Associative Memory, Beliefs and Market Interactions*

Benjamin Enke Frederik Schwerter Florian Zimmermann

June 12, 2023

Abstract

Recent theories and narratives highlight the potential role of associative recall in driving overreaction in expectations and market behavior. Based on a simple model, we test this idea through a series of experiments in which news are communicated with memorable contexts. Because participants predominantly remember those past signals that get cued by the current one, their beliefs about fundamentals strongly overreact to news and exhibit systematic history-dependence. In a betting market experiment, associative recall translates into overreaction in market prices, which makes realized prices too extreme. Our results highlight the importance of associative memory for beliefs and economic decisions.

JEL classification: D01

Keywords: Beliefs, expectations, memory, associations, bounded rationality

*This paper supersedes our working paper "Associative Memory and Belief Formation" that was first circulated in 2019. Relative to that paper, all experiments in this paper are new. They were pre-registered in the AEA RCT Registry as AEARCTR-0009215. Schwerter thanks the Sloan Foundation for post-doctoral funding. Enke: Harvard University and NBER; enke@fas.harvard.edu. Schwerter: Frankfurt School of Finance and Management and University of Cologne; frederik.schwerter@uni-koeln.de. Zimmermann: University of Bonn and briq; florian.zimmermann@briq-institute.org.

1 Introduction

In economic textbook models of belief formation, memory imperfections play no role: people combine prior knowledge with current information, and yesterday's posterior equals today's prior. In reality, people of course do not constantly have access to their beliefs about every potentially relevant state of the world. Rather, when people are prodded to act on or update their beliefs, they need to reconstruct prior knowledge and beliefs from memory. Both psychological and economic theories model this retrieval process as being *associative* in nature, which refers to the idea that people are more likely to recollect information that is cued by (similar to) what they observe today (Mullainathan, 2002; Kahana, 2012; Bordalo et al., 2020a, 2023). Oftentimes, such cues may be given by intrinsically uninformative contextual features such as stories, narratives and images.

The associative nature of memory potentially has at least two central implications for the economic study of belief formation and corresponding market behavior. First, associative recall could lead to systematic *overreaction*: after receipt of a piece of news, people reconstruct past knowledge from memory, yet predominantly remember those past news that are similar to today's news. As a consequence, beliefs might look like they overreact to recent news, purely as a result of how prior knowledge is reconstructed. Various popular accounts have hypothesized that the broad idea of associations-driven overreaction may be a driver of aggregate economic events. For example, in influential writings on the role of narratives, Shiller and co-authors appeal to the role of associative recall for expectation formation and overreaction in markets by observing that "[o]ne new narrative may remind of another that has been lying fairly dormant... there is cue-dependent forgetting" (Shiller, 2017, p. 975, also see Shiller, 2019, Goetzmann et al., 2022). Similarly, Gennaioli and Shleifer's (2018) treatment asserts that associative memory may underlie overreaction to news in the context of the 2007-2008 financial crisis.

A second potential implication of associative recall for the study of belief formation is that it implies a violation of *history-independence*. As we review below, the vast majority of economic models of belief updating – both rational and behavioral – make the strong prediction that today's posterior is fully characterized by a combination of yesterday's posterior and today's information. However, with associative recall, this is no longer the case because today's news may asymmetrically cue the selective retrieval of similar past information. Thus, with associative recall, the precise signal history matters for today's beliefs even holding fixed yesterday's beliefs.

A major hurdle in identifying the role of associative recall for expectation formation and market behavior is that, in field contexts, information signals and associated contexts are often inseparably intertwined, which makes it difficult to identify the *causal* effect of narratives, images and other associations above and beyond the normatively relevant informational content of the signals themselves. In laboratory experiments, on the other hand, signals and contexts can be decoupled and exogenously varied.

This paper reports on a series of such experiments to make three contributions. First, we translate classical psychological paradigms on cued recall into economic belief formation problems to document that people indeed have a pronounced tendency to asymmetrically remember past information that gets cued by contextual features, which generates systematic history-dependence of beliefs. Second, we show that these effects lead to strong overreaction of expectations with respect to current information. Third, using a betting market experiment, we document that associations-driven overreaction in beliefs translates into overreaction in market prices, even when people have an opportunity to fully or partially select out of the market.

Our experiments are structured around the predictions of a simple formal framework that applies the idea of associative recall to belief formation, based on the formulations in Bordalo et al. (2020a) and Mullainathan (2002). In this model, decision-makers (i) have imperfect memory and (ii) are more likely to recollect a piece of news if the context in which it was experienced is similar to today's context. The context could be a narrative, but also a picture, an emotion or a sound. This stylized model predicts both overreaction in beliefs and a specific form of history-dependence that is absent in Bayesian or most behavioral models. The model makes comparative statics predictions about how overreaction depends on the presence of associations and the signal history. Our treatments are tightly designed around these predictions.

Experimental design. In our experiment, participants form beliefs about whether each of multiple hypothetical companies is "good" or "bad." The experiment comprises two periods that we think of as "past" and "present." Across both periods, a subject sequentially observes noisy binary signals that are informative about a company's true quality. These news are communicated in a context, which consists of a story (narrative) and an image that relate to the news. For example, for one company, a positive signal would be shown with an (intrinsically uninformative) story about the company having launched a successful advertisement campaign with a celebrity, accompanied by a picture of that celebrity. Subjects' financial incentives are such that their second-period beliefs about a company's quality should incorporate both first- and second-period signals. Our main object of interest is whether the second-period signal produces overreaction of second-period beliefs through the logic of associative recall.

We deploy two types of random variation to causally identify the role of associative recall. First, we manipulate the scope for associative memory in a within-subjects treatment variation. Each subject forms beliefs about each of 14 companies, seven of which belong to treatment *Cue* and seven of which belong to treatment *NoCue*. For companies in treatment *Cue*, identical news are embedded in identical contexts. In other words, for each company in treatment *Cue*, all positive news are communicated with the same context, and all negative news are communicated with the same context. Thus, in this treatment, the second-period signal could cue the asymmetric retrieval of identical first-period signals. While it is rarely the case that real market participants experience multiple signals in exactly the same context, this simple setup is reflective of many applications in which similar signals are consistently associated with similar contexts. For example, whenever good news prevail in the stock market, people are disproportionately exposed to bulls, upward-sloping trend lines, and good-times stories.

For companies in treatment *NoCue,* on the other hand, each signal is communicated with a different context. Thus, subjects never observe the same story or image twice. As a result, the scope for associative recall is exogenously reduced.

A second dimension of random variation in our experiments is that, within treatment *Cue*, the number of first-period signals that equal the second-period signal (the number of signals that "get cued") differs randomly across subjects and companies. This is relevant because our stylized model predicts that overreaction of second-period beliefs should systematically depend on the number of cued first-period signals, even though – conditional on first-period beliefs – the signal history is irrelevant from a normative perspective. This within-treatment variation allows us both to provide causal evidence for the role of associative recall that does not rely on switching the presence of associations on and off, and to document that beliefs are indeed history-dependent in a way that violates the predictions of models that do not rely on memory limitations.

Results. We report two main results. First, second-period beliefs overreact substantially to the second-period signal in *Cue*, an effect that is not present in *NoCue*. This identifies a causal effect of the presence of associations on overreaction. Second, the magnitude of overreaction strongly increases in the number of first-period signals that get cued by the second-period signal. Thus, associations cause overreaction, and this overreaction is history-dependent in specific ways predicted by the model.

Our interpretation of overreaction in beliefs in *Cue* is that subjects asymmetrically remember those first-period signals that equal the second-period one. To supply direct evidence for this interpretation, we implement a second experiment in which we directly elicit people's recall of first-period signals in an incentivized fashion. We find that subjects in *Cue* indeed only recall more first-period signals than subjects in *NoCue* when those first-period signals get cued by the second-period signal.

Market experiments. An immediate question is whether this associations-driven overreaction in beliefs also impacts market behavior. One reason why this need not be the case is people's potential doubts about their memory technology. For instance, people may overreact when they are forced to state their beliefs, but they may worry that they are somehow getting the problem wrong (even if they don't know how). As a result, people might not be willing to act on their beliefs in market settings but instead self-select out (Camerer and Lovallo, 1999; Enke et al., 2021a). Yet, the psychological literature provides little empirical guidance on whether we should expect people with stronger associative memory to act less aggressively on their beliefs, perhaps because self-selection into or out of markets is of more concern to economists than psychologists.

To study this, we embed our individual belief elicitation paradigm into a parimutuel betting market experiment. In this market, groups of three subjects each receive public signals and then place bets on whether the company is good or bad. The parimutuel price mechanism redistributes the money bet among the market participants according to whose bet was right and the amount of money bet. A crucial feature of this market is that subjects can self-select in or out in a continuous fashion, by choosing the total amount of money they would like to bet. For instance, even if associative recall induces an excessively strong belief that the company is good, the subject may not bet much money on this proposition if s/he is partly unsure about her belief updating strategy.

Despite this potential for self-selection, we find that market prices react about twice as strongly to information when associative recall is facilitated (repeated contexts) compared to when associations are removed. As a result, just like memorable contexts induce individual beliefs to be too extreme, they also lead market prices to be too high (too low) following a positive (negative) second-period signal. Indeed, we find that the magnitude of associations-driven overreaction in market prices is very similar to the magnitude of overreaction in beliefs.

Contribution and related literature. Overall, our contribution to the literature is (i) to provide evidence that associative recall shapes the formation of probabilistic beliefs; (ii) that this produces systematic overraction; and (iii) that these effects affect experimental market prices, despite scope for self-selection. These results tie into a growing theory literature that has argued for the importance of associative memory for economics. Mullainathan (2002) and Bordalo et al. (2020a, 2023) present models of how cued recall shapes economic decision-making. Related theoretical work has investigated the implications of similarity and associations in various applications, ranging from decision making under uncertainty to finance to consumption to self-esteem (Gilboa and Schmeidler, 1995; Laibson, 2001; Gennaioli and Shleifer, 2010; Noor, 2019; Bodoh-Creed, 2020; Wachter and Kahana, 2019; Koszegi et al., 2019). As the simple formalism that structures our experiments directly draws from this literature, we view our experiments as providing the first direct evidence from economic decision making tasks in relation to this emerging body of theoretical work.

On the experimental side, there is a large psychology literature but scant evidence from economic decision problems. In a canonical psychological task on associative recall, subjects are asked to memorize words, and are subsequently more likely to remember a word if it was shown in conjunction with another word that is currently being displayed (see Schacter, 2008; Kahana, 2012, for overviews). As we review in Section 3.5 and Appendix A, our experimental design applies a variant of canonical psychological paradigms to economic decision tasks. This application to economic decisions is necessary because psychological experiments have at least two features that make them less-than-directly applicable to the questions that we are interested in here. First, these experiments are pure recall tasks that do not involve information-processing, belief updating or overreaction (e.g., Bordalo et al., 2021).[1] Second, psychologists have not paid attention to how associative recall affects market behavior, which is arguably of first-order relevance for understanding whether associative memory meaningfully impacts economic outcomes.

Within economics, a small number of lab experimental papers that were circulated subsequent to our work provide additional evidence on the role of bounded rationality in memory for belief formation.[2] Afrouzi et al. (2023) argue for a role of working memory in forecasting, but do not look at the role of associative recall or market behavior. Bordalo et al. (2023) study the role of interference in pure recall tasks. Graeber et al. (2023) show that recall of stories is more persistent than recall of statistical information.[3]

Finally, our paper also relates to a small but growing experimental literature on narratives and mental models in economics (Andre et al., 2021, 2022; Bursztyn et al., 2023; Barron and Fries, 2022; Esponda et al., 2020). Consistent with Shiller's (2017) intuition, we document that narratives determine the strength of recall, causing overreaction to information.

The remainder of the paper proceeds as follows. Section 2 offers a stylized formal framework that motivates the experimental design and structures the analysis. Section 3 describes the experimental design and pre-registration. Sections 4 and 5 present the results of the individual belief elicitation and market experiments. Section 6 concludes.

[1]Bhatia (2017) studies the role of associations for probabilistic beliefs in vignette studies but – because he works with ecological data or ambiguous vignettes – does not speak to how people update beliefs relative to the Bayesian benchmark, to potential overreaction nor to market behavior.

[2]Some recent observational and survey studies also provide field evidence suggesting that associative memory may be a driver of investment behavior (Charles, 2022a,b; Jiang et al., 2023). While these studies provide creative complementary evidence to our causally identified experiments, they do not afford the possibility to directly and exogenously manipulate the presence of associations and, hence, offer less control in identifying memory effects.

[3]In a study of attention, Hartzmark et al. (2021) also document that ownership of assets affects recall of return profiles. Work on motivated memory is reviewed in Amelio and Zimmermann (2023).

2 Theoretical Framework

2.1 Setup

This section presents a stylized model to guide the design of the experiments and to structure the empirical analysis. The mechanics of the model build on some of the formulations in Mullainathan (2002) and Bordalo et al. (2020a, 2023). The framework rests on two key assumptions: (i) people may forget prior knowledge, so that they need to reconstruct it from memory; and (ii) this recollection process is subject to associative recall. We abstract away from additional behavioral assumptions that the literature on associative memory has incorporated, such as salience or rehearsal.

Consider a decision-maker (DM) who forms beliefs about the realization of a binary variable θ with possible states denoted by $\theta = G(\text{ood})$ and $\theta = B(\text{ad})$. The prior probability is $P(G) = P(B) = 0.5$. The DM receives a series of i.i.d. binary signals s_x that take on the realizations $p(\text{ositive})$ and $n(\text{egative})$. The signal diagnosticity is given by $P(p|G) = P(n|B) = q > 0.5$. In what follows, we will use the terms "news" and "signal" interchangeably. With a slight abuse of notation, we will write $s_x = 1$ for positive and $s_x = -1$ for negative signals.

There are two periods. In the first ("past"), the DM receives potentially multiple signals, $s_1, \ldots, s_k$. Denote by N_p and N_n the number of positive and negative first-period signals. In the second period ("present"), the DM receives one additional signal, s_{k+1}. We call a first-period signal s_x *congruent* with the second-period signal if $s_x = s_{k+1}$. It is helpful to introduce a shorthand for the number of first-period signals that are congruent with the second-period signal: $z := \sum_{x=1}^{k} \mathbb{1}_{s_x = s_{k+1}}$.

We assume that each signal s_x is experienced in a context, c_x. By context we mean all environmental features that co-occur with a signal, except the signal realization itself. Loosely speaking, contexts are characterized by two aspects. First, conditional on the signal, they are uncorrelated with the state, meaning that they are intrinsically uninformative. Second, contexts are memorable in the sense that observing them today reminds people of occurrences of the same environmental features in the past.

We introduce two different counterfactual conditions (that will correspond to experimental treatment conditions), across which the mapping between signals and contexts differs. First, in a *Cue* condition ($\mathbb{T} = 1$), there is a one-to-one mapping between type of news (positive or negative) and context: $c_x = c_y \Leftrightarrow s_x = s_y$. Thus, all positive news appear in the same context and all negative news in the same (yet different) context. Second, in a *NoCue* condition ($\mathbb{T} = 0$), the same context never appears twice, regardless of the signal realizations: $c_x \neq c_y \; \forall x, y$. We say that a first-period signal gets "cued" by a second-period signal when both the signals and the contexts are identical.

2.2 Memory and Beliefs

First-period beliefs. Denote by $b_t(G|S_t)$ the DM's posterior belief in period t that the state is good, following signal history S_t. By standard arguments, the first-period Bayesian posterior belief odds can be expressed as a function of the likelihood ratio and the prior odds:

$$\frac{b_1(G|S_1)}{1-b_1(G|S_1)} = \left(\frac{q}{1-q}\right)^{(\sum_{x=1}^{k} s_x)} \frac{p(G)}{p(B)} \tag{1}$$

where the likelihood ratio consists of the diagnosticity odds to the power of the number of positive minus negative signals. The prior odds drop out because we assumed $P(G) = P(B) = 0.5$. A popular transformation of this expression in the literature is the so-called Grether (1980) decomposition. Taking logs and re-arranging, we get a linear expression for the DM's first-period *normalized log posterior odds (lpo)*:

$$lpo_1 := \frac{ln\left(\frac{b_1(G|S_1)}{1-b_1(G|S_1)}\right)}{ln(\frac{q}{1-q})} = \sum_{x=1}^{k} s_x = N_p - N_n \tag{2}$$

The normalized log posterior odds vary one-for-one with changes in the net number of positive signals. This property of Bayesian beliefs is well-understood.

Second-period beliefs: perfect memory benchmark. By a simple extension of the above, the normalized second-period Bayesian log posterior odds can be expressed as

$$lpo_2^{bayes} := \frac{ln\left(\frac{b_2(G|S_2)}{1-b_2(G|S_2)}\right)}{ln(\frac{q}{1-q})} = s_{k+1} + (N_p - N_n) \tag{3}$$

This expression is analytically very convenient because (i) it can be estimated using simple OLS regressions and (ii) the perfect-memory benchmark coefficient of the second-period signal is simple and given by one.

Second-period beliefs: the case of associative recall. Now consider a DM who potentially forgets some or all signals going from $t = 1$ to $t = 2$. Whether or not the DM remembers a signal is determined by two factors. First, irrespective of the piece of news, there is some probability $r \in [0, 1)$ that the DM will remember. Second, reflecting the logic of associative recall, the probability of recalling a past signal is higher if its context is identical to today's context. We assume that recall $\hat{s}_x$ of s_x is given by

$$\hat{s}_x = \begin{cases} s_x & \text{with probability} \quad r + (1-r)a\mathbb{1}_{c_x = c_{k+1}} \\ \emptyset & \text{else} \end{cases} \tag{4}$$

Thus, the probability of remembering a first-period signal s_x is r whenever the context of the first-period signal does not match the context of the second-period news, $c_x \neq c_{k+1}$. If, on the other hand, the context of the first-period signal equals the context of the second-period signal, the probability of recall receives an "associations boost" parameterized by $a \in (0,1]$.[4]

Following Mullainathan (2002) and Bordalo et al. (2023), we assume that the DM applies Bayes' rule to the signals she retrieves from memory. Thus, we derive the DM's posterior odds following equation (1), except that we replace the actual signals, $\sum s_x$, with the recalled ones, $\sum \hat{s}_x$.

Using the Grether decomposition again and doing a bit of algebra delivers:

$$\begin{aligned} lpo_2^{assoc} &= s_{k+1} + \sum_{x=1}^{k} \hat{s}_x \\ &= s_{k+1} + \sum_{x=1}^{k} E[\hat{s}_x | s_x, s_{k+1}] + \underbrace{\sum_{x=1}^{k} (\hat{s}_x - E[\hat{s}_x | s_x, s_{k+1}])}_{:=\epsilon} \\ &= [1 + \underbrace{(1-r)az\mathbb{T}}_{\text{Overreaction}}] s_{k+1} + r \underbrace{(N_p - N_n)}_{\text{1st-period lpo}} + \epsilon \end{aligned} \tag{5}$$

where the mean-zero noise term ϵ reflects that the memory technology in (4) is random. In this Grether decomposition, the second-period log posterior odds are expressed as a function of the second-period signal and the first-period log posterior odds. Here, beliefs *look like* they overreact to the second-period signal because the overall coefficient is potentially strictly larger than one. Intuitively, the second-period signal has both a direct effect on beliefs and an indirect effect through the asymmetric recall that it generates. Indeed, when the stable contexts are removed ($\mathbb{T} = 0$), equation (5) does not predict overreaction. Also observe that (5) clarifies that associative recall only distorts beliefs if $z > 0$. This is intuitive: if no first-period signal equals the second-period signal, then nothing gets cued and no asymmetric retrieval takes place.

Our experiments focus on testing the distinctive (comparative statics) predictions that arise from equation (5) relative to equation (3). In particular, our experiments will exogenously manipulate the experimental analogues of the parameters $\mathbb{T}$ and z.

[4]In a more general model, associativeness is formalized via a continuous similarity function (Bordalo et al., 2020a). Our formulation corresponds to a simplification in which similarity is either 0 or 1.

Model predictions.

1. *If a strictly positive number of first-period signals are congruent with the second-period signal ($z > 0$), overreaction of second-period beliefs to the second-period signal is larger in the presence of associations:* $\frac{\partial lpo_2}{\partial s_{k+1}}|_{\mathbb{T}=1,z>0} > \frac{\partial lpo_2}{\partial s_{k+1}}|_{\mathbb{T}=0,z>0}$.

2. *If no first-period signals are congruent with the second-period signal ($z = 0$), there is no differential overeaction across treatments:* $\frac{\partial lpo_2}{\partial s_{k+1}}|_{\mathbb{T}=1,z=0} = \frac{\partial lpo_2}{\partial s_{k+1}}|_{\mathbb{T}=0,z=0}$.

3. *In the presence of associations, overreaction increases in the number of congruent first-period signals, z, even holding fixed first-period beliefs:* $\frac{\partial^2 lpo_2}{\partial s_{k+1}\partial z}|_{\mathbb{T}=1,lpo_1} > 0$.

In a nutshell, these predictions can be summarized with two themes. First, if at least one first-period signal "gets cued" by the second-period context, associative recall produces systematic overreaction of beliefs to the second-period signal, which makes beliefs too extreme, on average. Second, associative recall implies a distinctive form of history-dependence: even holding fixed first-period beliefs, the signal history matters for second-period beliefs. This second prediction stands in stark contrast with the vast majority of economic models of belief formation, in which today's posterior beliefs are fully characterized by yesterday's posterior beliefs and today's signals. While in behavioral models people may make mistakes in weighting yesterday's beliefs and today's signals, the signal history never matters for how people process new information once yesterday's posterior belief is fixed. The history-dependence in Prediction 3. is, hence, a distinctive prediction of associative recall.

It is also worth emphasizing that the model predictions rely on the presence of associations. Models of recency bias (Fudenberg et al., 2014) or optimized responses to imperfect memory (Wilson, 2014) do not generate this joint set of predictions. For example, recency bias predicts overreaction, but not that overreaction depends on the history of news, or that it disappears once associative recall is shut down.

3 Experimental Design

Our experiment builds a bridge between the tightly-controlled and quantitative designs that dominate modern experimental economics research and psychological paradigms on cued recall problems. We particularly focused on the following design objectives: (i) a decision setup that is closely tied to the model in Section 2; (ii) tight control over signal structure and associations; (iii) a framed environment that is intuitive for participants; (iv) exogenous variation in the key model parameters; and (v) incentive-compatible belief elicitation.

3.1 Experimental Setup

Task overview. We implement a standard binary-state balls-and-urns (or bookbag-and-pokerchips) experiment that is widely studied in behavioral economics. The main difference to the standard implementation is that we include a memory component. To aid subject understanding, the experiment was framed as estimating the probability that each of 14 hypothetical companies is of a good rather than a bad type. Each company is denoted by a capital letter. The experiment consists of two periods (Figure 1 provides a timeline of the experiment). In both periods, subjects receive noisy information about each of the companies and subsequently state probabilistic beliefs, where first-period signals are also relevant for second-period beliefs. The information is embedded in memorable contexts that potentially facilitate associative recall. These contexts are intrinsically distinct from the signals itself, which allows us to identify the causal effect of associations above and beyond the normatively relevant informational content of the signals. This latter aspect is a clear advantage of a lab experiment relative to field contexts, where contexts and signals are often inseparably intertwined.

Signal structure. The objective type of each company, G(ood) or B(ad), is independently drawn according to $P(G) = P(B) = 0.5$. Subjects receive potentially multiple binary signals that can be positive, p, or negative, n. The signal diagnosticity is given by $P(p|G) = P(n|B) = 0.65$. In the experiment, good companies are represented by a box that comprises 65 positive and 35 negative news, while bad companies are represented by a box that comprises 35 positive and 65 negative news (see Appendix D for a picture). The computer draws at random from these boxes.

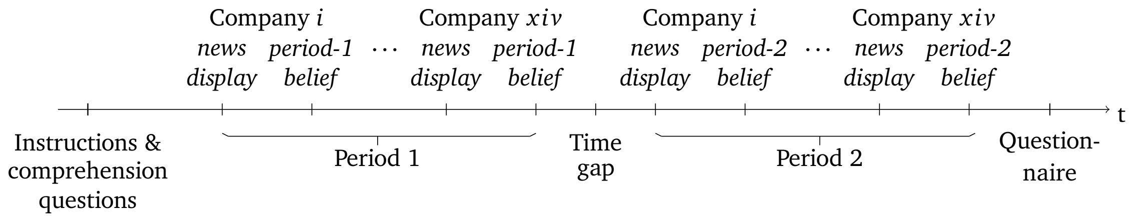
First period. In the first period, subjects complete the updating task for each of the 14 companies sequentially and in random order. For a given company f, subjects first observe k_f i.i.d signals on separate screens, with $k_f \in \{0, \ldots, 4\}$.[5] On a final screen directly thereafter, subjects state their first-period posterior belief about whether the company is good (0–100%). The same procedure is repeated for all companies.

After the first period, we implement a time gap in which subjects work on an unrelated real effort task, which requires subjects to type multiple combinations of letters and numbers into the keyboard. Subjects have 8 minutes to type in as many combinations as they can. For each correctly solved task, they receive 5 cents.

Second period. In the second period, subjects are again tasked with stating probabilistic beliefs about whether each of the 14 companies is of a good or a bad type. The true

[5]For each subject. $k = 0, 1, 3$ for two companies each and $k = 2, 4$ for four companies each.

Figure 1: Experimental Timeline



state for each company is the same as in the first period, such that all first-period signals are still relevant in the second period.

For each company, subjects receive one additional signal and immediately after state their second-period posterior belief. This procedure is repeated for each company, in random order. The experimental instructions and comprehension checks emphasize that first- and second-period signals are equally relevant for second-period beliefs.

To summarize, as depicted in Figure 1, the timeline of the experiment is as follows. Initially, subjects receive instructions and complete comprehension checks. This material covers both periods. In the first period of the updating task, a subject first receives all first-period signals for a company and immediately after states a first-period belief. Then, the subject receives all first-period signals for the next company and states a first-period belief. This process is repeated for all 14 companies, after which an 8-minutes real effort task follows. Then, the subject receives a second-period signal for a company and immediately after states a second-period belief. This procedure is then again repeated for all 14 companies.

Communication of news and contexts. Signals are communicated on subjects' computer screens, one per screen. The signal itself is communicated as "The news for company [X] is positive [negative]." In addition, this signal is embedded in an intrinsically uninformative context. In our experiment, we implement these contexts as events that explain the occurrence of positive or negative news. We chose this implementation of contexts because it is arguably intuitive for participants: all that happens is that they do not just receive an abstract piece of information about whether the company is good or bad, but that the computer also explains to them why the news are positive or negative. Our treatment variation (to be explained below) manipulates how events are linked to signals, as captured by the parameter $\mathbb{T}$ in the model.

All events are represented by a story and an image. For example, a positive signal may be shown along with a story about a successful hire and a picture of the new employee. Another example is a positive signal that is communicated to subjects with a short story about a successful advertising campaign with a celebrity, along with a picture of the celebrity. All stories were constructed to be of similar length and structure. See Appendix Figures 7 and 8 for examples.

The written instructions clarify to subjects that the images and stories have no purpose other than to provide a rationale for the positive or negative news. Conditional on the signal ("positive news" or "negative news"), they are uninformative about the true state of a company. The signal, image and story are displayed on subjects' computer screens for 15 seconds. The time was calibrated such that subjects had sufficient time to process the news, as well as to fully grasp the content of the image and the story.

3.2 Sources of Exogenous Variation

***Cue* and *NoCue* companies.** To exogenously manipulate the presence of associations ($\mathbb{T}$ in the model), our design employs a within-subjects treatment variation. For each subject, seven companies are assigned to be in the *Cue* condition, while the remaining seven companies are assigned to the *NoCue* condition. To counterbalance potential differences in news events across companies, the treatment assignment of companies was randomized across subjects, such that any given company was a *Cue* company for some and a *NoCue* company for other subjects.

The only difference between *Cue* and *NoCue* companies is the mapping between events (contexts) and signals. In *Cue*, every positive signal for a given company is communicated with the same story and image, and every negative signal for a given company is communicated with the same (but different) story and image. For example, if a subject received three positive and one negative signals for a company, then all three positive signals would be communicated with the same story and image, and the negative signal with a different story and image. Thus, for these companies, the second-period signal potentially triggers associative recall of congruent first-period signals through the identical contexts.

For *NoCue* companies, on the other hand, each piece of news is communicated with a unique context. Any given image and story never appear twice, even if the company and type of news are identical. Continuing the above example, if a subject received three positive and one negative signals for a company, then each signal would be communicated with a different story and image.

Subjects did not know ex ante which company was in the *Cue* or *NoCue* condition. In fact, subjects did not even know they were being subjected to a within-subjects treatment variation. Instead, we simply instructed them that the events that generate positive and negative news can potentially occur multiple times. For instance, the negative event that a factory burns down can occur multiple times and cause multiple negative news that should all (independently) be taken into account. The instructions emphasized that while the same event can occur multiple times, it can only occur for the same company. Likewise, we emphasized that the same event can only be associated with positive or negative news but never with both. Thus, subjects knew that if the second-period signal for company A is negative because a factory burned down, and the subject remembers having read this story before, then they know that they must have received at least one negative signal about company A in the first period.

We further instructed subjects to treat each piece of news as independent and in an identical fashion, regardless of which events are associated with these news. For instance, we emphasized that the following two signal histories are equally informative: (i) three positive news about a company, all of which are triggered by the same event and (ii) three positive news about a company, each of which is triggered by a different event. We verified subjects' understanding of the intrinsic irrelevance of whether the same event occurs repeatedly through comprehension questions (see Appendix D.3.2).

In summary, a within-subjects treatment design is particularly natural in our context because the entire treatment variation boils down to whether, for a given company, a subject receives multiple signals that are triggered by the same event or by multiple different events. The order of companies was randomized at the subject level, such that subjects (unknowingly) repeatedly alternated between *Cue* and *NoCue* companies. Thus, potential order or contrast effects – sometimes a concern in within-subjects designs – are implausible in our context.

Signal histories. On top of the within-subjects-across-company variation in the presence of associations, we also causally identify the role of associative recall by exogenously varying the number (and realizations) of first-period signals at the subject-company level. We leverage this source of exogenous variation to test the predictions derived in Section 2 about how the presence or magnitude of overreaction depends on the number of congruent first-period signals. This layer of randomization is directly built into the design because (i) the number of first-period signals for each company randomly varies between one and four, and (ii) conditional on the number of signals, both first- and second-period signals are randomly generated.

Interpretation of treatment comparison. Our *Cue* condition is admittedly extreme in the sense that signals and contexts are perfectly correlated. We chose this implementation to keep the experimental design as simple and transparent as possible. While in reality people likely do not repeatedly experience the same signals in *exactly* the same context, the *Cue* condition is arguably more reflective of reality than the *NoCue* condition. This is because in many contexts similar signals will be associated with similar contexts. For example, whenever good news prevail in the stock market, people are disproportionately exposed to bulls, upward-sloping trend lines, and good-times stories (see, e.g., Shiller, 2019).

3.3 Incentives

Subjects stated their beliefs about whether a company is good vs. bad using a slider (0–100%). Beliefs were incentivized using a binarized scoring rule (Hossain and Okui, 2013). Under this scoring rule, subjects could potentially earn a prize of 10 euros.[6] The probability of receiving the prize is given by $p = 1-(b-t)^2$, where b is the belief that a company is good and t the truth.[7] In order to avoid hedging motives, at the end of the experiment one of the 28 beliefs was randomly selected for payment. Since second-period beliefs are our main outcome measure, we incentivized them more heavily, in expectation: with 90% probability a second-period belief was randomly selected for payment, and with 10% probability a first-period belief.

3.4 Serial Independence of Signals

A key element of the theoretical framework in Section 2 and our experimental design is that signals are conditionally independent. Under serial dependence (positive autocorrelation), it would be "rational" for subjects who forgot the first-period signals to "overreact" to the second-period signal even without any associative recall, simply because they would rationally infer from a positive second-period signal that the first-period signals were likely positive. This would be a potential concern for our design (only) if subjects assumed a higher level of autocorrelation for the *Cue* than the *NoCue* companies.

To address such concerns, we took two steps. First, the instructions used intuitive language to emphasize that the signals are serially (conditionally) independent. We augmented these explanations with a comprehension check question that specifically asked subjects whether a positive signal becomes more likely after a positive signal was drawn.

Second, an account of overreaction that is based on assumed autocorrelation does not generate the additional prediction that overreaction depends in nuanced ways on the signal history. This is because assumed autocorrelation predicts that subjects always infer from a positive second-period signal that the first-period signals were likely also positive, irrespective of the actual realizations of the first-period signals. In contrast, our model

[6] Recent experimental work finds that the presence of cognitive biases is generally robust to the stake size employed (Enke et al., 2021b).

[7] Danz et al. (2022) provide evidence that the binarized scoring rule can lead to a tendency to state less extreme beliefs. Even if such bias was present in our experiment, it would not confound our causal identification, which holds the belief elicitation constant between treatments. If anything, it would lead to an under-estimation of the effect of associative recall on overreaction.

predicts that overreaction depends on the company-specific *random realizations* of the first-period signals. It seems implausible that subjects mentally impute (and remember) different degrees of autocorrelation for each company based on the first-period signals, especially given how salient our instructions are about the absence of autocorrelation.

3.5 Relationship to Psychology Paradigms

Our experimental design builds on the main ideas of a well-known paradigm in memory research, namely lists of word-pairs (Kahana, 2012). Subjects first sequentially observe word pairs, consisting of a "target" and a "cue". At a later stage, subjects' recall of target words is greater when they are provided with the cue word during recall elicitation (Tulving and Thomson, 1973). The analogy to our experimental design is that the signal is the target and the context serves as cue.

The technique we use to generate partial "forgetting" of first-period signals is a variant of the word-pairs paradigm that is called AB/AC in the psychology literature (see chapters 4–5 in Kahana, 2012). Subjects first memorize word pairs ("A" and "B"). Then, in a second step, they memorize new word pairs, some of which involve one of the words from the first set ("A" and "C"). The main finding is that recall of the A-B pair is significantly impaired after subjects learn the A-C pair. This is commonly referred to as "interference". Building on this paradigm, our experimental design creates partial forgetting of first-period signals through: (i) a time lag (distraction task) and (ii) interference that results from the presence of 14 companies with identical news ("positive" and "negative"). We believe the distinction between whether interference and / or the time lag generates forgetting is not of first-order concern for economists because in real life any meaningful time lag will usually induce new experiences that, in turn, also contribute to interference. See Appendix A for a more detailed discussion of the psychology literature on associative recall.

3.6 Econometric Specifications and Predictions

Following the theoretical framework in Section 2, for most analyses we transform subjects' raw beliefs into normalized log posterior odds. Equation (5) directly suggests the following estimating equation for a potential treatment difference in subject i's normalized second-period log posterior odds about whether company f is good:

$$lpo^{i,f} = \beta_1 s_{k+1}^{i,f} + \beta_2 s_{k+1}^{i,f} T^{i,f} + \beta_3 T^{i,f} + \beta_4 (N_p^{i,f} - N_n^{i,f}) + \epsilon^{i,f} \tag{6}$$

where $T^{i,f}$ is a binary treatment indicator that equals one if, for a given subject, company f was in the *Cue* condition. In words, we regress a subject's normalized log posterior odds on the second-period signal, a treatment indicator, their interaction and the net number of positive first-period signals. We predict that the interaction effect is positive, $\beta_2 > 0$, and that this positive interaction effect is only driven by cases with $z > 0$.

Furthermore, within the set of *Cue* companies, we test how overreaction depends on the number of cued first-period signals:

$$lpo^{i,f} = \beta_5 s^{i,f}_{k+1} + \beta_6 s^{i,f}_{k+1} z^{i,f} + \beta_7 z^{i,f} + \beta_8 (N^{i,f}_p - N^{i,f}_n) + \epsilon^{i,f} \tag{7}$$

Here, we predict a form of history-dependence, which is that the interaction between the second-period signal and the number of cued first-period signals is positive: $\beta_6 > 0$.

3.7 Procedures and Logistics

In total, we implemented three experiments, all of which were conducted at the same time and randomized within experimental sessions. The first experiment is the one described above, which we refer to as *Beliefs* experiment. In addition, we also implemented a *Recall* experiment (to directly elicit which first-period signals subjects remember in the second period) and a *Market* experiment (to study market behavior based on associative recall). We discuss these additional experiments in Sections 4.4 and 5, respectively.

All three experiments were conducted as Zoom online experiments based on the subject pool of the BonnEconLab of the University of Bonn. The experiments were computerized using Qualtrics and lasted up to 90 minutes. Subjects met with an experimenter via Zoom and received a participation link to the experimental software via Zoom chat. Subjects were told not to use any material (such as pen and paper) during the experiment. Appendix D contains the full set of instructions, translated into English. Subjects were given unlimited time to read the instructions and could ask questions at any point in time using the Zoom chat.

After subjects finished the instructions, they completed computerized comprehension check questions, see Appendix D. Whenever a subject did not solve a control question correctly, a computer screen pointed out the mistake and explained the correct solution. As we pre-registered, we exclude subjects that answered more than one comprehension check question incorrectly (5% of potential participants). As we pre-specified, 100 valid completes were collected for the *Beliefs* experiment. Average earnings were 16.50 euros, which includes a participation payment of 10 euros.

All experiments in this paper were pre-registered in the AEA RCT registry, see `https://www.socialscienceregistry.org/trials/9215`. The pre-registration includes the design of all experiments reported in this paper, predictions, sample sizes, and that subjects would be dropped from the sample (and replaced) if they answer more than one comprehension check question incorrectly.

4 Results

4.1 Preliminaries: First-Period Beliefs

Model equation (2) posits that the normalized log posterior odds in the first period move one-for-one with variation in the first-period signals. Appendix Table 3 shows that, in our data, this is indeed the case, for both *Cue* and *NoCue* companies. While our treatment comparions of second-period beliefs do not hinge on first-period beliefs being close to Bayesian, this piece of information is helpful because it shows that it is largely irrelevant whether our second-period regressions control for first-period log posterior odds or the number of positive minus negative first-period signals (see eq. (5)).

Our treatment comparison of second-period beliefs is, however, only valid if first-period beliefs do not differ from each other across treatments in a way that would spuriously generate a treatment difference also in second-period beliefs. While the experimental design offers no ex-ante reason for why first-period beliefs should differ across *Cue* and *NoCue* companies, Appendix Table 4 formally tests this. Reassuringly, we find that the difference in first-period beliefs across treatments is very small and statistically insignificant.

4.2 Second-Period Beliefs: A Look at the Raw Data

As derived in Section 2, an immediate implication of associative recall is that second-period beliefs are on average more extreme (further away from 50%) in the presence of associations. Figure 2 provides a first test of this, by showing kernel density plots of the distribution of second-period beliefs, separately for *Cue* and *NoCue* companies. The left panel shows beliefs following a negative second-period signal, while the right panel shows beliefs following a positive second-period signal. Recall that for each subject-company combination the signal realizations in the first and second period were randomly generated. The belief heterogeneity in Figure 2 therefore captures a combination of (i) variation in first-period signal realizations and (ii) variation in beliefs across subjects conditional on the same signal realizations.

We see that beliefs in *Cue* are substantially more extreme, following both a positive and a negative second-period signal. While we conduct more sophisticated regression analyses below, we note that this treatment difference in average beliefs is statistically highly significant in both panels, see Appendix Table 4.

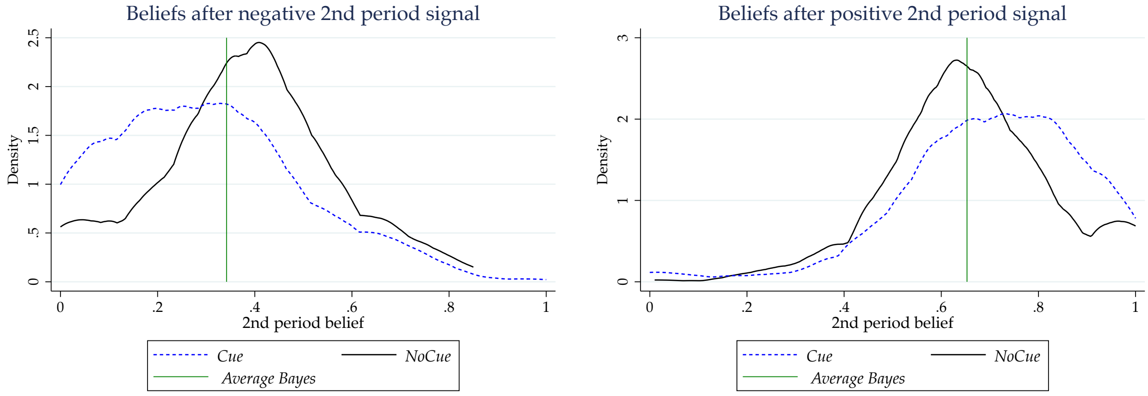


Figure 2: Kernel density estimates of second-period beliefs as a function of treatment and second-period signal. The horizontal green line indicates the average Bayesian posterior across all belief formation problems. Kernel is Epanechnikov.

4.3 Econometric Analysis

A main advantage of visualizing the data as in Figure 2 is that the analysis is very transparent as it does not require any transformations of the raw data. A disadvantage of working with the raw beliefs data, however, is that it does not allow for quantitative analyses of overreaction in which empirical results can be compared against normative benchmarks. The reason is that over- vs. underreaction is typically defined through the Grether (1980) regressions that our theoretical framework also directly motivates.[8]

As shown in equation (6), these Grether regressions relate the (normalized) log posterior odds to the second-period signal, controlling for the number of positive minus negative signals in the first period. Figure 3 visualizes the results of these OLS regressions by displaying the coefficient of the second-period signal. Recall that the second-period posterior log odds are transformed such that the Bayesian benchmark coefficient is one. We conduct this analysis separately by condition and by looking at random variation in the number of first-period signals that are congruent with the second-period signal. The figure shows point estimates along with 95% confidence intervals.

As predicted by the theoretical framework, we observe three patterns. First, when no first-period signal is congruent with the second-period signal ($z = 0$), subjects state identical beliefs across treatments. Second, for any strictly positive number of congruent first-period signals ($z > 0$), the effect of the second-period signal is significantly larger in *Cue* than in *NoCue*. This documents that associations generate overreaction.

[8]As is well-known in the literature, a slight challenge in directly estimating Grether regressions on real data is that people occasionally state beliefs of 0% or 100%, which makes them undefined under the log odds transformation. In our data, this is the case for 93 second-period beliefs (6.6% of all data). To avoid a loss of observations, we recode observations of 0% as 1% and 100% as 99%. Appendix Table 5 shows that our results are quantitatively virtually identical if we do not replace these observations but instead lose them through the log odds transformation.

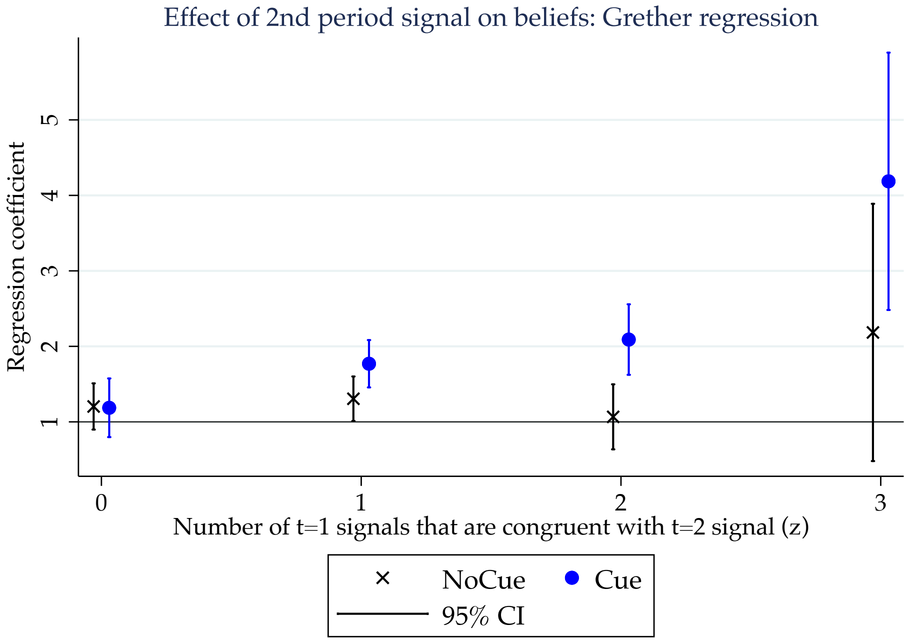


Figure 3: Effect of second-period signal on (normalized) second-period log posterior odds, as a function of the number of congruent first-period signals in *Cue* and *NoCue*. The point estimates are derived from the OLS regression equation (6), which is run separately for each value of z. The figure plots $\hat{\beta}_1$ for *NoCue* and $\hat{\beta}_1 + \hat{\beta}_3$ for *Cue*. The figure does not include $z = 4$ because there are very few observations with such a signal history. All regression analyses that are reported in tables include these cases. Whiskers show 95% confidence intervals, computed based on clustering at the subject level.

Third, looking within treatment *Cue*, the effect of the second-period signal monotonically increases in the number of congruent signals. Note that – consistent with the results from the *NoCue* condition – in the absence of associative recall the regression coefficient should not at all depend on the first-period signal history. Thus, this third result provides evidence for a form of history dependence that is absent in standard economic models of belief formation without associative recall.

Table 1 provides the regression results. The table notes contain detailed explanations about the construction of each variable, where the construction always follows the logic from the model in Section 2. The regressions again directly correspond to the estimating equations (6) and (7) that our model motivates. We construct the table such that the most relevant independent variables are listed at the top. First, columns (1)–(3) focus on across-treatment differences. Column (1) shows that, in the full sample of 1,400 second-period beliefs (100 subjects, 14 companies each), the effect of the second-period signal is indeed significantly larger in *Cue* than in *NoCue*. Columns (2) and (3) decompose this treatment difference into cases with $z = 0$ and $z > 0$, where the model only predicts a treatment difference in the latter case. Consistent with the visual impression from Figure 3, this is indeed what the regressions show. In terms of quantitative magnitude, column (3) shows that in the theoretically-relevant case with $z > 0$, the effect of the second-period signal is more than 80% larger for the *Cue* companies.

Columns (4) and (5) report regression results that only leverage variation in the number of congruent signals within the *Cue* condition. Here, the coefficient of interest is the interaction between the second-period signal and the number of congruent first-period signals. Consistent with what we saw in Figure 3, both regression specifications show that the effect of the second-period signal is significantly stronger when there are more congruent signals. This result on history-dependence of beliefs holds both when we control for the first-period signals (column (4)) and when we directly control for the subject's normalized first-period log posterior odds (column (5)).[9]

Result 1. *Overreaction in beliefs is significantly larger in Cue than in NoCue. This treatment difference only exists when the number of congruent first-period signals is strictly positive.*

Result 2. *Within condition Cue, overreaction increases significantly in the number of congruent first-period signals.*

4.4 Mechanism: Asymmetric Recall of Cued Signals

The model in Section 2 posits that the overreaction of beliefs reflects that subjects in *Cue* asymmetrically remember those first-period signals that are congruent with the second-period signal. In this subsection, we provide two pieces of causal evidence that the treatment difference between *Cue* and *NoCue* indeed reflects the asymmetric retrieval of specific signals that do / do not get cued by the second-period contexts.

Differential responsiveness of beliefs to congruent and incongruent signals. Re-consider the model in Section 2. Because our model of asymmetric recall focuses on whether or not a first-period signal is congruent with the second-period signal, it is useful to define by $N^z := zs_{k+1}$ the sum of congruent first-period signals and by $N^u := -(k-z)s_{k+1}$ the sum of incongruent first-period signals. Note that one of these quantities is positive, while the other is negative. To take a simple example, suppose that for a given company a subject observed three positive and one negative first-period signals and then a negative second-period signal. In this case, the sum of congruent first-period signals is (-1) and the incongruent sum is three.

[9]Our model posits that what drives the magnitude of overreaction is the number of congruent first-period signals, irrespective of the specific order in which signals were received. For example, from the perspective of the model, signal histories of pos-pos-neg and neg-pos-pos are identical. Appendix Table 6 provides a tentative analysis of this issue. While these analyses generally suffer from very low power (because there is a large number of distinct possible signal histories), the results are indicative that the order of signals indeed does not affect the magnitude of overreaction.

Table 1: Overreaction in *Cue* and *NoCue*

	Dependent variable: 2nd period normalized log posterior odds						
	Cue vs. *NoCue*			*Cue*		*Cue* vs. *NoCue*	
Sample:	Full	$z = 0$	$z > 0$	Full	Full	Full	k > z > 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
t=2 signal	1.20***	1.20***	1.13***	1.42***	1.40***	1.29***	1.38***
	(0.14)	(0.15)	(0.15)	(0.18)	(0.14)	(0.16)	(0.27)
t=2 signal × 1 if *Cue*	0.61***	-0.017	0.88***				
	(0.15)	(0.17)	(0.19)				
t=2 signal × # of congruent t=1 signals				0.32**	0.35***		
				(0.13)	(0.10)		
Sum of congruent t=1 signals × 1 if *Cue*						0.50***	0.64***
						(0.11)	(0.15)
Sum of incongruent t=1 signals × 1 if *Cue*						0.086	0.19
						(0.08)	(0.12)
1 if *Cue*	-0.097	0.0093	-0.16			-0.10	-0.14
	(0.12)	(0.14)	(0.12)			(0.11)	(0.14)
Sum of t=1 signals (pos. minus neg.)	0.42***	0.26***	0.41***	0.40***			
	(0.05)	(0.09)	(0.06)	(0.08)			
# of congruent t=1 signals				-0.074	-0.068		
				(0.06)	(0.05)		
t=1 normalized log posterior odds					0.38***		
					(0.07)		
Sum of congruent t=1 signals						0.26***	0.27**
						(0.07)	(0.13)
Sum of incongruent t=1 signals						0.28***	0.35***
						(0.09)	(0.11)
Observations	1400	418	982	700	700	1400	630
Adjusted R^2	0.41	0.20	0.47	0.49	0.54	0.42	0.46

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Following the estimating equations (6) and (7) that we derived from the model, the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. Both the first- and the second-period log posterior odds are normalized by the log diagnosticity odds as described by equation (2). In column (7), the sample is restricted to signal histories where $k > z > 0$, i.e., (i) with at least one congruent first-period signal and (ii) at least one incongruent first-period signal. Variable labels: "t=2 signal" equals 1 if signal positive and (-1) if negative. "# of (in)congruent t=1 signals" captures the number of 1st period signals that do (don't) equal the second-period signal. "Sum of (in)congruent t=1 signals" captures the number of positive minus negative 1st period signals that are (in)congruent with the 2nd period signal. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Using this notation, the main model equation (5) can equivalently be expressed as

$$lpo_2^{assoc} = s_{k+1} + \sum_{x=1}^{k} \hat{s}_x = s_{k+1} + [r + (1-r)a\mathbb{T}]N^z + rN^u + \epsilon \tag{8}$$

This alternative expression for the Grether decomposition is helpful because – unlike the regressions reported above – it includes the sum of congruent and the sum of incongruent first-period signals as separate regressors. Crucially, the straightforward implication of associative recall in our model is that the associations boost is asymmetric and only applies to the congruent-first period signals rather than all first-period signals. This implication of our model is distinctively different from a potential alternative model, according to which a second-period context (if it was also experienced in the first period) cues an improved recall of *all* first-period signals for a company, regardless of whether they are congruent or incongruent.

To test this, column (6) of Table 1 reports the results of a regression that interacts the sums of congruent and incongruent first-period signals separately with a treatment indicator. We find that the interaction effect of our treatment dummy ($= 1$ if *Cue*) with the sum of congruent first-period signals is quantitatively very large (twice as large as the raw coefficient of the congruent signals) and statistically highly significant. Meanwhile, the interaction effect with the sum of incongruent first-period signals is statistically insignificant and the point estimate close to zero. This shows that subjects' beliefs about *Cue* companies overreact to congruent first-period signals, as predicted by equation (8). Column (7) of Table 1 shows this result also holds when we restrict attention to cases where at least one first-period signal was congruent and at least one incongruent with the second-period signal ($k > z > 0$). This is a useful robustness check because in these cases the second-period signal definitely acts as a cue for either positive or negative signals (in column (6) we also include cases where potentially no first-period signal gets cued, such that treatment *Cue* cannot have an effect).

In summary, the results from columns (7) and (8) clarify that the entire treatment difference between *Cue* and *NoCue* is driven by an asymmetric responsiveness to congruent first-period signals, rather than an improved overall responsiveness to first-period signals in *Cue*.

Differential recall of congruent and incongruent signals. A second test of our mechanism is to directly gather data on which first-period signals subjects remember. According to the model, subjects should remember more congruent (but not incongruent) signals in *Cue* than in *NoCue*.

To test this, we implemented experiment *Recall*, which again randomized companies to treatments *Recall Cue* and *Recall NoCue* within subject. This experiment was randomized within experimental sessions with the *Beliefs* experiments described above. The experiment was identical to the *Beliefs* experiment, except that after receiving a second-period signal, subjects were asked to directly report the number of positive and negative signals they recall for a company. Subjects answered 28 such questions (recall of positive and negative signals for 14 companies each). For a randomly selected recall question, subjects received 10 euros if their answer was within +/-1 of the truth. Except for the recall component, the experiments and the underlying instructions were identical to those in *Beliefs*. To maximize similarity with the *Beliefs* experiment, the initial instructions only explained the belief elicitation task, and that first-period signals would also be relevant for second-period beliefs. Then, after subjects had concluded the first period as well as the distraction task, the recall task was announced as a surprise. As in the *Beliefs* experiment, subjects received one additional signal for each company and immediately after indicated their recall of positive and negative signals. Appendix D provides the experimental instructions. As we pre-registered, 70 subjects participated in this experiment. Average earnings were 18.50 euros, which includes a participation payment of 10 euros.

Figure 4 summarizes the results by reporting subjects' effective recall of first-period signals as a function of the truth, separately for congruent first-period signals (left panel) and incongruent first-period signals (right panel).[10] The left panel shows a large and statistically highly significant treatment difference in the recall of congruent signals: subjects remember substantially more congruent first-period signals for *Cue* companies than for *NoCue* ones. In contrast, the right panel shows that for incongruent first-period signals (those that differ in signal type from the second-period one), this treatment difference is much smaller and not statistically significant. Indeed, Appendix Table 7 shows that the relevant difference-in-difference effect (treatment condition times congruent / incongruent signals) is statistically highly significant. These results again show that the associations that are present in *Cue* primarily induce *asymmetric* recall of congruent first-period signals rather than improved recall in general.[11]

Result 3. *Overreaction in second-period beliefs is driven by asymmetric recall of those first-period signals that get cued by the second-period context.*

5 Betting Market Experiment

5.1 Design

The basic structure of the *Market* experiment is identical to the *Beliefs* experiment, except that the belief elicitation task is embedded in a parimutuel betting market. While the canonical application of parimutuel markets is horse race betting, there are direct analogies to financial markets, where betters bet on mutually exclusive states of the world,

[10]In our experiments, we elicited subjects' *total* recall of signals in the entire experiment, including of those in the second period. For example, in cases in which we elicit recall of positive signals and the subject observed a positive second-period signal, effective recall of first-period signals is given by the reported recall minus one. This corresponds to the arguably very plausible assumption that subjects do not forget the second-period signal that they saw a few seconds ago on the previous screen.

[11]Appendix Figure 9 shows that we replicate this recall pattern also when we restrict attention to cases in which at least one first-period signal is congruent with and at least one incongruent with the second-period signal, analogous to column (7) of Table 1.

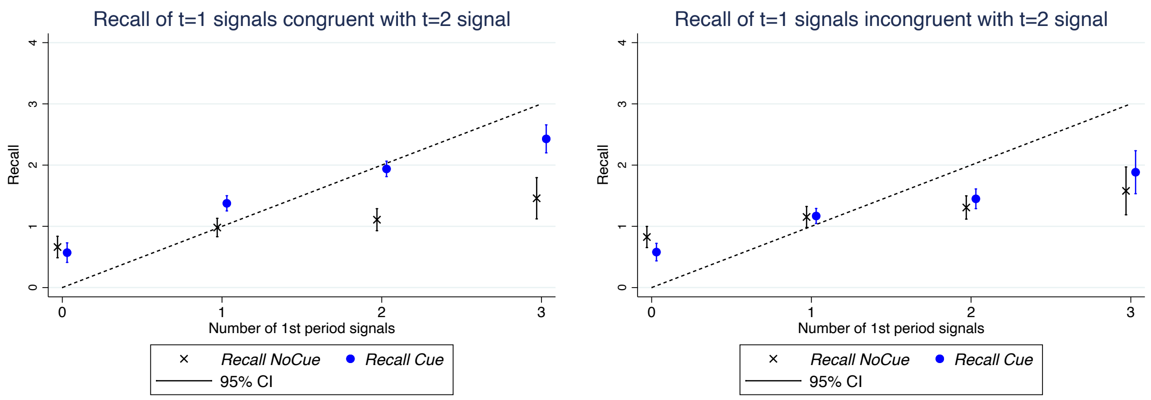


Figure 4: Effective recall of congruent (left panel) and incongruent (right panel) first-period signals in *Recall Cue* and *Recall NoCue*. Effective recall equals reported recall for signals that differ from the second-period signal, and reported recall minus one for signals that equal it. The point estimates stem from an OLS regression of effective recall on a treatment dummy. The figures plot the coefficient of the constant for *NoCue* and the sum of the coefficients of the constant and the treatment dummy for *Cue*. The figure does not include the case of four first-period signals because there are very few observations with such a signal history. Whiskers show 95% confidence intervals, computed based on clustering at the subject level.

such as whether an asset will increase or decrease in value. Indeed, parimutuel betting markets are frequently implemented in laboratory experiments because of their simplicity and appealing resemblance of real-world markets (e.g., Plott et al., 2003; Kendall and Oprea, 2018; Enke et al., 2021a).

Parimutuel betting and payoffs. In our implementation, subjects are again asked to state probabilistic beliefs about whether each of 14 hypothetical companies is of a good type, after receiving a series of binary signals. The prior probabilities and signal structure are identical to those in the *Beliefs* experiment. Subjects are matched into groups of three and know that all participants in their market group receive the same public signals. In both part 1 and part 2 of the experiment, after observing signals as in the *Beliefs* experiment, each market participant privately states their subjective percent chance that the company is good. Immediately after, in both parts of the experiment, the three subjects interact in a parimutuel betting market. For each of their 28 betting decisions (14 companies and two periods each), subjects receive a budget of 10 euros. This money can be fully or partly bet on one or both of two propositions: that the company is good and that it is bad. The bets are implemented in two steps (on the same decision screen):

1. *Betting amount:* Subjects state the total amount they want to bet (maximum 10 euros, minimum 0 euros). We denote this amount by m.

2. *Betting proportion:* Subjects state which fraction (denoted by w) of m they bet on the proposition that the company is good. Thus, the total amount bet on the event

that the company is good is wm and the total amount bet on the event that the company is bad is $(1-w)m$.

With subjects $i = 1, 2, 3$, the parimutuel market price in period t in group g for the asset that company f is good is defined as:[12]

$$\omega_t^{g,f} = \frac{\sum_{i=1}^{3} w_t^{i,g,f} m_t^{i,g,f}}{\sum_{i=1}^{3} m_t^{i,g,f}} \in [0,1] \tag{9}$$

This has a simple interpretation, according to which the price is given by a weighted average of the betting proportions, where the weights are given by the betting amounts. Thus, the parimutuel price for an asset increases in the fraction of the total money in the market that is bet on the respective state. To intuitively relate this market price back to subjects' beliefs, consider the hypothetical scenario that the betting proportion for the good state is directly given by each subject's belief that the company is good (under expected utility, this will be the case when utility is $u(x) = ln(x)$, see Wolfers and Zitzewitz, 2006). Under this scenario, the market price would be given by the weighted average belief, where the weights are given by how much each subject bets.

The payoffs of subject i in period t in market group g for company f are given by

$$\pi_t^{i,g,f} = (10 - m_t^{i,g,f}) + \frac{w_t^{i,g,f} m_t^{i,g,f}}{\omega_t^{g,f}} \mathbb{1}_{\theta=G} + \frac{(1-w_t^{i,g,f}) m_t^{i,g,f}}{(1-\omega_t^{g,f})} \mathbb{1}_{\theta=B} \tag{10}$$

In words, the subject keeps the part of the endowment that is not bet. In addition, the subject loses all money bet on the wrong state. Money bet on the right state yields a positive return whose magnitude depends on the market price (9). Intuitively, the subject earns more money the more the subject bets on the right state and the more other subjects bet on the wrong state. The parimutuel price mechanism in our implementation fully redistributes all money that is bet; there is no efficiency loss or transaction cost.

Treatments and randomization. Analogously to the *Beliefs* experiment, we conduct two treatment conditions (within-subject-across-companies) that exogenously manipulate the presence of associations. For companies in *Market Cue*, each positive / negative signal is again associated with the same context / event. In contrast, for companies in *Market NoCue,* each signal is communicated with a different event. As in the individual belief elicitation treatments, for each subject the computer randomly selected seven companies to be in *Cue* and seven to be in *NoCue*. In addition to the across-treatment

[12]When none of the subjects in a market group bets money in a given round, the market price is missing. This occurs in only two out of 1,120 second-period observations.

variation in the relevance of associative recall, the experiment again features random variation in the number of cued first-period signals (z).

Logistics and payoffs. As in the *Beliefs* experiment, there are a total of 14 hypothetical companies. The number of signals subjects see in part 1, the signal realizations as well as the order in which subjects see the companies is fully randomized across market groups. To avoid hedging, at the end of the experiment, for each subject, one of the two parts of the experiment and one company are randomly selected to be payout-relevant. For the randomly-selected company and part, either the subject's belief or the betting decision are randomly chosen and implemented for payment.[13]

The *Market* experiment was randomized within experimental sessions with the *Beliefs* experiment. The procedures and the subject pool (BonnEconLab, conducted over Zoom) were identical. The sample size was 240 subjects (80 groups). Average earnings were 19.80 euros, which includes a participation payment of 10 euros. Subjects remained in the same group throughout the experiment. No feedback was provided at any point. The *Market* experiment was also part of the pre-registration mentioned previously, including the experimental design, predictions and sample size.

5.2 Predictions

As is well-known, depending on assumptions on utility functions, betting market prices need not necessarily reflect average beliefs in the market.[14] Our main interest here, however, is not in understanding how exactly beliefs aggregate to market prices but, instead, in the comparative statics effects regarding the role of associative recall: whether second-period market prices react more strongly to second-period signals in the presence of associations, and how this "overreaction" varies as a function of the signal history. Accordingly, we here only heuristically discuss our pre-registered predictions for the experiment.

To begin, consider again the definition of the parimutuel market price in equation (9). Note that subjects' betting proportion on the proposition that a company is good (w) will usually increase in their belief that the company is good. Then, if associative recall generates overreaction in beliefs, we might also expect to see overreaction in market prices. To take a particularly simple example, suppose again that each subject's betting proportion is directly given by his / her belief. Then, the market price is given by the weighted average belief in the market, and the degree of overreaction in market prices

[13]Individual beliefs were incentivized with the same binarized scoring rule as in Section 3.

[14]For discussions of whether and how betting / pediction markets generally aggregate beliefs, see Manski (2006) and Wolfers and Zitzewitz (2004, 2006).

will be given by the average amount of overreaction in beliefs, weighted by each subject's betting amount.

At the same time, there are also reasons to expect that the degree of associations-driven overreaction in market prices may be attenuated relative to overreaction in beliefs. The reason is that – just like almost all real market environments – our betting market entails a strong element of *self-selection*, which is given by the amount of money a subject is willing to bet, captured by m. In particular, it is conceivable that those people who are more susceptible to associative recall have a loose awareness that their beliefs may be biased (but don't know how specifically, such that they cannot correct for it). If this is the case, then these people may be less inclined to bet aggressively on their beliefs and therefore influence the price less. Thus, heterogeneity in betting amounts that reflects heterogeneity in people's confidence in their belief updating rule "re-weights" individual beliefs as far as the market price is concerned. This re-weighting of beliefs through self-selection is similar to how wealth heterogeneity re-weights individual beliefs in classical models of betting markets (Wolfers and Zitzewitz, 2004).

To illustrate, take the extreme example that out of the three subjects in a market group, one has no memory limitations, while the other two succumb to associative recall. Further suppose that the two associative recall types have sufficient doubts about the rationality of their beliefs that they do not bet at all in the market. Then, average beliefs in the market will exhibit overreaction, but the market price will not, purely as a result of differential self-selection. Of course, by an analogous logic, the market price could also reflect more overreaction than individual beliefs if those subjects that have a stronger tendency for associative recall bet more money in the parimutuel market.

Enke et al. (2021a) study this type of self-selection mechanism in betting markets for various cognitive biases, but they do not consider memory. However, this is important to do because while much psychological research has documented the existence of associative recall, much less is known about whether people are willing to actually act on beliefs that are derived from associative recall, such that they become relevant when multiple individuals interact in markets.

Naturally, our main object of interest in the analysis of market prices will not be the *level* of over- or underreaction (as it could be affected by various considerations of how betting markets aggregate beliefs), but instead the causal effects of the random components of our experimental design: (i) the presence of associations; and (ii) the number of congruent first-period signals. For comparability, we analyze the data from the market experiments using the same methodology as the individual beliefs data. We first transform second-period market prices into normalized log market price odds[15]

[15]Similarly to the beliefs data, when a market price is 0% or 100% we replace it by 1% and 99%, respectively, to avoid a loff of observations from the log odds definition. This occurs in 11 out of 1,118

($lmpo$) in market g for company f following equation (2) and then link these to the second-period signal:

$$lmpo^{g,f} = \tilde{\beta}_1 s_{k+1}^{g,f} + \tilde{\beta}_2 s_{k+1}^{g,f} T^{g,f} + \tilde{\beta}_3 T^{g,f} + \tilde{\beta}_4 (N_p^{g,f} - N_n^{g,f}) + \epsilon^{g,f} \tag{11}$$

where $T^{g,f}$ is a binary treatment indicator that equals one if, for a given market group, company f is in the *Market Cue* condition. Here, we again predict and pre-registered that $\tilde{\beta}_2 > 0$, and that this treatment difference is only driven by cases with $z > 0$.

Furthermore, within the set of *Cue* companies, we again test for an interaction effect of the second-period signal with the number of congruent first-period signals:

$$lmpo^{g,f} = \tilde{\beta}_5 s_{k+1}^{g,f} + \tilde{\beta}_6 s_{k+1}^{g,f} z^{g,f} + \tilde{\beta}_7 z^{g,f} + \tilde{\beta}_8 (N_p^{g,f} - N_n^{g,f}) + \epsilon^{g,f} \tag{12}$$

We pre-registered the prediction that $\tilde{\beta}_6 > 0$.

5.3 Results

Replication of patterns on beliefs. Because we also elicited subjects' beliefs in the market experiments, we can use these data to replicate all patterns from the individual belief elicitation treatments. This is done in Appendix Table 8. The results are almost identical to those reported above: (i) there is more overreaction for *Cue* than *NoCue* companies; (ii) this treatment difference only exists when the number of congruent first-period signals (z) is strictly positive; (iii) overreaction in *Cue* significantly increases in z; and (iv) this overreaction reflects asymmetric recall of cued signals in the sense that the stronger responsiveness of second-period beliefs to first-period signals in *Cue* is only present for congruent signals.

Raw market prices data. Figure 5 shows that, very similarly to the belief elicitation experiments, second-period market prices in *Market Cue* are more extreme than those in *Market NoCue*, following both a positive and a negative second-period signal. Indeed, we see that market prices are typically more extreme than the average Bayesian belief, though we reiterate that our primary interest is the across-treatment comparison rather than the test against the Bayesian point prediction.

Econometric analysis. More formally, we resort to Grether-style regressions in which the dependent variable consists of the (normalized) second-period log market price

cases. We have verified that the results are quantitatively almost identical when we instead drop these observations.

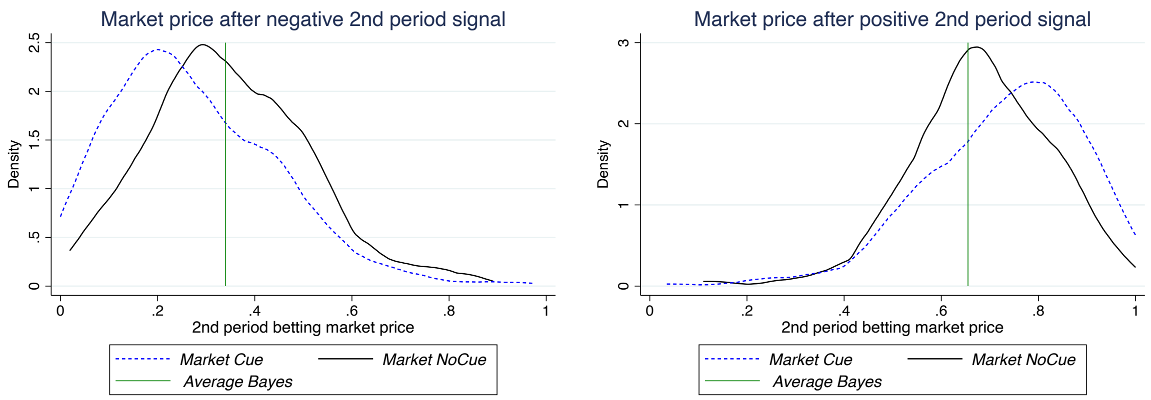


Figure 5: Kernel density estimates of second-period market prices as a function of treatment and second-period signal. The horizontal green line indicates the average Bayesian posterior across all rounds. Kernel is Epanechnikov.

odds. If the market price reflected Bayesian beliefs, the OLS regression coefficient of the second-period signal would equal one.

Figure 6 summarizes the results. Similarly to the belief elicitation experiments, there are three main takeaways. First, when $z > 0$, the responsiveness of second-period market prices to the second-period signal is significantly more pronounced in *Market Cue* than in *Market NoCue*. Second, when $z = 0$, this treatment difference disappears. Third, within *Market Cue*, overreaction of second-period beliefs monotonically increases in the number of congruent first-period signals.

Table 2 provides the regression estimates. The results confirm the statistical significance of the three main patterns that were evident from Figure 6: (i) higher responsiveness of market prices to the second-period signal in *Market Cue* than in *Market NoCue* (column (1)); (ii) this occurs only when $z > 0$ (columns (2)–(3)); and (iii) responsiveness that significantly increases in the number of congruent signals (columns (4)–(5)).

Result 4. *Market prices react significantly more to the second-period signal in Market Cue than in Market NoCue. This treatment difference only exists when the number of congruent first-period signals is strictly positive.*

Result 5. *Within condition Market Cue, the responsiveness of market prices to the second-period signal increases significantly in the number of congruent first-period signals.*

Do markets attenuate associations-based overreaction? In light of the discussion about self-selection in markets potentially attenuating the effect of associative recall on overreaction, it is of interest to compare the quantitative magnitude of associations-driven overreaction in the betting market (Table 2) with that in individual beliefs (Ta-

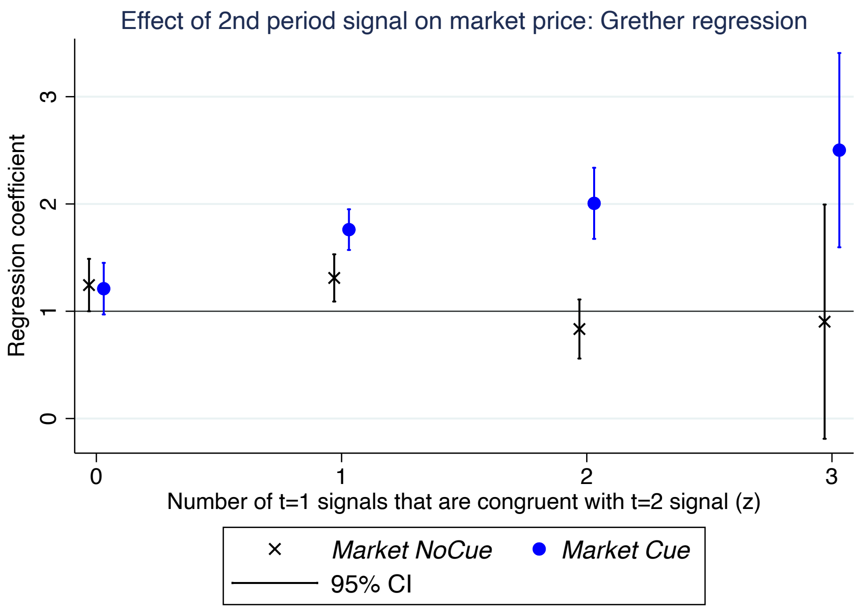


Figure 6: Effect of second-period signal on (normalized) second-period log market price odds, as a function of the number of congruent first-period signals. The point estimates are derived from regression equation 11, which is run separately for each value of z. We do not show $z = 4$ because there are very few observations with such a signal history. The figure plots $\hat{\hat{\beta}}_1$ for *Market NoCue* and $\hat{\hat{\beta}}_1 + \hat{\hat{\beta}}_3$ for *Market Cue*. Whiskers show 95% confidence intervals, computed based on clustering at the market group level.

ble 1).[16] The relevant quantities of interest here are the *causal effects* of the treatment and of the number of cued signals, rather than the *level* of overreaction across treatments. This is because only the causal effects reflect the impacts of associative recall, while the baseline level of overreaction in markets and beliefs may differ for various reasons. Comparing column (1) of Tables 1 and 2, we see that the causal effect of the treatment on overreaction is 0.61 in individual beliefs ($s.e. = 0.15$) and 0.65 in market prices ($s.e. = 0.09$). Similarly, comparing column (4) in Table 1 with column (4) in Table 2, we see that the causal effect of the number of cued signals is 0.32 ($s.e. = 0.10$) in the case of individual beliefs, while it is 0.34 ($s.e. = 0.08$) in the market experiment. These differences are not statistically significant and suggest that subjects who rely on associative recall in forming beliefs do not select out of the market and hence affect market prices.In sum, we believe that these results further underscore the economic relevance of associative memory.

Result 6. *Associative-recall-based overreaction in market prices is as large as overreaction in average individual beliefs, despite the scope for self-selection.*

[16]The *Beliefs* and *Market* experiments were conducted using within-session, random assignment to experiments.

Table 2: Overreaction in market prices as a function of treatment and signal history

	Dependent variable: 2nd period normalized log market price odds				
	Cue vs. *NoCue*			*Cue*	
Sample:	Full	$z = 0$	$z > 0$	Full	Full
	(1)	(2)	(3)	(4)	(5)
t=2 signal	1.17***	1.24***	1.10***	1.39***	1.33***
	(0.07)	(0.12)	(0.08)	(0.11)	(0.10)
t=2 signal × 1 if *Cue*	0.65***	-0.034	0.94***		
	(0.09)	(0.14)	(0.11)		
t=2 signal × # of congruent t=1 signals				0.34***	0.41***
				(0.08)	(0.07)
1 if *Cue*	-0.034	-0.063	-0.050		
	(0.06)	(0.11)	(0.08)		
t=1 signals (pos. minus neg.)	0.45***	0.33***	0.44***	0.47***	
	(0.04)	(0.09)	(0.04)	(0.08)	
# of congruent t=1 signals				-0.026	-0.032
				(0.04)	(0.04)
t=1 normalized log market price odds					0.35***
					(0.05)
Observations	1118	330	788	560	559
Adjusted R^2	0.63	0.37	0.69	0.72	0.73

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the market group level. Following the estimating equations (11) and (12), the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. Both the first- and the second-period log market price odds are normalized by the log diagnosticity odds as described by equation (2). See Table 1 for details on the construction of each variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6 Discussion

Associative recall is one of the most important principles underlying psychological research on memory. Despite growing theoretical interest, experimental and empirical behavioral economists have so far largely neglected the potential role of associative memory in shaping decision-making. By presenting the first set of theory-driven experiments that build a bridge between psychological paradigms on cued recall and structured, quantitative economic decision tasks, this paper has provided a causal analysis of the role of associative memory for belief formation and betting market behavior.

In doing so, we have provided three pieces of evidence that speak to the economic relevance of associative memory in beliefs. First, in contrast to the vast majority of economic models of belief formation, we have seen that associative recall implies a *history-dependence* of beliefs, meaning that – because of asymmetric cueing effects – the precise structure of past signals matters for today's beliefs even conditional on yesterday's beliefs. Our second contribution is to document that this history dependence generates systematic overreaction of beliefs when context and news are correlated in a consistent fashion over time. Third, we have shown that this overreaction in beliefs leads to systematic overreaction of market prices in a betting market environment.

We believe that by offering a new experimental paradigm in which questions related to associative memory can be studied, our paper provides a stepping stone for further experimental research in an agenda on memory imperfections. We here highlight two open questions and links to more applied literatures.

First, our experiments are potentially related to an active literature that documents overreaction in survey expectations about economic variables (e.g., Bordalo et al., 2020b). The result of overreaction in field data is often considered to be a slight puzzle from the perspective of laboratory research on belief formation. This is because structured laboratory belief updating problems almost always find underreaction. However, in these laboratory experiments, memory imperfections are by design ruled out. We do not intend to claim that associative recall can explain the entire pattern of over- and underreaction identified in the literature. However, it is conceivable that *part* of the reason why the laboratory and field literatures identify such different patterns is that memory constraints and memorable contexts likely play a more important role in the field, as exemplified by Shiller's (2017; 2019) discussion of the role of memorable narratives and "cue-dependent forgetting." Second, we conjecture that political economy applications of the role of associate recall are ripe for exploration. For example, valent narratives about immigrants or the costs of taxation could induce people to asymmetrically remember past information that is congruent with current news, and therefore to overreact.

References

Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar, "Overreaction in expectations: Evidence and theory," *Quarterly Journal of Economics*, 2023.

Amelio, Andrea and Florian Zimmermann, "Motivated Memory in Economics – A Review," *Games*, 2023, *14* (1), 15.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart, "Subjective Models of the Macroeconomy: Evidence from Experts and Representative Samples," *Review of Economic Studies*, 02 2022, *89* (6), 2958–2991.

__ **, Ingar Haaland, Christopher Roth, and Johannes Wohlfart**, "Inflation Narratives," *Working Paper*, 2021.

Barron, Kai and Tilman Fries, "Narrative Persuasion," *Working Paper*, 2022.

Bhatia, Sudeep, "Associative judgment and vector space semantics.," *Psychological review*, 2017, *124* (1), 1.

Bodoh-Creed, Aaron, "Mood, memory, and the evaluation of asset prices," *Review of Finance*, 2020.

Bordalo, Pedro, John J Conlon, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer, "Memory and probability," *Quarterly Journal of Economics*, 2023.

__ **, Katherine Coffman, Nicola Gennaioli, Frederik Schwerter, and Andrei Shleifer**, "Memory and representativeness," *Psychological Review*, 2021, *128* (1), 71–85.

__ **, Nicola Gennaioli, and Andrei Shleifer**, "Memory, attention, and choice," *The Quarterly Journal of Economics*, 2020, *135* (3), 1399–1442.

__ **,** __ **, Yueran Ma, and Andrei Shleifer**, "Overreaction in macroeconomic expectations," *American Economic Review*, 2020, *110* (9), 2748–82.

Bursztyn, Leonardo, Georgy Egorov, Ingar Haaland, Aakaash Rao, and Christopher Roth, "Justifying Dissent," *Quarterly Journal of Economics*, 2023. qjad007.

Camerer, Colin and Dan Lovallo, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 1999, *89*, 306–318.

Charles, Constantin, "Memory and Trading," *Working Paper*, 2022.

__ , "Memory Moves Markets," *Available at SSRN 4019728*, 2022.

Danz, David, Lise Vesterlund, and Alistair Wilson, "Belief Elicitation and Behavioral Incentive Compatibility," *American Economic Review*, 2022, *112* (9), 2851–2883.

Dougherty, Michael RP, Charles F Gettys, and Eve E Ogden, "MINERVA-DM: A memory processes model for judgments of likelihood.," *Psychological Review*, 1999, *106* (1), 180.

Ebbinghaus, Hermann, *Über das gedächtnis: untersuchungen zur experimentellen psychologie*, Duncker & Humblot, 1885.

Enke, Benjamin, Thomas Graeber, and Ryan Oprea, "Confidence, Self-Selection and Bias in the Aggregate," *Working Paper*, 2021.

__ , Uri Gneezy, Brian Hall, David Martin, Vadim Nelidov, Theo Offerman, and Jeroen van de Ven, "Cognitive biases: Mistakes or missing stakes?," *The Review of Economics and Statistics*, 2021, pp. 1–45.

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel, "Mental models and learning: The case of base-rate neglect," Technical Report, Tech. rep 2020.

Evers, Ellen and Alex Imas, "Mental Accounting, Similarity, and Preferences Over the Timing of Outcomes," *Working Paper*, 2019.

Fudenberg, Drew, David K Levine et al., "Learning with recency bias," *Proceedings of the National Academy of Sciences*, 2014, *111*, 10826–10829.

Gennaioli, Nicola and Andrei Shleifer, "What Comes to Mind," *Quarterly Journal of Economics*, 2010, *125* (4), 1399–1433.

__ and __ , *A crisis of beliefs: Investor psychology and financial fragility*, Princeton University Press, 2018.

Gilboa, Itzhak and David Schmeidler, "Case-based decision theory," *Quarterly Journal of Economics*, 1995, *110* (3), 605–639.

Goetzmann, William, Dasol Kim, and Robert Shiller, "Crash Narratives," *NBER Working Paper*, 2022.

Graeber, Thomas, Christopher Roth, and Florian Zimmermann, "Stories, Statistics and Memory," *Working Paper*, 2023.

Grether, David M., "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, 1980, *95*, 537–557.

Hartzmark, Samuel M, Samuel Hirshman, and Alex Imas, "Ownership, Learning, and Beliefs," *The Quarterly Journal of Economics*, 2021, *136* (3), 1665–1717.

Hossain, Tanjim and Ryo Okui, "The binarized scoring rule," *Review of Economic Studies*, 2013, *80* (3), 984–1001.

Jiang, Zhengyang, Hongqi Liu, Cameron Peng, and Hongjun Yan, "Investor Memory and Biased Beliefs: Evidence from the Field," *Working Paper*, 2023.

Juslin, Peter and Magnus Persson, "PROBabilities from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge," *Cognitive science*, 2002, *26* (5), 563–607.

Kahana, Michael Jacob, *Foundations of human memory*, OUP USA, 2012.

Kendall, Chad and Ryan Oprea, "Are biased beliefs fit to survive? An experimental test of the market selection hypothesis," *Journal of Economic Theory*, 2018, *176*, 342–371.

Koszegi, Botond, George Loewenstein, and Takeshi Murooka, "Fragile Self-Esteem," *Working Paper*, 2019.

Laibson, David, "A cue-theory of consumption," *The Quarterly Journal of Economics*, 2001, *116* (1), 81–119.

Manski, Charles F, "Interpreting the predictions of prediction markets," *economics letters*, 2006, *91* (3), 425–429.

Mullainathan, Sendhil, "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2002, pp. 735–774.

Noor, Jawwad, "Intuitive Beliefs," *Cowles Foundation Discussion Paper*, 2019.

Plott, Charles R, Jorgen Wit, and Winston C Yang, "Parimutuel betting markets as information aggregation devices: experimental results," *Economic Theory*, 2003, *22* (2), 311–351.

Sanborn, Adam N and Nick Chater, "Bayesian brains without probabilities," *Trends in cognitive sciences*, 2016, *20* (12), 883–893.

Schacter, Daniel L, *Searching for memory: The brain, the mind, and the past*, Basic books, 2008.

Shiller, Robert J, "Narrative economics," *American Economic Review: Papers and Proceedings*, 2017, *107* (4), 967–1004.

— , *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press, 2019.

Tulving, Endel, "Episodic and semantic memory.," 1972.

— **and Donald M Thomson**, "Encoding specificity and retrieval processes in episodic memory.," *Psychological review*, 1973, *80* (5), 352.

Tversky, Amos, "Features of similarity.," *Psychological review*, 1977, *84* (4), 327.

— **and Daniel Kahneman**, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment.," *Psychological review*, 1983, *90* (4), 293.

Underwood, Benton J, "Interference and forgetting," *Psychological review*, 1957, *64* (1), 49–60.

Wachter, Jessica A and Michael J Kahana, "A retrieved-context theory of financial decisions," *Available at SSRN*, 2019.

Wilson, Andrea, "Bounded memory and biases in information processing," *Econometrica*, 2014, *82* (6), 2257–2294.

Wolfers, Justin and Eric Zitzewitz, "Prediction markets," *Journal of economic perspectives*, 2004, *18* (2), 107–126.

— **and** — , "Prediction markets in theory and practice," 2006.

ONLINE APPENDIX

A Relation to Psychological Literature

By investigating the role of associative recall for belief updating and economic decision making, we relate to two literatures in psychology. First, psychological research on episodic memory studies the determinants of whether a temporally-dated and spatially-located personal experience will be remembered or forgotten (Kahana, 2012). Experiences are stored in long-term memory in the form of episodic memories or traces following the processes of encoding and consolidation. The retrieval of traces describes the process in which stored information is recalled (made cognitively present). Since the pioneering work of Ebbinghaus (1885) at the end of the 19th century, recall of target traces is shown to be highly context-dependent. If a context contains cues that are associated with the target trace, recall of the target trace is more likely. A common experimental paradigm to study the associativeness of recall involves lists of word-pairs (Kahana, 2012). Subjects first sequentially observe word pairs, consisting of a “target" word and a “cue” word. Observing the target word and cue word at the same time allows subjects to encode the words and form an association between them. At a later stage, subjects’ recall of target words is greater when subjects were provided with the cue word during recall elicitation (Tulving and Thomson, 1973). Our experimental design builds on this setup to study associative recall. When subjects update their beliefs in the second period of our experiment, they may try to recall first period signals – though we never explicitly ask subjects to recall first-period signals. Successfully recalling these target traces should then depend on whether or not they appear in a context that cues their recall.

A core concept in the psychological literature on memory cues is that of item similarity (Tversky, 1977; Evers and Imas, 2019). In our experiments, we deliberately designed the contexts such that – across signal realizations – they are either identical or very different from each other. In principle, one could easily imagine experiments in which similarity is varied in a more nuanced and continuous fashion.

The second literature in psychology that we relate to deals with how inaccurate probabilistic judgements in vignette-type studies (for instance, Tversky and Kahneman’s (1983) Linda situation) depends on recall. Dougherty et al. (1999); Juslin and Persson (2002); Sanborn and Chater (2016) provide theoretical work on this by looking at noisy recall (either due to noisy encoding or due to noisy retrieval). In addition, Bhatia (2017) presents empirical evidence that semantic memory could explain some of the evidence for the representativeness heuristic.[17] When social stereotypes are disproportionally shared in news/social media, individuals may encode semantic associations involving stereotypes that are later on cued by the specific language of vignettes. Note that our experimental setup rules out that such semantic memory accounts play a role, as associations between signals and context are created within our experiment and they are unlikely part of potential pre-existing semantic associations subjects may hold. Bordalo et al. (2021) also relates to our wok. They present a link between episodic memory and the representativeness heuristic. They conduct an experiment on selective recall that directly induces a memory database (unlike in vignette-type studies) of abstract images and shows that interference affects the similarity between target traces and contextual cues. Interference captures competition between different memory traces in the retrieval process (Underwood, 1957) and hence explains why sometimes target traces cannot be recalled, i.e., they are interfered with by other memory traces.

[17] Semantic memory is another component of long-term memory, distinct from episodic memory, which captures the storage and retrieval of knowledge about words and concepts, and their properties (Tulving, 1972).

B Additional Figures

Company N tries to advertise its products though commercials with German celebrities like, for instance, Boris Becker, Helene Fischer or Til Schweiger. Recently, a new advertisement campaign with a celebrity worked extremely well.

The news for Company N are positive.

Figure 7: Example of good news and corresponding event

Company K is located next to a lake. The local textile industry sometimes pollutes the lake. In that case customers avoid the lake for several weeks and Company K's revenue takes a hit. Recently, a new incident at a local textile company occurred and the water of the lake is polluted again.

The news for company K are negative.

Figure 8: Example of bad news and corresponding event

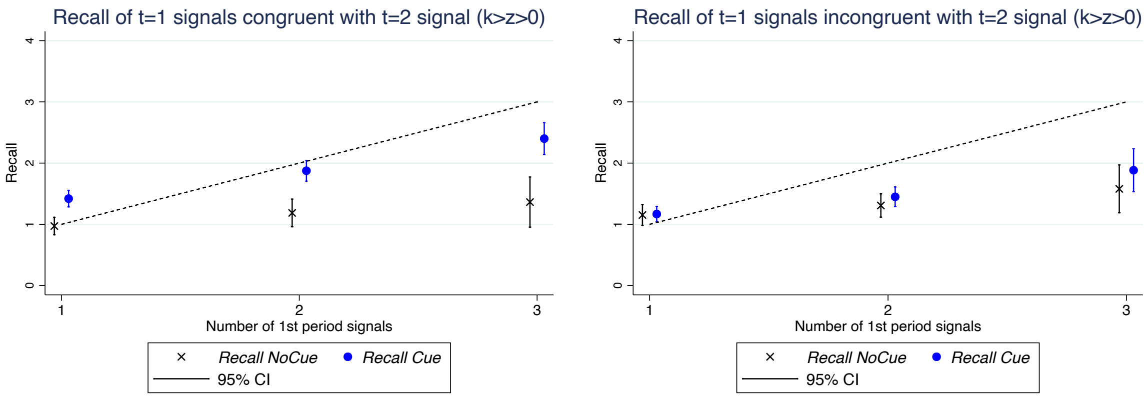


Figure 9: Effective recall of congruent (left panel) and incongruent (right panel) first-period signals in *Recall Cue* and *Recall NoCue*. In this robustness check, the sample is restricted to signal histories where $k > z > 0$, i.e., (i) where at least one first-period signal equals the second-period signal and (ii) at least one first-period signal does not equal the second-period signal. Effective recall equals reported recall for signals that differ from the second-period signal, and reported recall minus one for signals that equal it. The point estimates stem from an OLS regression of effective recall on a treatment dummy. The figures plot the coefficient of the constant for *NoCue* and the sum of the coefficients of the constant and the treatment dummy for *Cue*. The figure does not include the case of four first-period signals because there are very few observations with such a signal history. Whiskers show 95% confidence intervals, computed based on clustering at the subject level.

C Additional Tables

Table 3: First-period log posterior odds in *Beliefs*

	Dependent variable: 1st period normalized log posterior odds	
	(1)	(2)
Sum of t=1 signals (pos. minus neg.)	1.08***	1.14***
	(0.07)	(0.08)
t=1 signals (pos. minus neg.) × 1 if *Cue*		-0.12
		(0.08)
1 if *Cue*		-0.018
		(0.09)
Observations	1400	1400
Adjusted R^2	0.46	0.46

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Following the estimating equation (6) that we derived from the model, the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: First- and second-period beliefs as a function of treatment in *Beliefs*

	Dependent variable:			
	1st period belief		2nd period belief	
Sample:	Last signal neg.	Last signal pos.	Last signal neg.	Last signal pos.
	(1)	(2)	(3)	(4)
1 if *Cue*	0.0077	-0.020	-0.089***	0.043***
	(0.01)	(0.01)	(0.01)	(0.02)
Sum of t=1 signals (pos. minus neg.)	0.12***	0.11***	0.042***	0.036***
	(0.01)	(0.01)	(0.00)	(0.01)
Constant	0.50***	0.51***	0.39***	0.65***
	(0.01)	(0.01)	(0.02)	(0.01)
Observations	688	712	688	712
Adjusted R^2	0.60	0.56	0.16	0.12

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Overreaction as a function of treatment and signal history: Robustness check without recoding boundary beliefs

	Dependent variable: 2nd period normalized log posterior odds						
	Cue vs. *NoCue*			*Cue*		*Cue* vs. *NoCue*	
Sample:	Full	$z = 0$	$z > 0$	Full	Full	Full	k > z > 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
t=2 signal	0.88***	0.86***	0.81***	1.12***	1.12***	0.95***	1.26***
	(0.09)	(0.10)	(0.10)	(0.14)	(0.10)	(0.10)	(0.22)
t=2 signal × 1 if *Cue*	0.56***	0.061	0.78***				
	(0.13)	(0.16)	(0.14)				
t=2 signal × # of congruent t=1 signals				0.26***	0.29***		
				(0.10)	(0.08)		
Sum of congruent t=1 signals × 1 if *Cue*						0.46***	0.49***
						(0.08)	(0.11)
Sum of incongruent t=1 signals × 1 if *Cue*						0.090	0.15
						(0.07)	(0.11)
1 if *Cue*	-0.015	0.064	-0.072			-0.032	-0.099
	(0.10)	(0.12)	(0.11)			(0.10)	(0.12)
Sum of t=1 signals (pos. minus neg.)	0.32***	0.18**	0.33***	0.33***			
	(0.03)	(0.07)	(0.05)	(0.06)			
# of congruent t=1 signals				-0.028	-0.022		
				(0.06)	(0.05)		
t=1 normalized log posterior odds					0.32***		
					(0.08)		
Sum of congruent t=1 signals						0.19***	0.17*
						(0.06)	(0.09)
Sum of incongruent t=1 signals						0.19***	0.33***
						(0.07)	(0.09)
Observations	1307	402	905	646	646	1307	590
Adjusted R^2	0.43	0.20	0.50	0.52	0.56	0.44	0.47

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Following the estimating equations (6) and (7) that we derived from the model, the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. Both the first- and the second-period log posterior odds are normalized by the log diagnosticity odds as described by equation (2). In column (7), the sample is restricted to signal histories where $k > z > 0$, i.e., (i) with at least one congruent first-period signal and (ii) at least one incongruent first-period signal. Variable labels: "t=2 signal" equals 1 if signal positive and (-1) if negative. "# of (in)congruent t=1 signals" captures the number of 1st period signals that do (don't) equal the second-period signal. "Sum of (in)congruent t=1 signals" captures the number of positive minus negative 1st period signals that are (in)congruent with the 2nd period signal. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Overreaction as a function of signal order in *Beliefs Cue*

	Dependent variable: 2nd period normalized log posterior odds															
	1 pos., 1 neg.		2 pos., 1 neg.			1 pos., 2 neg.			3 pos., 1 neg.				1 pos., 3 neg.			
Signal history:	p-n	n-p	p-p-n	p-n-p	n-p-p	p-n-n	n-p-n	n-n-p	p-p-p-n	p-p-n-p	p-n-p-p	n-p-p-p	n-n-n-p	n-n-p-n	n-p-n-n	p-n-n-n
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
t=2 signal	1.51^{***}	1.84^{***}	2.97^{***}	1.72^{**}	2.28^{***}	2.03^{**}	2.03^{***}	2.60^{**}	2.88^{***}	2.51^{***}	0.64	1.85^{*}	2.14^{***}	2.43^{**}	4.34^{***}	2.91^{***}
	(0.22)	(0.40)	(0.74)	(0.61)	(0.43)	(0.80)	(0.58)	(0.80)	(0.79)	(0.43)	(0.63)	(0.82)	(0.55)	(0.97)	(0.92)	(0.65)
Observations	61	32	13	8	15	10	17	8	13	14	9	11	19	12	12	16
Adjusted R^2	0.43	0.39	0.54	0.46	0.64	0.35	0.39	0.54	0.49	0.71	0.00	0.27	0.41	0.30	0.68	0.54

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Following the estimating equation (6) that we derived from the model, the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Recall of first-period signals as a function of cue and treatment in *Recall*

	Dependent variable: Effective recall of 1st period signals					
	At least one $t = 1$ signal of this type			At least one $t = 1$ signal of this type and at least one t=1 congruent with t=2 signal		
Sample:	Congruent signals	Incongruent signals	All	Congruent signals	Incongruent signals	All
	(1)	(2)	(3)	(4)	(5)	(6)
1 if *Recall Cue*	0.67***	0.11*	0.11*	0.67***	0.077	0.077
	(0.07)	(0.06)	(0.07)	(0.07)	(0.08)	(0.08)
1 if *Recall Cue* × 1 if congruent signal			0.56***			0.60***
			(0.09)			(0.09)
1 if congruent signal			-0.16**			-0.26***
			(0.06)			(0.08)
Constant	0.84***	1.10***	1.05***	0.84***	1.13***	1.11***
	(0.06)	(0.07)	(0.07)	(0.06)	(0.08)	(0.07)
True number of 1st period signals FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	683	602	1285	683	445	1128
Adjusted R^2	0.25	0.06	0.17	0.25	0.06	0.19

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Effective recall is defined as reported recall for signals that differ from second-period signal, and as reported recall minus one for signals that equal the second-period signal. The analysis is restricted to signal histories in which the true number of signals that need to be recalled is strictly positive. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Second-period beliefs in *Markets Cue* and *Markets NoCue*

	Dependent variable: 2nd period normalized log posterior odds						
	Cue vs. *NoCue*			*Cue*		*Cue* vs. *NoCue*	
Sample:	Full	$z = 0$	$z > 0$	Full	Full	Full	$k > z > 0$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
t=2 signal	0.94***	0.89***	0.85***	1.20***	1.18***	1.04***	0.91***
	(0.06)	(0.09)	(0.07)	(0.10)	(0.11)	(0.08)	(0.21)
t=2 signal × 1 if *Cue*	0.76***	0.11	1.03***				
	(0.08)	(0.11)	(0.11)				
t=2 signal × # of congruent t=1 signals				0.41***	0.43***		
				(0.07)	(0.07)		
Sum of t=1 signals × 1 if *Cue*							
Sum if congruent t=1 signals × 1 if *Cue*						0.58***	0.56***
						(0.06)	(0.10)
Sum of incongruent t=1 signals × 1 if *Cue*						0.067	0.075
						(0.05)	(0.09)
1 if *Cue*	-0.053	-0.076	-0.069			-0.078	-0.025
	(0.06)	(0.08)	(0.07)			(0.06)	(0.09)
Sum of t=1 signals (pos. minus neg.)	0.46***	0.23***	0.48***	0.41***			
	(0.03)	(0.05)	(0.04)	(0.05)			
# of congruent t=1 signals				-0.033	-0.032		
				(0.04)	(0.04)		
t=1 log posterior odds					0.36***		
					(0.04)		
Sum of congruent t=1 signals						0.29***	0.31***
						(0.05)	(0.09)
Sum of incongruent t=1 signals						0.29***	0.26***
						(0.04)	(0.08)
Observations	3360	990	2370	1680	1680	3360	1521
Adjusted R^2	0.39	0.16	0.45	0.47	0.51	0.41	0.39

Notes. OLS estimates, robust standard errors (in parentheses) are clustered at the subject level. Following the estimating equations (6) and (7) that we derived from the model, the estimations do not include a constant. However, we have verified that including a constant delivers almost identical results. Both the first- and the second-period log posterior odds are normalized by the log diagnosticity odds as described by equation (2). In column (7), the sample is restricted to signal histories where $k > z > 0$, i.e., (i) with at least one congruent first-period signal and (ii) at least one incongruent first-period signal. Variable labels: "t=2 signal" equals 1 if signal positive and (-1) if negative. "# of (in)congruent t=1 signals" captures the number of 1st period signals that do (don't) equal the second-period signal. "Sum of (in)congruent t=1 signals" captures the number of positive minus negative 1st period signals that are (in)congruent with the 2nd period signal. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Experimental Instructions and Comprehension Checks

D.1 Experiment *Beliefs*

D.1.1 Instructions

General Information

We ask that you participate in this online study without taking breaks in between. Please enter the full screen mode of your browser. Please switch on your Zoom video for the entire duration of the experiment. Switch off your Zoom audio. In case you have a question, please contact one of the experimenters using the Zoom chat.

The main part of this experiment consists of two parts that belong together. On the next page you will receive information about both parts. Afterwards, we will ask you to complete a few comprehension check questions.

You can only participate in the experiment if you diligently complete all comprehension check questions.

PART 1 OF THE EXPERIMENT

In this experiment, there are 14 hypothetical companies that have no connection to real companies. These companies are denoted by letters.

Each company is either good or bad. Good companies have businesses that go well. Bad companies have businesses that go poorly. For each company, the computer determines through a coin flip whether the company is good or bad:

- If Heads comes up the company is good (probability 50%)
- If Tails comes up the company is bad (probability 50%)

The computer implements this coin flip separately for each company, such that the coin flips are completely independent of each other.

Important: You will not find out the result of the coin flips. This means that you will not know whether a company is good or bad. Rather, you will receive news about the companies. These news will help you to assess whether a company is likely to be good or bad.

The news

We represent each company through a box that is filled with 100 news. The computer will show you news by randomly drawing them from the company's box.

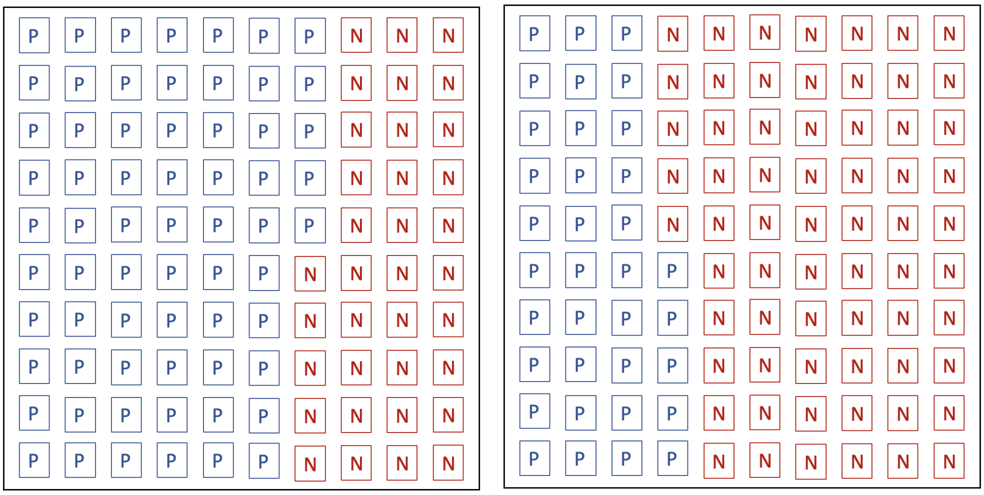


Figure 10: Good company on the left and bad company on the right

A good company consists of a box with 65 positive (P) and 35 negative (N) news. A bad company consists of a box with 35 positive (P) and 65 negative (N) news. Thus, the probability of good news is 65% for a good company but only 35% for a bad company.

Once the experiment begins, you will initially sequentially receive news for the different companies. If you don't receive news about a given company, this will also be communicated to you on your screen. How many news you receive about a company is randomly determined by the computer and has nothing to do with whether the company is good or bad.

If you receive more than one piece of news, the computer randomly draws from the box each time, where news that were drawn earlier get placed back into the box before the next draw. **This means that all news are independent of each other and equally informative.**

Summary: In this experiment, you will never know with certainty whether a company is good or bad. However, you will receive helpful information. While the news are prone to some error, they are on average helpful in estimating whether a company is good or bad. **For example, if you receive multiple positive news for a company, this is of course a stronger indication that the company is actually good than if you only receive one positive news.**

Communication of the news

For each of the 14 companies, you will receive the news sequentially on separate screens. You will first receive the news for one particular company, then for the next one, and so on.

Each news (positive / negative) will be communicated with an "event". An event consists of a story and an image that explain why positive or negative news occurred. These events fit the company and type of news (positive and negative). Thus, the same event CANNOT appear for different companies. Furthermore, an event CANNOT appear for both positive and negative news.

Example: For some company, you might receive one negative piece of news, which appears with the event that a factory of the company burned down. You would then see a short text that describes the fire and a picture thereof. The event "Factory burns down" could then only appear for this particular company and only for negative news.

Important:

- **Any given event and the corresponding news could appear MULTIPLE TIMES for the same company. In such cases, you would have to take into account these news multiple times (and equally strongly).** For example, it could happen that a company's factory burns down twice. You would then have to take into account these news twice.
- As mentioned earlier, it can happen that for some company you will receive multiple positive news and that all of them are communicated with the same event. At the same time, it can also happen that for some company you will receive multiple positive news and that each of them will be communicated with different events. You should treat these two cases equally.
- Example: Suppose that you receive two negative news for a company. In a first case, both of these news are communicated with the same event, while in a second case they are communicated with different events. You should treat these two cases equally.

Ultimately, all that matters for you is how many positive and negative news you receive, not which events cause these news and whether the same event appears multiple times.

Reminder:

- **Each event (story and image) is assigned to only one type of news (positive or negative) for only one company.**
- **It can never happen that a given event appears for both positive and negative news.**
- **It can also never happen that a given event appears for different companies.** In the example above, this would mean that the factory of only one company could burn down. For the other companies, negative news would have different underlying reasons.

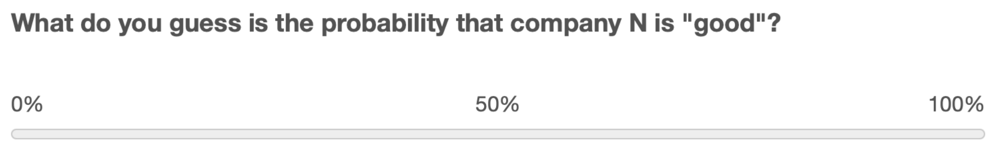
This means for you:

- **If you have already seen an event (story and image) for positive news, then you know that you have already received a positive piece of news for this company.**
- **If you have already seen an event (story and image) for negative news, then you know that you have already received a negative piece of news for this company.**
- **If news appear multiple times (regardless of whether this happens with different or the same events), you need to take these events into account multiple times (and equally strongly).**

Your decisions

After you have possibly received news about a company, you will be asked to make a **guess** (in percent) about whether the company is good or bad. This guess has to be between 0% and 100%. Here, 0% means that you are completely certain that the company is bad, while 100% means that you are completely certain that the company is good.

Your guess will appear on your screen as follows:



Your payment

You can earn up to 10 euros with your guess. The probability that you receive 10 euros is higher the closer your guess to the truth. To maximize your payment, you should therefore simply state your best guess. In case you're interested in the specific equation that we use to determine your payment, please hover here:

[If subject hovers:] The probability that you receive 10 euros for your guess is determined by the following equation:

$$\text{Probability of 10 euros (in \%)} = 100 - [0.01 \times (\text{Guess} - \text{Truth})^2]$$

Please also note: Because your guess is payout-relevant for each company, there is no point for you "strategizing." You should simply try to make the best guess you can.

Example

On the subsequent screens, we will go through an example. This example concerns company B. This is a company that will not appear in the actual experiment. Please note that the stories and images are just examples and do not correspond to those in the actual experiment.

You will now sequentially receive four news about company B. Just like in the actual experiment, these news will be shown to you on separate screens. Just like in the previous instruction screens, you will be able to go back and forth between the different screens. This feature will be disabled in the actual experiment.

Example, news 1

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, news 2

Company B acquired another company. The acquisition was supposed to yield new business opportunities. The acquired company, however, was much more in debt than previously expected by Company B.

The news for Company B are negative.

Example, news 3

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, news 4

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, your decision

What do you guess is the probability that company B is "good"?

0% 50% 100%

SECOND PART OF THE EXPERIMENT

In the second part, you will receive one additional piece of news for each company. Subsequently, you will again be asked to make exactly the same type of guess as in the first part.

Important: For your guess in the second part, all news are relevant that you've received throughout the entire experiment.

You should thus take all news from the first and the second part into account when you make your decisions in the second part.

TIMELINE OF THE EXPERIMENT

1. You first complete a few comprehension check questions.
2. You complete the first part of the experiment:
   - We first inform you about the hypothetical company that the news will be about.
   - You then sequentially receive news for this company.
   - You make a guess.
   - We repeat the same procedure for each of the 14 companies.
3. You complete a few unrelated tasks.
4. You complete the second part of the experiment:
   - We first inform you about the hypothetical company that the news will be about.
   - You then receive another piece of news for this company.
   - You make a guess. All news that you receive in the entire experiment (Part 1 and Part 2) are relevant for this guess.
   - We repeat the same procedure for each of the 14 companies.

Another reminder about the events:

- Each piece of news will be communicated with an event (story and image).

- Each event is only linked to one type of news (positive or negative) for one particular company.
  - It can never happen that an event appears with both positive and negative news.
  - Likewise, it can never happen that an event appears for multiple companies.
- Thus: If you have already seen an event (a story and an image), then you know that for this company you have already received the same type of news before. In a case like this, you need to take into account these news multiple times (and equally strongly).

YOUR PAYMENT

In addition to a fixed participation payment of 4 euros, you will receive 6 euros if you diligently answer the comprehension check questions and complete both parts of the study. In addition, you can earn money with your guesses, as described above. In total, you will take 28 decisions in this experiment: two for each of the 14 companies. At the end of the experiment, the computer will randomly select one of the 14 companies and one of the two corresponding decisions to be payout-relevant. Here, the probability that a decision from Part 2 will be selected is 90% and the probability that a decision from Part 1 will be selected is 10%. You will then be paid according to your earnings for this particular guess about a company. You will realize that this means that each of your decisions potentially determines your payment. But because only one decision will be paid out in the end, it doesn't make sense for you to "strategize".

D.1.2 Comprehension Check

There are seven comprehension check questions that will check your understanding of the instructions. You can only participate in the experiment if you diligently answer these questions.

1. Is the following statement correct: "The more positive news I receive for a given company, the more likely it is that the company is good."

   - Yes
   - No

2. Suppose that one event is given by a factory that burns down. Which of the following statements is correct?

    - The event (factory burns down) can never occur multiple times.
    - The event (factory burns down) can occur multiple times but only for the same company.
    - The event (factory burns down) can occur for different companies.

3. Which of the following statements is correct?

    - If I have already seen an event that is associated with positive news, then I know that I have already received a positive piece of news for this company.
    - The same event is associated with both positive and negative news.

4. Suppose that you have received two negative news for a company that were communicated with the same event. Which of the following statements is correct?

    - Only one of the two positive news should be taken into account because they were communicated with the same event.
    - I need to take into account both pieces of news because both were randomly and independently determined.

5. Suppose that you have received two positive news for a company. Which statement is correct?

    - I should take these two news into account more strongly if they were communicated with the same event than if they were communicated with different events.
    - I should take these two news into account less strongly if they were communicated with the same event than if they were communicated with different events.
    - For how strongly I should take these news into account it doesn't matter which events caused these news and whether they were different or identical.

6. Suppose that a company is good and that the first news about this company was positive. What does this imply for the probability that the next piece of news that the computer draws for this company will also be positive?

    - The computer randomly and independently draws from the box each time, where previously drawn balls get put back before the next draw. Thus, the probability that the next piece of news is positive is the same as before.

- If the computer draws a positive piece of news, it is more likely that the next piece of news is also positive.
- If the computer draws a positive piece of news, it is less likely that the next piece of news is also positive.

7. In the second part, you will make another guess about each company. Which news should you take into account for your decision in the second part of the experiment?

    - Only the news from the first part of the experiment.
    - Only the news from the second part of the experiment.
    - The news from the first and the second part of the experiment.

D.2 Experiment *Recall*

D.2.1 Instructions

The instructions of this experiment are identical to the *Beliefs* experiment up to the following screens.

TIMELINE OF THE EXPERIMENT

1. You first complete a few comprehension check questions.
2. You complete the first part of the experiment:
    - We first inform you about the company that the news will be about.
    - You then sequentially receive news for this company.
    - You make a guess.
    - We repeat the same procedure for each of the 14 companies.
3. In an interim part, you complete a few unrelated tasks. Among these are 28 estimation tasks that will be explained to you at the beginning of this interim task.
4. You complete the second part of the experiment:
    - We first inform you about the company that the news will be about.
    - You then receive another piece of news for this company.
    - You make a guess. All news that you receive in the entire experiment (Part 1 and Part 2) are relevant for this guess.
    - We repeat the same procedure for each of the 14 companies.

Another reminder about the events:

- Each piece of news will be communicated with an event (story and image).
    - Each event is only linked to one type of news (positive or negative) for one particular company.
    - It can never happen that an event appears with both positive and negative news.
    - Likewise, it can never happen that an event appears for multiple companies.

- Thus: If you have already seen an event (a story and an image), then you know that for this company you have already received the same type of news before. In a case like this, you need to take into account these news multiple times (and equally strongly).

YOUR PAYMENT

In addition to a fixed participation payment of 4 euros, you will receive 6 euros if you diligently answer the comprehension check questions and complete both parts of the study. In addition, you can earn money with your guesses, as described above. In total, you will take 56 decisions in this experiment. In part 1, these are 14 decisions (one guess per company) and in Part 2 it is 14 additional decisions (again one guess per company). In addition, there are 28 decisions in the interim task that will be explained to you later. At the end of the experiment, the computer will randomly select one of the 56 decisions to be payout-relevant. You will then be paid according to your earnings for this particular guess about a company. You will realize that this means that each of your decisions potentially determines your payment. But because only one decision will be paid out in the end, it doesn't make sense for you to "strategize".

After the first period of the experiment as well as the distraction (adding) task, subjects were introduced to the surprise recall task as follows.

In what follows, we will show you one additional piece of news for each company. These news will be determined following the same procedures as described to you at the beginning of the experiment. After you have received another piece of news, you will be asked to make two guesses: how many positive and how many negative news you were shown for this company throughout the experiment (including the news that will be shown to you in this interim part)?

In case one of your 28 decisions in this part is payout-relevant, you have a chance of earning 10 euros. You will receive 10 euros if your guess is within +/- 1 of the correct answer. You will receive 0 euros if your guess deviates more than +/- 1 from the correct answer.

D.2.2 Comprehension Check

Same as in *Beliefs* experiment.

D.3 Experiment *Market*

D.3.1 Instructions

General Information

We ask that you participate in this online study without taking breaks in between. Please enter the full screen mode of your browser. Please switch on your Zoom video for the entire duration of the experiment. Switch off your Zoom audio. In case you have a question, please contact one of the experimenters using the Zoom chat.

The main part of this experiment consists of two parts that belong together. On the next page you will receive information about both parts. Afterwards, we will ask you to complete a few comprehension check questions.

You can only participate in the experiment if you diligently complete all comprehension check questions.

PART 1 OF THE EXPERIMENT

In this experiment, there are 14 hypothetical companies that have no connection to real companies. These companies are denoted by letters.

Each company is either good or bad. Good companies have businesses that go well. Bad companies have businesses that go poorly. For each company, the computer determines through a coin flip whether the company is good or bad:

- If Heads comes up the company is good (probability 50%)
- If Tails comes up the company is bad (probability 50%)

The computer implements this coin flip separately for each company, such that the coin flips are completely independent of each other.

Important: You will not find out the result of the coin flips. This means that you will not know whether a company is good or bad. Rather, you will receive news about the companies. These news will help you to assess whether a company is likely to be good or bad.

The news

We represent each company through a box that is filled with 100 news. The computer will show you news by randomly drawing them from the company's box.

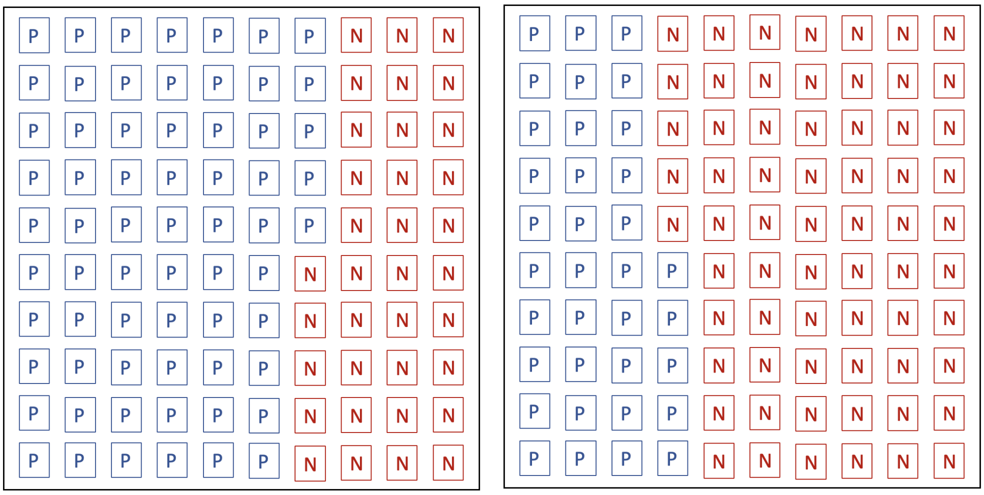


Figure 11: Good company on the left and bad company on the right

A good company consists of a box with 65 positive (P) and 35 negative (N) news. A bad company consists of a box with 35 positive (P) and 65 negative (N) news. Thus, the probability of good news is 65% for a good company but only 35% for a bad company.

Once the experiment begins, you will initially sequentially receive news for the different companies. If you don't receive news about a given company, this will also be communicated to you on your screen. How many news you receive about a company is randomly determined by the computer and has nothing to do with whether the company is good or bad.

If you receive more than one piece of news, the computer randomly draws from the box each time, where news that were drawn earlier get placed back into the box before the next draw. **This means that all news are independent of each other and equally informative.**

Summary: In this experiment, you will never know with certainty whether a company is good or bad. However, you will receive helpful information. While the news are prone to some error, they are on average helpful in estimating whether a company is good or bad. **For example, if you receive multiple positive news for a company, this is of course a stronger indication that the company is actually good than if you only receive one positive news.**

Communication of the news

For each of the 14 companies, you will receive the news sequentially on separate screens. You will first receive the news for one particular company, then for the next one, and so on.

Each news (positive / negative) will be communicated with an "event". An event consists of a story and an image that explain why positive or negative news occurred. These events fit the company and type of news (positive and negative). Thus, the same event CANNOT appear for different companies. Furthermore, an event CANNOT appear for both positive and negative news.

Example: For some company, you might receive one negative piece of news, which appears with the event that a factory of the company burned down. You would then see a short text that describes the fire and a picture thereof. The event "Factory burns down" could then only appear for this particular company and only for negative news.

Important:

- **Any given event and the corresponding news could appear MULTIPLE TIMES for the same company. In such cases, you would have to take into account these news multiple times (and equally strongly).** For example, it could happen that a company's factory burns down twice. You would then have to take into account these news twice.
- As mentioned earlier, it can happen that for some company you will receive multiple positive news and that all of them are communicated with the same event. At the same time, it can also happen that for some company you will receive multiple positive news and that each of them will be communicated with different events. You should treat these two cases equally.
- Example: Suppose that you receive two negative news for a company. In a first case, both of these news are communicated with the same event, while in a second case they are communicated with different events. You should treat these two cases equally.

Ultimately, all that matters for you is how many positive and negative news you receive, not which events cause these news and whether the same event appears multiple times.

Reminder:

- **Each event (story and image) is assigned to only one type of news (positive or negative) for only one company.**
- **It can never happen that a given event appears for both positive and negative news.**
- **It can also never happen that a given event appears for different companies.** In the example above, this would mean that the factory of only one company could burn down. For the other companies, negative news would have different underlying reasons.

This means for you:

- **If you have already seen an event (story and image) for positive news, then you know that you have already received a positive piece of news for this company.**
- **If you have already seen an event (story and image) for negative news, then you know that you have already received a negative piece of news for this company.**
- **If news appear multiple times (regardless of whether this happens with different or the same events), you need to take these events into account multiple times (and equally strongly).**

Your decisions

You will be randomly matched with two other participants of this study. **These participants will receive the same instructions as you. They will also receive exactly the same news about the companies as you.** This means that when you receive one positive and two negative news about a company, then the other two participants will also receive one positive and two negative news about this company.

Each of the three participants then takes multiple decisions that are payoff-relevant, i.e., a guess and a betting decision.

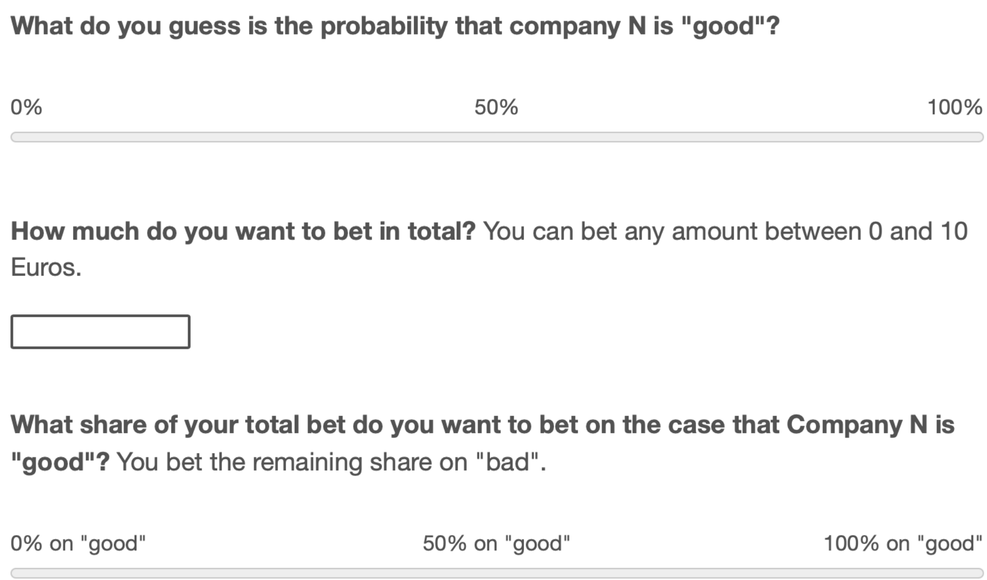
First, you will be asked to make a **guess** (in percent) about whether the company is good or bad. This guess has to be between 0% and 100%. Here, 0% means that you are completely certain that the company is bad, while 100% means that you are completely certain that the company is good.

Subsequently, you will receive a budget of 10 euros that you can fully or partly bet. You can bet that a company is good, that it is bad, or you can make a mixed bet and partly bet on both. You will make your bet in two steps:

1. We will ask you which **absolute amount (in euros) you would like to bet, in total**. You can bet at most your budget of 10 euros, but you can also bet less or nothing.

2. We will ask you which **proportion** (in percent) of your betting amount you would like to bet that the company is good. Here, 0% means that you bet your entire betting amount that the company is bad, while 100% means that you bet your entire betting amount that the company is good. For example, if you indicate that you bet 60% of your betting amount on "Company is good" and 40% of your betting amount on "Company is bad", and you bet 5 euros in total, then this means that you bet 3 euros that the company is good and 2 euros that it is bad.

EXAMPLE OF A DECISION SCREEN: Company N


What do you guess is the probability that company N is "good"?

0% 50% 100%

How much do you want to bet in total? You can bet any amount between 0 and 10 Euros.

What share of your total bet do you want to bet on the case that Company N is "good"? You bet the remaining share on "bad".

0% on "good" 50% on "good" 100% on "good"

Your payment

For each company, we determine with probability 50:50 whether your guess or the betting decision is payout-relevant.

If your **guess** is payout-relevant:

- You can earn up to 10 euros with your guess. The probability that you receive 10 euros is higher the closer your guess to the truth. To maximize your payment, you should therefore simply state your best guess. In case you're interested in the specific equation that we use to determine your payment, please hover here:

  [If subject hovers:] The probability that you receive 10 euros for your guess is determined by the following equation:

$$\text{Probability of 10 euros (in\%)} = 100 - [0.01 \times (\text{Guess} - \text{Truth})^2]$$

If your **betting decision** is payout-relevant:

- You keep the part of your budget that you don't bet.
- The part of your budget that you bet can either multiply or you can lose it.
  - The amount that you bet on the WRONG event (e.g. you bet that the company is good but it is actually bad) is lost.
  - The amount that you bet on the CORRECT event (e.g. you bet that the company is good and it is actually good) will deliver a profit for you. In case you're interested in the specific equation, please hover here:

    [If subject hovers:] Your earnings from the amount that you bet on the correct event will be determined as follows:

$$\begin{aligned} \text{Your earnings} = &\ \text{Amount that you bet on correct event} \\ &\times \text{Amount bet by all participants} \\ &/\text{Amount bet on correct event by all participants} \end{aligned}$$

    While this equation may look complicated, the underlying principle is very simple: in case you bet on the correct event, you will get at least the money back that you bet, and probably more. The earnings of the other two participants are determined in the same fashion.

Important: You obtain a new budget of 10 euros for each betting decision. Therefore it does not make sense for you to "strategically" save money in the first decisions to have more money for the later decisions. Since you obtain a new budget for each decision, you should always bet the amount that you deem appropriate for that decision.

In addition, please notice the following: Since for each company, either your guess or your betting decision will be relevant for your payment, it does not make sense for you to "strategize". You should simply try to make every decision as best as you can.

Example

On the subsequent screens, we will go through an example. This example concerns company B. This is a company that will not appear in the actual experiment. Please note that the stories and images are just examples and do not correspond to those in the actual experiment.

You will now sequentially receive four news about company B. Just like in the actual experiment, these news will be shown to you on separate screens. Just like in the previous instruction screens, you will be able to go back and forth between the different screens. This feature will be disabled in the actual experiment.

Example, news 1

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, news 2

Company B acquired another company. The acquisition was supposed to yield new business opportunities. The acquired company, however, was much more in debt than previously expected by Company B.

The news for Company B are negative.

Example, news 3

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, news 4

Company B's marketing head is the award-winning Tom Stark. The newest marketing campaign again pulls out all the stops. In particular, environmentally minded customers were targeted and their demand increased substantially.

The news for Company B is positive.

Example, your decision

What do you guess is the probability that company B is "good"?

0% 50% 100%

How much do you want to bet in total? You can bet any amount between 0 and 10 Euros.

What share of your total bet do you want to bet on the case that Company B is "good"? You bet the remaining share on "bad".

0% on "good" 50% on "good" 100% on "good"

SECOND PART OF THE EXPERIMENT

As mentioned earlier, this experiment consists of two parts. In the second part, you will receive one additional piece of news for each company. The other participants in your group will again receive exactly the news as you.

Subsequently, you will again be asked to make exactly the same decisions as in the first part, i.e., a guess and a betting decision.

Important: For your decisions in the second part, all news are relevant that you've received throughout the entire experiment.

You should thus take all news from the first and the second part into account when you make your decisions in the second part.

TIMELINE OF THE EXPERIMENT

1. You first complete a few comprehension check questions.
2. You complete the first part of the experiment:
    - We first inform you about the hypothetical company that the news will be about.
    - You and the other two participants then sequentially receive news for this company, where the other participants receive exactly the same news as you.
    - You make a guess and a betting decision. The two other participants take the same decisions.
    - We repeat the same procedure for each of the 14 companies.
3. You complete a few unrelated tasks.
4. You complete the second part of the experiment:
    - We first inform you about the hypothetical company that the news will be about.
    - You and the other two participants then receive another piece of news for this company, where again everyone receives the same news.
    - You again make a guess and a betting decision. All news that you receive in the entire experiment (Part 1 and Part 2) are relevant for these decisions. The two other participants take the same decisions.

- We repeat the same procedure for each of the 14 companies.

Another reminder about the events:

- Each piece of news will be communicated with an event (story and image).
  - Each event is only linked to one type of news (positive or negative) for one particular company.
  - It can never happen that an event appears with both positive and negative news.
  - Likewise, it can never happen that an event appears for multiple companies.
- Thus: If you have already seen an event (a story and an image), then you know that for this company you have already received the same type of news before. In a case like this, you need to take into account these news multiple times (and equally strongly).

YOUR PAYMENT

In addition to a fixed participation payment of 4 euros, you will receive 6 euros if you diligently answer the comprehension check questions and complete both parts of the study. In addition, you can earn money with your guesses, as described above. In total, you will take 56 decisions in this experiment: four for each of 14 companies. At the end of the experiment, the computer will randomly select one of the 14 companies and one of the four corresponding decisions to be payout-relevant. Here, the probability that a decision from Part 2 will be selected is 90% and the probability that a decision from Part 1 will be selected is 10%. You will then be paid according to your earnings for this particular guess about a company. You will realize that this means that each of your decisions potentially determines your payment. But because only one decision will be paid out in the end, it doesn't make sense for you to “strategize”.

D.3.2 Comprehension Check

There are nine comprehension check questions that will check your understanding of the instructions. You can only participate in the experiment if you diligently answer these questions.

1. Is the following statement correct: “The more positive news I receive for a given company, the more likely it is that the company is good.”

   - Yes
   - No

2. Suppose that one event is given by a factory that burns down. Which of the following statements is correct?

   - The event (factory burns down) can never occur multiple times.
   - The event (factory burns down) can occur multiple times but only for the same company.
   - The event (factory burns down) can occur for different companies.

3. Which of the following statements is correct?

   - If I have already seen an event that is associated with positive news, then I know that I have already received a positive piece of news for this company.
   - The same event is associated with both positive and negative news.

4. Suppose that you have received two negative news for a company that were communicated with the same event. Which of the following statements is correct?

   - Only one of the two positive news should be taken into account because they were communicated with the same event.
   - I need to take into account both pieces of news because both were randomly and independently determined.

5. Suppose that you have received two positive news for a company. Which statement is correct?

   - I should take these two news into account more strongly if they were communicated with the same event than if they were communicated with different events.
   - I should take these two news into account less strongly if they were communicated with the same event than if they were communicated with different events.
   - For how strongly I should take these news into account it doesn't matter which events caused these news and whether they were different or identical.

6. Suppose that a company is good and that the first news about this company was positive. What does this imply for the probability that the next piece of news that the computer draws for this company will also be positive?
    - The computer randomly and independently draws from the box each time, where previously drawn balls get put back before the next draw. Thus, the probability that the next piece of news is positive is the same as before.
    - If the computer draws a positive piece of news, it is more likely that the next piece of news is also positive.
    - If the computer draws a positive piece of news, it is less likely that the next piece of news is also positive.
7. In the second part, you can decide for each company how much money you want to bet on the company being good or bad. Which news should you take into account for your decision in the second part of the experiment?
    - Only the news from the first part of the experiment.
    - Only the news from the second part of the experiment.
    - The news from the first and the second part of the experiment.
8. You and the other two participants in your group each decide for each company how much money they would like to bet that the company is good or bad. Which statement is correct?
    - The other two participants potentially receive different news than I do.
    - The other two participants receive exactly the same news as I do.
9. Which of the following two statements is correct?
    - You receive a new budget of 10 euros for each decision. Thus, how much money you bet in one decision does not affect how much money you have available for the other decisions.
    - You receive a total budget of 10 euros. Thus, if for example you bet 5 euros that a company is good, then you only have 5 euros available for the other decisions.