MOOD, MEMORY, AND BIASED BELIEFS AND DECISIONS

AARON L. BODOH-CREED
UNIVERSITY OF CALIFORNIA - BERKELEY

Abstract. I first provide a static model of associative memory, a mnemonic process wherein cues in the environment bias information recalled from memory. I apply the model in a principal-agent setting to analyze contradictory predictions for the relationship between employee morale and productivity from the organizational behavior literature. I then develop a dynamic model that incorporates rehearsal and a model of mood dynamics. Applied in an asset pricing setting, the theory provides explanations for short-run continued overreaction to news, long-run correction of the overreaction, and excess price volatility. The model predicts that these effects are stronger for more volatile assets.

1. Introduction

Human memory is a complex psychological phenomena, but at a basic level memory can be divided into short-term and long-term memory systems (Atkinson and Shiffrin [2]). Short-term memory is used to hold a small amount of information in an active, available state for immediate use. Long-term memory has an effectively unbounded size and can store information over long periods of time in an inactive state. The goal of this paper is to provide a model for how data are recalled from long-term memory and to use this model to trace out the implications for economic decision making in a principal-agent setting and an asset pricing application.

Data is recalled from long-term memory through an *associative memory* process. Cues present in the environment or one's state of mind are associated with data in memory, and the appearance of these cues can prompt the recollection of the associated information. While these cues provide an efficient process for focusing on important data, the cues can cause beliefs (and the resulting decisions) to be based on a systematically biased, incomplete set of information.

Although I provide a general model of how cues influence recall from memory, in my applications I focus on *mood congruent memory.* Mood congruent memory is a psychological phenomenon wherein the subjectively experienced goodness of an agent's affective state, referred to as the *valence* of the affective state, serves as a cue for information stored in long-term memory (Isen [39]). For example,

I would like to thank Ilya Segal, B. Douglas Bernheim, Marcello Miccoli, Juuso Toikka, Alex Hirsch, Carlos Lever, M. Bumin Yenmez, Edward Augenblick, Xiaochen Fan, Annika Todd, and Albert Bollard for invaluable discussions and feedback over the course of this project. Bodoh-Creed gratefully aknowledges the support of the Jacob K. Javits Fellowship program and the Bradley Research Fellowship. Correspondence: acreed@berkeley.edu.

if an individual is in a happy and relaxed state of mind, then that affect will serve as a cue for positive information. Appraisal theories of emotion predict that the valence of the emotion elicited by and associated with a piece of information, situation, or a cognitive process is determined by the implications for the individual's well-being (Smith et al. [75]). Since one's well being is presumably correlated with the satiation of an agent's preferences, the appraisal theory of emotion implies that positively valenced emotional states are associated with high utility in the moment and with information that suggests high utility in the future. In summary, the psychology of mood and memory predicts that agents will be optimistically biased regarding future utility levels when presently in a high utility state and pessimistically biased when in a low utility state.

As a first step, I study a static model that incorporates associative recall from long-term memory. Prior to the time of decision, the agent observes a history of informative signals and stores this complete history in long-term memory. At the time of decision, the agent recollects a sample of data from memory in order to update his prior beliefs and form a posterior. While the entire set of previously observed information is recalled and employed to form a posterior belief in a traditional model of decision making, in my model the information is recalled incompletely and the set of data recollected is biased by mnemonic cues. Given the posterior beliefs, the agent makes an optimal decision given his preferences and the feasible set he faces. I provide sufficient conditions for robust comparative statics on the relationship between the presence and intensity of mnemonic cues, the information recalled, and the resulting choices of the agent.

In my first application, I consider the debate in the organizational behavior literature between two schools of thought regarding the link between employee morale and productivity (see Staw and Barsade [76] for a literature review). The "Happier-and-Smarter" school of thought suggests that optimistic workers will prove more productive through increased effort and heightened creativity. The "Sadder-but-Wiser" hypothesis emphasizes that experimental subjects in a depressed mood tend to have more accurate beliefs regarding uncertain events than their happier peers, and as a result those in a depressed mood make better decisions. If one assumes an employee's morale is correlated with mood or affective disposition, one can use my model to study the relationship between morale and productivity in an otherwise traditional principal-agent model.

My model predicts that whether happy or depressed individuals are more productive depends crucially on whether the marginal product of effort is correlated with high utility states. Specifically, if the agent believes that a "good" state of the world, one wherein the agent reaps a high utility, has a high marginal productivity of effort, then a positive affective state generates an optimistic bias regarding the marginal productivity of effort. High morale thus yields an increase in agent effort. Conversely, if the "good" state of the world has a low marginal productivity of effort, exactly the opposite effect occurs and a positive affective state reduces agent effort. This second effect can explain *defensive pessimism*, a phenomenon where an individuals hold pessimistic views to motivate effort. By clarifying the interactions between morale, beliefs, and effort choice, my model sheds light on the conflicting empirical work in this area and points to the correlation between utility level and the marginal productivity of effort as an important factor for future studies to assess.

Next, I provide an extension of the model to dynamic settings that includes two intertemporal elements of mood and memory. The first element I incorporate is *rehearsal*, a phenomenon whereby information that was recalled in the recent past is more easily recalled in the present. The second dynamic component of the model is the evolution of affect. I assume, in line with appraisal theories of emotion (Smith et al. [75]), that good news increases the valence of affect while bad news reduces the valence.

Association and rehearsal represent the essence of long-term memory, which leads me to conclude that including these effects is required for a realistic dynamic model of long-term memory. Since the model focuses on the intersection of memory and affect, it is also important to leverage what the psychology literature can tell me about the dynamics of affect. Immune neglect and the influence of news on affect are core determinants of affective dynamics. Again, any serious model of affect has to incorporate both of these effects to make realistic predictions.

My second application is to asset pricing theory in a dynamic setting. I first use my model to study short-run overreaction to news and long-run correction of the initial overreaction. Short-run overreaction refers to positive (negative) price drift in response to positive (negative) news over horizons of less than one year (Cutler, Poterba and Summers [16], Jegadeesh and Titman [40], Rouwenhorst [66]). Long-run correction denotes the negative (positive) responses of asset prices to positive (negative) news over time scales longer than a year, which is sometimes described as asset prices reacting too much to news in the short-run (De Bondt and Thaler [18], Lakonishok, Shleifer and Vishny [54]). Both of these patterns imply that current information is predictive of future price changes.

My model generates short-run continued overreaction to news through the autocorrelation of the asset price biases. Conditional on good news in the current period, the data recalled from memory through association are biased towards positive information about the asset, which generates an optimistic bias regarding the asset's value. Although the affective state begins to return towards its ergodic distribution in the next period, the asset price changes in future periods are expected to be positive if the rehearsal effect is sufficiently strong. However, since rehearsal effects fade with time, a long-run correction effect occurs. The model also predicts that asset price changes are autocorrelated, which I refer to as asset price momentum.[1] I also find that the price biases are stronger for assets that have a higher fundamental volatility.

[1]Since most papers on asset price momentum and momentum strategies (e.g., Jegadeesh and Titman [40] and [41]) do not condition on whether the recent news regarding the asset has been positive or negative, the assessment of the unconditional autocorrelation is required to tie the predictions to the existing empirical literature.

Finally, I find that the dynamics of the mnemonic cues can cause excess volatility. I decompose asset price volatility into three components: the fundamental volatility, volatility caused by changes in affective state over time, and the influence of news on the current affective state. Immune neglect suggests that affective state (and hence asset price) is less volatile when the affect is positive - there is less room to increase affect, and negative information causes small decreases in affect. This asymmetry between the evolution of positive and negative affect leads the model to predict that excess volatility is higher during market downturns.

I begin with a survey of related literature in economics, psychology and finance in Section 2. Section 3 presents the static model as well as the application to a principal-agent setting. Section 4 presents the dynamic model and the application to asset pricing problems. I conclude in Section 5. Proofs are contained in Appendix A, while the other appendices include robustness checks as well as additional numerically generated predictions.

2. Related Literature - Emotions and Memory in Economics

2.1. **Psychology of Emotion and Belief.** Affective state has been identified as a cause of a variety of changes in the cognitive processes that drive decisions. A concise summary of these effects is given in Clore et al. [13] and Isen [38].[2] Isen et al. [39] is the seminal study of mood congruent recall. The authors employed a field experiment to demonstrate a large effect on consumer product evaluations through the use of low cost affect manipulations and a supplementary laboratory study to isolate the cognitive processes underlying the phenomenon. In the field experiment, shoppers in a suburban mall were assigned randomly to treatment and control groups, and subjects in the treatment group received free samples of products valued at $0.29 in 1977 dollars. The participants in the treatment group registered significantly higher product satisfaction ratings than those in the control group. The core idea is that the subjects recall past experiences with a product when forming opinions about their product satisfaction, and it is easier for subjects to recall positive experiences once their mood has been improved by the free sample.

The laboratory experiment involved inducing a random affective state in the participants, requiring the subjects to memorize a list of words, inducing a (possibly different) random affective state in the subjects, and then determining the number and affective valence of the words that could be recollected in the new affective state.[3] The basic idea behind the experiment is that subjects will have an easier time recalling words with the same valence as their current emotional state. The affective state was created by having the subject play a computer game developed by

[2]Hirshleifer and Shumway [32] provides an excellent summary of the research on mood and individual judgement. Elster [25], Loewenstein [56], and Loewenstein and Lerner [57] provide overviews of the role of emotion in decision making.

[3]Kida, Smith and Maletta [46] provides a series of experiments that reveal that affective reactions to information are related to the valence of the information in a straightforward fashion, encoded into memory, and have significant impact on the recall of the information at later times.

the experimenters. Winning and losing were randomly determined by the experimenter for each repetition of the game. After playing the computer game once, the participants were provided a tape recording containing 36 words. 18 of the words were traits with 6 words each of positive, negative, and neutral valence, while 18 additional non-trait words were added as a control. After playing the computer game a second time, the participants were given 5 minutes to recall as many words as possible. Consistent with mood congruent recall, recollection was improved most when the valence of the material memorized matched the valence of the affective state at the time of recollection.

2.2. **Prior Economic Literature.** Few prior studies in the behavioral finance or economics literatures explicitly focus on long-term memory. Mullainathan [62] provides a model of long-term memory to explain deviations from the consumption paths predicted by the permanent income hypothesis. Sarafidis [67] applies Mullainathan [62] in a setting where politicians take advantage of the polity's associative memory by strategically releasing information. Hirshleifer and Welch [33] show decision-making can exhibit inertia or impulsiveness when beliefs are driven by perfectly recalled actions combined with imperfectly recollected signals. Ericson [26] studies the interaction of memory, reminders, and procrastination in a setting where agents can forget to complete a task by a deadline. Although one can think of the reminders as cues, the main focus of Ericson [26] is on the effect of imperfect memory on estimates of present bias and the subtle effects of reminders on agent behavior. Benabou and Tirole ([6], [7], [8], [9]) develop models of malleable, imperfect memory in the context of agents with self-control problems in order to explain the use of intrapersonal rules for self-regulation. Gottlieb [28] uses a model of agents with malleable memory and preferences over their own attributes to explain anomalies in the literature on choice under risk.

Models of memory of limited volume have also been developed, and these models are most easily interpreted as models of short-term memory. Wilson [79] and Hellman and Cover [31] derive optimal memory processes and decision strategies in the context of a decision problem featuring an infinite sequence of informative signals, but where memory is of a fixed, finite length. Dow [22] discusses optimal memory schemes in the short-run context of a two period decision problem where the agent stores information from period to period in an optimal, but limited, fashion.

An alternative view of my application to a principal-agent relationship is that my model provides microfoundations for the confidence enhanced productivity model of Compte et al. [15], which has two crucial behavioral components. First, Compte et al. [15] assumes a reduced form model for agent memory wherein successes are more likely to be remembered than failures, which results in agents possessing optimistic views about their future performance.[4] My model of memory provides an avenue through which these optimistic beliefs could be formed and suggests when the influence of

[4]This is an example of a self-affirming bias, which is significantly different than the associative memory model I am using. Compte et al. [15] supposes that failures are less likely to be recorded in long-term memory. I assume that all data is in long-term memory and that data is recollected in an imperfect fashion.

optimistic beliefs will be prevalent. Second, the Compte et al. [15] model assumes that confidence increases the chance of successfully executing the task. My model of morale provides a simple structure based on the complementarity of effort choice and optimism that founds these effects. In addition, I highlight the possibility that defensive pessimism may make holding pessimistic beliefs optimal given the appropriate economic primitives, which would reverse the headline result of Compte et al. [15].

Very little economic theory has addressed the issue of emotion in decision-making. Elster [25] and Loewenstein and Lerner [57] provide qualitative overviews of the role of emotion in economic decision making. Hermalin and Isen [30] provides a study of preferences over emotional states, although the model and application of their study is unrelated to my work.

2.3. **Prior Finance Literature.** Several prior works have provided behavioral models to explain some of the asset pricing puzzles that I study. Barberis, Shleifer and Vishny [4] explains short-run continued overreaction and long-run correction of asset prices to news with a model of investor sentiment.[5] Daniel, Hirshleifer and Subrahmanyam [17] develops a model of investor sentiment driven by overconfidence and positively biased self-attribution in order to explain these same anomalies. Hong and Stein [34] provides a model wherein traders are boundedly rational and limited to using simple forecasting models. Hong, Stein and Yu [36] provide a model of investors vacillating between different simple models and the resultant effect on asset prices. DeLong et al. [20] assumes that noise traders are subject to exogenous sentiment shocks and study how these shocks create risk that prevents arbitrageurs from eliminating mispricings and allows the noise traders to earn supernormal profits.

Of particular interest for this paper is the branch of the behavioral finance literature that searches for correlations between asset prices and exogenous mood shifters such as weather, length of the day, and sporting event outcomes. In an early contribution, Saunders [68] finds that the cloud cover in Manhattan has a negative effect on the prices of stocks traded on the New York Stock Exchange. Hirshleifer and Shumway [32] extends this analysis to cover markets across the globe and finds the same negative correlation between cloud-cover and asset prices. Kamstra, Kramer and Levi [42] provides an analysis of seasonal effects around the world under the assumption that the depressed affect caused by short days during the winter in turn depresses stock prices, which is particularly striking since changing day length is predictable years in advance. Edmans, Garci and Norli [24] shows stock market valuations decline following a loss in important sports matches and that the magnitude of the decline is greater for smaller stocks and more important events. Yuan, Zheng and Zhu [80] and Keef and Khaled [45] argue that the lunar cycle influences asset prices through mood. Krivelyova and Robotti [50] finds evidence that solar events influence asset prices

[5] I have adapted the terminology of Barberis, Shleifer and Vishny [4] to the nomenclature used in this study. Several prior works refer to short-run continued overreaction and "underreaction" and long-run correction as "overreaction."

through the geomagnetic storm's effect on mood. My model supplies foundations for these asset price anomalies in a Bayesian framework with the optimism caused by the biased recollection of information salient for estimating asset values, and I use the model to tease out the effects of these biases beyond simple increases or decreases in price levels on the day of the mood influencing event.

While I study the effect of mood on beliefs, there is growing evidence that affect may influence preferences with a particular focus on the effect of mood on risk preferences. Kamstra et al. [43] models these effects under the assumption that seasonal mood swings cause predictable changes in risk aversion. Although most of the predictions are non-overlapping with ours, Kamstra et al. [43] predicts heightened volatility during seasons with low affective states.

Ljungqvist, Nanda and Singh [55] and Derrien [21] provide models wherein IPOs are timed to take advantage of the optimism of traders to explain why IPOs appear underpriced in the short-run and overpriced in the long-run. My model provides microfoundations for the belief formation process that leads the initial beliefs of investors to be either bearish or bullish and the factors that cause these beliefs to change over time. My predictions are supported by Purnanandam and Swaminathan [65] and Cook, Kieschnick and Van Ness [14].[6]

Tetlock [77] studies the influence of the Wall Street Journal's (WSJ) "Abreast of the Market" section on market affect. Included in this WSJ section are summaries of market events, explanations of the market behavior from third parties, and predictions about future market behavior. Tetlock conducts time-series analysis of the relationship between the number of negatively valenced words in the WSJ section and the performance of the Dow Jones Industrial Average (DJIA) as well as a variety of other financial market statistics. Tetlock [77] finds that negatively valenced words in the WSJ predict depressed performance of the DJIA in the following week. Tetlock [77] argues that the WSJ column does not solely have an informational effect since the price changes he studies partially revert with time, which contradicts the random walk pattern predicted by a purely informational model.

3. Static Model

In this section I develop a model of the effects of mnemonic cues, such as agent affective state, within a generic static decision problem under uncertainty. Prior to the time of decision, the agent is assumed to have observed a history of informative signals and stored this history in long-term memory. At the time of decision, the agent recollects a sample of data from memory in order to update his prior beliefs and form a posterior. Given his posterior beliefs, the agent makes an optimal decision given his preferences and the feasible set he faces. After laying out the general

[6]I do not claim that these effects would not be realized if other models of asset price overreaction and correction were applied in this context. To isolate my model's predictions, I would need to understand whether the IPO promotion efforts involved heightening positive affect or cueing positive information from memory (as my story suggests) or if the provision of new information played a crucial role in pre-IPO promotion (which might imply other models are salient).

model, I provide sufficient conditions for comparative statics for both beliefs and decisions under uncertainty. I then apply my results to a principal-agent problem.

3.1. **Basic Model and Notation.** A *history*, $H_N = \{\omega_1, \omega_2, ..., \omega_N\}$, is a set of signals observed by the agent prior to decision and stored in long term memory. The signals ω_i are members of a space Ω that can be represented as a finite Cartesian product of totally ordered subspaces. The agent uses these signals to update a prior belief about the distribution of a parameter $\theta_0 \in \Theta$ where Θ is a partially ordered parameter space. The agent's prior beliefs are described by the cumulative distribution function (CDF) $G(\theta)$. I assume the signals are identically and independently distributed according to probability density function (PDF) $f(\omega|\theta)$. I focus on the case where $\Theta \subseteq \mathbb{R}$ and $\Omega \subseteq \mathbb{R}$, although the results below generalize readily.

Assumption 1. *$f(\omega|\theta)$ is log-supermodular (log-spm) in (ω,θ).*

Assumption 2. *$f(\omega|\theta)$ has full support over Ω.*

If $\theta \geq \theta'$, Assumption 1 is equivalent to $\frac{f(\omega|\theta)}{f(\omega|\theta')}$ being increasing in ω, which implies that higher values of ω are evidence of higher values of θ regardless of any other information observed. Let the order relation $\succcurlyeq$ refer to the strong stochastic order, which in a one-dimensional setting is equivalent to first order stochastic dominance. Where the strong stochastic ordering is strict, I use the symbol $\succ$.

The state variable $\varphi \in [-1, 1]$ represents the memory cues facing the agent during the decision problem. The parameter is meant to capture both the intensity of cues present at the time of decision as well as the strength of the association between the cues and the relevant data stored in memory. I use the convention that positive (negative) values of φ are associated with higher (lower) values of ω stored in memory. In the context of emotion, φ indexes the valence of the agent's affective state with higher values associated with more positively valenced moods.[7] The bounds on φ reflects the idea that there exists a natural upper and lower bound on the intensity of affective states.[8]

A *recollected history* is a random variable that consists of those signals that are recalled by the agent at the time of decision. Let the random variable $\mathbf{H}^R(\varphi)$ represent the recollected history with a typical realization denoted $H^R \subseteq H_N$. $\mathbf{H}^R(\varphi)$ takes φ as an argument to reflect the fact that the distribution of H^R depends on the state variable. The length of a recollected history is denoted by $|H^R| \in \{0, ..., N\}$. I let $\mathbb{H}(H_N, n)$ denote the set of all possible length n recollected histories given the true history H_N. I assume naivety on the part of the agent in the sense that the agent does not use knowledge of the cue state to correct for the biases induced in the recollection and

[7]I do not model the source of these associations, but the psychology literature has shown that cues become associated with data that have similar features (such as valence) through conscious and unconscious processes of association (Smith et al. [75]).

[8]See Rayo and Becker [58] for arguments supporting this assumption.

Bayesian updating process.[9] I assume throughout that the prior beliefs, $G(\theta)$, and every possible posterior belief, $G(\theta|H^R)$, admit bounded PDFs $g(\theta)$ and $g(\theta|H^R)$.

3.2. **Recall Probability Functions.** A simple model for associative memory would be to treat the recollection of each datum in memory as an independent random event with the recollection probability a function of both cue state and the value of the datum. However, this generates a model in which both the length and content of the recollected history are influenced by the cue state.[10] For general stochastic processes it is impossible to stochastically order the posteriors generated by data series of different lengths, and without such an ordering it is impossible to generate robust comparative statics results.

Let the *recall probability function*, $\rho\left(H^R|\varphi, H_N, n\right)$, be the probability that a set of data H^R is recollected at the time of decision given a complete history of data H_N, a cue state φ, and a restriction that the recollection contain exactly $n \leq |H_N|$ elements. The length of the recollected memory is a random variable $l(N)$ that can be conditional on N, but must be independent of φ. The distribution of $l(N)$ is:[11]

$$k(n, N) = \Pr\{l(N) = n\} = \Pr\{|H^R| = n \text{ given } |H_N| = N\}$$

I assume that, conditional on recollecting n events, the contents of the recollected history are generated by a process of independent sampling without replacement from the history in memory. Denote the relative sampling probability of recalling event ω in a recollected history of length n from a true history H_N as $\rho_\omega(\omega|\varphi, H_N, n)$. In order to generate associative memory effects, the relative sampling probabilities are assumed to be functions of the cue state and the signals in memory. To interpret $\rho_\omega(\omega|\varphi, H_N, n)$, consider two signals $\omega, \omega' \in H_N$. If

$$\frac{\rho_\omega(\omega|\varphi, H_N, n)}{\rho_\omega(\omega'|\varphi, H_N, n)} = \beta$$

then ω is β times more likely to be recalled than ω' given affective state φ_t. Note that I do not require that ρ_ω be a sampling probability with values in $[0, 1]$ — ρ_ω only represents the relative probabilities that different data are sampled from memory.

---

[9] The agent is assumed to be either unaware of the cue's effect on his memory or of the cue's existence. Alternately, the problem of removing the bias may suffer from an identification problem. Although I prefer the notion that the agent is unaware of the bias as most psychological subjects seem to be ignorant of their own mnemonic biases, either of these explanations suffices to support the behavior modeled herein.

[10] One of the important assumptions of Mullainathan [62] is the use of a random walk information process. For a random walk, a datum consists of an innovation that is unambiguously good or bad news, in the sense of Milgrom [60], regardless of the number or magnitude of prior innovations. Therefore, beliefs contingent on memories of different lengths can be compared in the strong stochastic order in the random walk setting. One of the contributions of the model is identifying restrictions on the model of memory sufficient for comparative statics predictions to be made outside of the random walk setting.

[11] Throughout this work $\Pr\{E\}$ denotes the probability of an event E under the true probability distribution over the states of the world.

For $H^R \in \mathbb{H}(H_N, n)$, ρ is defined as as:

$$\rho\left(H^R|\varphi, H_N, n\right) = \frac{\prod_{\omega \in H^R} \rho_\omega(\omega|\varphi, H_N, n)}{\sum_{H^R \in \mathbb{H}(H_N,n)} \left[\prod_{\omega \in H^R} \rho_\omega(\omega|\varphi, H_N, n)\right]} \propto \prod_{\omega \in H^R} \rho_\omega(\omega|\varphi, H_N, n) \tag{1}$$

and $\rho\left(H^R|\varphi, H_N, n\right) = 0$ if $H^R \notin \mathbb{H}(H_N, n)$.

Assumption 3. *$\rho_\omega(\omega|\varphi, H_N, n)$ is log supermodular in (ω, φ)*

Assumption 4. *$\rho_\omega(\omega|\varphi, H_N, n) > 0$ if and only if $\omega \in H_N$*

Assumption 3 states that high values of φ increase the probability of recollecting good news events (high ω) relative to bad news events (low ω). Assumption 4 defines the support of $\rho_\omega(\omega|\varphi, H_N, n)$. One example of a relative sampling function that fits the assumptions is $\rho_\omega(\omega|\varphi, H_N, n) = e^{\varphi\omega}$. As an example, in Figure 1 the solid line represents the probability density function (PDF) of an unbiased recollection of a large number of realizations of a normal distribution, while the dashed line represents the distribution of points recollected through association by the representative agent when $\varphi > 0$ and $\rho_\omega(\omega|\varphi, H_N, n) = e^{\varphi\omega}$.

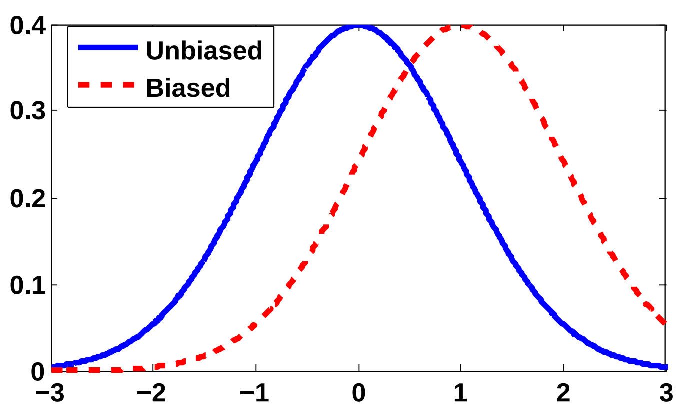

FIGURE 1. Biases in Recollection

Combining all of these elements, the probability that $\mathbf{H}^R(\varphi) = H^R$ where $H^R \in \mathbb{H}(H_N, n)$ is:

$$k(n, N) \frac{\prod_{\omega \in H^R} \rho_\omega(\omega|\varphi, H_N, n)}{\sum_{H^R \in \mathbb{H}(H_N,n)} \prod_{\omega \in H^R} \rho_\omega(\omega|\varphi, H_N, n)}$$

3.3. **Memory Cues and Estimation.** In this section I study the effect of cues on judgments about the expected values of functions of random variables. Let $q : \Theta \to \mathbb{R}$ denote an increasing function of θ. For example, $q(\theta) = U(x;\theta)$ when the agent is estimating the expected utility of a choice $x \in X$ given an unknown state θ. I assume that the the true history of signals is stored in the agent's long-term memory, and the agent must update his beliefs based on a potentially biased set of recalled information when forming his expectation.

Let $G(\circ|\mathbf{H}^R(\varphi))$ denote the CDF of the agent's posterior beliefs after recalling history $\mathbf{H}^R(\varphi)$. Since the recollected information is random, the judgment about the mean of $q(\theta)$ based on the recollected information will also be a random variable. Denote this random variable as $\mathbf{q}(\varphi)$, which is defined as follows:

$$\mathbf{q}(\varphi) = \int q(\theta) G(d\theta|\mathbf{H}^R(\varphi)) \tag{2}$$

$\mathbf{q}(\varphi)$ takes φ as an argument since φ biases the set of information on which the judgment is based. Since $\mathbf{q}(\varphi)$ is a random variable, the comparative statics take the form of stochastic dominance relations. Theorem 1 proves that changes in φ lead to increases in $\mathbf{q}(\varphi)$ in the sense of first order stochastic dominance.

Theorem 1. *$\varphi_1 \geq \varphi_2$ implies $\mathbf{q}(\varphi_1) \succcurlyeq \mathbf{q}(\varphi_2)$.*

For any given length of memory n, the log-supermodularity of ρ_ω weights the recalled data towards more positive values of ω as φ increases. Since I have assumed that the signal distribution $f(\omega|\theta)$ is log-supermodular, high signal values are good news. This good news property implies that higher expectations will result when a more positive set of data is recollected. Hence, for each fixed n, the distribution of expected values under cue φ_1 will first order stochastically dominate the distribution under φ_2. Since the distribution of n is independent of the cue state, these relations hold when I consider randomizations over n.

The theorem above uses a relation in the strong stochastic order that is not strict, whereas in certain circumstances I may require a strict stochastic ordering. In order to achieve this, I have to strengthen the assumptions on the distributions made above.[12]

Corollary 1. *Assume that:*

(1) *$f(\omega|\theta)$ is strictly log-spm in (ω, θ) and $\rho(\omega|\varphi, H_N, n)$ is strictly log-spm in (ω, φ)*
(2) *There exist $\omega_1, \omega_2 \in H_N$ such that $\omega_1 \neq \omega_2$*
(3) *$q(\circ)$ is strictly increasing*

Then $\varphi_1 > \varphi_2$ implies $\mathbf{q}(\varphi_1) \succ \mathbf{q}(\varphi_2)$

[12]The proof of the corollary follows from a straightforward modification of Theorem 1 and is omitted.

3.4. **Stochastic Optimization and Cues.** Now I relate changes in mnemonic cues to changes in the decisions made by an agent under the cues' influence. Suppose the agent faces a stochastic optimization problem of the form:

$$x^*(\mathbf{H}^R(\varphi)) = \underset{x \in X}{\arg\max} \int u(x;\theta) G(d\theta | \mathbf{H}^R(\varphi)) \tag{3}$$

where X is a lattice and the utility function $u(x;\theta)$ is a function of the choice variable x and an unknown parameter θ. Since the recollection of the agents, $\mathbf{H}^R(\varphi)$, is a random variable parameterized by φ, $x^*(\mathbf{H}^R(\varphi))$ is a random variable whose distribution is induced by $\mathbf{H}^R(\varphi)$. Assume sufficient regularity conditions that the maximizer is unique for all H^R.[13] The agent observes the true history prior to the decision and engages in recollection before the choice is made.

In order to prove the monotone comparative statics result relating cue state, φ, to agent choices, $x^*(\mathbf{H}^R(\varphi))$, I require that $u(x;\theta)$ satisfies a single cross condition with respect to (x,θ).

Definition 1. *(from Athey [1]) The function $u(x;\theta)$ satisfies the two-dimensional single crossing property (2-SCP) if for each $x_H > x_L$ pair there exists θ_0 such that for all $\theta < \theta_0$ I have $u(x_H;\theta) - u(x_L;\theta) \leq 0$ and for all $\theta > \theta_0$ I have $u(x_H;\theta) - u(x_L;\theta) \geq 0$*

The two dimensional single crossing condition, when combined with our assumptions on the information and memory processes, allows me to show that increases in the agent affective state induce the agent to choose higher values from X. This is captured by the following theorem.

Theorem 2. *Assume that $u(x;\theta)$ satisfies the two-dimensional single crossing condition. Then $\varphi_1 \geq \varphi_2$ implies $x^*(\mathbf{H}^R(\varphi_1)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2))$.*

For any given length of memory n, the log-supermodularity of ρ_ω implies that a higher value of φ weights the recalled data towards more positive values of ω. Since I have assumed that the signal distribution $f(\omega|\theta)$ is also log-supermodular, high signal values are good news. This good news property translates into posteriors that increase in the strong stochastic order with increasing values of the cue state for fixed n. I prove that $x^*(H^R)$ is weakly increasing in H^R, which implies the desired result conditional on n. Since the distribution of n is independent of the cue state, the orderings for a fixed n holds when I consider randomizations over n.[14]

---

[13] I refrain from studying set valued $x^*(H^R)$ for notational convenience. If the maximizer is not unique, the results of Athey [1] reveal that $|H^R| = |\widehat{H^R}|$ and $H^R \geq \widehat{H^R}$ imply $x^*(H^R) \geq x^*(\widehat{H^R})$ where $x^*(H^R)$ is a set valued function and the order refers to the strong set order. In this case, the notation $x^*(\mathbf{H}^R(\varphi_1)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2))$ implies that the distribution of the random set increases in the strong-set order when $\varphi_1 \geq \varphi_2$.

[14] Although stated in terms of individual decisions, this result implies that the distribution of judgments in an economy will be shifted by an increase in a public cue state. Consider a large economy where supply and demand are equated to form market prices. Assume that all of the agents are in the same cue state and that the law of large numbers holds so that the demand curve in the market is equal to the distribution of demands formed as a function of the random recollection process. The price is then formed by the activity of the marginal agent in the system, and the theorem above shows that the marginal agent's choice will shift with changes in cue state regardless of the identity of the marginal agent in the economy.

3.5. **Mood, Motivation, and Morale.** The organizational behavior literature has seen a debate between two schools of thought regarding the link between employee morale and productivity. The "Happier-and-Smarter" school of thought suggests that optimistic workers will prove more productive through increased effort choice and heightened creativity. The "Sadder-but-Wiser" hypothesis emphasizes that experimental subjects in a depressed mood tend to have more accurate beliefs regarding uncertain events than their happier, more optimistic peers, which in turn leads subjects in a depressed mood to make better decisions than happy, optimistic subjects.

Interpreting high (low) morale as positive (negative) affect allows me to use my model to analyze the relationship between employee morale and productivity in an otherwise traditional principal-agent model.[15] It turns out that whether or not the level of productivity is positively or negatively correlated with the marginal productivity of effort determines which school of thought correctly predicts employee behavior. In order to understand when and how firms might exploit mood congruent memory, I then embed these insights into a model of the firm's decision to encourage either a positive, high affect culture or a negative, low affect work environment.

I start with a standard model of agent effort choice in a context of imperfect information. Employee productivity, y, is a random variable with a probability density, $f(y|e,\theta)$, that is a function of both an unknown parameter, $\theta \in \mathbb{R}$, and the agent's choice of costly effort, e, at associated cost $c(e)$. The parameter θ describes the marginal productivity of agent effort. The agent has prior $G(\theta)$ over the parameter space and has observed a history of signals distributed as $f(\overrightarrow{\omega}|\theta)$ prior to making her effort choice. I assume that the agent has a strictly increasing utility function for wealth, $v(w)$, and chooses her effort given a strictly increasing wage function, $w : \mathbb{R} \to \mathbb{R}_+$, that maps y to income. As my focus is on the role of affect in motivating the worker, I assume that the wage contract is given exogenously.

Given a stochastic recollected history $\mathbf{H}^R$, the agent's problem is then:

$$\underset{e\in\mathbb{R}_+}{Max}\ E[v(w(y))|e;\mathbf{H}^R(\varphi)] - c(e) = \underset{e\in\mathbb{R}_+}{Max} \int_{\mathbb{R}} \int_{\mathbb{R}} v(w(y))f(y|e,\theta)dyG(d\theta|\mathbf{H}^R(\varphi)) - c(e) \tag{4}$$

I assume that for any θ there is a unique e that solves Equation 4. Given an effort choice e and parameter θ, the agent's expected utility from wages is :

$$W(e,\theta) = \int_{\mathbb{R}} v(w(y))f(y|e,\theta)dy \tag{5}$$

Assumption 5. *$W(e,\theta)$ is supermodular in (e,θ).*

The following example illustrates a setting in which Assumption 5 holds.

[15] Koszegi ([47], [48], [49]) provides studies of ego utility in principal-agent settings. Koszegi interprets mood and morale as preferences over ego or as anticipatory utility, whereas the model is a purely informational account of the effect of mood.

Example 1. *Consider a two point distribution for* y *such that* $y \in \{0,1\}$, *states of the world* $\theta \in \{0,1\}$, *and stochastic output function* $f(y=1|\theta,e) = \theta g(e)$ *where* $g(e)$ *is chosen so that* $g(e) \in (0,1)$, $g'(e) > 0$, $g''(e) < 0$. *This example represents an extreme case of complementarities between the marginal productivity of effort and state wherein effort is only effective in state* $\theta = 1$ *and failure is assured if* $\theta = 0$. *I then have* $W(e,\theta) = v(w(1))\theta g(e)$, *which is supermodular.*

Mood regulated memory implies that positive mood states encourage the recollection of data from memory that supports a belief that θ is a high wage state. Assumption 5 implies that the high wage states are also states with high marginal productivity of effort. In this case, affective cues bias the agent towards optimistic beliefs regarding the productivity of effort. Given his biased beliefs, the agent chooses a level of effort that is increasing in his affective state, as reflected in the following theorem.

Proposition 1. *Let Assumption 5 hold. If* $\varphi_1 \geq \varphi_2$, *then* $\mathbf{e}(\varphi_1) \succeq \mathbf{e}(\varphi_2)$.

If Assumption 5 holds, a firm can achieve higher effort levels from the agents through manipulations that induce a positive affect in the agents.[16] For simplicity, assume the firm can choose the agent's affect level φ by spending $m(\varphi)$. I think of $m(\varphi)$ as including costly ancillary benefits, mood enhancing improvements to the work-place, and time spent by managers and other figures investing in employee morale and encouraging a positive work-culture.

Suppose the firm has a reduced form profit function gross of wages $\Pi(y,\kappa)$, where y denotes the employee productivity level and $\kappa \in \mathbb{R}$ is a parameter that indexes the sensitivity of the firm's revenues to employee productivity. $\Pi(y,\kappa)$ is more sensitive to employee productivity as κ increases if $\frac{\partial^2 \Pi(y,\kappa)}{\partial \kappa \partial y} > 0$. For example, creative and service industries are relatively sensitive to the effort level of the employees,[17] whereas a capital-intensive manufacturing firm's output is relatively insensitive to employee effort.[18]

The firm's problem is then:

$$\varphi(\kappa) = \underset{\varphi \in \mathbb{R}}{argmax}\ E[\Pi(\mathbf{y}(\varphi),\kappa) - w(\mathbf{y}(\varphi))] - m(\varphi) \tag{6}$$

[16]Since the effort choice is based on the stochastic recollection of data, I mean the term *higher* to be interpreted as a greater distribution of effort choices in the strong stochastic order.

[17]Although I do not take the analysis very seriously, an inspection of the 2017 *Fortune* "100 Best Places to Work" list has nine of the top ten workplaces within the high-tech or business services industries. These industries are very human capital intensive and firm profits are likely to be sensitive to the effort exerted by the firm's staff.

[18]Suppose the firm operates in a competitive market and the production function takes the Cobb-Douglas form:

$$y = A * K^{\alpha} L^{1-\alpha}$$

The comparative statics parameter κ is analogous to the parameter $1-\alpha$ of the production function. The theory predicts a correlation between $1-\alpha$, the amount of money and effort firms spend on employee morale, and proxies for the happiness and morale of firm employees.

where $\mathbf{y}(\varphi)$ refers to the distribution of productivity induced by the distribution of employee effort levels $\mathbf{e}(\varphi)$. The objective is supermodular in (φ, κ) and so by Theorem 4 of Milgrom and Shannon [61] I know $\varphi(\kappa)$ is weakly increasing in κ. Therefore, if Assumption 5 holds, firms spend more to increase employee morale in industries in which corporate revenues are relatively sensitive to employee productivity. To the author's knowledge, a comparative static of this form has not been tested in the organizational behavior literature.

The flip side to confidence enhanced performance is *defensive pessimism*, the mental habit of cultivating pessimistic moods (and hence beliefs) in order to induce greater effort. Consider a graduate student studying for qualifying exams. A student that is optimistic about the difficulty of the exam believes that there is little marginal benefit to studying. However, by focusing attention on the difficulty of past exams and the consequences of failure, the student can induce a negative affective state that encourages the recall of negative information about his preparedness for the upcoming exam, and these negative beliefs motivate the student to study.

Modifying the structure above, suppose that $W(e, \theta)$ is supermodular in $(e, -\theta)$. Therefore, states with a high utility are associated with a low marginal productivity of effort. I now provide an example of such a function.

Example 2. *Consider a two point distribution of output, $y \in \{0, 1\}$. Suppose there are two possible states of the world, $\theta \in \{0, 1\}$, and stochastic output function $f(y = 1|\theta, e) = \theta + (1 - \theta)g(e)$ where $g(e)$ is concave and chosen so that $g(e) \in (0, 1)$, $g'(e) > 0$, $g''(e) < 0$. I then have $W(e, \theta) = v(w(1))(\theta + (1 - \theta)g(e))$. Then $W(e, \theta)$ is supermodular in $(e, -\theta)$.*

A straightforward adaptation of the proof of Proposition 1 implies that $\varphi_1 > \varphi_2$ yields $\mathbf{e}(\varphi_2) \succcurlyeq \mathbf{e}(\varphi_1)$. Again, assuming that κ indexes the sensitivity of firm revenues to employee productivity, the logic above implies that $\varphi(\kappa)$ is weakly decreasing in κ. Firms facing such a production function can take actions to reduce employee affective state by, for example, cultivating a high-stress corporate-culture that emphasizes the costs of failure rather than the benefits of success. The negative affect causes the agents to hold pessimistic beliefs and therefore increase work effort. Again, the incentive to cultivate a work culture that generates a negatively valenced affect is strongest for firms that are sensitive to employee productivity.

Left out of this discussion is the employee's individual rationality constraint. There are three channels through which negative affect might influence the individual rationality constraint of an employee. First, a negative affect presumably directly reduces an employee's present utility — one need only assume that negative affects are aversive. Second, the negative affect makes the employee pessimistic about future utility from employment with the principal. Third, a negative affect may make the employee more pessimistic about his outside options.

The first channel could be modeled as a direct cost that enters the agent's utility function. The second channel is a distortion of the agent's beliefs about the future utility that is captured by my model. Both of these effects tighten the individual rationality constraint by making continued employment a lower utility option. However, if the agent also has pessimistic beliefs about his outside options due to his negative affect (third channel), then the individual rationality constraint will be slackened. If cultivating a negative affect does tighten the individual rationality constraint, then the firm will have to compensate the employee through improved wages or limit the degree of pessimism that is cultivated.

Interestingly, if firms find it profitable to reduce φ, the third channel implies that a firm's choice to encourage a negative affect is a strategic complement to the choice of other firms to encourage a negative affect. To see this, note that if all of the other firms in an industry cultivate negative affects in their employees, then an employee cannot escape a negative working environment by switching firms. The industry-wide culture reduces the utility the employee receives from his outside options, which in turn loosens the limited liability constraint of the employee. The loosened limited liability constraint makes it more appealing for an individual firm to encourage a negative affect in its employees, which closes the loop on my argument that the choice of a negative corporate-culture is a strategic complement across firms.

Finally, the theory outlined above focuses on individual morale and effort choice, but mood and morale are social phenomena and can spread between members of a team. Barsade [5] provides evidence of emotional contagion from a confederate of the experimenter exhibiting either positive or negative emotions to the other participants in the experiment. Although outside of the scope of this paper, this suggests the potential for complementarities of affective manipulations across employees that work together. Barsade [5] surveys evidence that negative emotions spread more effectively than positive emotional states, which implies that firms ought to take care to limit the interactions between unmotivated, demoralized workers and the remainder of the work force in situations where positive affect increases employee effort. Conversely, if a pessimistic outlook encourages effort, the easy contagion of negative affect might amplify the effect of a firm's efforts to cultivate a negative affect amongst it's workers.

4. Dynamic Model of Mood and Memory

In this section I develop a dynamic version of my model of mood and memory. I incorporate rehearsal into the model of memory and provide a model of the dynamics of affective state so that I can endogenize the links between beliefs, outcomes, and emotions. I close this section by providing an extended asset pricing application.

4.1. **Associative Recall and Rehearsal.** Time is indexed as $t \in \{..., -1, 0, 1, 2, ...\}$, and I assume that a new signal, ω_t, is observed and stored in long-term memory each period. As in the static model, each period the agent is required to form beliefs about a random variable θ. To form these beliefs, the agent recollects a subset of previously observed signals. A *history*,

$H_t = \{..., \omega_1, \omega_2, ..., \omega_t\}$, is a set of data observed by the agent prior to period t and stored in long term memory.[19] A *recollected history* is a random set that consists of those events that are recalled by the agent in a particular period and used to form a posterior. Let a typical realization of the recollected history be denoted $H_t^R \subseteq H_t$. I again assume naïveté on the part of the agent in the sense that he does not use knowledge of the affective state to correct for the biases in his recollection.

I incorporate two channels of recall from memory into the model. The first channel is the associative recall featured in the static model of Section 3. I assume that $l(N)$ data are recalled in each period from the last N elements of history H_t through associative recall. Since the history of data is infinite, I use limits as $N \to \infty$ and $l(N) \to \infty$ as a technical device to model the agent's beliefs when he samples from the infinite set H_t. The contents of the recollected history are determined by the recall probability function $\rho\left(H^R|\varphi_t, H_t, l(N)\right)$ given affective state φ_t. For the duration of this section I assume that $l(N)$ is deterministic, although this assumption can be weakened at the cost of extra notation. The full history of affective states is represented by a vector indexed by time, $\overrightarrow{\varphi}_t = (..., \varphi_{-1}, \varphi_0, \varphi_1, ..., \varphi_{t-1}, \varphi_t)$. I discuss the model of the evolution of φ_t in Section 4.2.

I assume that the relative sampling probability has the form:

$$\rho_\omega(\omega|\varphi, H_t, l(N)) = \begin{cases} e^{m\omega\varphi} \text{ if } \omega \in H_t \\ 0 \text{ otherwise} \end{cases} \tag{7}$$

The variable m parameterizes the intensity of the association effect. If m is large, then mood strongly drives the recall of information of a similar valence, whereas if $m = 0$ then each datum is equally likely to be recalled regardless of the agent's mood. Since $e^{m\omega\varphi} > 1$ if $\omega > 0$, clearly ρ_w denotes a relative sampling probability as opposed to a probability.

The second component of recall I model is *rehearsal*, a phenomenon whereby information that was recalled in the recent past is more easily recalled in the present. The data recalled through rehearsal depends on the past affective states that determined the data recalled via association in prior periods. Because of this, the rehearsal process generates autocorrelations in the biased beliefs of the agent.

To capture rehearsal I assume that the agent recollects each piece of information recalled in the prior period, H_{t-1}^R, with probability $\phi \in \mathbb{R}_+$. Note that the data recalled from period $t-1$ includes data retrieved through rehearsal from period $t-2$. Therefore data from period $t-2$ may be recalled through rehearsal, albeit indirectly, in period t. Since the data recalled by association in periods $t-1$ and $t-2$ was influenced by the respective affective states, φ_{t-1} and φ_{t-2}, rehearsal causes the biases in prior periods to influence the data used for inference in period t.

[19]I redefine my notation for a history and a recollected history because time now stretches from $-\infty$ to the present period t.

I make two simplifying assumptions at this point. First, I assume that ω_i is a normally distributed random variable with an unknown mean θ and a known variance of 1. The assumption of a normal distribution for θ in conjunction with my choice of ρ_ω generates a convenient closed form for the expected value of the recalled ω_i. If one combines Equations 1 and 7 and ignores the sampling without replacement, the distribution of recollected histories is the same as in a situation where the data is drawn from the following distribution:

$$\frac{\rho_\omega(\omega|\varphi, H_t, l(N))f(\omega|\theta)}{\int_{-\infty}^{+\infty}\rho_\omega(s|\varphi, H_t, l(N))f(s|\theta)ds} \propto e^{-\omega.^2/2}e^{m\omega\varphi} = e^{\left(-\frac{(\omega_i - m\varphi_t)^2}{2}\right)}e^{\left(\frac{m^2\varphi_t^2}{2}\right)}$$

In other words, the recollected data has an identical distribution to a normal variable with a mean of $m\varphi_t$ and a variance of 1. While my assumption of a normal distribution for θ and an exponential for ρ_ω generates an easy-to-work-with closed form, there is nothing essential to my results that relies on this closed form.

I focus on situations where the agent recollects a large sample of data from memory ($l(N) \to \infty$), but this sample is biased and incomplete ($l(N)/N \to 0$). In effect, the agent is confident in his beliefs regarding θ, but these beliefs exhibit a bias that depends on $\overrightarrow{\varphi}_t$. Given these assumptions, the following proposition implies that I can make use of a simple autocorrelation formula for the biased beliefs given the affective state in the current and all prior periods.

Proposition 2. *Suppose that as $N \to \infty$ both $l(N) \to \infty$ and $l(N)/N \to 0$. Then as $N \to \infty$:*

$$E_t[\theta|\overrightarrow{\varphi}_t] \to m(1-\phi)\sum_{\tau=0}^{\infty}\phi^\tau\varphi_{t-\tau} \tag{8}$$

In order to generate the closed form presented in Equation 19, I need to account for the fact that data could be recalled through association in multiple periods in the recent past. By requiring $l(N)/N \to 0$, I insure that a vanishing fraction of the data recalled in the present period was recalled in the recent past.

For expositional purposes, I now unpack Equation 19 to identify the sources of each term. In period t data is recalled through association, and these data have a (biased) average of $m\varphi_t$ and a volume of $l(N)$. The data recalled through rehearsal consists of $\phi l(N)$ data recalled through association in period $t-1$, $\phi^2 l(N)$ data recalled associatively in period $t-2$ that are recalled via rehearsal in period $t-1$, $\phi^3 l(N)$ such data points from $t-3$, etc. Note that the data recalled through association in period $t-\tau$ have an expectation equal to $m\varphi_{t-\tau}$. Averaging over all of these data (as Bayesian inference requires) yields:

$$E_t[\theta|\overrightarrow{\varphi}_t] = m(1-\phi)\sum_{\tau=0}^{\infty}\phi^\tau\varphi_{t-\tau}$$

A more direct model of mood's influence on beliefs would be to assume the conclusion of Proposition 2 and discard the model of memory entirely. However some of my most striking predictions, such as the interaction of the degree of bias and the volatility of the underlying asset (section 4.4.3), would not have followed from directly assuming Proposition 2. These additional predictions stem from the fact the bias is informational rather than a reduced form shift of expectations.

4.2. **Dynamics of Affect.** I now describe how the valence of the sentiment of the agent evolves over time. The law of motion for the valence of the agent's affect is:

$$\varphi_{t+1} = \varphi_t + \sigma_{t+1}$$

where σ_{t+1} is an innovation to the agent's affect. σ_{t+1} is determined by news in the current period, ω_{t+1}, and last period's affective state, φ_t. In the asset pricing application below, ω_{t+1} denotes a change in the dividend in the current period.[20] I employ the following formula for innovations to affect:

$$\sigma_{t+1} = g(\varphi_t, \omega_{t+1}) = \begin{cases} (1-\varphi_t)(1-e^{-\omega_{t+1}}) \text{ if } \omega_{t+1} \geq 0 \\ (1+\varphi_t)(e^{\alpha\omega_{t+1}}-1),\ \alpha \in (0,1) \text{ if } \omega_{t+1} < 0 \end{cases} \tag{9}$$

I insure that $\varphi_{t+1} \in (-1, 1)$ by making the innovation proportional to $1-\varphi_t$ or $1+\varphi_t$ as appropriate. When $\alpha < 1$ the innovations described by g have the property that bad news ($\omega_{t+1} < 0$) causes smaller changes in φ_{t+1} than an equivalent[21] good news event, which follows naturally from the theory of *immune neglect*. Immune neglect refers to the empirical regularity that bad news has little effect on affect (Gilbert et al. [27]) relative to the effect of good news. In Figure 2 I plot g for 3 different affective states φ_t to illustrate the effects of bounded affect and immune neglect. The kink at $\omega = 0$ is a result of the asymmetric treatment of positive and negative news on affect caused by immune neglect.

An important effect of the boundedness of the affective state and the unbounded support of the affect innovations is that φ_t is ergodic. This implies that even the most extreme affective states are transient. The transient nature of the affective impact of the information contained in dividend innovations also differentiates the transient psychological component of the model from the permanent, informational effect of dividend innovations.

Lemma 1. *φ_t is ergodic.*

[20] The choice of the determinants of σ_{t+1} reflects an assumption as to what kinds of events are salient with respect to changes in affective state. Since there is little research in psychology that shines light on the financial model, it remains an empirical matter to use predictions based on different assumptions regarding the saliency of different market variables to choose the theory that makes superior empirical predictions. A leading alternative as a driver for changes in the representative agent's affect is total returns, $\sigma_{t+1} = P_t + d_t - (1+r)P_{t-1}$. Under this alternate theory I do not find short-run continued overreaction, but some of the other findings hold qualitatively. I provide a brief analysis of a model where returns drive mood changes in appendix C.

[21] Given φ_t and $\varepsilon_{t+1} \geq 0$, an equivalent bad news event is $1-\varphi_t$ and $-\varepsilon_{t+1}$.

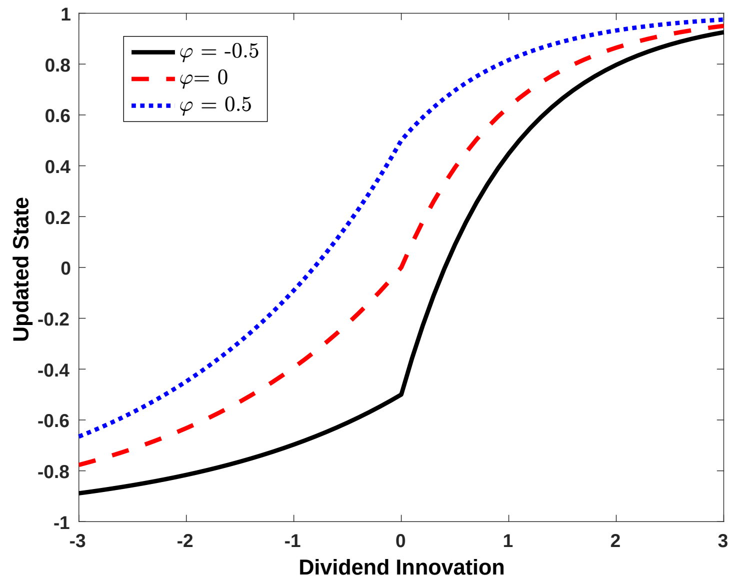

FIGURE 2. Evolution of Affective State

4.3. **Asset Pricing.** I assume throughout that the beliefs under study are those of a risk-neutral representative agent, which implies that an asset's price equals the net present value of expected future dividends with a discount factor $r > 0$. I assume that dividends at time $t \in \{.., -1, 0, 1, 2, ...\}$, denoted d_t, are described by a stochastic process of the form $d_{t+1} = d_t + \omega_{t+1}$ where the ω_t are independently and identically distributed normal random variables with a known variance of 1 and an unknown mean of θ. For notational ease I assume the true value of θ is 0.

Prior to the time at which beliefs about asset valuations are formed, the agent is assumed to have observed a history of dividend innovations and stored this history in long-term memory. For the representative agent to compute the net present value of the asset, and hence determine the asset's price, the representative agent needs to form an expectation regarding future dividend growth (i.e., θ). When forming a belief about dividend growth, the agent recollects a sample of data from memory in order to update his prior beliefs and form a posterior about the mean of the dividend innovation in each period.[22] Throughout the analysis, I assume that the representative agent has the same mnemonic biases as the individual agents. The representative agent in the model fully and correctly incorporates recollected information into beliefs about asset valuation using Bayes's rule, but the agent does not correct for his faulty memory. The affective state of the representative agent is influenced by news in the present period with positive (negative) dividend innovations

[22] One could also interpret the recalled dividend innovations as other forms of information regarding the asset's fundamental value such as the opinions of equity analysts, information gleaned from contacting industry insiders, and interviews with corporate officers.

encouraging positive (negative) affective states.[23] The affect of the representative agent should be interpreted as the average affect of the market participants.

A risk-neutral representative agent has an expectation about the net present value of a security (and hence the market price) equal to:

$$P_t(\overrightarrow{\varphi}_t) = E_t\left[\sum_{\tau=1}^{\infty} \frac{d_{t+\tau}}{(1+r)^{\tau}} \middle| \overrightarrow{\varphi}_t\right] = \frac{d_t}{r} + \frac{1+r}{r^2} E_t[\theta|\overrightarrow{\varphi}_t] \tag{10}$$

where $E_t[\circ|\overrightarrow{\varphi}_t]$ represents the time t expectation of the agent given a history of affective valences $\overrightarrow{\varphi}_t$. Since the true mean of ω_{t+1} is 0, I can describe the bias relative to a perfectly informed observer as:

$$\delta(\overrightarrow{\varphi}_t) = \frac{1+r}{r^2} E_t[\theta|\overrightarrow{\varphi}_t] = \frac{1+r}{r^2} m(1-\phi) \sum_{\tau=0}^{\infty} \phi^{\tau} \varphi_{t-\tau} \tag{11}$$

where Equation 11 uses Proposition 2.

A note of interpretation is required to distinguish the two roles that the dividend innovation ω_t plays in the model. As in a classical model of asset pricing, ω_t provides information about future dividend payouts. The second role of the dividend innovations is to determine the evolution of the representative agent's affect. In this second role ω_t provides a context for the formation of beliefs through the biased recall caused by the affective state.

Many of the predictions of this section turn on comparing various aspects of asset price performance with indices of market affect. Prior work has used measures as varied as weather over the relevant exchange (Saunders [68], Hirshleifer and Shumway [32]), seasonal changes in the length of

---

[23] There are significant reasons to question whether biases, such as imperfect recall, that are plausible on the individual level aggregate to the level of a representative agent and have a market-wide impact. First, since the model is fundamentally one of biased beliefs, one might conjecture that arbitrageurs would have significant power to mitigate the impact of biased beliefs on prices. However, there is a rich literature on the limits of arbitrageurs to correct mispricings within the market (Shleifer [74] provides an early survey). Factors that limit arbitrage include risk aversion on the part of arbitrageurs and constraints on and costs of short sales. These limits are exacerbated by the potentially long time it can take for asset prices to correct when the ability to arbitrage is bounded.

Second, one might object to the model by arguing that although imperfect recall has plausible and significant effects on the individual level, market participants who are significantly affected by the resulting biased beliefs will be forced out of the market or choose not to participate. Chapman and Polkovnichenko [12] points out that in an economy with heterogeneous agents, potential participants with non-standard preferences who are aware of their preferences may choose to exit the market. The net effect of this selection is that the agents with non-standard preferences have a very limited effect on equilibrium prices. Most market participants recognize their own imperfect recall, but few of these agents are aware of the biases caused by associative memory. Therefore imperfect recall provides little motivation to exit the market.

Third, involuntary selection effects can also occur as agents are forced out of the market over time due to their poor decisions. Blume and Easley [11] and Easley and Yang [23] model this selection and prove that, when the selection occurs, the price impact of agents with incorrect beliefs vanishes. However, even if agents with severe biases due to imperfect recall are forced out of the market, new agents with these biases who enter over time would continue to have an impact on market prices.

the day (Kamstra, Kramer and Levi [42]), and sporting match outcomes (Edmans, Garci and Norli [24]). I close this subsection with suggestions for new measures of affect.

One novel mood measure is the Gallup Daily: U.S. Mood Poll.[24] In the context of the theory, the advantage of examining mood as a predictor is that it highlights the exact psychological channel I am studying whereas weather is an indirect measure of affect. Two issues arise with interpreting a naive regression using the Gallup poll as a measure of affect. First, the U.S. Mood Poll is presumably correlated with diffuse information regarding future market performance (e.g., the poll is correlated with informative consumer confidence). As suggested in Tetlock [77], if a correlation between the poll and asset prices is purely informational, one would expect a permanent shift in asset prices once the poll is released. On the contrary, if the poll's ability to predict asset price changes (at least partially) reflects the influence of mood, then the theory predicts partial reversion of these changes due to the transient nature of the affective biases. The second issue is that reverse causality could be present if positive stock market moves induce poll respondents to express a more positive mood. One could use lagged poll responses to limit effects of this nature.

Another possible index for market affect is the Michigan Consumer Sentiment Index (MCSI). The questions on which the index are based query consumer expectations about current and future economic performance both at the personal and the aggregate level. I believe that the MCSI holds promise as a measure of market affect, but the caveats regarding the potential information contained within the index noted above for the Gallup Poll also apply to the MCSI. Therefore, care must be taken to isolate the purely affective component of the index.

4.4. **Market Effects.** In this section I trace out linkages between affective state, beliefs, and market prices for financial assets. I show that the effects of dividend announcements on mood can explain short-run continued overreaction and long-run correction of securities prices to the release of news, price momentum, and excess volatility of asset prices (especially during market downturns). I close by considering two risky assets with different volatilities, and I argue the high volatility assets exhibit stronger price biases than low volatility assets.

4.4.1. *Dynamics of Overreaction and Correction.* Let E_t^* refer to an expectation with respect to the true dividend process, and I interpret E_t^* as reflecting the unbiased estimates of an outside observer such as an econometrician. I study short-run continued overreaction to news and long-run correction by examining the difference between the market price and the outside observer's estimate of the net present value of the asset following valenced news events. Since shocks to fundamental value, ω_t, influence both the fundamental component of prices, d_t, and the evolution of the affective state, φ_t, the response to such a shock will evolve over several periods. Recall that in a model with

[24]Available at http://www.gallup.com/poll/106915/gallup-daily-us-mood.aspx

unbiased recall the future price increments are unpredictable, so:

$$E_t^*[P_{t+1} - P_t|\omega_t \geq 0] = E_t^*[P_{t+1} - P_t|\omega_t < 0]$$

since all of the information from ω_t is incorporated into P_t

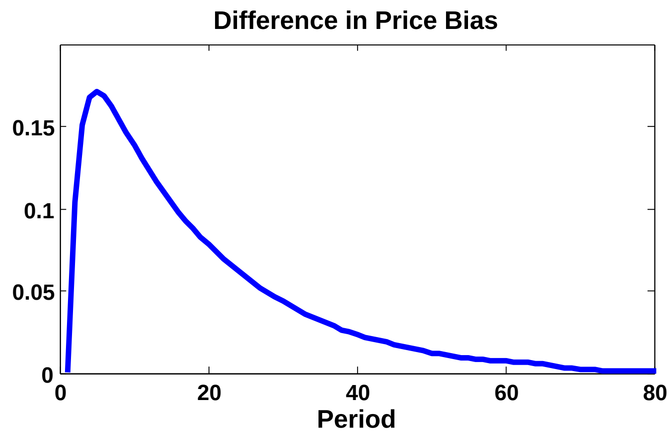

FIGURE 3. Price Impulse Response Difference

Short-run continued overreaction to dividend innovations reflects a continued reaction to old information generated by the autocorrelation produced by rehearsal. To illustrate these effects, I plot the following in Figure 3:

$$E_t^*\left[\delta_{t+\tau}(\overrightarrow{\varphi}_{t+\tau})|\omega_t > 0\right] - E_t^*\left[\delta_{t+\tau}(\overrightarrow{\varphi}_{t+\tau}))|\omega_t < 0\right]$$

The impulse response to dividend innovations increases over the short-run due to the autocorrelation in the evolution of $\varphi_{t+\tau}$ in response to ω_t, but these effects fade as the affective state returns to its long-run ergodic distribution.

I can write the price changes as:

$$\begin{aligned} P_{t+1} - P_t &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2} m(1-\phi)\left[\sum_{\tau=0}^{\infty}\phi^\tau \varphi_{t+1-\tau} - \sum_{\tau=0}^{\infty}\phi^\tau \varphi_{t-\tau}\right] \\ &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2} m(1-\phi)\sum_{\tau=0}^{\infty}\phi^\tau \sigma_{t+1-\tau} \end{aligned}$$

where $E_t^*\left[\omega_{t+1}\right] = 0$. Two countervailing effects are at work. First, I have that $E_t^*[\sigma_t|\omega_t \geq 0] > 0 > E_t^*[\sigma_t|\omega_t < 0]$, which reflects the changes in the affective state driven by news in the current period. The autocorrelation of price biases implies that this effect is carried forward to period $t+1$ and discounted by the rehearsal probability ϕ. Second, since the affective state has been moved away from the ergodic distribution and converges back to the ergodic distribution monotonically in expectation I have $E_t^*\left[\sigma_{t+1}|\omega_t \geq 0\right] < E_t^*\left[\sigma_{t+1}|\omega_t \leq 0\right]$. I predict short-run continued overreaction if ϕ is sufficiently large that the first effect dominates the second effect in the short-run.

Proposition 3. *For $\phi \in (0,1)$ sufficiently large, asset prices exhibit short-run continued overreaction to news*

$$E_t^*[P_{t+1} - P_t|\omega_t \geq 0] > E_t^*[P_{t+1} - P_t|\omega_t < 0] \tag{12}$$

My second prediction concerns long-run correction of short-run continued overreaction. The model predicts that an econometrician studying asset price time-series data finds that the effects of memory biases fade over time as the distribution of affective states asymptotically returns to the ergodic distribution. This convergence implies the affect component of any price change fades with time, so price changes become unpredictable in the long-run.

Proposition 4. *There is long-run correction of the price effect of changes in affect in the current period:*

$$\underset{\tau\to\infty}{Lim}\; E_t^*[P_{t+\tau} - P_t|\omega_t \geq 0] \leq \underset{\tau\to\infty}{Lim}\; E_t^*[P_{t+\tau} - P_t|\omega_t < 0]$$

Most empirical studies of short-run continued overreaction to news use price changes as the conditioning variable. However, conditioning on price changes is problematic for several reasons. First, it is unclear whether $P_t - P_{t-1} > 0$ implies $\omega_t > 0$ as it could be that the price is recovering from low affect in past periods despite bad news in the current period. This leads to the second difficulty, the correlation of price changes and φ in previous periods, which makes the computations more difficult. I have been unable to produce an analogue of Proposition 3 that conditions on price changes in period t. To show the results of Propositions 3 and 4 extend to the case where I condition on price changes in period t, I numerically compute:

$$E_t^*[P_{t+\tau+1} - P_{t+\tau}|P_t - P_{t-1} \geq 0] - E_t^*[P_{t+\tau+1} - P_{t+\tau}|P_t - P_{t-1} < 0]$$

The result is displayed in Figure 4. The difference in the response of future price changes to the price change in the current period is positive in the short-run (similar to the predictions of Proposition 3), which I refer to as *asset price momentum.* The sign of the price changes become negative in the medium-run, which represents the overcorrection of the short-run. At longer horizons the effect of mnemonic biases fades entirely as the affective state returns to the ergodic distribution.

4.4.2. *Excess Volatility.* Empirical studies find that asset prices exhibit volatility greater than can be explained by the fundamental contribution alone (see, for example, Shiller [72], [73]). Changes in asset prices are driven by three effects in my model. Information about future dividends is revealed by dividend innovations in the current period. This information is incorporated into the representative agent's beliefs about the net present value of the asset, and the price movement caused by this updating is referred to as *fundamental volatility.* The movement of affective state

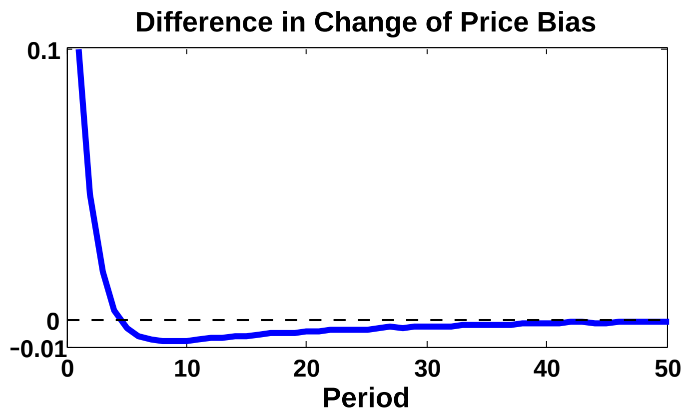

FIGURE 4. Autocorrelation of Bias

will cause price volatility that I denote as *affective volatility.* Furthermore, information that drives fundamental changes in price also drives changes in affect, resulting in an *affective volatility of information.* I make two predictions regarding the effect of associative memory on excess volatility. First (and unsurprisingly), mnemonic biases increase the volatility of asset prices. Second, I predict that excess volatility is highest when the representative agent's affect is depressed (i.e., during market downturns).

From period t to period $t+1$ the price increment is:

$$\begin{aligned} P_{t+1} - P_t &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2} m(1-\phi) \sum_{\tau=0}^{\infty} \phi^{\tau} \left(\varphi_{t+1-\tau} - \varphi_{t-\tau} \right) \\ &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2} m(1-\phi) \sum_{\tau=0}^{\infty} \phi^{\tau} \sigma_{t+1-\tau} \end{aligned}$$

I interpret the variance of this price increment as price volatility. The fundamental contribution to volatility is:

$$\left[\frac{1}{r}\right]^2 Var\left(\omega_{t+1}\right) > 0$$

The affective volatility is: proportional to:

$$Var\left(\sum_{\tau=0}^{\infty} \phi^{\tau} \sigma_{t+1-\tau}\right) = \frac{Var(\sigma_{t+1})}{(1-\phi)^2} + \frac{2}{1-\phi} \sum_{\tau=1}^{\infty} \phi^{\tau} Cov(\sigma_{t+1}, \sigma_{t+1-\tau}) > 0$$

The final component of volatility, the affective volatility of information, is:

$$\frac{1+r}{r^3} m(1-\phi) Cov\left(\omega_{t+1}, \sigma_{t+1}\right) > 0$$

The affective volatility of information is positive due to the fact that ω_{t+1} and σ_{t+1} have the same sign by definition (i.e., Equation 9).

I use numerical simulations to compute the volatility of σ_t as a function of φ_t in Figure 5. It is

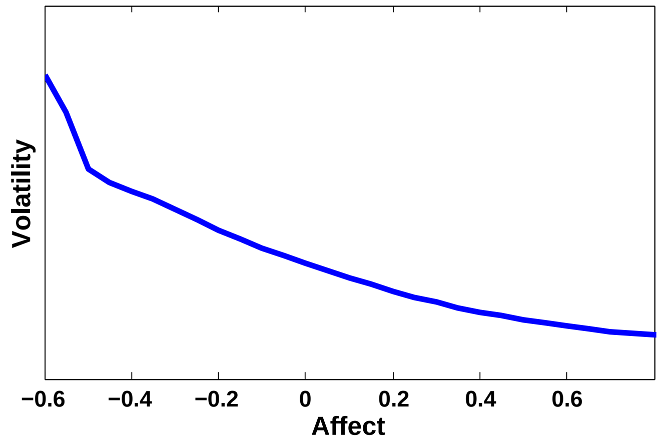

FIGURE 5. Volatility and Affect

clear from the figure that the effect of φ_t is asymmetric, a feature that is generated by two forces. First, the effect of positive shocks to affect is highest when affective state is low. Second, the complementary effect of negative shocks to affect when φ is large is dampened by immune neglect. Together these effects generate an asymmetry between the volatility in high and low affective states.

My predictions regarding the asymmetry of excess volatility as a function of the market state mirror findings in the ARCH literature dating back to Black [10]. In particular, the earlier literature finds that volatility rises following bad news and falls following good news. My predictions are somewhat different since I focus on excess volatility generated by the behavioral bias I study rather than the total return volatility. One could test my predictions by analyzing the conditional volatility using a GARCH model with an affect measure incorporated into the volatility equation alongside a traditional set of controls.

4.4.3. *Effect of Volatility of Dividends.* The log-supermodularity of ρ_ω implies that the probability of recollecting an event of similar valence to the representative agent's affect is increasing in the extremity of the event. When an agent in affective state φ_t recollects data about a highly volatile asset through associative memory, the data recalled is dominated by the extreme events of the same sign as φ_t. In contrast, the data recalled by an agent (in the same affective state) about a low volatility asset, while biased towards recollecting extreme events, is dominated by the moderate dividend innovations generated by the low volatility asset. These effects makes the price biases for assets with high fundamental volatility larger than the corresponding biases for assets with a low fundamental volatility. This provides an alternative explanation for why measures of fundamental volatility, as used in Kumar [51] and Zhang [81], predict the magnitude of price biases.

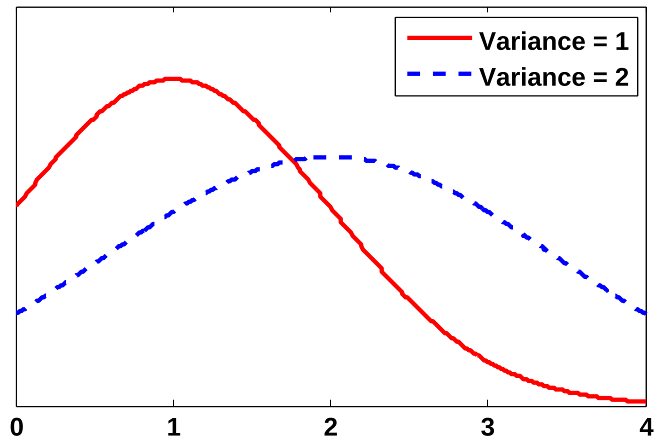

Figure 6. Volatility and Magnitude of Biases in Recollection

The prediction is illustrated in Figure 6. The dividend innovations are normally distributed with mean 0 and a variance equal to 1 for the low volatility asset and a variance of 2 for the higher volatility asset. have plotted the distribution of data recollected from the representative agent's memory through association. The mean of the recalled data for the high variance asset is approximately twice the mean of the recalled data for the lower variance asset, which implies that the bias in the representative agent's beliefs about the net present value of the high volatility asset is twice as large as the corresponding bias for the low volatility asset. Although I have held fixed the affective valence when computing Figure 6, high volatility of ω_t would also cause high volatility in φ_t, which would amplify these effects.

In order to test the prediction, the econometrician would need to regress a cross-section of asset price changes on a standard set of controls, a metric for the representative agent's affective state, and an interaction between fundamental volatility and affective state. The results of this section predict that the interaction term is positive, while the results of section 4.4 suggest the coefficient on the affective state changes is positive as well. While I stated this result as a cross-sectional prediction, one could also interpret this as a time series prediction for a single risky asset that exhibits periods of low and high fundamental volatility.

5. Conclusion

The goal of my paper is to provide a foundation for modeling the effect of mood and memory on beliefs and decisions. I began by providing a general model of the effect of arbitrary mnemonic cues on decisions under uncertainty, and I derived robust comparative statics that can be employed in a variety of economic settings. I then used my theory to model employee morale. The application to morale provides a framework for interpreting the conflicting results in the empirical literature on morale and employee productivity. One conclusion of my analysis is that the influence of affect on employee effort and resulting firm profitability depends on whether optimistic employee beliefs about productivity level are correlated with optimism about the marginal value of effort. By identifying this relationship between mood, biased beliefs, and the marginal value of effort, I am able to generate novel predictions regarding correlations between a firm's industry, the employee's job function, and the amount of firm resources expended on morale enhancing activities. Additional care must be taken when interpreting the prior literature since, as a rule, these studies do not present evidence sufficient to answer the question as to whether beliefs and effort are complementary. Testing my predictions in either a laboratory or field setting remains a subject for future research.

I then extend the model to dynamic environments. To close my model of mood and memory in dynamic settings, I provide a model of the dynamics of affect that incorporates two important stylized facts from the psychology of affect. First, the future prospects of the agent as signaled through the receipt of good and bad news influence the agent's affect. Second, the agent's mood is moderated by immune neglect, which refers to the empirical regularity that bad news has less effect on mood compared to good news. Given my model of the dynamics of affect, I can explore how the coevolution of mood and beliefs influences market outcomes.

I then apply the dynamic model to an asset pricing problem. The model provides a unified explanation for short-run continued overreaction to news and long-run correction of prices and price momentum. The model also predicts excess volatility in asset prices and decomposes the volatility into three components: fundamental volatility, price volatility driven by volatility of market affect, and volatility driven by the correlation of news and affect. The model also makes the prediction that excess volatility will be highest in market downturns. Finally, my model predicts that the effect of price biases is largest for assets with a high degree of fundamental volatility.

Memory errors could also play a significant role in how shareholders evaluate executive competency. If firm performance has been exceptional, shareholders will be in a positive mood and optimistic about future performance. Shareholders might conclude that firm management is of higher quality than the evidence warrants and as a result fail to hold corporate officers accountable for relative performance. The timing of IPOs and other strategic market events could also be related to mnemonic biases. Market makers may time the sale of securities to take advantage of transient shifts in market affect, much as the bankers managing IPO publicity in Cook, Kieschnick, and Van Ness [14]. I hope to explore these applications in future work.

In addition to the predictions presented above, a number of related results can be explored. For example, changes in the representative agent's affect will influence the prices of all of the assets in the market, which implies that changes in prices will be correlated even if the fundamentals determining future dividends are not. This prediction is difficult to test because one would have to control for the many potential ways in which the fundamental values of two seemingly unrelated assets could be correlated (e.g., the value of any two assets is presumably fundamentally correlated through the response to the general health of the economy).

My theory also provides an explanation for the effect of repeated exposure to the same information. While there are other explanations for some of my predictions (e.g., short-run continued overreaction to news), one of two explanations is most plausible for why repeated exposure to information has an effect on market prices. First, it could be that some market participants were not exposed to the first release of information due to the information's limited circulation. This is unlikely in the EntreMed case studied by Huberman and Regev [37] as both information releases were reported in at least one widely-read source, the *New York Times*. The second explanation is imperfect memory. The effects of repeated exposure in the model are mediated through two channels. First, a market participant may simply forget the data between exposures. Second, exposure to a datum a second time can cue the recall of related data, which can in turn bias beliefs. The theory predicts that the second exposure to the news ought to have an outsized effect on the market's assessment of EntreMed's stock price due to associative recall and that these effects fade (in expectation) with the gradually decreasing effects of rehearsal.

There are a number of additional applications of the model in settings such as advertising and persuasive speech where affect and cued memory are important influences on agent behavior. One example includes the choice of firms to employ either affective or informative advertising strategies. By incorporating my theory of how consumers respond to affective cues into a model of advertising and market competition, one could derive predictions as to which types of firms use each form of advertising and how these advertising strategies segment the customer base. Persuasive speech often contains emotional cues with little novel information. This form of speech is important within the political sphere where politicians face a trade-off between revealing policy information to voters or spending their valuable media exposure opportunities attempting to influence the voters with affective appeals. It would be of interest to study both the determinants of the politicians' choice of the nature of their public speeches and how this influences the decision processes of the voting public.

References

[1] Athey, S. (2002) "Monotone Comparative Statics Under Uncertainty," *The Quarterly Journal of Economics*, 117, pp. 187 - 223.

[2] Atkinson, R. and R. Shiffrin(1968) "Chapter: Human memory: A proposed system and its control processes" in *The psychology of learning and motivation (Volume 2)*, New York: Academic Press. pp. 89–195.

[3] Au, K.; F. Chan; D. Wang; and I. Vertinsky (2003) "Mood in foreign exchange trading: Cognitive processes and performance," *Organizational Behavior and Human Decision Processes*, 91 (2), pp. 322-338.

[4] Barberis, N; A. Shleifer; and R. Vishny (1998) "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, pp. 307 - 343.

[5] Barsade, S. (2002) "The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior," *Administrative Science Quarterly*, 47, pp. 644-675.

[6] Benabou, R. and J. Tirole. (2004) "Willpower and Personal Rules," *Journal of Political Economy*, Vol. 112. pp. 848-886.

[7] Benabou, R. and J. Tirole. (2006a) "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 117, pp. 871-915.

[8] Benabou, R. and J. Tirole. (2006b) "Incentives and Prosocial Behavior," *American Economic Review*, 96, pp. 1652 - 1678.

[9] Benabou, R. and J. Tirole. (2011) "Identity, Dignity, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics,* 126, pp. 805-855.

[10] Black, F. (1976) "Studies of Stock Market Volatility Changes," *1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 177-181.

[11] Blume, L. and D. Easley (2006) "If You're So Smart, Why Aren't You Rich? Belief Selection in Complete and Incomplete Markets," *Econometrica*, 74 (4), pp. 929 - 966.

[12] Chapman, D. and V. Polkovnichenko (2009) "First-Order Risk Aversion, Heterogeneity, and Asset Market Outcomes," *The Journal of Finance*, 64 (4) pp. 1863 - 1887.

[13] Clore, G.L.; N. Schwarz; and M. Conway (1984) "Affective Causes and Consequences of Social Information Processing" in *Handbook of Social Cognition*, Vol. 1, pp. 323-415.

[14] Cook, D.; R. Kieschnick; and R. Van Ness (2006) "On the Marketing of IPOs," *Journal of Financial Economics*, 82, pp. 35-61.

[15] Compte, O. and A. Postlewaite (2004) "Confidence-Enhanced Performance," *American Economic Review,* 94, pp. 1536 - 1557.

[16] Cutler, D.; J. Poterba; and L. Summers (1991) "Speculative Dynamics," *Review of Economic Studies*, 58, pp. 529 - 546.

[17] Daniel, K.; Hirshleifer, D.; and Subrahmanyam, A. (1998) "Investor Psychology and Security Market Under- and Overreactions," *The Journal of Finance*, 6, pp. 1839 - 1885.

[18] De Bondt, W. and R. Thaler, "Does the Stock Market Overreact?" *The Journal of Finance*, 40, pp. 793 - 805.

[19] DeGroot, M. (2004) *Optimal Statistical Decisions,* Hoboken, NJ: John Wiley & Sons, Inc.

[20] DeLong, J; A. Shleifer; L. Summers; and R. Waldmann (1990) "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, pp. 703-738.

[21] Derrien, F. (2005) "IPO Pricing in 'Hot' Market Conditions: Who Leaves Money on the Table?" *The Journal of Finance*, 60, pp. 487 - 521.

[22] Dow, J. (1991) "Search Decisions with Limited Memory," *The Review of Economic Studies*, 58, 1-14.

[23] Easley, D. and L. Yang (2015) "Loss Aversion, Survival and Asset Prices," *Journal of Economic Theory*, 160, pp. 494 - 516.

[24] Edmans, A.; D. Garcia; and O. Norli (2007) "Sports Sentiment and Stock Returns," *The Journal of Finance*, pp. 1967 - 1998.

[25] Elster, J. (1998) "Emotions and Economic Theory," *Journal of Economic Literature*, 36, 47-74.

[26] K. Ericson (2017) "On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation," *Journal of the European Economic Association*, 15 (3), pp. 692 - 719.

[27] Gilbert, D.; E. Pinel; T. Wilson; S. Blumberg; and T. Wheatley (1998) "Immune Neglect: A Source of Durability Bias in Affective Forecasting," *Journal of Personality and Social Psychology,* 75 (3), pp. 617 - 638.

[28] Gottlieb, D. (2015) "Imperfect Memory and Choice Under Risk," *Games and Economic Behavior*, 85, pp. 127 - 158.

[29] Hanif, M. and K. Brewer (1980) "Sampling with Unequal Probabilities without Replacement: A Review," *International Statistical Review*, 48 (3), pp. 317 - 335.

[30] Hermalin, B.E. and A.M. Isen (2006) "A Model of the Effect of Affect on Economic Decision Making," *Quantitative Marketing and Economics,* 6, pp. 17-40.

[31] Hellman, M. and T. Cover (1970) "Learning with Finite Memory," *The Annals of Mathematical Statistics*, Vol. 41, pp. 765-782.

[32] Hirshleifer, D. and T. Shumway (2003) "Good Day Sunshine: Stock Returns and the Weather," *The Journal of Finance*, pp. 1009 - 1032.

[33] Hirshleifer, D. and I. Welch (2002) "An Economic Approach to the Psychology of Change: Amnesia, Inertia, and Impulsiveness," *Journal of Economics and Management Strategy*, 11 (3), pp. 379-421.

[34] Hong, H. and J. Stein (1999) "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *The Journal of Finance*, 54, pp. 2143 - 2184.

[35] Hong, H. and J. Stein (2007) "Disagreement and the Stock Market," *Journal of Economic Perspectives*, 21, pp. 109 - 128.

[36] Hong, H.; J. Stein; and J. Yu (2007) "Simple Forecasts and Paradigm Shifts," *The Journal of Finance*, 62, pp. 1207 - 1242.

[37] Huberman, G. and T, Regev (2001) "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar," *The Journal of Finance*, 56 (1), pp. 387 - 396.

[38] Isen, A. (1993) "Positive Affect and Decision Making" in *Handbook of Emotions*, pp. 261 - 277.

[39] Isen, A.; T. Shalker; M. Clark; and L. Carp (1978) "Affect, Accessibility of Material in Memory and Behavior: A Cognitive Loop?" *Journal of Personality and Social Psychology*, 36, 1-12.

[40] Jegadeesh, N. and S. Titman (1993) "Returns to Buying Winner and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, 48, pp. 65 - 91.

[41] Jegadeesh, N. and S. Titman (2001) "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *The Journal of Finance*, 56, pp. 699 - 720.

[42] Kamstra, M.; L. Kramer; and M. Levi (2003) "Winter Blues: A SAD Stock Market Cycle," *American Economic Review*, Vol. 93, pp. 324-343.

[43] Kamstra, M.; L. Kramer; M. Levi; and T. Wang (2014) "Seasonally Varying Preferences: Theoretical Foundations for an Empirical Regularity," *The Review of Asset Pricing Studies,* 4, pp. 39 - 77.

[44] Karlin, S. and Y. Rinott (1980) "Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions," *Journal of Multivariate Analysis*, 10, 499-516.

[45] Keef, S. and M. Khaled (2011) "Are investors moonstruck? Further evidence on lunar phases and stock returns," *Journal of Empirical Finance*, 18 (1), pp. 56 -63.

[46] Kida, T; J. Smith; and M. Maletta (1998) "The Effects of Encoded Memory Traces for Numerical Data on Accounting Decision Making," *Accounting, Organizations, and Society*, 23, pp. 451-466.

[47] Koszegi, B. (2000) "Ego Utility and Information Acquisition," *mimeo.*

[48] Koszegi, B. (2006) "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association,* 4, pp. 673 - 707.

[49] Koszegi, B. (2006) "Emotional Agency," *Quarterly Journal of Economics,* 121, pp. 121 - 156.

[50] Krivelyova, A. and C. Robotti (2003) "Playing the Field: Geomagnetic Storms and the Stock Market," *mimeo.*

[51] Kumar, A. (2009) "Hard-to-Value Stocks, Behavioral Biases, and Informed Trading," *Journal of Financial and Quantitative Analysis*, 44 (6), pp. 1375 - 1401.

[52] La Porta, R. (1996) "Expectations and the Cross-Section of Stock Returns," *The Journal of Finance*, 49, pp. 1541 - 1578.

[53] Lahav, Y. and S. Meer (2012) "The Effect of Induced Mood on Prices in Asset Markets - Experimental Evidence," *mimeo.*

[54] Lakonishok, J; A. Shleifer; and R. Vishny (1994) "Contrarian Investment, Extrapolation, and Risk," *The Journal of Finance*, 51, pp. 1715 - 1742.

[55] Ljungqvist, A; V. Nanda; and R. Singh (2006) "Hot Markets, Investor Sentiment, and IPO Pricing," *Journal of Business*, 79, pp. 1667 - 1702.

[56] Loewenstein, G. (2000) "Emotions in Economics Theory and Economic Behavior," *AEA Papers and Proceedings*, pp.426-432.

[57] Loewenstein, G. and J. Lerner (2003) "The Role of Affect in Decision Making" in *Handbook of Affective Sciences*, pp. 619 - 642.

[58] Luis, L. and G. Becker (2007) "Evolutionary Efficiency and Happiness," *Journal of Political Economy*, 115 (2), pp. 302 - 337.

[59] Meyn, S. and Tweedie R. (1993) *Markov Chains and Stochastic Stability*, New York, NY: Springer-Verlag.

[60] Milgrom, P.R. (1981) "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, Vol. 12, pp. 380-391.

[61] Milgrom, P. and C. Shannon (1994) "Monotone Comparative Statics." *Econometrica*, Vol. 62, pp. 157-180.

[62] Mullainathan, S. (1998) "A Memory Based Model of Bounded Rationality," *The Quarterly Journal of Economics*, Vol. 117, pp. 735-774.

[63] Muller, A. and D. Stoyan (2002) *Comparison Methods for Stochastic Models and Risks.* New York: John Wiley and Sons, Inc.

[64] Officer, R. (1973) "The Variability of the Market Factor of the New York Stock Exchange," *The Journal of Business,* 46 (3), pp. 434 - 453.

[65] Purnanandam, A. and B. Swaminathan (2004) "Are IPOs really Underpriced?" *The Review of Financial Studies*, 17, pp. 811 - 848.

[66] Rouwenhorst, K. (1998) "International Momentum Strategies," *The Journal of Finance*, 53, pp. 267 - 284.

[67] Sarafidis, Y. (2007) "What Have you Done for me Lately? Release of Information and Strategic Manipulation of Memories," *The Economic Journal*, 117 (518), pp. 307 - 326.

[68] Saunders, E. (1993) "Stock Prices and Wall Street Weather," *American Economic Review*, Vol. 83, pp. 1337–1345.

[69] Schwarz, N. and G.L. Clore (1983) "Mood, misattribution and judgments of well-being: Informative and directive functions of affective states," *Journal of Personality and Social Psychology*, 45 (3), pp. 513-523.

[70] Schwarz, N. and G.L. Clore (1988) "How Do I Feel About It? Informative Functions of Affective States," in *Affect, Cognition, and Social Behavior*, pp. 44-62.

[71] Schwert, G.W. (1989) "Why Does Stock Market Volatility Change Over Time?" *The Journal of Finance*, 44 (5), pp.1115 - 1153.

[72] Shiller, R. (1981) "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71, 421-436.

[73] Shiller, R. (1981) "The Use of Volatility Measure in Assessing Market Efficiency," *The Journal of Finance*, 36, pp. 291-304.

[74] Shleifer, A. (2000) *Inefficient Markets: An Introduction to Behavioral Finance.* Oxford: Oxford University Press.

[75] Smith, C. A.; B. David; and L.D. Kirby (2006) "Emotion-eliciting appraisals of social situations" in *Affect in Social Thinking and Behavior*. Psychology Press: New York.

[76] Staw, B. and S. Barsade (1993) "Affect and Managerial Performance: A Test of the Sadder-But-Wiser vs. Happier-and-Smarter Hypotheses," *Administrative Science Quarterly*, Vol. 38, pp. 304-331.

[77] Tetlock, P. (2007) "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance*, pp. 1139 - 1168.

[78] Shleifer, A. (2000) *Inefficient Markets: An Introduction to Behavioral Finance.* Oxford: Oxford University Press.

[79] Wilson, A. (2014) "Bounded Memory and Biases in Information Processing," *Econometrica,* 82, pp. 2257 - 2294

[80] Yuan, K.; L. Zheng; and Q. Zhu (2006) "Are investors moonstruck? Lunar phases and stock returns," *Journal of Empirical Finance*, 13 (1), pp. 1-23.

[81] Zhang, X.F. (2006) "Information Uncertainty and Stock Returns," *The Journal of Finance*, 61 (1), pp. 105-137.

Appendix A. Proofs

The analysis makes use of multivariate stochastic orders. Of principal interest are distributions that obey the condition of *multivariate total positivity of order 2 (MTP-2).* A probability density function (PDF) $f(\mathbf{x})$, $x \in \mathbb{R}^N$, is *MTP-2* if for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ I have

$$f(\mathbf{x})f(\mathbf{y}) \leq f(\mathbf{x} \wedge \mathbf{y})f(\mathbf{x} \vee \mathbf{y}) \tag{13}$$

where $\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), ..., \max(x_N, y_N))$ and $\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), ..., \min(x_N, y_N))$ are lattice operations. This condition is identical to the notion of log-supermodularity (log-spm) used in economics.is larger than $\mathbf{Y}$ in the *strong likelihood ratio order* or *tp-2 order* (denoted $\mathbf{X} \succcurlyeq_{tp2} \mathbf{Y}$) if the respective PDFs obey for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$,

$$f_{\mathbf{X}}(\mathbf{x})f_{\mathbf{Y}}(\mathbf{y}) \leq f_{\mathbf{Y}}(\mathbf{x} \wedge \mathbf{y})f_{\mathbf{X}}(\mathbf{x} \vee \mathbf{y}) \tag{14}$$

Random variable $\mathbf{X}$ is larger than $\mathbf{Y}$ in the *strong stochastic order* (denoted $\mathbf{X} \succcurlyeq \mathbf{Y}$) if for all increasing functions $u : \mathbb{R}^N \to \mathbb{R}$, $Eu(\mathbf{X}) \geq Eu(\mathbf{Y})$. The strong stochastic order is a multivariate generalization of the one dimensional first order stochastic dominance. Mueller and Stoyan [63] show that $\mathbf{X} \succcurlyeq_{tp2} \mathbf{Y}$ implies $\mathbf{X} \succcurlyeq \mathbf{Y}$.

A.1. **Proofs from Section 3.** I now show that the recall of sequences of signals inherits the log-spm properties assumed for the functions ρ_ω defining the relative recall probabilities of the individual signals.

Lemma 2. *Assumption 3 implies that* $\rho\left(H^R|\varphi, H_N, n\right)$ *is log-spm in* (H^R, φ).

Proof. I proceed by showing the log-spm relation holds between any pair of variables in (H^R, φ) (See Karlin and Rinott [44] for a proof of the sufficiency of this argument), which then implies log-spm amongst all of the variables.

Consider two recollected histories $H \cup \{\omega_i, \omega_j\}$ and $H \cup \{\widehat{\omega}_i, \widehat{\omega}_j\}$ where $\omega_i \geq \widehat{\omega}_i$, $\omega_j \geq \widehat{\omega}_j$ and $H \cup \{\omega_i, \omega_j, \widehat{\omega}_i, \widehat{\omega}_j\} \subseteq H_N$. To show log-spm of the recall probability function holds between pairs of signal values, note

$$\begin{aligned} \frac{\rho\left(H \cup \{\omega_i, \omega_j\} | \varphi, H_N, n\right)}{\rho\left(H \cup \{\omega_i, \widehat{\omega_j}\} | \varphi, H_N, n\right)} &= \frac{\rho_\omega(\omega_j | \varphi, H_N, n)}{\rho_\omega(\widehat{\omega_j} | \varphi, H_N, n)} \\ &= \frac{\rho\left(H \cup \{\widehat{\omega}_i, \omega_j\} | \varphi, H_N, n\right)}{\rho\left(H \cup \{\widehat{\omega}_i, \widehat{\omega_j}\} | \varphi, H_N, n\right)} \end{aligned} \tag{15}$$

Therefore I have log-spm in $\{\omega_i, \omega_j\}$.

Suppose that $\omega \geq \omega'$,$\varphi \geq \varphi'$, and $H_1^R, H_2^R \in \mathbb{H}(H_N, n)$ differ only in that $\omega \in H_1^R$, $\omega' \notin H_1^R$ and $\omega' \in H_2^R$, $\omega \notin H_2^R$. Therefore

$$\frac{\rho\left(H_1^R | \varphi_1, H_N, n\right)}{\rho\left(H_2^R | \varphi_1, H_N, n\right)} = \frac{\rho_\omega(\omega | \varphi_1, H_N, n)}{\rho_\omega(\omega' | \varphi_1, H_N, n)} \tag{16}$$

From Assumption 3 it follows that

$$\begin{aligned} \frac{\rho\left(H_1^R | \varphi_1, H_N, n\right)}{\rho\left(H_2^R | \varphi_1, H_N, n\right)} &= \frac{\rho_\omega(\omega | \varphi_1, H_N, n)}{\rho_\omega(\omega' | \varphi_1, H_N, n)} \\ &\geq \frac{\rho_\omega(\omega | \varphi_2, H_N, n)}{\rho_\omega(\omega' | \varphi_2, H_N, n)} \\ &= \frac{\rho\left(H_1^R | \varphi_2, H_N, n\right)}{\rho\left(H_2^R | \varphi_2, H_N, n\right)} \end{aligned} \tag{17}$$

Therefore the log-spm holds in (ω_i, φ) pairwise. It follows then that $\rho\left(H^R | \varphi, H_N, n\right)$ is log-spm in (H^R, φ). □

The following lemma is a restatement of the fact that if two random variables are ordered in the strong likelihood ratio order, then they are also ordered in the strong multivariate stochastic order. Two technical conditions are required in addition to the log-supermodularity of $\rho\left(H^R | \varphi, H_N, n\right)$. First, $\mathbb{H}(H_N, n)$ must form a lattice, which is obviously true. Second, the distributions $\rho\left(H^R | \varphi, H_N, n\right)$ (for different values of φ) must be absolutely continuous with respect to some σ-finite measure. In this case, this measure can be taken to be the counting measure over the elements of H_N. Finally, I note that although I have established that the $\rho\left(H^R | \varphi, H_N, n\right)$ are log-supermodular in (H^R, φ) for a given value of n, this is a sufficient but not necessary condition for $\mathbf{H}^R(\varphi, n)$ to be ordered in the strong stochastic order with respect to parameter φ.

Lemma 3. *(Theorem 3.11.4,* [63]*) If $\rho\left(H^R | \varphi, H_N, n\right)$ is log-spm in (H^R, φ), then for $\varphi_1 \geq \varphi_2$, $\mathbf{H}^R(\varphi_1, n)$ dominates $\mathbf{H}^R(\varphi_2, n)$ in the strong multivariate stochastic order.*

I now extend Milgrom's [60] representation result to the case of multiple signals, which is important in the proof of Theorem 1. Assume that the agent has recollected a set of signals

$H^R = \{\omega_1, ..., \omega_n\}$. Given a prior belief about the distribution of the parameter θ, denoted $G(\theta)$ with density $g(\theta)$, the agent forms a Bayesian posterior equal to $G(\theta|H^R)$.

The following lemma captures the notion of a multi-dimensional signal as either good or bad news. Since $\mathbb{R}^N$ is only partially ordered, there are many multi-dimensional signal comparisons for which this result remains silent. In addition, the result only holds for comparisons between signals of the same length. However, when both of these conditions are satisfied, I establish that $\overrightarrow{\omega}_1 \geq \overrightarrow{\omega}_2$ implies $G(\theta|\overrightarrow{\omega}_1)$ first order stochastic dominates $G(\theta|\overrightarrow{\omega}_2)$. The interpretation of the first order stochastic dominance relation as optimism, and hence $\overrightarrow{\omega}_1 \geq \overrightarrow{\omega}_2$ implying $\overrightarrow{\omega}_1$ is good news relative to $\overrightarrow{\omega}_2$, follows if the agent prefers high realizations of θ.

Lemma 4. *If the condition density of signals, $f(\omega|\theta)$, is log-spm in (ω, θ), then $H_1^R = \overrightarrow{\omega}_1 = (\omega_1, ..., \omega_n) \geq H_2^R = \overrightarrow{\omega}_2 = (\widehat{\omega_1}, ..., \widehat{\omega_n})$ implies $G(\theta|\overrightarrow{\omega}_1)$ first order stochastic dominates $G(\theta|\overrightarrow{\omega}_2)$.*

Proof. Choose θ^* such that $0 < G(\theta^*) < 1$. Our assumption of the log-spm of $f(\omega|\theta)$ implies that $f(\omega_1, ..., \omega_n|\theta) = \prod_{i=1}^{n} f(\omega_i|\theta)$ has the log-supermodularity property in $(\omega_1, ..., \omega_n; \theta)$ as products of log-spm functions are log-spm. This then implies

$$\frac{\int_{\overline{\theta} \geq \theta^*} f(\omega_1, ..., \omega_n|\overline{\theta}) G(d\overline{\theta})}{f(\omega_1, ..., \omega_n|\theta)} \geq \frac{\int_{\overline{\theta} \geq \theta^*} f(\widehat{\omega_1}, ..., \widehat{\omega_n}|\overline{\theta}) G(d\overline{\theta})}{f(\widehat{\omega_1}, ..., \widehat{\omega_n}|\theta)}$$

$$\begin{aligned}
&\Longrightarrow \quad \frac{f(\omega_1, ..., \omega_n|\theta)}{\int_{\overline{\theta} \geq \theta^*} f(\omega_1, ..., \omega_n|\overline{\theta}) G(d\overline{\theta})} \leq \frac{f(\widehat{\omega_1}, ..., \widehat{\omega_n}|\theta)}{\int_{\overline{\theta} \geq \theta^*} f(\widehat{\omega_1}, ..., \widehat{\omega n}|\overline{\theta}) G(d\overline{\theta})} \\
&\Longrightarrow \quad \frac{\int_{\overline{\theta} \leq \theta^*} f(\omega_1, ..., \omega_n|\theta) G(d\overline{\theta})}{\int_{\overline{\theta} \geq \theta^*} f(\omega_1, ..., \omega_n|\overline{\theta}) G(d\overline{\theta})} \leq \frac{\int_{\overline{\theta} \leq \theta^*} f(\widehat{\omega_1}, ..., \widehat{\omega_n}|\theta) G(d\overline{\theta})}{\int_{\overline{\theta} \geq \theta^*} f(\widehat{\omega_1}, ..., \widehat{\omega_n}|\overline{\theta}) G(d\overline{\theta})} \\
&\Rightarrow \quad \frac{G(\theta^*|\omega_1, ..., \omega_n)}{1 - G(\theta^*|\omega_1, ..., \omega_n)} \leq \frac{G(\theta^*|\widehat{\omega_1}, ..., \widehat{\omega n})}{1 - G(\theta^*|\widehat{\omega_1}, ..., \widehat{\omega_n})} \\
&\Longrightarrow \quad G(\theta^*|\omega_1, ..., \omega_n) \leq G(\theta^*|\widehat{\omega_1}, ..., \widehat{\omega_n})
\end{aligned}$$

as required. □

The condition that $|H_1^R| = |H_2^R|$ is an important requirement of the result. As noted earlier, if the recollected histories are of different lengths, no clear stochastic ordering may be possible. To see this, consider the following simple example of updating a normal distributed prior with normally distributed signals.

Example 3. *Consider prior beliefs $g(\theta) = N(0, 1)$ and signals distributed as $\omega_i \sim N(\theta_0, 1)$. Consider the two signal histories $H_1^R = (-1, 1)$ and $H_2^R = (0.5)$ drawn from true history $H = (-1, 0.5, 1)$. The posterior beliefs are then $g(\theta|H_1^R) = N(0, \frac{1}{3})$ and $g(\theta|H_2^R) = N(\frac{1}{4}, \frac{1}{2})$. Although*

$g(\theta|H_2^R)$ has a mean than $g(\theta|H_1^R)$, the lower variance of $g(\theta|H_2^R)$ implies that neither a first nor a second order stochastic dominance relation can be established between these posteriors.

I now prove Theorem 1.

Theorem 1. *$\varphi_1 \geq \varphi_2$ implies $\mathbf{q}(\varphi_1) \succcurlyeq \mathbf{q}(\varphi_2)$.*

Proof. By Lemma 4, I have that for $\overrightarrow{\omega_1}, \overrightarrow{\omega_2} \in \mathbb{H}(H_N, n)$, $\overrightarrow{\omega_1} \geq \overrightarrow{\omega_2}$ implies $G(\circ|\overrightarrow{\omega_1}) \succcurlyeq G(\circ|\overrightarrow{\omega_2})$. Therefore, $\int q(\theta)G(d\theta|\overrightarrow{\omega_1}) \geq \int q(\theta)G(d\theta|\overrightarrow{\omega_2})$ as $q(\theta)$ is increasing in θ. Since the distribution $\rho\left(\overrightarrow{\omega} = H^R|\varphi, H_N, n\right)$ is log-spm in $(\overrightarrow{\omega}, \varphi)$ and $\int q(\theta)G(d\theta|\overrightarrow{\omega})$ is increasing in each element of $\overrightarrow{\omega}$, I have that the distribution of $\mathbf{q}(\varphi)$ contingent on n is stochastically ordered by φ (follows from Theorem 3.3.11 of [63]). Since the distribution of n is independent of cue state, the unconditional distribution of $\mathbf{q}(\varphi)$ is ordered by φ in the strong stochastic order (so $\mathbf{q}(\varphi_1) \succcurlyeq \mathbf{q}(\varphi_2)$). ☐

In the context of decision problems under uncertainty, the ability to use monotone comparative statics theorems turns on showing that the distributions of the salient random variables obey the requisite log-spm properties. Let $g(\theta|H^R)$ denote the probability density function of $G(\theta|H^R)$. Now I show that $g(\theta|H^R)$ possesses the log-spm property required to apply Athey's monotone comparative statics results.

Lemma 5. *$g(\theta|H^R, n)$ is log-supermodular in (θ, H^R) where n is fixed.*

Proof. Note that

$$g(\theta|H^R, n) = \frac{g(\theta)}{\widetilde{f}(H^R)} f(H^R|\theta) = \frac{g(\theta)}{\widetilde{f}(H^R)} \prod_{\omega \in H^R} f(\omega|\theta) \tag{18}$$

where $\widetilde{f}(H^R)$ and $f(H^R|\theta)$ denote the unconditional and conditional distributions of signals. As log-supermodularity is preserved by multiplication, this implies that $\prod_{\omega \in H^R} f(\omega|\theta)$ is log-supermodular in (θ, H^R). Also, $g(\theta)$ and $\frac{1}{\widetilde{f}(H^R)}$ are separately and trivially log-supermodular in (θ, H^R). Therefore, $\frac{g(\theta)}{\widetilde{f}(H^R)} \prod_{\omega \in H^R} f(\omega|\theta)$ is log-supermodular in (θ, H^R). From the above equalities, I then have that $g(\theta|H^R)$ is log-supermodular in (θ, H^R). ☐

As I have assumed the two-dimensional single crossing condition holds and have proven that $g(\theta|H^R, n)$ is log-supermodular in (θ, H^R) for fixed n, it is straightforward to show using Athey's monotone comparative statics results that for histories of fixed length, higher values recollected from memory lead to increasing choices.

Lemma 6. *If $|H^R| = |\widehat{H^R}|$ and $H^R \geq \widehat{H^R}$, then I have $x^*(H^R) \geq x^*(\widehat{H^R})$*

Proof. This follows directly from Theorem 2 of Athey [1] since I have, by assumption, that $u(x;\theta)$ obeys the single crossing property and I have shown in Lemma 5 that $g(\theta|H^R)$ is log-supermodular in

(θ, H^R). Therefore $x^*(H^R) \in \arg\max_{x \in X} \int u(x;\theta) G(d\theta | H^R)$ is increasing in H^R element by element.

□

I now provide a proof of Theorem 2.

Theorem 2. *Assume that $u(x;\theta)$ satisfies the two-dimensional single crossing condition. Then $\varphi_1 \geq \varphi_2$ implies $x^*(\mathbf{H}^R(\varphi_1)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2))$.*

Proof. From Lemma 6 I have that $x^*(H^R)$ is an increasing function of H^R. Denote the random variable $\mathbf{H}^R(\varphi)$ contingent on $|H^R| = n$ as $\mathbf{H}^R(\varphi; n)$. Since $\rho\left(\overrightarrow{\omega} = H^R | \varphi, H_N, n\right)$ is log-spm in $(\overrightarrow{\omega}, \varphi)$, I have that $\mathbf{H}^R(\varphi_1; n) \succcurlyeq \mathbf{H}^R(\varphi_2; n)$. Therefore, from Theorem 3.3.11 of [63] and Lemma 6, I have that $x^*(\mathbf{H}^R(\varphi_1; n)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2; n))$. Since the length of recollected history is independent of N, I have that the same result holds for the unconditional distribution of $x^*(\mathbf{H}^R(\varphi_1; n))$, so $x^*(\mathbf{H}^R(\varphi_1)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2))$. □

Finally, I prove the result on the comparative statics of employee affect and effort choice.

Proposition 1. *Let Assumption 5 hold. If $\varphi_1 \geq \varphi_2$, then $\mathbf{e}(\varphi_1) \succeq \mathbf{e}(\varphi_2)$.*

Proof. The agent's total expected utility is

$$\underset{e \in \mathbb{R}_+}{Max} \int_{\mathbb{R}} W(e,\theta) G(d\theta | \mathbf{H}^R(\varphi)) - c(e)$$

Since θ does not enter the agent's cost of effort function, Assumption 5 implies that $W(e,\theta) - c(e)$ is supermodular. My desired result then follows immediately from Theorem 2. □

A.2. **Proofs from Section 4.**

Proposition 2. *Suppose that as $N \to \infty$ both $l(N) \to \infty$ and $l(N)/N \to 0$. Then as $N \to \infty$:*

$$E_t[\theta | \overrightarrow{\varphi}_t] \to m(1-\phi) \sum_{\tau=0}^{\infty} \phi^\tau \varphi_{t-\tau} \tag{19}$$

Proof. The representative agent uses a set of recalled data $\{\omega_1, ..., \omega_R\}$ to form beliefs regarding the true value of θ, where R denotes the length of the data set recalled. In the limit as $l(N) \to \infty$ (and hence $R \to \infty$) I can neglect the prior beliefs term since $g(\theta)$ is bounded. Since the representative agent is attempting to estimate the mean of a normal variable, I can write the agent's estimate of θ as:

$$\widehat{\theta}_{l(N)}(\overrightarrow{\varphi}_t) = \frac{1}{R} \sum_{i=1}^{R} \omega_i \tag{20}$$

The set $\{\omega_1, ..., \omega_R\}$ consists of two separate components: data recalled through association and data recalled through rehearsal.

I first deal with the $l(N)$ data recalled by association in the current period by proving a weak law of large numbers for this data set. Since my model is based on sampling without replacement from the data in memory, I need to account for the covariance between the data recalled. To prove the weak law of large numbers, I use Chebyshev's inequality as follows:

$$\Pr\left\{\frac{1}{l(N)}\sum_{i=1}^{l(N)}\omega_i - E\left[\omega_i|\theta,\varphi_t\right] \geq k\sigma\right\} \leq \frac{1}{k^2}$$

where:

$$\begin{aligned}\sigma^2 &= Var\left(\frac{1}{l(N)}\sum_{i=1}^{l(N)}\omega_i\right)\\ &= \frac{Var(\omega_i)}{l(N)} + \frac{1}{l(N)^2}\sum_{i=1}^{l(N)}\sum_{j\neq i}Cov(\omega_i,\omega_j|\varphi_t)\\ &= \frac{Var(\varepsilon_i)}{l(N)} + \frac{l(N)(l(N)-1)}{l(N)^2}Cov(\omega_i,\omega_j|\varphi_t)\end{aligned}$$

The first term goes to 0 since $l(N) \to \infty$. The covariance term arises solely because of the sampling without replacement procedure for associatively recalling data from memory. Despite the relative sampling probabilities being different, the covariance vanishes at a rate of $O\left(\frac{l(N)}{N}\right)$. Therefore $\sigma^2 \to 0$ and $\frac{1}{l(N)}\sum_{i=1}^{l(N)}\omega_i \to E\left[\omega_i|\theta,\varphi_t\right]$ in probability as $N \to \infty$ where the expectation is taken with respect to the distribution of recollected signals.

The remainder of the data recalled is generated by rehearsal. If I simply ignore the issue of what happens when the same datum is recalled through associative recall in the current period and rehearsal from a prior period, then the claim of Proposition 2 would be true. My goal is to prove that this "double counting" vanishes in the limit as $N \to \infty$.

Consider a uniform gridding of $\mathbb{R}$ with grid points $\{..., -2d, -d, 0, d, 2d, ...\}$ where $d > 0$. From the Law of Large Numbers, for large N there are approximately $N(\Phi(kd) - \Phi((k-1)d))$ signals with realizations in the k^{th} bin, and I can make this approximation arbitrarily accurate with arbitrarily high probability as $N \to \infty$. Since ρ_ω is continuous, the probability of recall of different $\omega \in ((k-1)d, kd]$ is approximately the same for small $d > 0$. Since the number of signals recalled from each $((k-1)d, d]$ vanishes at a rate of $O(l(N)/N)$ as $N \to \infty$, the probability of any given datum being recalled from the bin vanishes at the same rate. Therefore, the probability that any particular datum that is recollected from $\omega \in ((k-1)d, kd]$ through associative recall in the current period has been recalled associatively within the last T periods vanishes as $N \to \infty$. Since the choice of k and d were arbitrary, this argument implies that for any $T < \infty$ and any fraction of recalled data $z > 0$ and probability $q < 1$, for N sufficiently large less than a fraction z of the data recalled associatively in the current period was recalled associatively in the past T periods with probability at least q. Since the probability of recall through rehearsal drops geometrically with T, in the limit as $N \to \infty$ I can ignore the effect of data recalled both associatively and through rehearsal.

To find the closed form for $\delta(\varphi)$, first note that the mean of the data recalled in each period through association (as opposed to rehearsal) is sampled according to the recall probability function $e^{m\omega_i\varphi_t}$. Therefore the data recalled through association in affective state φ_t will be distributed with a probability density function proportional to

$$\begin{aligned}\exp\left(-\frac{{\omega_i}^2}{2}\right)\exp\left(m\omega_i\varphi_t\right) &= \exp\left(-\frac{{\omega_i}^2 - 2m\omega_i\varphi_t + m^2\varphi_t^2}{2}\right)\exp\left(-\frac{m^2\varphi_t^2}{2}\right) \\ &= \exp\left(-\frac{(\omega_i - m\varphi_t)^2}{2}\right)\exp\left(-\frac{m^2\varphi_t^2}{2}\right)\end{aligned}$$

Once this is normalized to a total probability of 1 I find the probability density function of a normal random variable with mean $m\varphi_t$. This logic for the associative recall applies in each period of the economy. When I average over the data recalled by association in period t and the data recalled through rehearsal (which is in effect a combination of data recalled by association in prior periods) I find the equation defined in the proposition. □

Lemma 1. *φ_t is ergodic.*

Proof. Since $g(\varphi_t, 0) = 0$ the process is strongly aperiodic. Since φ_{t+1} has full support over $(-1, 1)$, the process is Harris recurrent. Combined with strong aperiodicity this implies that the affect process has an invariant measure and is irreducible (Theorem 10.4.2 of Meyn and Tweedie [59]). Together the properties of irreducibility, existence of an invariant measure, and Harris recurrence imply the process is ergodic and converges to the invariant measure π in the total variation norm (Theorem 13.3.1 of Meyn and Tweedie [59]) □

A.2.1. *Proofs of Price Time-Series Effects.*

Proposition 3. *For $\phi \in (0, 1)$ sufficiently large asset prices exhibit overreaction in the short-run*

$$E_t^*[P_{t+1} - P_t|\omega_t \geq 0] > E_t^*[P_{t+1} - P_t|\omega_t \leq 0] \tag{21}$$

Proof. I can write the price change formula as:

$$\begin{aligned}P_{t+1} - P_t &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2}m(1-\phi)\left[\sum_{\tau=0}^{\infty}\phi^\tau\varphi_{t+1-\tau} - \sum_{\tau=0}^{\infty}\phi^\tau\varphi_{t-\tau}\right] \\ &= \frac{\omega_{t+1}}{r} + \frac{1+r}{r^2}m(1-\phi)\sum_{\tau=0}^{\infty}\phi^\tau\sigma_{t+1-\tau}\end{aligned}$$

Note that $E_t^*[\omega_{t+1}] = 0$. Since φ_t is ergodic, I have $E_t^*[\varphi_{t-i}] = E_t^*[\varphi_{t-j}]$ for all $i, j > 0$. Using these two facts I can write:

$$\begin{aligned} E_t^*[P_{t+1} - P_t|\omega_t &\geq 0, \overrightarrow{\varphi}_t] - E_t^*[P_{t+1} - P_t|\omega_t \leq 0, \overrightarrow{\varphi}_t] \\ &= \frac{1+r}{r^2} m(1-\phi)(E_t^*[\sigma_{t+1} + \phi\sigma_t|\omega_t \geq 0, \varphi_t] - E_t^*[\sigma_{t+1} + \phi\sigma_t|\omega_t \leq 0, \varphi_t]) \\ &= \frac{1+r}{r^2} m(1-\phi)\left(E_t^*[\sigma_{t+1}|\omega_t \geq 0, \varphi_t] - E_t^*[\sigma_{t+1}|\omega_t \leq 0, \varphi_t]\right) + \\ &\qquad \frac{1+r}{r^2} m(1-\phi)\phi\left(E_t^*[\sigma_t|\omega_t \geq 0, \varphi_t] - E_t^*[\sigma_t|\omega_t \leq 0, \varphi_t]\right) \end{aligned}$$

To sign this equation, I must analyze the expectations. The events on which the two expectations are conditioned can be broken down into individual events where $\omega \in \mathbb{R}_+$ and $\widetilde{\omega}_t = -\omega_t$. Under this notation, let σ_t denote the innovation to affect given ω_t and $\widetilde{\sigma}_t$ denote the innovation to affect given $\widetilde{\omega}_t$ Consider a particular realization $\omega_t, \omega_{t+1} > 0$, which yields:

$$\begin{aligned} \sigma_t - \widetilde{\sigma}_t &= (1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t}) \\ \sigma_{t+1} - \widetilde{\sigma}_{t+1} &= -(1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t})(1 - e^{-\omega_{t+1}}) \end{aligned}$$

Combining these I have:

$$(\sigma_{t+1} - \widetilde{\sigma}_{t+1}) + \phi\left(\sigma_t - \widetilde{\sigma}_t\right) = (1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t})\left(\phi - (1 - e^{-\omega_{t+1}})\right)$$

which is positive in expectation for $\phi < 1$ sufficiently large. If I instead consider realizations $\omega_t > 0, \omega_{t+1} < 0$ I find

$$\begin{aligned} \sigma_t - \widetilde{\sigma}_t &= (1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t}) > 0 \\ \sigma_{t+1} - \widetilde{\sigma}_{t+1} &= -(1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t})(1 - e^{\alpha\omega_{t+1}}) \end{aligned}$$

which in turn yields:

$$(\sigma_{t+1} - \widetilde{\sigma}_{t+1}) + \phi\left(\sigma_t - \widetilde{\sigma}_t\right) = (1-\varphi_{t-1})(e^{-\alpha\omega_t} - e^{-\omega_t})\left(\phi - (1 - e^{\alpha\omega_{t+1}})\right)$$

which is again positive in expectation for $\phi < 1$ sufficiently large. A symmetric argument holds for the case where $\omega_t < 0$, and together these facts yield:

$$E_t^*[P_{t+1} - P_t|\omega_t \geq 0] > E_t^*[P_{t+1} - P_t|\omega_t < 0]$$

□

Proposition 4. *There is long-run correction of the price effect of changes in affect in the current period:*

$$\underset{\tau\to\infty}{Lim}\ E_t^*[P_{t+\tau} - P_t|\omega_t \geq 0] \leq \underset{\tau\to\infty}{Lim}\ E_t^*[P_{t+\tau} - P_t|\omega_t < 0]$$

Proof. Note that I can write the price in each period as:

$$P_t = \frac{d_t}{r} + \frac{1+r}{r^2} m(1-\phi) \sum_{k=0}^{\infty} \phi^{\tau} \varphi_{t-k} \tag{22}$$

$$P_{t+\tau} = \frac{d_{t+\tau}}{r} + \frac{1+r}{r^2} m(1-\phi) \sum_{k=0}^{\infty} \phi^{\tau} \varphi_{t+\tau-k} \tag{23}$$

Since $E_t^* [d_t] = E_t^* [d_{t+\tau}]$ and thus cancels out when computing $E_t^* [P_{t+\tau} - P_t]$, I focus on the second term in each of Equations 22 and 23. Let $\overline{\varphi}$ denote the mean of the ergodic distribution of φ_t. Since the expectation is not conditional on price changes, in all periods prior to t the expected value of φ_t is equal to the mean of the ergodic distribution, which implies:

$$(1-\phi)E_t^* \left[\sum_{k=0}^{\infty} \phi^{\tau} \varphi_{t-k} | \omega_t \geq 0 \right] = (1-\phi)E_t^* [\varphi_t | \omega_t \geq 0] + \phi \overline{\varphi}$$

Since φ_t is ergodic, I have as $\tau \to \infty$:

$$(1-\phi)E_t^* \left[\sum_{k=0}^{\infty} \phi^{\tau} \varphi_{t+\tau-k} | \omega_t \geq 0 \right] \to \overline{\varphi}$$

Since φ_{t-1} is drawn from the ergodic distribution, I have:

$$E_t^* [\varphi_t | \omega_t \geq 0] = \overline{\varphi} + E_t^* [\sigma_t | \omega_t \geq 0] > \overline{\varphi}$$

Noting that $E_t^* [\sigma_t | \omega_t \geq 0] > 0$, I find:

$$\underset{\tau \to \infty}{Lim} \; E_t^*[P_{t+\tau} - P_t | \omega_t \geq 0] < 0 \tag{24}$$

Symmetrically, since $E_t^* [\sigma_t | \omega_t \leq 0] < 0$ I have:

$$E_t^* [\varphi_t | \omega_t \leq 0] < \overline{\varphi}$$

so I have:

$$\underset{\tau \to \infty}{Lim} \; E_t^*[P_{t+\tau} - P_t | \omega_t \leq 0] > 0 \tag{25}$$

Combining equations 24 and 25 yields the result. □

Appendix B. For Online Publication: Cross-sectional Market Effects

In this section, I consider agents that have different levels of knowledge about an asset. For example, I could compare securities analysts at investment banks (high-knowledge agents) with casual investors (low-knowledge agents). A common intuition is that, ceteris paribus, knowledgeable investors are less biased than casual investors due to either experience or selection. The theory suggests that the opposite may be true since the level of knowledge interacts with the mnemonic biases - knowledgeable investors are more likely to have stored extreme dividend realizations in memory, and these easily recollected extreme events amplify the bias suffered by knowledgeable agents. While both countervailing effects are likely at play in reality, this provides a sharp (if one-sided) prediction that delineates my theory from other psychological models.

I provide an example to illustrate this effect. There are two kinds of agents, inexperienced and experienced investors. Inexperienced investors have two realization of ω_t stored in memory, and experienced investors have $N \gg 2$ realizations in long-term memory. Agents recollect one piece of data prior to their decision (i.e., $l(1, N) = 1$) and use the data to update a diffuse prior regarding the mean of θ.[25] Agents can be either in a neutral mood with unbiased recall ($\varphi = 0$) or in a positive mood with optimistic recall ($\varphi = 1$). Note that since the agents recall a single datum, their expectation is equal to the value recalled.

Under a neutral affect, the distribution of the estimates within both populations are standard normal distributions. Under optimistic recall, agents are biased towards recalling (a single) higher signal stored in memory. For experienced investors the log-supermodularity of ρ_ω insures that a high realization of ω_t is more likely to be recalled (relative to the datum recalled by a less knowledgeable agent). Figure 7 depicts the distribution of beliefs of the agents with little knowledge (solid line) and those agents with a large store of knowledge in long-term memory (dashed line). Note that the population of less knowledgeable agents is biased by approximately $\frac{1}{3}$ of a standard deviation on average, while the knowledgeable agents are biased by more than 1 standard deviation on average.

The predictions contrast with the usual conjecture that more experienced agents ought to be *less* subject to biases in belief.[26] One justification for the usual conjecture is that experienced agents are typically sophisticated, and the sophistication may render these agents less susceptible to biases. Another justification for this assumption is that agents become aware of the bias and take measures to correct it. To the best of the knowledge, there is no evidence that suggests agents are aware of the bias in their recall, although most people appreciate that recall is only partial.

---

[25]These results are robust to having experienced agents recall more data from memory than less experienced agents so long as the experienced agents recollect a smaller fraction of the total information in memory.

[26]Agents subject to different degrees of bias could be easily captured by attributing high values of m to very biased agents and values of m close to 0 to nearly unbiased agents.

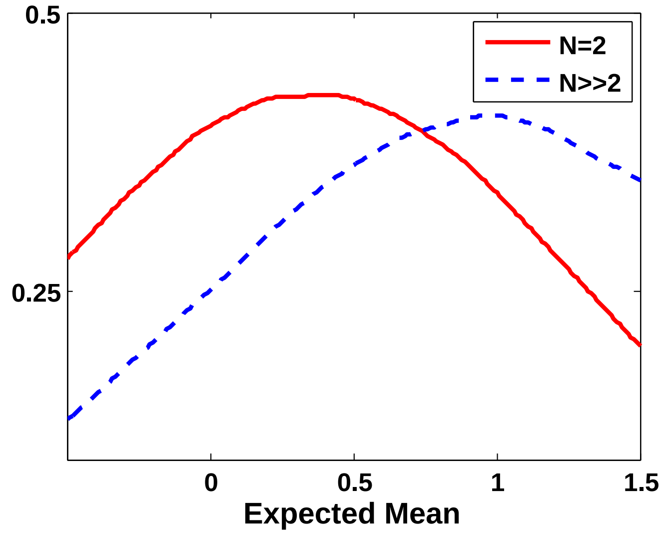

Figure 7. Biases in Recollection and Knowledge

Yet a third justification for the usual conjecture that experience ameliorates biases in judgement is that selection should drive traders with biased expectations from the market, although these results are sensitive to assumptions on preference parameters such as time discount rates and issues such as market completeness (Blume and Easley [11], Easley and Yang [23]).

The goal in this section is not to claim that the usual conjecture must be wrong, but to point out that, unlike in many behavioral models, the theory suggests this intuition *may* be wrong. The predictions of this section provide a strong one-sided test for the theory.[27] Suggestive correlations are provided by examining the beliefs of agents with different levels of knowledge. La Porta [52] shows that forecasts by stock market analysts reflect the pattern of long-run correction, which suggests that the beliefs of experienced market participants may be subject to biases, but I do not claim that the result of La Porta [52] is more than suggestive. To test our prediction would require a differences-in-differences study of informed and uninformed investors influenced (or not influenced) by affective cues, which is substantially different than La Porta [52]. Controlling for selection would require variation in the experience of the agents or some other metric for the agents' exposure to selection pressures.

[27]In addition to testing against the usual conjecture that experience eliminates biases in judgement, the test also delineates the theory from a model that treats agent mood as an informative signal (Schwarz and Clore [69] and [70]). Under this theory market participants with large amounts of information at their disposal would be less influenced (relative to their less informed peers) by the effect of a single erroneous datum.

Appendix C. For Online Publication: Returns as Determinant of Affect

In this appendix I provide a brief discussion of an alternative model that uses returns to define the innovations to affect. I assume the following form for the innovations to affect:

$$\begin{aligned} \sigma_{t+1} &= P_{t+1} + d_{t+1} - (1+r)P_t \\ \varphi_{t+1} &= \varphi_t + \sigma_{t+1} \end{aligned} \tag{26}$$

This formulation does not incorporate immune neglect and does not presume an upper or lower bound to affect. I am required to abandon these aspects of the model to find a closed form solution (the necessity of which is clear below). The pricing equations still apply - all that has changed is the process governing the evolution of the affect of the representative agent. Since m is a free parameter, I normalize it so that I can write the pricing equation as:

$$P_t(\varphi) = \frac{d_t}{r} + m \sum_{\tau=0}^{\infty} \phi^{\tau} \varphi_{t-\tau} \tag{27}$$

Note that a positive return in period t is driven in part by the change in the agents affect in period t, which is in turn driven by the return in period t. The change in affect is, in this sense, endogenous. One might worry that the endogeneity of innovations to affect might lead to indeterminacy of predictions. For example, intuition suggests that, especially for mild dividend innovations, changes to affective state may become a self-fulfilling prophecy in that downward changes in affect in period t reduce the price in period t, which in turn reduces the period t return, which (closing the loop) implies a downward change in affect. If this intuition were correct, then asset prices would, to the extent that they are determined by affect, be arbitrary.

I show that my model makes a single price prediction. I use the linearity of Equation 26 to express the evolution of affect (and hence market prices) through a closed form solution for the model. Without assuming linearity (or an equally convenient functional form) I would not have been able to "close the endogenous loop" caused by returns in the current period influencing affect in the current period, which in turn determine the returns in the current period.

First, I combine the affect innovation and price equations to find:

$$\begin{aligned} \sigma_{t+1} &= P_{t+1} + d_{t+1} - (1+r)P_t \\ &= \frac{d_{t+1}}{r} + d_{t+1} - \frac{1+r}{r} d_t + m \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\varphi_{t-\tau+1} - (1+r)\varphi_{t-\tau} \right] \\ &= \frac{1+r}{r} \left(d_t + \omega_{t+1} - d_t \right) + m \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\sigma_{t-\tau+1} - r\varphi_{t-\tau} \right] \\ &= \frac{1+r}{r} \omega_{t+1} + m \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\sigma_{t-\tau+1} - r\varphi_{t-\tau} \right] \end{aligned}$$

I bring together the σ_{t+1} terms to find the autocorrelation process of innovations to affective state.

$$\sigma_{t+1} = \frac{(1+r)}{r(1-m)}\omega_{t+1} - \frac{mr}{1-m}\varphi_t + \frac{m}{1-m}\sum_{\tau=1}^{\infty}\phi^{\tau}\left[\sigma_{t-\tau+1} - r\varphi_{t-\tau}\right] \tag{28}$$

Inserting equation 28 into the formula for affective valence, I find that the representative agent's affect is:

$$\begin{aligned} \varphi_{t+1} &= \varphi_t + \sigma_{t+1} \\ &= \left(\frac{1-m-mr}{1-m}\right)\varphi_t + \frac{(1+r)}{r(1-m)}\omega_{t+1} + \frac{m}{1-m}\sum_{\tau=1}^{\infty}\phi^{\tau}\left[\sigma_{t-\tau+1} - r\varphi_{t-\tau}\right] \end{aligned} \tag{29}$$

The next step is to state the equation solely in terms of affective state. First I have:

$$\sigma_{t-\tau+1} - r\varphi_{t-\tau} = \varphi_{t-\tau+1} - (1+r)\varphi_{t-\tau}$$

which implies I can write equation 29 as:

$$\varphi_{t+1} = \left(\frac{1-m-mr}{1-m} + \phi\right)\varphi_t + \frac{(1+r)}{r(1-m)}\omega_{t+1} - \frac{m}{1-m}\sum_{\tau=1}^{\infty}\phi^{\tau}\left((1+r) - \phi\right)\varphi_{t-\tau} \tag{30}$$

Equation 30 has a number of unappealing properties that lead me to reject it on both psychological and economic grounds. Psychology research on emotion (as well as common intuition) suggests that affective states are transient. However, the dynamics described by Equation 30 are not necessarily ergodic,[28] which implies that the model would not reflect the transient nature of emotion. In addition, I find it counter-intuitive that affective state today should be negatively correlated with past affective states. Equation 30 implies that affect in period $t+1$ is negatively correlated with the affective state in period $t-1$ (and all earlier periods). Intuition suggests that, if anything, the valence of one's affect should be positively correlated with past affective states.

Finally, I note that numerical simulations of the alternative model using plausible values for m and ϕ show that while the model evinces ergodic behavior (and so exhibits long-run correction), there is no short-run continued overreaction to news. In other words, the model of affective state driven by returns predicts that positive (negative) price changes in the present period predict negative (positive) price changes in all future periods.

Appendix D. For Online Publication: Analysis of Returns Instead of Price Changes

The bulk of the time-series analysis in section 4.4.1 studies price changes rather than returns. The reason for this is simple: it is exceedingly difficult to obtain closed form results in terms of

[28] If m is small and ρ is large, then φ_t resembles a process with a root exceeding 1. This would imply that affective state ought to grow unboundedly in either the positive or negative direction.

returns. To see why, it suffices to examine the equation for asset returns

$$\begin{aligned} P_{t+1} + d_{t+1} - (1+r)P_t &= \frac{1+r}{r} d_{t+1} + \delta(\varphi_{t+1}) - (1+r)\left(\frac{d_t}{r} + \delta(\varphi_t)\right) \\ &= \frac{1+r}{r}\omega_{t+1} + \big(\delta(\varphi_{t+1}) - \delta(\varphi_t)\big) - r\delta(\varphi_t) \end{aligned}$$

The first term has a mean of 0, while the second term reflects the time path of price biases that drove the results in the main body of the paper. The final term, $-r\delta(\varphi_t)$, reflects the fact that the bias will not be reflected in the actual return (on average).

If I (for example) wanted to study short-run overreaction of returns, we would need to compute

$$E_t^* \left[P_{t+1} + d_{t+1} - (1+r)P_t | \omega_t > 0\right] = E_t^* \left[\delta(\varphi_{t+1}) - \delta(\varphi_t) | \omega_t > 0\right] - rE_t^* \left[\delta(\varphi_t) | \omega_t > 0\right]$$

The proof for the short-run overreaction of price changes was based on the fact that

$$E_t^* \left[\delta(\varphi_{t+1}) - \delta(\varphi_t) | \omega_t > 0\right] > 0 > E_t^* \left[\delta(\varphi_{t+1}) - \delta(\varphi_t) | \omega_t < 0\right]$$

To complete an analysis of the short-run overreaction of returns, I would need to compute

$$\begin{aligned} -rE_t^* \left[\delta(\varphi_t) | \omega_t > 0\right] + rE_t^* \left[\delta(\varphi_t)) | \omega_t < 0\right] &= m(1-\phi) \left(E_t^* \left[\varphi_t | \omega_t > 0\right] - E_t^* \left[\varphi_t | \omega_t < 0\right]\right) && (31) \\ &= m(1-\phi)(E_t^* \left[(1 - e^{-\omega_t})(1 - \varphi_{t-1}) | \omega_t > 0\right] \\ &\quad - E_t^* \left[(1 + \varphi_{t-1})(e^{\alpha\omega_t} - 1) | \omega_t < 0\right]) && (32) \end{aligned}$$

Since φ_{t-1} is unknown to the econometrician, this expectation involves an integration over the ergodic distribution of φ_{t-1} as well as the distribution of ω_t. To understand this difficulty, note that if the ergodic distribution was focused around $\varphi_{t-1} = 1$, then equation 31 would clearly be negative and short-run overreaction of returns would hold. On the other hand, if the ergodic distribution is more diffuse and α is not too large, then assessing whether short-run overreaction holds would depend on r and the magnitude of the short-run overreaction of prices to news.

Although I cannot provide analytical results on the impact of mnemonic biases on returns, I have conducted numerical simulations which show that for all of the parameters considered the results of Section 4.4.1 carry over to returns.[29] In other words, the numerical simulations imply that returns also exhibit short-run overreaction, long-run correction, and momentum.

[29] I suggested above that if α is small and the ergodic distribution is diffuse, then the results may not carry-over to returns. However, our simulations seem to imply that low values of α cause the ergodic distribution to place a large amount of probability mass near $\varphi_t = 1$, which implies that returns do exhibit the described time-series effects.