesis testing for the same reason I do not do it when assessing the return predictability of the predictors to narrow down my dataset for the mispricing analysis: false negatives and false positives both negatively affect the quality of my dataset. Hence, it does not make sense to focus excessively on controlling the rate of false positives.

Table 4.7 shows the results of a regression of forecast error on the day of portfolio formation on the long dummy. A positive long dummy indicates that analysts are more pessimistic about stocks in the long portfolio, which is predicted for resolution predictors. The results indicate that between 41.89% and 56.76% of the predictors are resolution predictors. In my preferred specification (6), in which I control for the return, include year/month fixed effects and adjust for multiply hypothesis testing, 45.95% of mispricings are classified as resolution. The share of strategies with a significant positive coefficient on the long dummy is low at between 4.05% and 1.35%, suggesting that there are few predictors where the long portfolio is already underpriced relative to the short portfolio at the time of portfolio formation.

Table 4.8 shows the results of a regression of forecast error at the time of the next resorting on the long dummy. A negative long dummy indicates a build-up predictor. A positive long dummy indicates a predictor for which initial mispricing has not been

Table 4.7: Share of Resolution Predictors



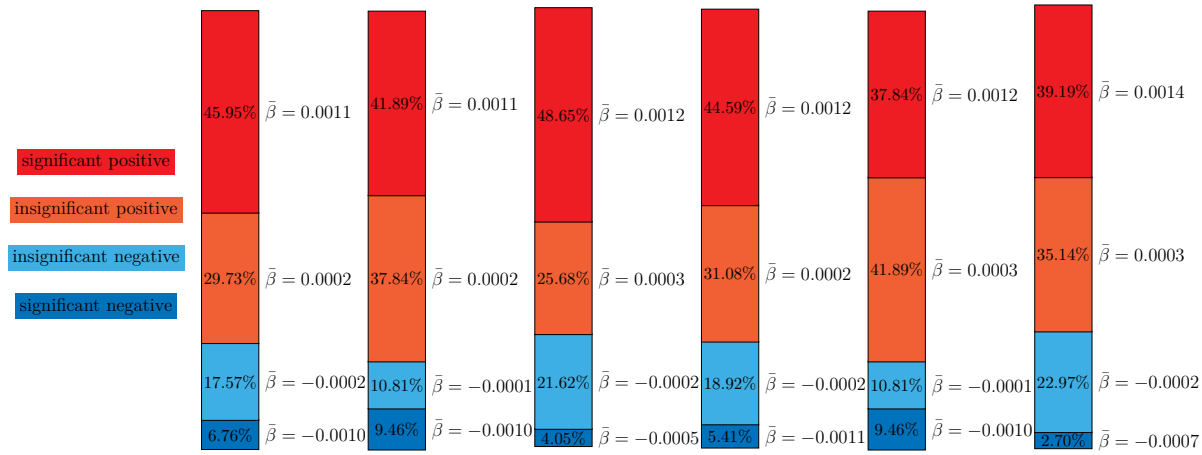
	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	61496	52356	52356	61496	52356	52356
Number of Regressions	74	74	74	74	74	74

The table shows the result of the regression $forecast\,error_{i,d,t} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t}$, where $forecast error_{i,d,t}$ indicates the consensus forecast error on the day of portfolio formation. The regression is run separately for all predictors identified as mispricing in the regression shown in Table 4.3 and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the 12 monthly returns before the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects.

fully resolved. The results suggest that there are few build-up predictors. Depending on the specification, only 2.7% to 9.46% of mispricings have a significant negative long dummy. Moreover, the initial mispricing is not resolved for a large share of predictors, suggesting that the resorting period may be shorter than necessary for at least some of the predictors, and more time is needed until the full effect of the predictor is realised. A potential reason for this is that around half of the significant mispricings are resorted monthly, leaving little time for mispricing to build up or resolve.

While the ideal approach to deal with this issue would be to examine forecast errors after the longest possible period for which the initial sort on the predictor characteristic predicts returns, determining this period for each predictor is beyond this project's scope. Instead, I examine errors of five quarters ahead forecasts made one year after the portfolio formation. Table 4.9 shows that around 15% of predictors still have a significantly positive

Table 4.8: Share of Build-up & Unresolved Predictors - Errors Before Resorting



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	28667	24501	24501	28667	24501	24501
Number of Regressions	74	74	74	74	74	74

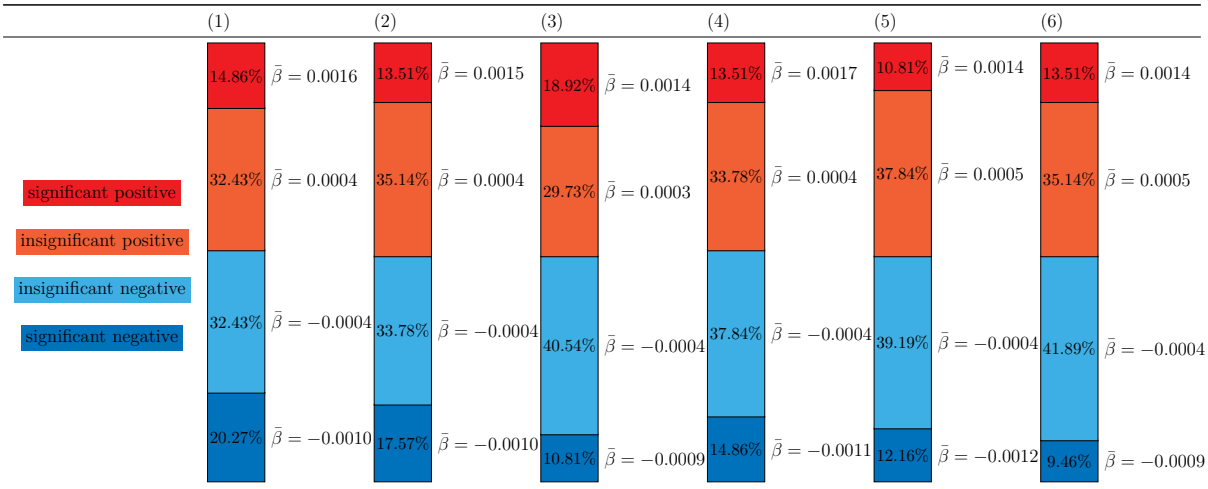
The table shows the result of the regression $forecast\,error_{i,d,t+x} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t}$, where $forecasterror_{i,d,t+x}$ indicates the consensus forecast error on the day of portfolio resorting. The regression is run separately for all predictors identified as mispricing in the regression shown in Table 4.3, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the 12 monthly returns before the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects.

long-dummy which suggests unresolved mispricing.

The share of predictors which have a significantly negative long dummy and are thus classified as build-up predictors varies between around 10% and around 20% between specifications and it is lowest in my preferred specification (6) with return controls, year/month fixed effects and MHT controls.

To test if one year is still a too short horizon for mispricing to build up, I repeat the analysis with forecasts made two years after portfolio formation (Table A4 in Appendix 4.C) and find that almost no predictors can be classified as build-up even without MHT controls. This suggests that the time over which mispricing is build-up is generally less than two years.

Table 4.9: Share of Build-up & Unresolved Predictors - Errors 1 Year After Portfolio Formation



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	24578	23275	23275	24578	23275	23275
Number of Regressions	74	74	74	74	74	74

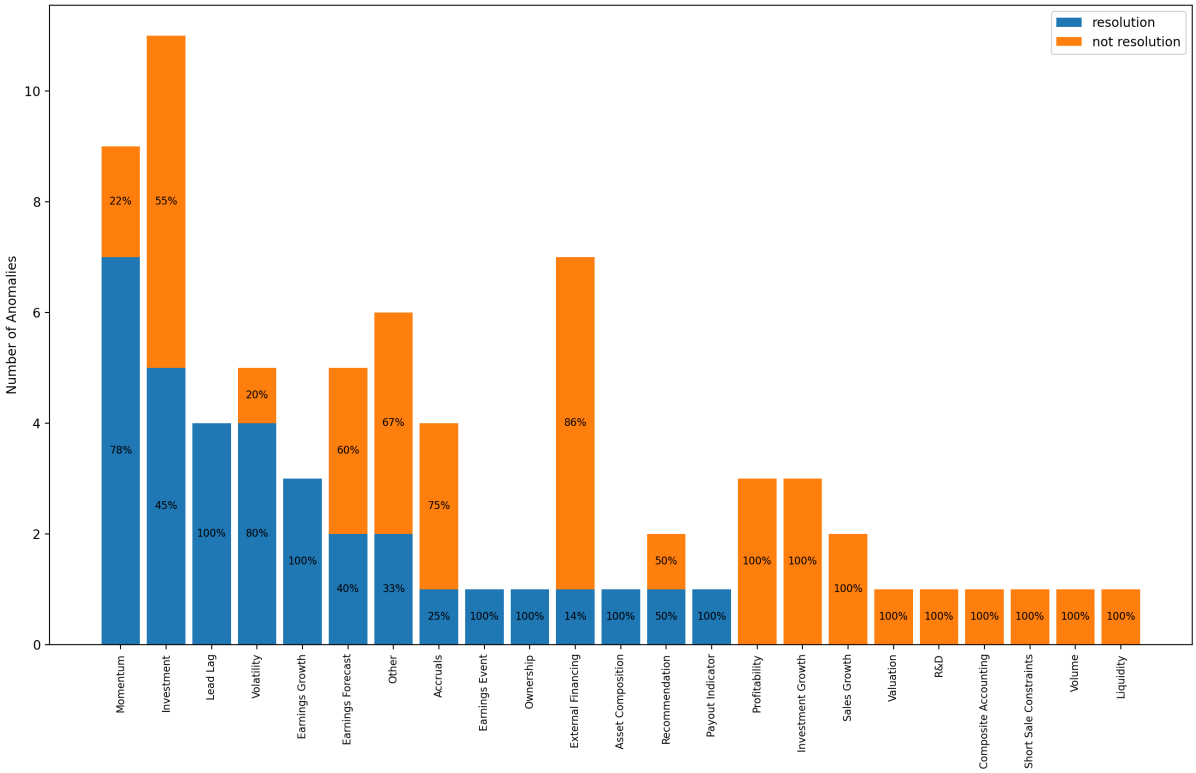
The table shows the result of the regression $forecast\ error_{i,d,t+12} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t}$, whereby $forecast\ error_{i,d,t+12}$ indicates the forecast error twelve months after portfolio formation. The regression is run for all predictors identified as mispricing in the regression shown in Table 4.3, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the return over the period between the formation of the predictor portfolio and the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects.

4.4.2 Which Predictors are Build-up and Which are Resolution?

Next, I separate the mispricing predictors by economic category using the classifications from A. Y. Chen and Zimmermann (2021) and examine each category's share of resolution predictors. I use the regression specification with return controls, fixed effects and MHT controls (Column 6 in Table 4.7) for this classification.

Reassuringly, all four *lead-lag* predictors, which aim to capture a delayed price response to information are classified as resolution. Moreover, 78% of *momentum* mispricings are resolution, suggesting that momentum returns are not delayed overreaction to news but instead underreaction to news. Another category with a relatively large share of resolution predictors is *investment*, with 45%. Finally, four out of five *volatility*

Figure 4.3: Share of Resolution Predictors by Category

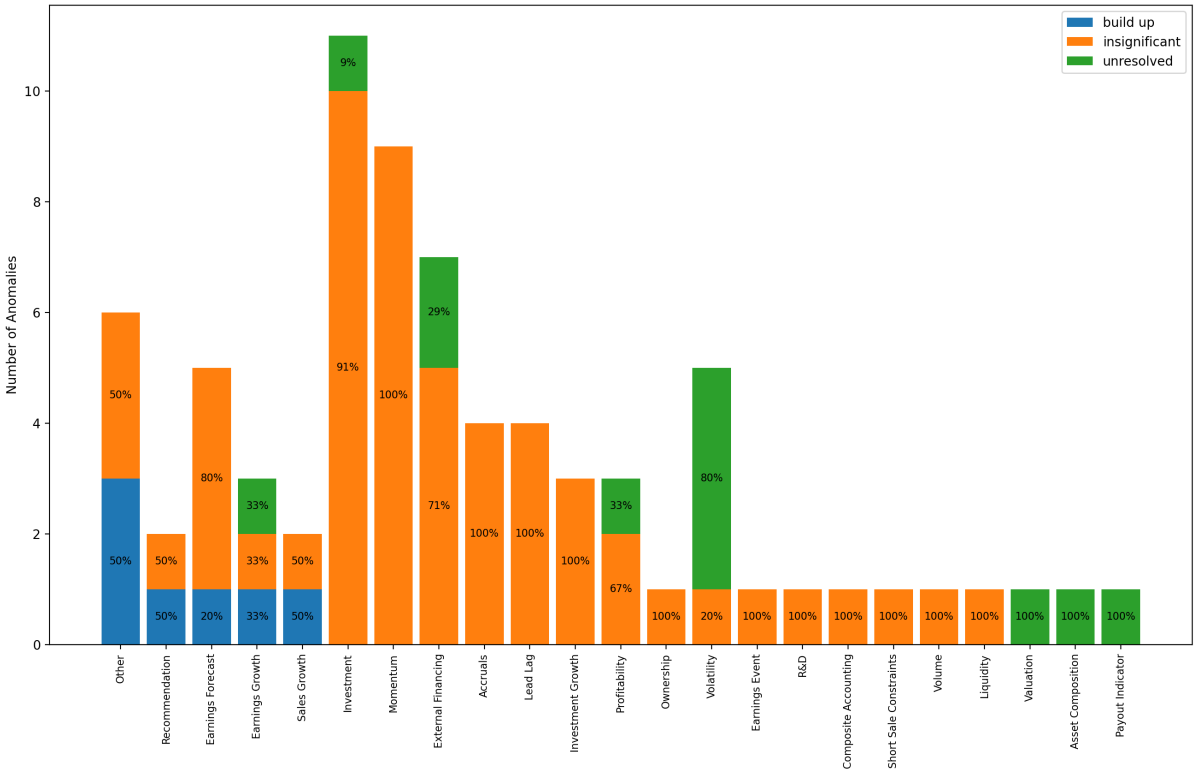


The figure shows the share of predictors classified as resolution by predictor category. The categories come from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification as resolution predictor is based on the regression of the forecast error at the time of portfolio formation on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Specification (6) from Table 4.7).

mispricings and all three *earnings growth* mispricings are resolution. The categories without any resolution predictors are *investment growth*, *profitability* and *sales growth* (as well as some categories with just one significant mispricing).

Figure 4.4 shows the share of build-up and unresolved predictors by category. There is no category with more than one build-up predictor except for the *other* category. Hence, it is hard to say which type of predictors are likely associated with the build-up of mispricing. However, an interesting finding is that the mispricing of the four *volatility* mispricings which were classified as resolution remains unresolved one year after the formation of the initial portfolio, suggesting that mispricing from volatility predictors may be exceptionally long-lasting.

Figure 4.4: Share of Build-up and Unresolved Predictors by Category



The figure shows the share of predictors classified as build-up and unresolved by predictor category. The categories come from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification is based on the regression of the forecast error one year after portfolio formation using forecasts for five quarters ahead on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Specification (6)) from Table 4.9).

4.5 Comparison to other Classification Methods

4.5.1 Mispricing Classification of the Original Authors

A. Y. Chen et al. (2022) hand-collect original authors' interpretation of whether the predictor they discovered represents mispricing or risk from the texts of the papers and find that "18% predictors are attributed to risk-based theory. 58% are attributed to mispricing and 24% have uncertain origins." They created a publicly available dataset including the predictor-level attributions, allowing me to compare them with my expectations-based classification. I find that 68.5% of mispricings according to my method (based on specification 6 in Table 4.3) were also attributed to mispricing by the original authors.

The original authors were agnostic for 17.1% of my mispricings, and they attributed 14.3% of them to risk. Among the predictors classified as risk factors by the original authors, 36.66% are classified as mispricing by my method, suggesting that several supposed risk factors are misclassified. This aligns with the results of A. Y. Chen et al. (2022), who find that supposed risk factors offer no better out-of-sample return predictability than supposed mispricings, even though, unlike mispricing, risk cannot be arbitraged away by trading post-publication. Together these results highlight the importance of formal classification methods in understanding the nature of a predictor.

4.5.2 Mispricing Classification of Holcblat et al. (2022)

Holcblat et al. (2022) classify predictors into mispricings and risk factors by testing if every risk-averse investor (defined as having a concave von Neumann-Morgenstern utility function) would prefer to invest in the predictor long over the predictor short portfolio. If they can reject this null hypothesis, at least some type of risk-averse investor would invest in the short portfolio. Therefore, it is possible that its risk profile can justify its lower returns. However, if the null hypothesis is true, no risk-averse investor would prefer the risk-return profile of the short portfolio over the long portfolio. Therefore, it must be mispriced.

Based on this argument, they classify a predictor as a **potential** risk factor if they reject the null hypothesis that every investor prefers the long over the short portfolio. Otherwise, they classify the predictor as mispricing. They provide an unconditional test and a test conditional on the market factor. Since I use unconditional portfolio sorts, the unconditional test is the correct comparison for my results.

Their test has the disadvantage that it can falsely classify a predictor as mispricing due to a lack of statistical power since the classification is based on failing to reject the null hypothesis. They show that their test asymptotically makes neither type one nor type two errors, so this would not be a concern with an infinite sample size. They examine the

finite sample properties with Monte Carlo simulations of normally distributed processes calibrated so that the null hypothesis is barely false. They find that, with more than 250 observations, almost all false hypotheses are rejected. However, if the two return processes are correlated, several false hypotheses are not rejected even with 500 data points. Therefore, their test may falsely classify variables as mispricing in finite samples. In contrast, my test classifies a variable as mispricing if I reject the null hypothesis that there is no difference in how forecasts were revised in the long compared to the short portfolio during the period for which the predictor predicts abnormal returns.

Both my test and the test by Holcblat et al. (2022) fail to detect certain types of mispricing. Their test does not classify a predictor as mispricing if a risk-averse utility function exists that rationalises a preference for the short over the long portfolio, regardless of whether any investor has such a utility function. In contrast, my test fails to detect mispricing from biased beliefs about future required returns rather than earnings. I also need to use analyst forecasts as a proxy for unobservable market expectations, which introduces noise. This highlights the complementarity of our tests.

I compare the classification of both tests and find that 66 of the 71 predictors that my test classifies as mispricing (using specification (6) of Table 4.3) are also mispricing according to their test, whereas only four are not classified as mispricing (*AbnormalAccruals*, *InvestPPEInv*, *Spinoff* & *TotalAccruals*) and one (*Recomm_ShortInterest*) is missing from their data. However, they classify 89 additional predictors as mispricing. This suggests that overall my test is more conservative. In Table A10 in Appendix 4.C, I show a predictor-level comparison between our tests for each predictor classified as mispricing by my test.

4.5.3 Build-up and Resolution Classification of van Binsbergen et al. (2023)

van Binsbergen et al. (2023) also classify predictors into build-up and resolution. Their method relies on measuring price wedges defined as the negative of the log of the fundamental value divided by the price. This means a negative price wedge implies underpricing, and a positive price wedge implies overpricing. The fundamental value is defined as the stream of future dividends discounted by a benchmark SDF. They estimate the price wedge at the portfolio level using 15 years of realised dividends. Their approach works for any factor-based SDF, and they use CAPM for their main result. A variable is considered to be a resolution predictor if the α and the long minus short price wedge have the opposite sign and a build-up predictor if they have the same sign. To determine the probability that a specific predictor portfolio is overpriced or underpriced, they use bootstrap simulations and study the fraction of simulated portfolios that are overpriced or underpriced. Because it is necessary to assume a specific SDF to apply the method, it is subject to the joint hypothesis problem. Thus, it may misclassify predictors if the wrong SDF is chosen.

van Binsbergen et al. (2023) do not provide a significance test for the long minus short price wedge, but they provide separate tests for the price wedges of the long and the short portfolio. I, therefore, consider a predictor to be a resolution (build-up) predictor if either the price wedge of the short portfolio is positive (negative) and significant or the price wedge of the long portfolio is negative (positive) and significant[11].

I cannot comprehensively compare our results as van Binsbergen et al. (2023) use a different set of predictors. However, there are 34 predictors from their sample for which I have the same or a closely related predictor in my sample. To keep coverage of as many variables as possible, I include variables from my sample even if they are not classified as mispricing in my first analysis step (Table 4.3. However, I still require a significant long minus short return using all returns up to the end of 2021, which reduces

[11]In theory it would be possible that a predictor is related to the build-up of mispricing in one leg of the sorted portfolio and the resolution of mispricing in the other leg, but in practice, this does not occur.

Figure 4.5: Comparison of expectation and price-wedge-based classifications

The figure shows the number of predictors classified as build-up/resolution by one method that is classified as build-up, resolution or neither by the other method.

the overlapping set of predictors to 25.

In general, the expectation-based and price-wedge-based approaches to classify predictors into build-up and resolution are complementary because they use different data and have orthogonal strengths and weaknesses. However, Figure 4.5 shows significant discrepancies between the two classification methods. While five predictors are classified as resolution by both methods, the expectations method classifies five additional predictors as resolution, and the price wedge method classifies six additional predictors as resolution. The disagreement is even more pronounced for build-up predictors, where not a single predictor is classified as build-up by both methods.

While it may seem like there is limited usefulness in combining both methods due to the large fraction of contradictions, this is potentially driven by the choice of CAPM as SDF. Future research using different SDFs could find less disagreement. Moreover, the comparison is based on just 25 predictors, and the overlap may be higher in a more complete sample of predictors.

4.6 Conclusion

Determining the degree of stock market efficiency is a fundamental question of finance with considerable implications for the efficiency of capital allocation a hence the real economy. Asset pricing researchers have discovered hundreds of characteristics that can predict the cross-section of stock returns, which challenges the efficient market hypothesis. However, this challenge crucially depends on whether these predictors represent mispricing or risk. My results suggest that for at least 40% of them, the predictable excess returns align with predictable changes in future earnings expectations suggesting that they represent mispricing. As my analysis does not capture mispricings stemming from biased beliefs about future required returns and uses an imperfect proxy for the market's future earnings expectations, the actual share of mispricings among the predictors is likely even higher. Moreover, my results suggest that the excess returns of some predictors capture the build-up rather than the resolution of mispricing, implying that traders who capitalize on these predictors worsen rather than correct mispricing. Overall my results suggest that the ample evidence for predictability of the cross-section of returns does not just mean that existing asset pricing models have not yet incorporated all relevant risk factors but that there is widespread mispricing in the stock market.

References

Adam, Klaus and Stefan Nagel (2023). "Chapter 16 - Expectations Data in Asset Pricing". In: *Handbook of Economic Expectations*. Ed. by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw. Academic Press, pp. 477–506. ISBN: 978-0-12-822927-9. DOI: 10.1016/B978-0-12-822927-9.00024-0.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang (2006). "The Cross-Section of Volatility and Expected Returns". In: *The Journal of Finance* 61.1, pp. 259–299. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2006.00836.x.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw (2011). "Maxing out: Stocks as Lotteries and the Cross-Section of Expected Returns". In: *Journal of Financial Economics* 99.2, pp. 427–446. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2010.08.014.

Ben-Rephael, Azi, Steffen Hitzemann, and Yuanyuan Xiao (2021). *Investor (Mis)Reaction, Biased Beliefs, and the Mispricing Cycle*. SSRN Scholarly Paper ID 3966758. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3966758.

Benjamini, Yoav and Yosef Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1, pp. 289–300. ISSN: 0035-9246.

Van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni (2023). "Dynamic Asset (Mis)Pricing: Build-up versus Resolution Anomalies". In: *Journal of Financial Economics* 147.2, pp. 406–431. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2022.11.005.

Van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira (2022). *Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases*. DOI: 10.2139/ssrn.3625279. preprint.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2022). *Belief Overreaction and Stock Market Puzzles*. preprint.

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar (2019). "Sticky Expectations and the Profitability Anomaly". In: *The Journal of Finance* 74.2, pp. 639–674. ISSN: 1540-6261. DOI: 10.1111/jofi.12734.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan (2006). "The Relation between Corporate Financing Activities, Analysts Forecasts and Stock Returns". In: *Journal of Accounting and Economics*. Conference Issue on Implications of Changing Financial Reporting Standards 42.1, pp. 53–85. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2006.03.004.

Campbell, John Y. (1991). "A Variance Decomposition for Stock Returns". In: *The Economic Journal* 101.405, pp. 157–179. ISSN: 0013-0133. DOI: 10.2307/2233809.

Campbell, John Y. and Robert J. Shiller (1988a). "Stock Prices, Earnings, and Expected Dividends". In: *The Journal of Finance* 43.3, pp. 661–676. ISSN: 0022-1082. DOI: 10.2307/2328190.

— (1988b). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". In: *The Review of Financial Studies* 1.3, pp. 195–228. ISSN: 0893-9454.

Carhart, Mark M. (1997). "On Persistence in Mutual Fund Performance". In: *The Journal of Finance* 52.1, pp. 57–82. ISSN: 0022-1082. DOI: 10.2307/2329556.

Charoenrook, Anchada Aida and Jennifer S. Conrad (2005). *Identifying Risk-Based Factors*. DOI: 10.2139/ssrn.757210. URL: https://papers.ssrn.com/abstract=757210 (visited on 09/15/2022). preprint.

Chen, Andrew Y., Alejandro Lopez-Lira, and Tom Zimmermann (2022). *Peer-Reviewed Theory Does Not Help Predict the Cross-section of Stock Returns*. DOI: 10.2139/ssrn.4308069. URL: https://papers.ssrn.com/abstract=4308069 (visited on 01/20/2023). preprint.

Chen, Andrew Y. and Tom Zimmermann (2021). *Open Source Cross-Sectional Asset Pricing*. SSRN Scholarly Paper 3604626. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3604626.

Chen, Long, Zhi Da, and Xinlei Zhao (2013). "What Drives Stock Price Movements?" In: *Review of Financial Studies* 26.4, pp. 841–876. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hht005.

Chordia, Tarun, Amit Goyal, and Alessio Saretto (2020). "Anomalies and False Rejections". In: *The Review of Financial Studies* 33.5, pp. 2134–2179. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa018.

De La O, Ricardo and Sean Myers (2021). "Subjective Cash Flow and Discount Rate Expectations". In: *The Journal of Finance* 76.3, pp. 1339–1387. ISSN: 1540-6261. DOI: 10.1111/jofi.13016.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina (2002). "ForecastDispersion - Differences of Opinion and the Cross Section of Stock Returns". In: *The Journal of Finance* 57.5, pp. 2113–2141. ISSN: 1540-6261. DOI: 10.1111/0022-1082.00490.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff (2020). "Analysts and Anomalies". In: *Journal of Accounting and Economics* 69.1, p. 101249. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2019.101249.

Engelberg, Joseph, R. David Mclean, and Jeffrey Pontiff (2018). "Anomalies and News". In: *The Journal of Finance* 73.5, pp. 1971–2001. ISSN: 1540-6261. DOI: 10.1111/jofi.12718.

Fama, Eugene F. and Kenneth R. French (1992). "The Cross-Section of Expected Stock Returns". In: *The Journal of Finance* 47.2, pp. 427–465. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1992.tb04398.x.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu (2020). "Taming the Factor Zoo: A Test of New Factors". In: *The Journal of Finance* 75.3, pp. 1327–1370. ISSN: 1540-6261. DOI: 10.1111/jofi.12883.

Gafka, Beata, Pavel G. Savor, and Mungo Ivor Wilson (2021). *Sources of Return Predictability*. DOI: 10.2139/ssrn.3937343. URL: https://papers.ssrn.com/abstract=3937343 (visited on 06/08/2023). preprint.

Giglio, Stefano, Yuan Liao, and Dacheng Xiu (2021). "Thousands of Alpha Tests". In: *The Review of Financial Studies* 34.7, pp. 3456–3496. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa111.

Grinblatt, Mark, Gergana Jostova, and Alexander Philipov (2018). "Analyst Bias and Mispricing". In: p. 49.

Guo, Li, Frank Weikai Li, and K. C. John Wei (2020). "Security Analysts and Capital Market Anomalies". In: *Journal of Financial Economics* 137.1, pp. 204–230. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2020.01.002.

Harvey, Campbell R. and Yan Liu (2021). "Lucky Factors". In: *Journal of Financial Economics* 141.2, pp. 413–435. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2021.04.014.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (2016). " and the Cross-Section of Expected Returns". In: *The Review of Financial Studies* 29.1, pp. 5–68. ISSN: 0893-9454. DOI: 10.1093/rfs/hhv059.

He, Ai, Dashan Huang, Jiaen Li, and Guofu Zhou (2022). "Shrinking Factor Dimension: A Reduced-Rank Approach". In: *Management Science*. ISSN: 0025-1909. DOI: 10.1287/mnsc.2022.4563.

Holcblat, Benjamin, Abraham Lioui, and Michael Weber (2022). *Anomaly or Possible Risk Factor? Simple-To-Use Tests*. DOI: 10.2139/ssrn.4064159. URL: https://papers.ssrn.com/abstract=4064159 (visited on 07/05/2022). preprint.

Hou, Kewei, Chen Xue, and Lu Zhang (2020). "Replicating Anomalies". In: *The Review of Financial Studies* 33.5, pp. 2019–2133. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy131.

Jackson, Andrew and Timothy Johnson (2006). "Unifying Underreaction Anomalies". In: *The Journal of Business* 79.1, pp. 75–114. ISSN: 0021-9398. DOI: 10.1086/497406.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen (2023). "Is There a Replication Crisis in Finance?" In: *The Journal of Finance* n/a.n/a. ISSN: 1540-6261. DOI: 10.1111/jofi.13249.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh (2018). "Interpreting Factor Models". In: *The Journal of Finance* 73.3, pp. 1183–1223. ISSN: 1540-6261. DOI: 10.1111/jofi.12612.

La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny (1997). "Good News for Value Stocks: Further Evidence on Market Efficiency". In: *The Journal of Finance* 52.2, pp. 859–874. ISSN: 0022-1082. DOI: 10.2307/2329502.

Lintner, John (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets". In: *The Review of Economics and Statistics* 47.1, pp. 13–37. ISSN: 0034-6535. DOI: 10.2307/1924119.

Pruitt, Seth (2023). *Dogs and Cats Living Together: A Defense of Cash-Flow Predictability.* DOI: 10.2139/ssrn.4343101. URL: https://papers.ssrn.com/abstract=4343101 (visited on 06/27/2023). preprint.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam (2019). "A Protocol for Factor Identification". In: *The Review of Financial Studies* 32.4, pp. 1573–1607. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy093.

Sharpe, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*". In: *The Journal of Finance* 19.3, pp. 425–442. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1964.tb02865.x.

Sloan, Richard G. (1996). "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?" In: *The Accounting Review* 71.3.

4.A Derivations of Equation 1 & 2

The log return of an asset between time t and time $t+1$ is given by the following equation, where P_t is the assets price at time t and D_{t+1} is the dividend paid between time t and $t+1$:

$$r_{t+1} = log(P_{t+1} + D_{t+1}) - log(P_t))$$

Based on this, Campbell and Shiller (1988a,b) showed that the following linear relationship between the log return the log dividend and the log price holds approximately:

$$r_{t+1} = k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t$$

In this equation, d_t is the log dividend paid at time t, and p_t is the log stock price. The parameter ρ is related to the price dividend ratio and is close to but smaller than 1, and k is a constant term. Each variable can refer to an individual stock or the average in a portfolio of stocks. Following Campbell (1991) this equation can be rewritten as follows:

$$r_{t+1} = k - \rho(d_{t+1} - p_{t+1}) + d_{t+1} - p_t$$

Defining $g_{t+1} = d_{t+1} - d_t$ and rearranging gives:

$$p_t - d_t = k + \rho(p_{t+1} - d_{t+1}) - r_{t+1} + g_{t+1}$$

Iterating the equation forward while imposing $\lim_{s\to\infty} \rho^s(d_{t+s} - p_{t+s}) = 0$ gives:

$$p_t - d_t = \frac{k}{1-\rho} + \sum_{s=0}^{\infty} \rho^s g_{t+1+s} - \sum_{s=0}^{\infty} \rho^s r_{t+1+s}$$

Bringing $p_t - d_t$ to the right side and r_{t+1} to the left side gives:

$$r_{t+1} = \frac{k}{1-\rho} + \sum_{s=0}^{\infty} \rho^s g_{t+s} - \sum_{s=1}^{\infty} \rho^s r_{t+s} - (p_t - d_t)$$

This holds in expectation

$$\mathbb{E}_t(r_{t+1}) = \frac{k}{1-\rho} + \mathbb{E}_t(\sum_{s=0}^{\infty} \rho^s g_{t+s}) - \mathbb{E}_t(\sum_{s=1}^{\infty} \rho^s r_{t+s}) + (d_t - p_t)$$

The realised return for a period is then the expected return plus the change in dividend expectations and minus the change in return expectations:

$$r_{t+1} = E_t(r_{t+1}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} \rho^s g_{t+1+s} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} \rho^s r_{t+1+s}$$

Campbell and Shiller (1988a,b) originally used the log linearisation to study the market portfolio. However, since it follows directly from the definition of the log return in a dividend discount model, it can be applied to other portfolios and individual stocks, as I do in Equation 4.1.

If it is applied to a portfolio, the (potentially weighted) average dividend and price of the portfolio need to be calculated before doing the log transformation since the average of a log is not the same as the log of the average. Notably, it is theoretically possible to aggregate the long-minus-short portfolio into a single asset by taking the average price (dividend) of the short portfolio and subtracting it from that of the long portfolio before applying the log-linearisation. However, this is impractical because it is not guaranteed that the resulting averages would all be positive which is necessary to take the log. Therefore, I average over the long and short portfolios separately and define:

$$d_t^L = log(\sum_{i \in L_t} (w_i \cdot D_{i,t}))$$

$$P_t^L = log(\sum_{i \in L_t} (w_i \cdot P_{i,t}))$$

In the equation above, L_t is the set of all stocks that are in the long portfolio at time t and w_i is the weight applied to a stock. Common weighting methods are equal weighting and value weighting. Applying the Cambell-Shiller log-linearisation and solving forward, as shown above, we can then write:

$$r_{t+1}^L = E_t(r_{t+1}^L) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^L)^s g_{t+1+s}^L - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^L)^s r_{t+1+s}^L$$

We can define expressions for S_t the set of all stocks in the short portfolio at time t accordingly. We can then write:

$$r_t^{LS} = r_t^L - r_t^S$$

$$g_t^{LS} = g_t^L - g_t^S$$

This gives us the equation for the the long-minus-short return (Equation 4.2):

$$r_{t+1}^{LS} \approx \mathbb{E}_t(r_{t+1}^{LS}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^{LS})^s g_{t+1+s}^{LS} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^{LS})^s r_{t+1+s}^{LS}$$

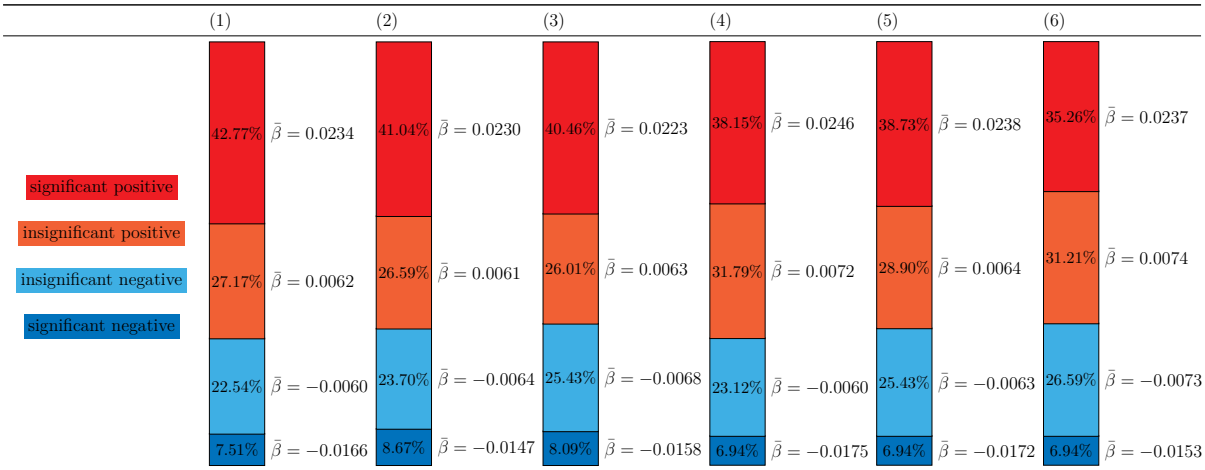
This equation requires a further approximation because ρ^L and ρ^S are not exactly the same, however, they are both close to 1 by construction so we can approximate them both by ρ^{LS}.

4.B Long-Term Growth Data

Table A1 shows the same regression analysis as Table 4.3, except that it uses long-term growth (LTG) forecasts from IBES, whereby LTG is defined as the growth rate over a complete business cycle. As for the quarterly forecast, I construct the forecast revisions by first calculating the consensus forecast for all firms in the predictor long and short portfolios in a given month and then subtracting it from the consensus LTG forecast made after the next portfolio resorting. Since LTG forecasts have no fixed target date, keeping the target date constant is impossible. Hence, the first and the second forecast may describe earnings growth rates over slightly different periods.

The share of firms with positive and negative coefficients is similar for quarterly and LTG forecasts. However, slightly fewer results are significant with LTG forecasts, likely due to worse data availability. Next, I study if the same predictors are classified as mispricing with both types of forecasts, focusing on the specification with return controls, fixed effects and MHT controls. Thirty-eight predictors have positive significant

Table A1: Share of Mispricings among the Predictors - Long-Term Growth Data



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	35460	35299	35299	35460	35299	35299
Number of Regressions	173	173	173	173	173	173

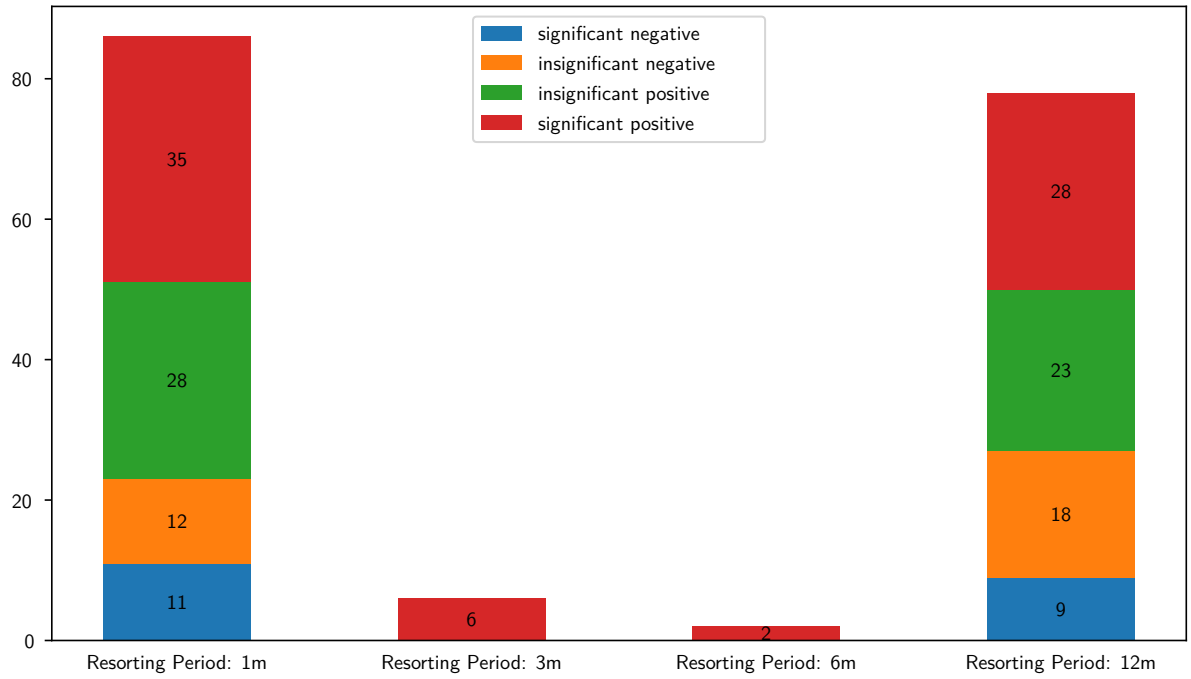
The table shows the result of the regression $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecastrevision_{i,d,t,x}$ is the revision of the consensus long-term earnings growth forecast. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

coefficients with both LTG and quarterly data. Twenty-three additional predictors have significant coefficients with LTG and only for one of those, the long-dummy is negatively significant using the quarterly forecast. Moreover, an additional 33 predictors have significant positive coefficients with quarterly forecasts, but only one has a significant negative long-dummy using LTG forecast.

The fact that some predictors are only significant for one of the two forecast horizons and that there are even a few cases with significant coefficients that go in opposite directions is not a challenge to the validity of my methodology. As Equation 4.1 shows, returns are not related to expectations for a specific horizon but to expectations about the sum of all future dividends. Therefore, a mispricing predictor may predict relatively higher returns and relatively more negative forecast revisions for a forecast horizon if it predicts more positive forecast revisions for other horizons. In contrast, risk factors can never predict earnings forecast revisions of any kind.

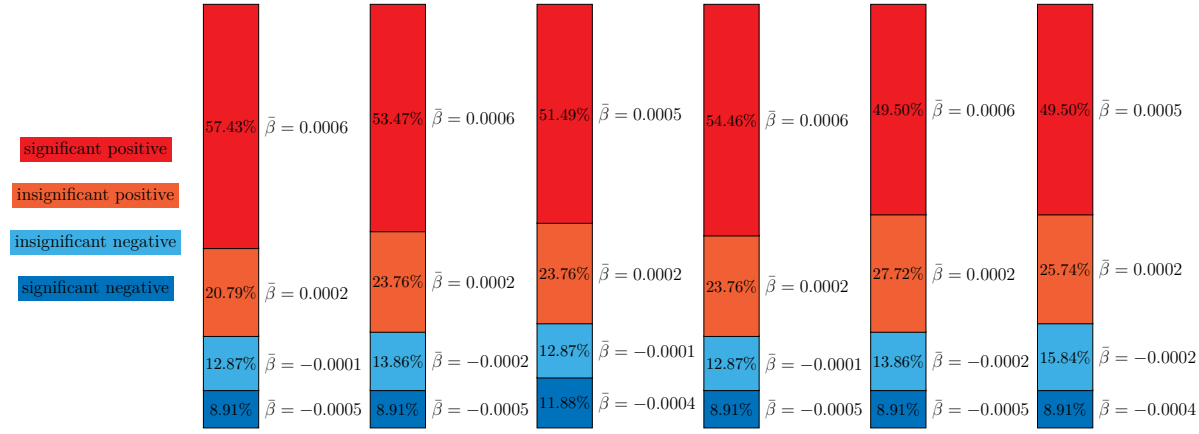
4.C Additional Figures & Tables

Figure A1: Share of Mispricings by Rebalancing Period



The figure shows the share of predictors that are classified as mispricing by rebalancing period. The classification as mispricing is based on the regression of the forecast revision on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Specification (6) from Table 4.3).

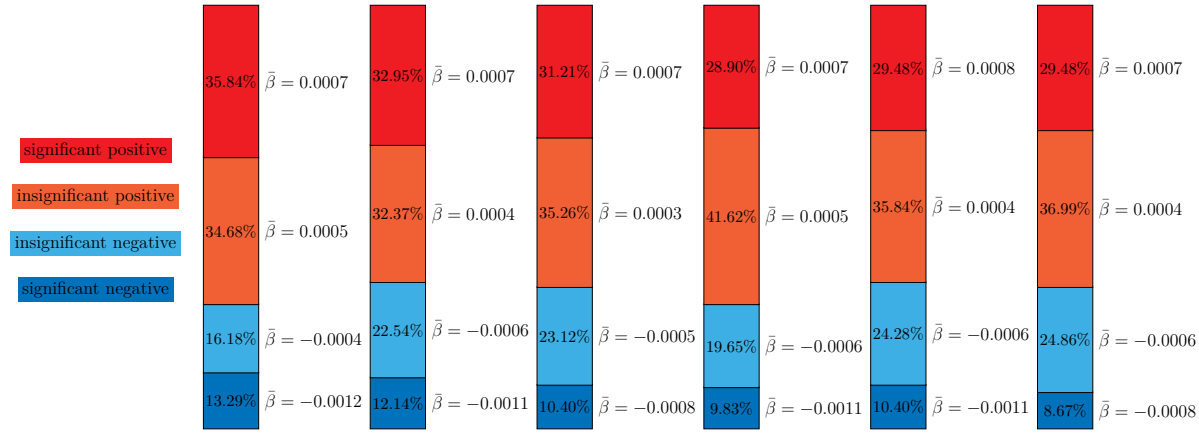
Table A2: Share of Mispricings among the Predictors - Using only Predictors with significant FDR controlled returns



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	55965	55492	55492	55965	55492	55492
Number of Regressions	101	101	101	101	101	101

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each of the 101 predictors with significant long minus short returns when controlling the false discovery rate to 5%. The table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

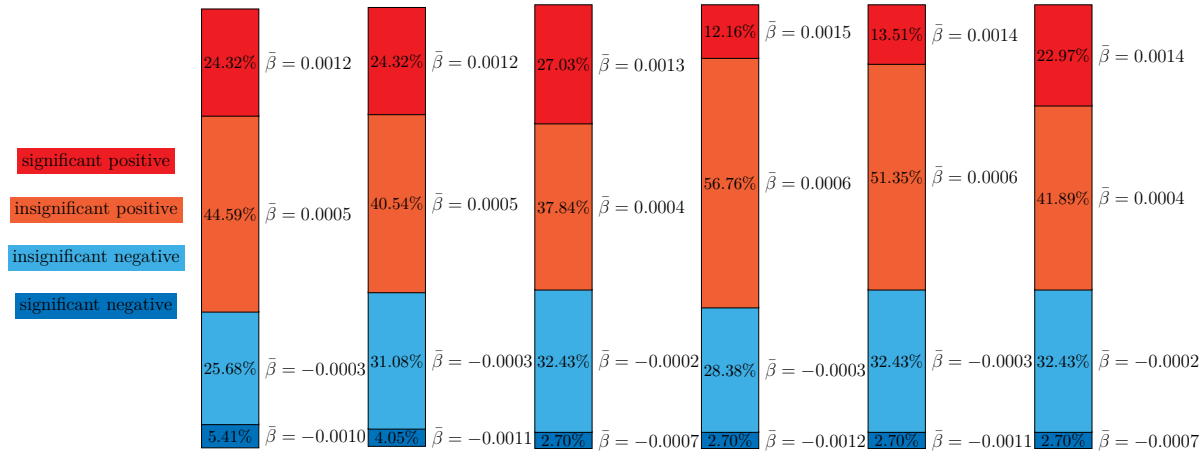
Table A3: Share of Mispricings among the Predictors - Using only 1 Data Point per Portfolio Resorting



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	31020	26545	26545	31020	26545	26545
Number of Regressions	173	173	173	173	173	173

The table shows the result of the regression $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\,revision$. $forecast\,revision_{i,d,t,x}$ is the revision of the consensus five quarter ahead earnings forecast between the time a portfolio is formed and the time of the next resorting. That is, the regression does not use the method to obtain forecast revisions at a monthly frequency described in Section 4.2.2 but has one data point per portfolio formation. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

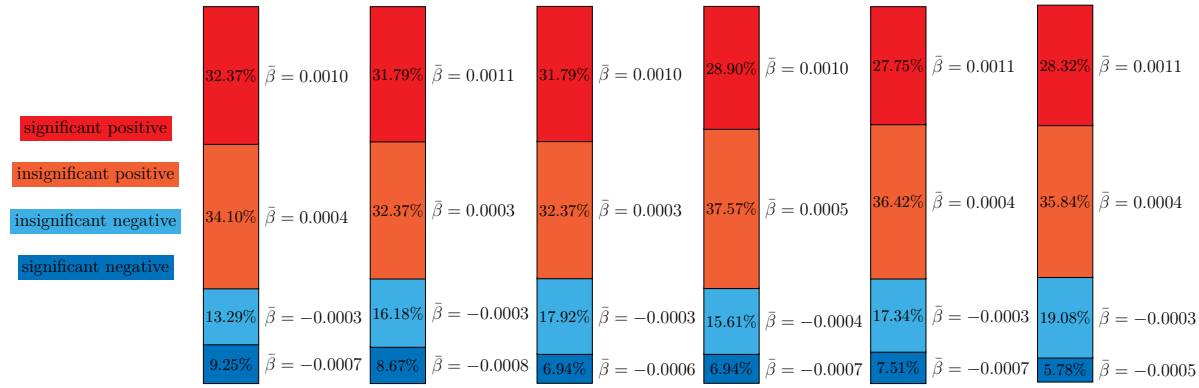
Table A4: Share of Build-up & Unresolved Predictors - Errors 2 Years After Portfolio Formation



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	21728	21186	21186	21728	21186	21186
Number of Regressions	74	74	74	74	74	74

The table shows the result of the regression $forecast\,error_{i,d,t+24} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t}$, whereby $forecast\,error_{i,d,t+24}$ indicates the forecast error 24 months after portfolio formation. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for all predictors identified as mispricing in the regression shown in Table 4.3, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the return over the period between the formation of the predictor portfolio and the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects.

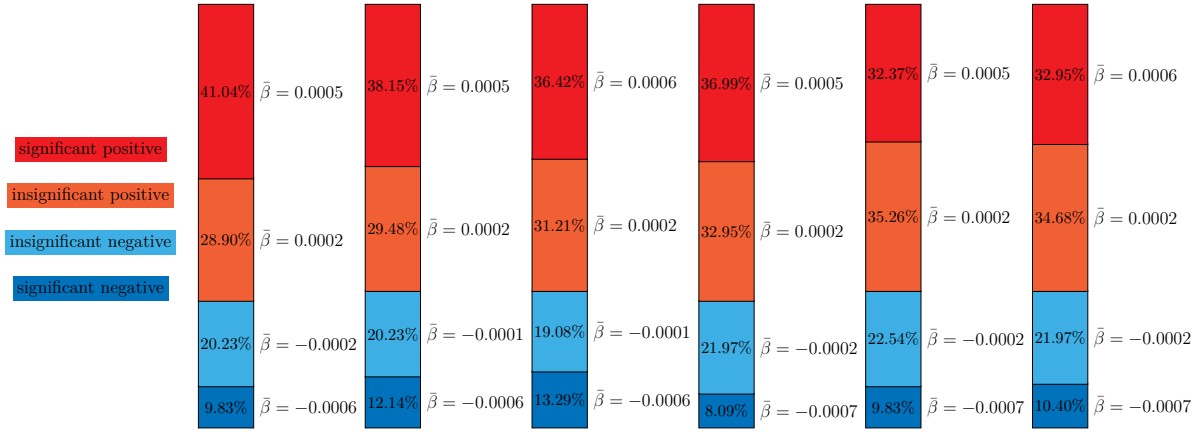
Table A5: Share of Mispricings among the Predictors - Pre-Publication Data



	(1)	(2)	(3)	(4)	(5)	(6)
significant positive	32.37% $\bar{\beta} = 0.0010$	31.79% $\bar{\beta} = 0.0011$	31.79% $\bar{\beta} = 0.0010$	28.90% $\bar{\beta} = 0.0010$	27.75% $\bar{\beta} = 0.0011$	28.32% $\bar{\beta} = 0.0011$
insignificant positive	34.10% $\bar{\beta} = 0.0004$	32.37% $\bar{\beta} = 0.0003$	32.37% $\bar{\beta} = 0.0003$	37.57% $\bar{\beta} = 0.0005$	36.42% $\bar{\beta} = 0.0004$	35.84% $\bar{\beta} = 0.0004$
insignificant negative	13.29% $\bar{\beta} = -0.0003$	16.18% $\bar{\beta} = -0.0003$	17.92% $\bar{\beta} = -0.0003$	15.61% $\bar{\beta} = -0.0004$	17.34% $\bar{\beta} = -0.0003$	19.08% $\bar{\beta} = -0.0003$
significant negative	9.25% $\bar{\beta} = -0.0007$	8.67% $\bar{\beta} = -0.0008$	6.94% $\bar{\beta} = -0.0006$	6.94% $\bar{\beta} = -0.0007$	7.51% $\bar{\beta} = -0.0007$	5.78% $\bar{\beta} = -0.0005$
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	16394	16393	16393	16394	16393	16393
Number of Regressions	173	173	173	173	173	173

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. *forecast revision*. $forecast\,revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

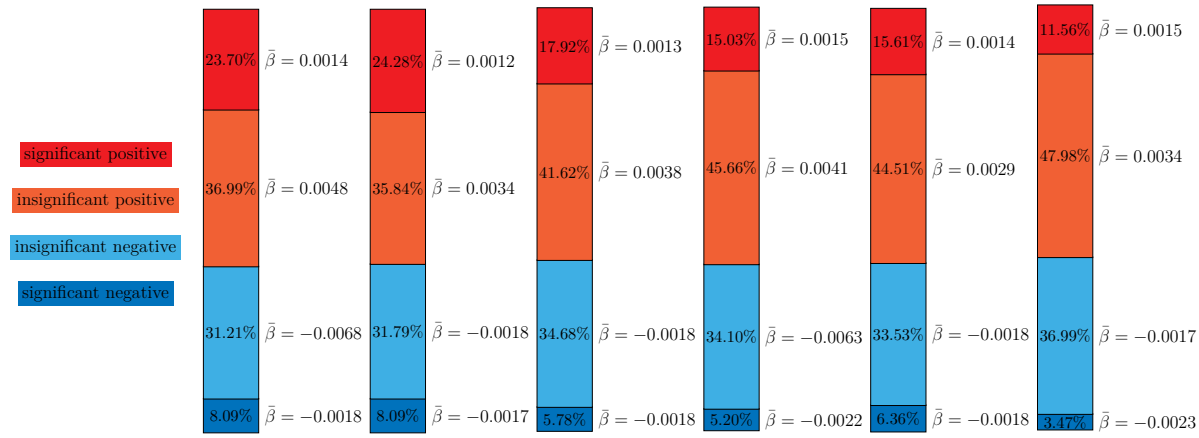
Table A6: Share of Mispricings among the Predictors - Post-Publication Data



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	38729	38274	38274	38729	38274	38274
Number of Regressions	173	173	173	173	173	173

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

Table A7: Share of Mispricings among the Predictors - Without Winsorization



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	53322	52867	52867	53322	52867	52867
Number of Regressions	173	173	173	173	173	173

The table shows the result of the following regression: $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 4.2.2. Unlike in the main analysis, the forecast revisions were not winsorized for this analysis. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects.

Table A8: Significant Mispricings with Categories

name	t value	parameter	standart error	Description	Economic Category
AnnouncementReturn	18.42	0.0006	0.000032	Earnings announcement return	Earnings Event
AnalystRevision	17.45	0.0007	0.000038	EPS forecast revision	Earnings Forecast
ResidualMomentum	15.25	0.0008	0.000055	Momentum based on FF3 residuals	Momentum
EarningsStreak	14.95	0.0007	0.000047	Earnings surprise streak	Earnings Growth
Mom6mJunk	14.55	0.0013	0.000088	Junk Stock Momentum	Momentum
Mom12mOffSeason	13.63	0.0014	0.000100	Momentum without the seasonal part	Other
Mom6m	13.27	0.0013	0.000096	Momentum (6 month)	Momentum
EarningsForecastDisparity	12.42	0.0006	0.000045	Long-vs-short EPS forecasts	Earnings Forecast
DelBreadth	11.47	0.0009	0.000075	Breadth of ownership	Ownership
FirmAgeMom	9.70	0.0010	0.000104	Firm Age - Momentum	Momentum
NumEarnIncrease	9.57	0.0002	0.000024	Earnings streak length	Earnings Growth
ForecastDispersion	9.54	0.0005	0.000051	EPS Forecast Dispersion	Volatility
MomRev	9.21	0.0011	0.000122	Momentum and LT Reversal	Momentum
MomVol	8.71	0.0018	0.000211	Momentum in high volume stocks	Momentum
Mom12m	8.60	0.0010	0.000119	Momentum (12 month)	Momentum
IdioVol3F	8.40	0.0006	0.000074	Idiosyncratic risk (3 factor)	Volatility
REV6	8.34	0.0008	0.000099	Earnings forecast revisions	Earnings Forecast
IdioRisk	8.30	0.0006	0.000077	Idiosyncratic risk	Volatility
FEPS	7.80	0.0004	0.000056	Analyst earnings per share	Profitability
RevenueSurprise	7.56	0.0003	0.000033	Revenue Surprise	Sales Growth
CompEquIss	7.11	0.0003	0.000037	Composite equity issuance	External Financing
ChTax	7.11	0.0004	0.000060	Change in Taxes	Other
EarningsSurprise	7.07	0.0003	0.000046	Earnings Surprise	Earnings Growth
InvestPPEInv	6.98	0.0002	0.000027	change in ppe and inv/assets	Investment
DelEqu	6.87	0.0008	0.000117	Change in equity to assets	Investment
dNoa	6.32	0.0002	0.000036	change in net operating assets	Investment
IntMom	6.01	0.0004	0.000062	Intermediate Momentum	Momentum
TotalAccruals	5.64	0.0006	0.000115	Total accruals	Investment Alt
IndRetBig	5.57	0.0003	0.000056	Industry return of big firms	Lead Lag
ChEQ	5.42	0.0006	0.000115	Growth in book equity	Investment
retConglomerate	5.38	0.0004	0.000077	Conglomerate return	Lead Lag
NetEquityFinance	5.20	0.0007	0.000139	Net equity financing	External Financing
DelCOA	5.16	0.0006	0.000120	Change in current operating assets	Investment Alt
ShareIss1Y	5.00	0.0005	0.000103	Share issuance (1 year)	External Financing
ShareRepurchase	4.99	0.0003	0.000069	Share repurchases	Payout Indicator
hire	4.86	0.0006	0.000117	Employment growth	Investment Alt
PctAcc	4.46	0.0007	0.000157	Percent Operating Accruals	Accruals
AssetGrowth	4.18	0.0008	0.000191	Asset growth	Investment
VolumeTrend	4.13	0.0005	0.000124	Volume Trend	Volume
IndMom	4.08	0.0002	0.000059	Industry Momentum	Momentum
MeanRankRevGrowth	4.00	0.0006	0.000152	Revenue Growth Rank	Sales Growth
zerotradeAlt1	3.90	0.0010	0.000249	Days with zero trades	Liquidity
NOA	3.81	0.0002	0.000055	Net Operating Assets	Asset Composition
grcapx	3.58	0.0005	0.000128	Change in capex (two years)	Investment Growth
DelFINL	3.57	0.0003	0.000096	Change in financial liabilities	External Financing
CompositeDebtIssuance	3.56	0.0004	0.000100	Composite debt issuance	External Financing
MaxRet	3.37	0.0003	0.000100	Maximum return over month	Volatility
XFIN	3.35	0.0006	0.000167	Net external financing	External Financing
ConsRecomm	3.35	0.0002	0.000057	Consensus Recommendation	Recommendation
PctTotAcc	3.30	0.0005	0.000147	Percent Total Accruals	Accruals
grcapx3y	3.21	0.0004	0.000129	Change in capex (three years)	Investment Growth
EarnSupBig	3.15	0.0002	0.000063	Earnings surprise of big firms	Lead Lag
ChInvIA	3.11	0.0004	0.000116	Change in capital inv (ind adj)	Investment Growth
roaq	3.10	0.0002	0.000067	Return on assets (qtrly)	Profitability
CredRatDG	3.07	0.0003	0.000081	Credit Rating Downgrade	Other
InvGrowth	2.95	0.0006	0.000213	Inventory Growth	Profitability
NetPayoutYield	2.88	0.0007	0.000249	Net Payout Yield	Valuation
Spinoff	2.74	0.0002	0.000059	Spinoffs	Other
ShareIss5Y	2.65	0.0003	0.000111	Share issuance (5 year)	External Financing
AbnormalAccruals	2.58	0.0004	0.000174	Abnormal Accruals	Accruals
MomOffSeason	2.57	0.0002	0.000087	Off season long-term reversal	Other
Recomm_ShortInterest	2.55	0.0004	0.000147	Analyst Recommendations and Short-Interest	Recommendation
CustomerMomentum	2.49	0.0002	0.000077	Customer momentum	Lead Lag
Accruals	2.43	0.0004	0.000157	Accruals	Accruals
ChInv	2.41	0.0004	0.000162	Inventory Growth	Investment Alt
OrgCap	2.39	0.0005	0.000189	Organizational capital	R&D
ChNNCOA	2.35	0.0002	0.000101	Change in Net Noncurrent Op Assets	Investment Alt
MomSeasonShort	2.34	0.0002	0.000081	Return seasonality last year	Other
DownRecomm	2.33	0.0001	0.000022	Down forecast EPS	Earnings Forecast
betaVIX	2.33	0.0001	0.000030	Systematic volatility	Volatility
ShortInterest	2.27	0.0002	0.000068	Short Interest	Short Sale Constraints
RDS	2.14	0.0002	0.000073	Real dirty surplus	Composite Accounting
ChNWC	2.04	0.0002	0.000094	Change in Net Working Capital	Investment Alt
ChForecastAccrual	2.03	0.0002	0.000106	Change in Forecast and Accrual	Earnings Forecast

The table shows the significant mispricings based on the regression of forecast revision on a long dummy with return controls and year/month fixed effects which is specification (3) of Table 4.3, as well as the economic category they belong to based on the classification by A. Y. Chen and Zimmermann (2021).

Table A9: Build-up and Resolution Predictors with Category

name	Economic Category	T Value Start	Significant Fdr Control Start	T Value 1 Year After	Significant Fdr Control 1 Year After
AbnormalAccruals	Accruals	2.33	No	0.60	No
Accruals	Accruals	1.75	No	-0.83	No
PctTotAcc	Accruals	1.24	No	1.09	No
PctAcc	Accruals	0.84	No	0.29	No
NOA	Asset Composition	7.98	Yes	3.96	Yes
RDS	Composite Accounting	-1.90	No	-1.42	No
AnnouncementReturn	Earnings Event	6.26	Yes	1.26	No
ChForecastAccrual	Earnings Forecast	-1.25	No	0.29	No
AnalystRevision	Earnings Forecast	15.64	Yes	-0.67	No
DownRecomm	Earnings Forecast	0.30	No	-1.67	No
REV6	Earnings Forecast	12.54	Yes	0.97	No
EarningsForecastDisparity	Earnings Forecast	0.56	No	-6.29	Yes
NumEarnIncrease	Earnings Growth	3.91	Yes	-0.53	No
EarningsSurprise	Earnings Growth	5.67	Yes	-2.77	Yes
EarningsStreak	Earnings Growth	14.09	Yes	3.27	Yes
ShareIss5Y	External Financing	0.85	No	1.07	No
CompEquIss	External Financing	1.58	No	-0.86	No
CompositeDebtIssuance	External Financing	0.38	No	1.52	No
ShareIss1Y	External Financing	1.17	No	2.01	No
NetEquityFinance	External Financing	2.28	No	2.91	Yes
XFIN	External Financing	0.69	No	3.02	Yes
DelFINL	External Financing	-0.18	No	0.11	No
InvestPPEInv	Investment	8.31	Yes	3.09	Yes
DelEqu	Investment	2.73	Yes	-0.92	No
AssetGrowth	Investment	0.70	No	-1.13	No
ChEQ	Investment	2.04	No	-0.98	No
dNoa	Investment	6.83	Yes	1.37	No
TotalAccruals	Investment Alt	2.80	Yes	-1.48	No
hire	Investment Alt	1.75	No	-0.09	No
ChInv	Investment Alt	1.12	No	-1.22	No
DelCOA	Investment Alt	3.90	Yes	-0.12	No
ChNNCOA	Investment Alt	0.80	No	0.88	No
ChNWC	Investment Alt	1.34	No	0.68	No
grcapx3y	Investment Growth	1.63	No	0.04	No
ChInvIA	Investment Growth	0.77	No	-1.62	No
grcapx	Investment Growth	1.25	No	-0.37	No
EarnSupBig	Lead Lag	2.62	Yes	-1.75	No
CustomerMomentum	Lead Lag	3.19	Yes	-0.57	No
retConglomerate	Lead Lag	2.68	Yes	0.33	No
IndRetBig	Lead Lag	2.74	Yes	-1.28	No
zerotradeAlt1	Liquidity	0.84	No	0.48	No
FirmAgeMom	Momentum	8.84	Yes	0.46	No
ResidualMomentum	Momentum	11.00	Yes	-0.50	No
IntMom	Momentum	3.18	Yes	-0.41	No
MomVol	Momentum	3.79	Yes	-0.53	No
MomRev	Momentum	1.88	No	-1.20	No
IndMom	Momentum	-0.14	No	-1.84	No
Mom6m	Momentum	5.57	Yes	-0.50	No
Mom12m	Momentum	5.12	Yes	-0.16	No
Mom6mJunk	Momentum	9.64	Yes	-0.60	No
MomOffSeason	Other	3.91	Yes	1.14	No
MomSeasonShort	Other	2.00	No	-1.68	No
Mom12mOffSeason	Other	7.17	Yes	-3.33	Yes
Spinoff	Other	2.19	No	0.48	No
ChTax	Other	0.01	No	-3.63	Yes
CredRatDG	Other	-1.54	No	-3.67	Yes
DelBreadth	Ownership	3.16	Yes	-2.43	No
ShareRepurchase	Payout Indicator	2.38	No	7.56	Yes
roaq	Profitability	-1.50	No	1.36	No
InvGrowth	Profitability	1.42	No	-0.32	No
FEPS	Profitability	0.52	No	2.40	No
OrgCap	R&D	0.30	No	-1.26	No
Recomm_ShortInterest	Recommendation	2.91	Yes	2.23	No
ConsRecomm	Recommendation	-4.67	Yes	-4.15	Yes
MeanRankRevGrowth	Sales Growth	0.30	No	1.39	No
RevenueSurprise	Sales Growth	1.65	No	-3.13	Yes
ShortInterest	Short Sale Constraints	-0.36	No	0.26	No
NetPayoutYield	Valuation	2.05	No	2.41	No
ForecastDispersion	Volatility	11.79	Yes	5.85	Yes
IdioRisk	Volatility	17.53	Yes	9.26	Yes
IdioVol3F	Volatility	17.38	Yes	9.62	Yes
betaVIX	Volatility	0.88	No	-1.77	No
MaxRet	Volatility	9.68	Yes	4.94	Yes
VolumeTrend	Volume	0.99	No	0.67	No

The table shows the results of a regression of the forecast errors of five quarters ahead forecasts made at the time of portfolio formation (Table 4.7 and the results of a regression of the forecast error of five quarters ahead forecasts made one year after portfolio formation (Table 4.9) on a long dummy, for each predictor that is classified as a mispricing based on the regression of forecast revision on a long-dummy (Table 4.3), as well as the economic category they belong to based on the classification by A. Y. Chen and Zimmermann (2021).

Table A10: Significant Mispricings - Comparison to Holcblat et al. (2022)

Name	P-value Expectations	P-value Risk
AbnormalAccruals	p=0.010	missing
Accruals	p=0.015	p=1.000
AnalystRevision	p<0.001	p=1.000
AnnouncementReturn	p<0.001	p=1.000
AssetGrowth	p<0.001	p=1.000
betaVIX	p=0.020	p=0.400
ChEQ	p<0.001	p=0.640
ChInv	p=0.016	p=1.000
ChInvIA	p=0.002	p=0.310
ChNNCOA	p=0.019	p=1.000
ChTax	p<0.001	p=0.580
CompEquIss	p<0.001	p=1.000
CompositeDebtIssuance	p<0.001	p=1.000
ConsRecomm	p<0.001	p=1.000
CredRatDG	p=0.002	p=1.000
CustomerMomentum	p=0.013	p=0.270
DelBreadth	p<0.001	p=0.590
DelCOA	p<0.001	p=1.000
DelEqu	p<0.001	p=0.730
DelFINL	p<0.001	p=1.000
dNoa	p<0.001	p=1.000
DownRecomm	p=0.020	p=1.000
EarningsForecastDisparity	p<0.001	p=0.370
EarningsStreak	p<0.001	p=1.000
EarningsSurprise	p<0.001	p=0.470
EarnSupBig	p=0.002	p=0.530
FEPS	p<0.001	p=1.000
FirmAgeMom	p<0.001	p=1.000
ForecastDispersion	p<0.001	p=1.000
grcapx	p<0.001	p=1.000
grcapx3y	p=0.001	p=0.640
hire	p<0.001	p=1.000
IdioRisk	p<0.001	p=1.000
IdioVol3F	p<0.001	p=1.000
IndMom	p<0.001	p=0.250
IndRetBig	p<0.001	p=1.000
IntMom	p<0.001	p=0.490
InvestPPEInv	p<0.001	missing
InvGrowth	p=0.003	p=1.000
MaxRet	p<0.001	p=0.420
MeanRankRevGrowth	p<0.001	p=0.500
Mom12m	p<0.001	p=1.000
Mom12mOffSeason	p<0.001	p=1.000
Mom6m	p<0.001	p=1.000
Mom6mJunk	p<0.001	p=0.530
MomOffSeason	p=0.010	p=0.390
MomRev	p<0.001	p=1.000
MomSeasonShort	p=0.019	p=1.000
MomVol	p<0.001	p=1.000
NetEquityFinance	p<0.001	p=1.000
NetPayoutYield	p=0.004	p=1.000
NOA	p<0.001	p=1.000
NumEarnIncrease	p<0.001	p=1.000
OrgCap	p=0.017	p=1.000
PctAcc	p<0.001	p=0.460
PctTotAcc	p<0.001	p=0.490
RecommShortInterest	p=0.011	missing
ResidualMomentum	p<0.001	p=1.000
retConglomerate	p<0.001	p=1.000
REV6	p<0.001	p=1.000
RevenueSurprise	p<0.001	p=0.300
roaq	p=0.002	p=1.000
ShareIss1Y	p<0.001	p=1.000
ShareIss5Y	p=0.008	p=0.360
ShareRepurchase	p<0.001	p=1.000
ShortInterest	p=0.023	p=1.000
Spinoff	p=0.006	missing
TotalAccruals	p<0.001	missing
VolumeTrend	p<0.001	p=1.000
XFIN	p<0.001	p=1.000
zerotradeAlt1	p<0.001	p=1.000

The table compares the expectation-based mispricing specification from this paper and the risk-based classification by Holcblat et al. (2022). The Column "P-value Expectations" shows the results of the regression shown in Equation 4.3 using specification (3) from Table 4.3. A low p-value indicates that the predictor is a mispricing. The column "P-value Risk" shows the result of their unconditional test (Table A.11 of the March 2022 version of their working paper). A high p-value indicates that the predictor is a mispricing.

4.D Detailed Predictor Definitions

This section contains more detailed variable definitions for the predictors mentioned by name in the main part of the paper. Definitions come from the Open Source Asset Pricing (OpenAP) data documentation by A. Y. Chen and Zimmermann (2021) and are in some cases augmented with details from the original paper or highlight differences between the OpenAP definition and the one in the original paper. The OpenAp definitions are in quotation marks. Anything not in quotation marks is additional information..

AbnormalAccruals "Define Accruals as net income (ib) minus operating cash flow (oancf), divided by average total assets (at) for years t-1 and t. If oancf is missing, replace operating cash flow with funds from operations (fopt) minus the annual change in total current assets (act) plus the annual change in cash and short-term investments (che) plus the annual change in current liabilities (lct) minus the annual change in debt in current liabilities (dlc). For each year t and 2-digit sic code, regress Accruals on: the inverse of average total assets for year t-1, the change in revenue (sale) from year t-1 to t divided by total assets for t-1, property plant and equipment (ppegt) divided by total assets for t-1. AbnormalAccrual is the residual from this cross-sectional regression." Original paper by Xie (2001).

Accruals "Annual change in current total assets (act) minus annual change in cash and short-term investments (che) minus annual change in current liabilities (lct) minus annual change in debt in current liabilities (dlc) minus change in income taxes (txp). All divided by average total assets (at) over this year and last year. Exclude if abs(prc) $<$ 5." Originally by Sloan (1996), who interprets this predictor as resolution of mispricing.

AdExp "Advertising expense (xad) over market value of equity (shrout*abs(prc))" originally from Chan et al. (2001).

AnalystRevision - EPS forecast revision "keep fpi == 1, last obs each month. Signal is meanest / last month's meanest." Take stocks that are followed by three or more analysts that have shown the greatest increase in their mean earnings forecasts since the prior month. The horizon of the forecasts is not specified in the paper. Original paper by Hawkins et al. (1984).

AnnouncementReturn - Earnings announcement return "Get announcement date for quarterly earnings from IBES (fpi = 6). AnnouncementReturn is the sum of (ret - mktrf + rf) from one day before an earnings announcement to 2 days after the announcement." Stocks with positive AnnouncementReturn are in the long portfolio. Originally by Chan et al. (1996).

AssetGrowth "Annual growth rate of total assets (at)". Firms that grow less are in the long portfolio. Originally by Cooper et al. (2008).

BetaLiquidityPS "Monthly excess return (ret -rf) regressed on innovations in liquidity from Pastor's website (`https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt`). Use 60 month rolling window regression, and require at least 36 non-missing observations." Originally from Pástor and Stambaugh (2003)

BMdec "BM using most recent December value of market equity." Originally from Fama and French (1992).

Hire - Employment growth "Change in number of employees (emp) between t -1 and t, scaled by average number of employees in t-1 and t. Replace hire with 0 if emp or lagged emp is missing. Firms with fewer hires are in the long portfolio." The original paper by Belo et al. (2014) has a model and offers a risk-based explanation for this effect.

ChEQ - Growth in book equity "Ratio of book equity (ceq) to book equity in the previous year. Include only if book equity is positive this year and last year." Originally by Lockwood and Prombutr (2010).

ChInv - Inventory Growth "12 month change in inventory (invt) divided by average total assets." Originally by Thomas and Zhang (2002). Change in inventory is a component of Accruals.

ChInvIA - Change in capital inv (ind adj) "Growth in capital expenditure (capx) minus average growth in capital expenditure in the same industry (two-digit SIC). If capx is missing, capital expenditure is defined as the annual change in property, plant and equipment (ppent). Capital expenditure growth is defined as the percentage growth of capx today relative to the average capx over the previous two years (.5*(capx$_{t-1}$ + capx$_{t-2}$), or as percentage growth relative to the previous year only if t-2 is missing." Stocks with low growth in capital expenditure are in the long portfolio. Originally by Abarbanell and Bushee (1998), who use a regression and initially start with the opposite hypothesis that more investment should mean higher returns.

ChTax - Change in Taxes "4-quarter change in quarterly total taxes (txtq), scaled by lagged total assets (at)." Stocks with an increase in tax expense are in the long portfolio. Originally by Thomas and Zhang (2002), whose interpretation is that an increase in tax, while it may seem bad, is a proxy for increased core profitability to which the market underreacts.

CompEquIss - Composite Equity Issuance "5 year growth rate of market value of equity minus 5 year stock return." The long portfolio contains the stocks with low Composite Equity Issuance. Originally by Daniel and Titman (2006).

DelCOA - Change in current operating assets "Difference in current operating assets (total current assets (act) minus cash and short-term investments (che)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005). The long portfolios are the stocks with a low difference in operating assets.

DelBreadth - Breadth of ownership "Quarterly change in the number of institutional owners (numinstowners) from 13F data. Exclude if in the lowest quintile of stocks by market value of equity (based on NYSE stocks only)." Stocks with positive changes in owners are in the long portfolio. Originally from J. Chen et al. (2002).

DelEqu - Change in equity to assets "Difference in book equity (ceq) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005).

DelFINL - Change in financial liabilities "Difference in financial liabilities (sum of long-term debt (dltt), current liabilities (dlc) and preferred stock (pstk)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005). They do Fama-MacBeth regression and find a significant positive effect on return if only the variable is included and a negative (insignificant) if it is included together with their other financing measures. The original paper defines the measure in the opposite way as the difference between years t and t-1, but then flips the sign for the relevant regression in table 8. Stocks with a decrease in liabilities are in the long portfolio.

dNoa - change in net operating assets "12-month growth in Net Operating Assets scaled by lagged total assets (at). Net Operating assets are operating assets minus operating liabilities. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dlc), book equity (ceq) and preferred stock (pstk), all items (except at and ceq) replaced with 0 if missing." Stocks with a decrease in NOA are in the long portfolio. Originally by Hirshleifer et al. (2004). They find a strong effect in a regression but only when not controlling for the level of net operating assets, which is also among the predictors (NOA)

EarningsForecastDisparity - Long-vs-short EPS forecasts "Analyst forecasted 5-year earnings growth (fgr5yr) minus 100 times the difference between mean earnings forecast (meanest) and fiscal year earnings expectations (fy0a). " The original paper (Da

and Warachka, 2011) uses the following definition, which is somewhat different: Compare 1-year growth forecast ISTG calculated by taking IBES 1 year ahead forecast and getting the change to the last announced value to LTG. The short portfolio is the one that has a high LTG and low ISTG, and the long portfolio has a Low LTG and high ISTG.

FEPS - Analyst Earnings Per Share "Using IBES unadjusted forecasts, keep fpi == 1, signal is meanest." Construct the mean 1 Year ahead earnings per share forecast in the previous month. The original paper (Cen et al., 2006) first sorts on firm characteristics (size, price BTM and 6-month return) and then within these sorts on FEPS. Stocks with high forecasts are in the long portfolio.

ForecastDispersion - EPS Forecast Dispersion "Keep fpi = 1 and fpedats > statpers + 30. Standard deviation of earnings estimates (stdev_est) scaled by mean earnings estimate." "Dispersion is defined as the standard deviation of earnings forecasts scaled by the absolute value of the mean earnings forecast. If the mean earnings forecast is zero, then the stock is assigned to the highest dispersion category." Use one-year-ahead forecasts. Original paper by Diether et al. (2002)

grcapx - Change in capex (two years) "Growth rate of capital expenditures (capx) relative to two years ago. If capx is missing, replace with annual change in property, plant and equipment (ppent)." Originally by Anderson and Garcia-Feijóo (2006).

grcapx3y - Change in capex (three years) "Capital expenditures (capx) divided by the sum of capital expenditures from year - 1, year -2, and year -3. If capx is missing, replace with annual change in property, plant and equipment (ppent)." Originally by Anderson and Garcia-Feijóo (2006).

Illiquidity "Past twelve month average of: daily return (abs(ret)) divided by turnover((abs(prc)*vol)" Originally from Amihud (2002).

InvestPPEInv - change in ppe and inv/assets "One-year change in property, plants and equipment (ppegt) plus one year change in inventory (invt), scaled by one-year lagged assets (at)." Originally by Lyandres et al. (2008).

Leverage - Market leverage "Total liabilities (lt) divided by market value of equity." Stocks with high leverage are in the long portfolio. Originally by Bhandari (1988)

MeanRankRevGrowth - Revenue Growth Rank "Rank firms by their annual revenue growth each year over the past 5 years. MeanRankRevGrowth is the weighted average of ranks over the past 5 years, that is, MeanRankRevGrowth = (5*Rank_{t-1} + 4*Rank_{t-2} + 3*Rank_{t-3} + 2*Rank_{t-4} + 1*Rank_{t-5})/15. Exclude NASDAQ stocks." Stocks with low growth are in the long portfolio. Originally from Lakonishok et al. (1994), who use a somewhat different methodology that sorts on multiple characteristics.

Mom6m - Momentum (6 month) "Stock return between months t-6 and t-1" following Jegadeesh and Titman (1993).

Mom6mJunk - Junk Stock Momentum "Mom6m. Include only stocks with a credit rating (splticrm) of BBB or lower" Originally by Avramov et al. (2007).

Mom12m - Momentum (12 month) "Stock return between months t-12 and t-1" Originally by Jegadeesh and Titman (1993).

MomSeason - Return seasonality years 2 to 5 "Average return in the same month over the preceding 2-5 years." Originally by Heston and Sadka (2008).

MomVol - Momentum in high volume stocks "Define momentum as Mom6m, and volume as the rolling average of the past 6 months of monthly turnover (minimum 5 months). Independent sort stocks into 10 momentum ports and 3 volume ports. Keep if volume is in the top port, and assign signal = momentum port. Drop if less than 2 years on CRSP." Originally by Lee and Swaminathan (2000).

MRreversal - Medium-run reversal "Stock return between months t-18 and t-13." Based on De Bondt and Thaler (1985) who do not do a statistical test for this horizon and just provide visual evidence.

NetEquityFinance "Sale of common stock (sstk) minus purchase of common stock (prstkc), scaled by average total assets (at) from years t and t-1. Exclude if absolute value of ratio is greater than 1." Stocks with high net equity financing are in the short portfolio, and those with low net equity finance are in the long portfolio. Originally by Bradshaw et al. (2006).

NetPayoutYield "Dividends (dvc) plus purchase of common and preferred stock (prstkc) minus sale of common and preferred stock (sstk), divided by market value of equity lagged 6 months. Exclude if NetPayoutYield is 0, financial firm based on SIC code, ceq <= 0, or less than 2 years in CRSP". High-yield stocks are in the long portfolio, and low-yield stocks are in the short portfolio. Originally by Boudoukh et al. (2007).

NOA - Net Operating Assets "Difference between operating assets and operating liabilities, scaled by lagged total assets. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dc) and book equity (ceq)." Stocks with low NOA are in the long portfolio. Originally by Hirshleifer et al. (2004).

NumEarnIncrease - Earnings streak length "Number if consecutive 4-quarter increases in ibq, up to 8." Stocks with a long positive streak are in the long portfolio. This is based on Loh and Warachka (2012). They do not study the same signal, but their Table 4 suggests this should work. Their definition of a streak is just that the surprise has the same sign (but it does not have to increase). Their Table 4 shows that earnings surprises have a stronger effect if they are part of a streak rather than a reversal of a streak.

PctTotAcc - Percent Total Accruals "Net income (ni) minus (purchase of common and preferred stock (prstkcc) minus sale of common and preferred stock (sstk) plus dividends (dvt), cash flow from operations (oancf), from financing (fincf) and investment (ivncf)). Scaled by absolute value of net income." Stocks with low values of the characteristic are in the long portfolio. Originally by Hafzalla et al. (2011).

PriceDelayRsq - Price delay r square "Each July regress stock return on day t on market return in $t, t-1, \ldots, t-4$ using observations from July 1 of the previous year to June 30 of the current year. Then regress again with no market return lags. PriceDelay Rsq = 1 - [Rsq from second regression]/[Rsq from first regression]" Firms with higher delay are in the long portfolio. Originally from Hou and Moskowitz (2005). Measures delayed market response to information, as it captures how much of the market level information is incorporated into the stock price on the same day compared to on later days. If everything is incorporated on the same day, adding the lagged returns does not improve R^2 subsequentially.

Recomm_ShortInterest - Analyst Recommendations and Short-Interest "Go long firms in lowest quintile of short interest (shortint/shrout) and lowest quintile of analyst recommendations (monthly consensus recommendation using the most recent analyst recommendation in the past 12 months). Go short firms in highest quintile of short interest and highest quintile of analyst recommendations." Originally by Drake et al. (2011). Note that the construction of the portfolios is different from the original paper due to a lack of data.

RDIPO - IPO and no R&D spending "Binary variable equal to 1 if R&D expense (xrd) = 0 and IndIPO = 1. 0 otherwise." Stocks with RnD are in the long portfolio, and stocks without are in the short portfolio.

RDS - Real Dirty Surplus "Define Dirty Surplus as annual change in marketable securities adjustment msa plus annual change in retained earnings adjustment (recta) + .65 times the annual change in min(Unrecognized prior service cost (pcupsu) - Pension

additional minimum liability (paddml),0). Real dirty surplus is the annual change in book equity (ceq) minus dirty surplus minus (net income (ni) minus dividends preferred (dvp)) + dividends (divamt) - end-of-fiscal-year-stock-price (prcc_f)*annual change in common shares outstanding (csho)." Originally by Landsman et al. (2011)

ShareIss1Y - Share issuance (1 year) "Growth in number of shares between t-18 and t-6. Number of shares is calculated as shrout/cfacshr to adjust for splits." Stocks with the most positive share issuance are on the short portfolio and those with the most negative share issuances are in the long portfolio. The original paper by Pontiff and Woodgate (2008) uses Fama-MacBeth regressions instead of portfolio sorts.

ShareRepurchase "Binary variable equal to 1 if stock repurchase indicated in cash flow statement (prstkc $>$ 0), and 0 if prstkc = 0." Stocks with share repurchases form the long portfolio, and stocks without form the short portfolio. The intuition from the original paper by Ikenberry et al. (1995) is that firms strategically buy back shares if they are undervalued. The original paper hand collects share repurchase data by looking at all share repurchases announced in the *Wall Street Journal*.

sinAlgo "Using Compustat Segment data, sinAlgo is defined as a binary variable equal to 1 if at least one segment of a firm is listed as being in at least one of the following industries: sic $\geq$ 2100 & sic $\leq$ 2199, sic $\geq$2080 & sic $\leq$ 2085, NAICS in {7132, 71312, 713210, 71329, 713290, 72112, 721120}. As in the original paper, we assume that the sin stock indicator applies to the entire history and future of the identified firm. sinAlgo is equal to 0 if the firm is not identified in the CS Segment data as a sin stock and if the firm is in one of the following industries: (sic $\geq$ 2000 & sic $\leq$ 2046) OR (sic $\geq$ 2050 & sic $\leq$ 2063) OR (sic $\geq$ 2070 & sic $\leq$ 2079) OR (sic $\geq$ 2090 & sic $\leq$ 2092) OR (sic $\geq$ 2095 & sic $\leq$ 2099) OR (sic $\geq$ 2064 & sic $\leq$ 2068) OR (sic $\geq$ 2086 & sic $\leq$ 2087) OR (sic $\geq$ 920 & sic $\leq$ 999) OR (sic $\geq$ 3650 & sic $\leq$ 3652) OR sic == 3732 OR (sic $\geq$ 3931 & sic $\leq$ 3932) OR (sic $\geq$ 3940 & sic $\leq$ 3949) OR (sic $\geq$ 7800 & sic $\leq$ 7833) OR (sic $\geq$ 7840 & sic $\leq$ 7841) OR (sic $\geq$ 7900 & sic $\leq$ 7911) OR (sic $\geq$ 7920 & sic $\leq$

7933) OR (sic $\geq$ 7940 & sic $\leq$ 7949) OR sic == 7980 OR (sic $\geq$ 7990 & sic $\leq$ 7999)" Originally from Hong and Kacperczyk (2009).

Size "Log of monthly market value of equity (abs(prc)*shrout))." Originally from Banz (1981).

SP "Ratio of annual sales (sale) to market value of equity." Originally from Barbee et al. (1996).

Tax "Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5." Originally from Lev and Nissim (2004).

TotalAccruals "Before 1988: Change in net working capital ((act - che) - (lct - dlc)) plus change in net noncurrent assets ((at - act - ivao) - (lt - dlc - dltt)) plus change in net financial assets ((ivst + ivao - (dltt + dlc + pstk)). Starting in 1988: net income (ni) minus total, operating and investment cashflows (oancf, ivncf, fincf) plus stock sales minus repurchases and dividends (sstk, prstkc, dv)). Scaled by lagged total assets (at). Replace missings in ivao, ivst, dltt, dlc, pstk sstk, prstkc, dv with 0." Originally from Richardson et al. (2005).

VolSD "Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net

income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5." Originally from Lev and Nissim (2004)

XFIN - Net external financing "Sale of common stock (sstk) minus dividends (dv) minus purchase of common stock (prstkc) plus long-term debt issuance (dltis) minus long-term debt reductions (dltr). Scaled by total assets (at)." Stocks with high net external financing are in the short portfolio and those with low net external finance are in the long portfolio. Originally form Bradshaw et al. (2006).

zerotrade - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48*10^5$), multiplied by 21/number of trading days per month. Zerotrade is the 6-month average of that variable." Originally from Liu (2006).

zerotradeAlt1 - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48*10^5$), multiplied by 21/number of trading days per month. Take 1-month average." Originally from Liu (2006).

zerotradeAlt12 - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48*10^5$), multiplied by 21/number of trading days per month. Take 12-month average." Originally from Liu (2006).

Chapter 5

Conclusion

In this thesis, I study the impact of beliefs and preferences that deviate from the standard assumptions of rational expectations and discounted expected utility on financial decision-making.

Chapter 2 extends salience theory (Bordalo et al., 2012), a static model of choice under risk, to a continuous-time dynamic context. The model incorporates the idea that states with high contrast in the outcomes between the different choices attract attention so that their probabilities aver overestimated. It predicts that agents have skewness preferences and will take risks with zero or even slightly negative expected values, despite being intrinsically risk averse. They do this even if the underlying risk is symmetric because they can endogenously create skewness through the choice of their stopping strategy. An experiment confirms the central predictions of the model. Most subjects take a risk with an expected value of zero and the share of subjects willing to take the risk monotonically decreases as its expected value decreases. If subjects take the risk, the majority of their stopping strategies lead to a positively skewed outcome distribution.

Chapter 3 studies if a contrast effect influences the earnings forecasts of professional analysts after a firm's earnings announcement. Hartzmark and Shue (2018) find that the return of a firm after its earnings announcement is more negative the more positive the earnings surprise of other firms, which announced their earnings on the previous day, was. I find that analyst earnings expectations are also more pessimistic relative to the actually announced value if the contrast surprise was more positive. This suggests that the abnormal returns found by Hartzmark and Shue are driven by biased expectations about future

earnings and not just a biased perception of the firm's current announcement. Moreover, it demonstrates that contrast effects can influence expectations, which potentially has wide-reaching consequences, given that contrasts are readily available across a wide range of domains in which the expectations of economic agents play an important role.

Chapter 4 develops a test to determine if a characteristic that can predict returns is linked to mispricing. While both risk and mispricing can explain return predictability, only mispricing-driven abnormal returns can be linked to biased expectations. If a return predictor, therefore, predicts more positive changes in earnings expectations for the stocks for which it predicts more positive returns, this implies that the return predictability is driven by mispricing. Applying this test to a dataset of 173 return predictors from Chen and Zimmermann (2021), I find that around 40% of them reflect mispricing from biased expectations.

As a second step, I follow van Binsbergen et al. (2023) and further classify mispricing predictors into those that capture the build-up and those that capture the resolution of mispricing. If a return predictor captures the resolution of mispricing, it should also predict that earnings expectations are initially too pessimistic for stocks with subsequently higher returns relative to those with lower returns. Similarly, if a return predictor captures the build-up of mispricing it should be able to predict that earnings expectations are too optimistic after stocks earned abnormally high returns, relative to those with lower returns. I apply these tests to the predictors identified as mispricing in the previous step and find that approximately 45% capture the resolution and 10% capture the build-up of mispricing, while the remaining cannot be classified.

Overall, the results of this thesis suggest two areas for future research. First, the results of Chapters 2 & 3 together with many other works discussed in Chapter 1 suggest that contrast effects are a fundamental feature of human perception and decision making. Arguably, in most situations where an economic actor needs to form beliefs and make decisions, information that can serve as a contrast to the problem at hand is available. However, contrast effects have only been studied in a few contexts involving economic or financial decision-making, leaving room for many future works. Moreover, since they impact both preferences and beliefs they could eventually become part of a unified model

of economic decision-making that describes deviations both from the standard assumption of expected utility preferences and the standard assumption of rational expectations based on a single psychological mechanism.

Second, the literature reviewed in Chapter 1 together with the results from Chapters 3 & 4 highlights the importance of data on expectations to understand financial decision-making. While considerable work has been done in this area there is room to study the expectation formation process of both households and institutional investors in more detail using lab or field experiments. Moreover, future research can expand on the idea that a significant part of stock market return predictability is driven by biased expectations by incorporating expectation biases directly into an asset pricing model and testing this model empirically.

References

Van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni (2023). "Dynamic Asset (Mis)Pricing: Build-up versus Resolution Anomalies". In: *Journal of Financial Economics* 147.2, pp. 406–431. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2022.11.005.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). "Salience Theory of Choice Under Risk". In: *The Quarterly Journal of Economics* 127.3, pp. 1243–1285. ISSN: 0033-5533. DOI: 10.1093/qje/qjs018.

Chen, Andrew Y. and Tom Zimmermann (2021). *Open Source Cross-Sectional Asset Pricing*. SSRN Scholarly Paper 3604626. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3604626.

Hartzmark, Samuel M. and Kelly Shue (2018). "A Tough Act to Follow: Contrast Effects in Financial Markets". In: *The Journal of Finance* 73.4, pp. 1567–1613. ISSN: 1540-6261. DOI: 10.1111/jofi.12685.