Mood, Associative Memory, and the Evaluation of Asset Prices

Aaron Bodoh-Creed, Cornell University *

February 13, 2012

Abstract

We provide a model of the effect of associative memory and mnemonic cues on agent beliefs to explain a variety of asset pricing puzzles. Our model is specialized to study affective state as a cue for information congruent with the affect of the agent, a phenomenon called **mood congruent memory**. We employ our model to provide novel explanations for phenomena such as short-run underreaction and long-run overreaction to news, excess volatility, and the influence of non-fundamental events. Unique to our model, we predict that excess volatility will be highest during market downturns and that knowledgeable agents will be more biased than agents who are relatively ignorant. We also provide a number of new cross-sectional and panel-data predictions regarding asset prices and portfolio choices.

*I would like to thank Ilya Segal, Paul Milgrom, Doug Bernheim, Jon Levin, C. Lanier Benkard, Larry Blume, Eric Budish, David Easley, Matthew Jackson, Ramesh Johari, Philip Reny, Robert Wilson, Marcello Miccoli, Juuso Toikka, Alex Hirsch, Carlos Lever, Edward Augenblick, Isabelle Sin, John Lazarev, and Bumin Yenmez for invaluable discussions regarding this topic. In additional, valuable comments were received during seminars at Stanford, the Stanford Graduate School of Business, Yale School of Management, Kellogg MEDS, Boston University, the University of Rochester, Cornell University, the Ohio State University, and the INFORMS Conference. I gratefully acknowledge the support of the Jacob K. Javits Fellowship program and the Bradley Research Fellowship"

1 Introduction

The efficient market hypothesis implies that all information is immediately incorporated into asset prices. However, studies in behavioral finance have found a variety of regularities such as asset price short-run underreaction and long-run overreaction to news and excess price volatility that suggest that available information has significant predictive power for future returns even after controlling for current price.[1] Furthermore, prices have also been shown to significantly respond to nonfundamental events such as weather and daylight length (Saunders [55], Hirshleifer and Shumway [26], Kamstra, Kramer and Levi [34]).

We provide a model of long-term memory and mood that provides a unified explanation for all of the phenomena mentioned above. In addition, our model provides a number of novel predictions that remain to be tested. First, our model suggests that excess volatility will be greatest during market downturns when mood is depressed. Second, our model suggests that investors with a great deal of knowledge (e.g. hedge fund managers, equity analysts) may be more biased by these effects than less knowledgeable investors. Third, our model suggests that market prices for securities will be more correlated than predicted by the underlying fundamentals. In addition, we provide novel predictions regarding cross-sectional portfolio choice and the power of direct assesments of mood to predict asset prices.

Long-term memory stores an enormous amount of information, most of which is irrelevant for any given decision. Cues present in our environment or state of mind prompt the recollection of information salient for the decision at hand. While these cues provide an efficient process for focusing on important data, the cues can cause our beliefs and decisions to be based on systematically biased, incomplete sets of information. Mood

[1] These findings are surveyed in Barberis, Shleifer and Vishny [2] and Daniel, Hirshleifer and Subrahmanyam [14].

regulated memory is a psychological phenomenon wherein agent affective states serve as cues for information stored in long-term memory (Isen [32]). We focus on the valence of the affective state[2] as a cue for information of the same valence.[3,4]

Asset prices are based on a representative consumer's beliefs regarding the net present value of the future dividends of an asset. Agent affective state is influenced by news in the present period with positive (negative) earnings shocks encouraging positive (negative) affective states. Mood-regulated memory causes positive (negative) affective states of a decision maker to cue the recollection of positive (negative) information from memory, and the beliefs founded on the recalled data will be optimistically (pessimistically) biased.[5] Our predictions and explanations are rooted in our model of the time-series relationship between market returns, mnemonic biases, and price changes.

We first use our model to study asset price under- and overreaction. Asset price underreaction is a phenomenon whereby assets prices exhibit positive autocorrelation patterns over horizons of less than one year (Cutler, Poterba and Summers. [13], Jegadeesh and Titman [33], Rouwenhorst [53]). Asset price overreaction studies suggest that this pattern is reversed over longer horizons and that asset prices are negatively correlated over time scales longer than a year (De Bondt and Thaler [15], Lakonishok, Shleifer and Vishny [39]). We capture both of these effects by assuming that news events are autocorrelated, which

[2]The valence of an affective state refers to the subjectively experienced goodness of the state. For example joy is a positively valanced affect, while anger and sadness are negatively valanced affective states.

[3]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. [62]). This suggests that positive affective states of an agent are correlated with the satiation or expected satiation of an agent's desires and goals, which makes affective valence a natural point of connection between the psychology of emotion and economic theory.

[4]Our model of memory can be easily extended to accommodate generic cues.

[5]We do no insist that all data is recalled from memory. For example when trying to form a belief about the net present value of a security, hard information is combined with soft data recalled from memory about industry performance, the efficacy of the firm's leadership, and the present and future product portfolio of the firm.

in turn implies that changes in affect (our mnemonic cue) are correlated. Good news that encourages a positive affective state this period implies that the news next period will also be good and further shift the distribution of affect in the market towards more positively valenced states. While the informational value of the cues is correctly incorporated into beliefs about net present value, the dynamics of agent affect cause the asset price to exhibit under- and overreaction dynamics.

Since changing cues in the environment induce price effects, we find that the dynamics of the mnemonic cues can cause excess volatility (Shiller [59] and [60] provide seminal discussions of this topic). When fundamental events act as cues for similar data from memory or the events are correlated with mnemonic cues (as in the case of affective states), then the memory biases serve as a multiplier of the volatility effect of these events. Our model provides microfoundations for the psychological processes underlying this volatility and predicts that the volatility will be higher during market downturns.

Another novel prediction of out paper is that that agents with more knowledge (e.g. equity analysts) may be more subject to mnemonic biases than less informed agents. Since knowledgeable agents have more prior knowledge stored in memory, biases in the recall of that information can have more effect on these agents than on market participants with little prior knowledge. If the set of arbitrageurs that help enforce market efficiency are also highly knowledgeable, then this effect may blunt the power of arbitrage to correct mispricings founded on our mnemonic biases.

We also argue that mnemonic cues can induce correlations amongst asset prices in excess of the correlation amongst the funamentals. Since affective cues have a general effect on cognitive processing, if an agent observes an affective cue regarding one element of his portfolio, then his affective state could bias his beliefs about unrelated segments of his portfolio. In other words, if good news about one firm or industry induces a positive affect that causes optimistically biased beliefs, these biased beliefs could spill over and affect beliefs about unrelated assets. Moreover, these non-informational effects are transient and fade with time.

Relatedly when non-fundamental events serve as cues for memory, prices may respond to these events. Saunders [55] and Hirshleifer and Shumway [26] provide evidence that market prices respond to meteorological events that have (presumably) little information regarding future market prices. Our model describes a clear link between the psychological foundations (mood regulated memory), beliefs about net present value, and market price dynamics and provides additional untested predictions. We can directly test our model by studying the linkage between the Gallup Daily: U.S. Mood Poll and asset price time-series data.

Finally, we provide an explanation for why repeated exposure to the same piece of information can have continued market price effects. The effects of repeated exposure in our model are mediated through two channels. First, a market participant may forget the data between exposures.[6] Second, exposure to a datum can cue the recall of related data, which can in turn bias beliefs. Tetlock [63] provides evidence for these effects through repeated exposure to events covered by the Wall Street Journal. Unlike in models of gradual information diffusion (e.g. Hong and Stein [28]), we predict that the effect of exposure to information will (at least partially) reverse as cues for the information fade.

One technical contribution of our model is to ground the biases in beliefs in a formally Bayesian decision problem, which allows for the analysis of welfare effects of potential interventions. Since our model is based on the imperfect recall of information, one intervention of first order importance is the timely provision of information that may be neglected by the market consensus. In addition, by providing a formal model we can integrate these

[6] In our model the agents need to freshly recall all data for each decision, and so data is always forgotten in this sense.

memory biases into a wide array of other finance models (e.g. IPO pricing).

1.1 Prior Finance and Economics Literature

Several prior works have provided behavioral models to explain some of the asset pricing puzzles that we study. Barberis, Shleifer and Vishny [2] explains over- and underreaction of asset prices to news with a model of investor sentiment. Daniel, Hirshleifer and Subrahmanyam [14] develops a model of investor sentiment driven by overconfidence and positively biased self-attribution in order to explain price over- and underreaction. Hong and Stein [28] provides a model wherein traders are boundedly rational and limited to using simple forecasting models that do not condition on prices. Hong, Stein and Yu [30] provide a model of investors vacillating between different simple models and the resultant effect on asset prices. DeLong et al. [16] assumes that noise traders are subject to exogenous sentiment shocks and study how these shocks create risk that prevents arbitrageurs from eliminating mispricings and allow the noise traders to earn supernormal profits.

While there is a great deal of overlap between the asset pricing predictions of these prior works and ours, three unique predictions distinguish our model. First, we predict that increases in agent knowledge will exacerbate the effects we find, which none of the prior works suggest.[7] Second, our model suggests that volatility will be highest during market downturns.[8] Third, our model suggests that price responses will be more correlated than market fundamentals suggest and that these correlations partially reverse with time.

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

[7] In the case of models relying on model misspecification (e.g. Hong et al. [30]), one might conjecture that the opposite occurs - knowledgeable, sophisticated investors are more likely to have the correct asset pricing model and serve as arbitrageurs.

[8] Prior work (e.g. Officer [50], Schwert [58]) focuses on the time-series behavior of volatility (rather than excess volatility). These studies find that volatility is higher during market downturns and is only partially explained by factors such as leverage and macroeconomic volatility (Schwert [58]).

weather, length of the day, and sporting event outcomes within the geographical region of a trading market. In an early contribution, Saunders [55] finds that the cloud cover in Manhattan has a significantly negative effect on the prices of stocks traded through the New York Stock Exchange. Hirshleifer and Shumway [26] extends this analysis to cover markets across the globe and find the same negative correlation between cloud-cover and asset prices. Kamstra, Kramer and Levi [34] 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. The Kamstra, Kramer and Levi [34] results are particularly striking since changing day length is not only uninformative, it is predictable months in advance. Edmans, Garci and Norli [19] shows stock market valuations decline following a loss in important sports matches and that the magnitude of the loss is greater for smaller stocks and more important events. Our 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.[9] The change in prices is due to the fact that these events cause the moods and the corresponding biases in belief of traders within the market to become correlated.

Ljungqvist, Nanda and Singh [40] and Derrien [17] provide models wherein IPOs are timed to take advantage of the optimism of individual traders in the market to explain why IPOs appear underpriced in the short run and overpriced in the long run. Our 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. Our predictions are supported by Purnanandam and Swaminathan [51] and

[9] We refer to our model as Bayesian since the agents employ the correct model of the world to incorporate recollected data into prior beliefs to form posteriors. However the agents are assumed to be naive in that they are unaware of their memory imperfections and do not account for this aspect of their cognition when updating their beliefs.

Cook, Kieschnick and Van Ness [11].[10]

Tetlock [63] studies the influence of the Wall Street Journal's (WSJ) "Abreast of the Market" section on market sentiment. 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 [63] finds that negatively valenced words in the WSJ predict depressed performance of the DJIA in the following week. Tetlock argues that the widely read WSJ does not just convey information, but instead influences mood and in turn causes market prices to become depressed.[11]

Few prior studies of memory exist in the behavioral finance and economics literature. Mullainathan [48] provides a model of long-term associative memory to explain deviations from the consumption paths predicted by the permanent income hypothesis. Sarafidis [54] applies Mullainathan [48] in a setting where politicians take advantage of the polity's associative memory by strategically releasing information. Hirshleifer and Welch [27] show decision-making can exhibit inertia or impulsiveness when beliefs are driven by perfectly recalled actions combined with imperfectly recollected signals.

Several models of memory of limited volume have appeared in the literature.[12] Wilson [65] and Hellman and Cover [25] derive optimal memory processes and decision strategies

[10] We cannot claim that these effects would not be realized in application of other models of asset price over- and underreaction in this context. To isolate out model we would need to understand whether the promotion efforts involved hightening positive affect or cueing positive information in the memories of potential buyers (as our story suggests), or if the provision of new information played a crucial role in pre-IPO promotion (which might imply other models are also salient).

[11] More specifically, Tetlock [63] argues that the WSJ column does not solely have an informational effect. As in some of our prediction, Tetlock argues that overinference is a sign that there is more to the phenomenon than just conveyance of information.

[12] Models of limited memory volume could be interpreted as capturing the effects of short-term memory. For a survey of these topics, please see Cowan [12].

in the context a of decision problem with an infinite repetition of informative signals, but where memory is of a fixed, finite length.[13] Dow [18] 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. Benabou and Tirole ([3], [4], [5], [6]) 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 [23] uses a model of agents with malleable memory and preferences over their own attributes to explain anomalies in the literature of choice under risk.

1.2 Psychology of Emotion and Belief

Affective state has a clear and significant effect on histories recollected by agents in experimental settings (see Isen [31] and Clore, Schwarz and Conway [10] for surveys).[14] Isen et al. [32] is the seminal study of mood congruent recall. The authors employed a field study 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 study 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 free-sample treatment registered significantly higher product satisfaction ratings than those in the control group.

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 va-

[13]The results of Miller and Rozen [47] can be interpreted as an optimal mnemonic process.

[14]Hirshleifer and Shumway [26] provides an excellent summary of the research on mood and individual judgement. Elster [20], Loewenstein [41], and Loewenstein and Lerner [42] provide overviews of the role of emotion in decision making.

lence of the words that could be recollected in the new affective state. The affective state was created by having the subject play a computer game developed by 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 recollection.[15]

Kida, Smith and Maletta [36] provides a series of experiments that reveal that affective reactions to information are related to information in a straightforward fashion, encoded into memory, and have significant impact on the recall of the information at later times. The subjects, experienced managers, were required to remember and utilize numerical data, a task that was familiar from their job functions. The Kida, Smith and Maletta [36] experiments show that when provided with a natural benchmark from which to establish the valence of experimental data (e.g. industry benchmarks), managers were markedly better able to recall their affective reactions to the data than the actual comparison to the benchmark. Further, the evidence suggests that the managers had difficulty recollecting data about a firm when the valence of the data did not conform to their overall impression of the firm. For example, if a firm's accounting data was typically below market benchmarks, subjects had difficulty recollecting the firm had an exceptionally high cash flow relative to industry standards. This suggests that trained managers, who are exposed to large volumes

[15] The psychology literature strongly supports the notion that positive affect generates a mood-congruent recall bias towards positive recollections, but the evidence regarding the effects of negative mood is more contentious (Isen [31]). The model we develop is agnostic to asymmetries between positive and negative emotions and their relative influence on memory.

of numerical data on a regular basis, are subject to affective influences on their recollection processes.

2 Model

In this section we present the asset pricing problem and develop a model of memory that describes the effects of mnemonic cues (such as agent affective state) on the belief updating process. We assume throughout that the beliefs under study are those of a risk neutral representative consumer, which implies that asset prices equal the net present value of expected future dividends with a discount factor $r > 0$.[16,17] Prior to the time at which beliefs about asset valuations are formed, the agent is assumed to have observed a history of informative signals and stored this history in long-term memory. When forming a belief about asset values, the agent recollects a sample of data from memory in order to update his prior belief and form a posterior. The representative agent in our model fully and correctly incorporates recollected information into beliefs about asset valuation, but he does not foresee how shifts in the affective state will influence future prices.

2.1 Asset Pricing

We assume that dividends at time $t \in \{.., -1, 0, 1, 2, ...\}$, denoted d_t, are described by a random walk of the form $d_{t+1} = d_t + \eta_{t+1}$ where $\eta_{t+1} = \varepsilon_{t+1} + \gamma\varepsilon_t$, $\gamma \in [0, 1]$, and ε_t is symmetrically, identically, and independently distributed with a mean of θ unknown to the market participants.[18] We denote the event $\{\eta_t > 0\}$as good news and $\{\eta_t < 0\}$ as bad

[16]We thank an anonymous referee for suggesting the cleaner formulation using a representative agent structure. Prior versions employed a heterogeneous agent model with a partial equilibrium price setting assumption. Details are available upon request from the author.

[17]By allowing asset prices to be determined by the biased beliefs of a representative agent, we implicitly assume that arbitrageurs are unable to completely correct these biased prices.

[18]The exact form of dividend innovation autocorrelation is not crucial for any of our results.

news. For notational ease we assume the true dividend process is driftless, implying the true mean of ε_t is 0. Denote the probability density function of ε_t as $f(\cdot|\theta)$. The signals recollected by the agents are past realizations of the unpredictable component of dividend innovations (ε_t).

We let the variable $\varphi \in [0,1]$ denote the valence of the representative consumer's sentiment. In other words, higher values of φ denote more positive moods. The parameter is meant to capture the intensity of cues present at the time of decision as well as the strength of association between the cues and the data in memory.[19] φ could also represent a current news event that directly cues the recall of similar past news events.

Given the sentiment valence φ, a risk-neutral representative agent forms beliefs about the net present value of a security (and hence the market price) equal to

$$P_t(\varphi) = E_t\left[\sum_{\tau=1}^{\infty} \frac{d_{t+\tau}}{(1+r)^{\tau}}|\varphi\right]$$

where $E_t[\circ|\varphi]$ represents the time t expectations of an agent in cue state $\varphi \in [0,1]$. We describe $E_t[\circ|\varphi]$ formally in Section 2.2, but stated informally the agent's expectations are based on a posterior belief formed by updating a prior belief using information about past dividend innovations that is recalled from memory.[20] The sentiment valence, φ, quantifies biases in memory that influence what data is recalled and what data is forgotten (and hence not used to compute the expectation above).

[19] We 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. [62]).

[20] This information includes not only dividend announcements and earnings reports, but also the opinions of equity analysts, information gleaned from contacting industry insiders, and interviews with corporate officers. Most of these data would be difficult to recover except through memory. A reinterpretation of the model below is that hard information is perfectly recalled and that the biases are induced by the imperfect recall of relevant soft information. The qualitative effects of such a model would be identical to those found below although the notation would be significantly more complicated.

The net present value equation can be written

$$P_t(\varphi) = \frac{d_t + \gamma\varepsilon_t}{r} + \frac{1+\gamma+r}{r(1+r)} E_t[\varepsilon_{t+1}|\varphi]$$

The predictions of our model turn on how changes in φ impact expectations regarding ε_{t+1}.

2.2 Static Model of Recall

The representative agent recollects a subset of previously observed values of the unpredictable component of dividend innovations, ε_t, that are stored in long-term memory. The agent uses the recollected signals to conduct Bayesian updating of a prior belief regarding the distribution of $\theta \in \mathbb{R}$ with associated cumulative distribution function (CDF) $G(\theta)$.

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

Assumption 2. *$\{f(\varepsilon|\cdot)\}_{\varepsilon\in\mathbb{R}}$ is equicontinuous*

Assumption 1 implies that higher values of ε_i represent evidence of higher values of θ regardless of any other information observed. This property is sufficient for the distribution of θ given ε_i to first order stochastic dominate the distribution of θ given ε_i' where $\varepsilon_i > \varepsilon_i'$. Therefore higher values of ε_i are referred to as "good news" (Milgrom [44]). Assumption 2 is a technical assumption required for Bayesian estimation to be well-behaved.

A *history*, $H_t = \{..., \varepsilon_1, \varepsilon_2, ..., \varepsilon_t\}$, is a set of data observed by the agent prior to decision and stored in long term memory. A *recollected history* is a random variable that consists of those events that are recalled by the agent at the time of decision. Let the random variable $\mathbf{H}^R(\varphi)$ represent the recollected history (where the random variable takes the cue state as a parameter) with a typical realization denoted $H^R \subseteq H_t$. We will let $\mathbb{H}(H_t, n)$ denote the set of all possible length n recollected histories given the history H_N. We 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 in the recollection and Bayesian updating process.

Let the *recall probability function*, $\rho\left(H^R|\varphi, H_t, n\right)$, be the probability that a set of data H^R is recollected at the time of decision given a complete history of data H_t, a cue state φ, and a restriction that the recollection contain precisely $n \leq |H_t|$ elements. Let the length of the recollected history given the true history, H_t, is of length N be denoted by the random variable $l(N)$. Note that we assume $l(N)$ is independent of φ.

Contingent on recollecting n events, the contents of the recollected history are generated by a process of independent sampling without replacement from the complete history. Denote the relative sampling probability of recalling ε_i in a recollected history of length n from a complete history H_t as $\rho_\varepsilon(\varepsilon_i|\varphi, H_t, n)$. In order to generate associative memory effects, the relative sampling probabilities are assumed to be functions of the cue state with high cue states encouraging recollection of high signal values.

Assumption 3. $\rho_\varepsilon(\varepsilon_i|\varphi, H_t, n)$ *is log supermodular in* (ε_i, φ) *and continuous in* φ.

Assumption 4. $\rho_\varepsilon(\varepsilon_i|\varphi, H_t, n) > 0$ *if and only if* $\varepsilon_i \in H_t$

Assumption 3 states that high values of φ increase the probability of recollecting events indicative of a high value of θ (high ε_i) relative to events suggesting a low θ value (low ε_i). Assumption 4 defines the support of $\rho_\varepsilon(\varepsilon_i|\varphi, H_t, n)$.

We then have

$$\rho\left(H^R|\varphi, H_t, n\right) = \frac{\prod_{\varepsilon_i \in H^R} \rho_\varepsilon(\varepsilon_i|\varphi, H_t, n) * \prod_{\varepsilon_i' \in H_N \setminus H^R} (1 - \rho_\varepsilon(\varepsilon_i'|\varphi, H_t, n))}{\sum_{H^R \in \mathbb{H}(H_t, n)} \left[\prod_{\varepsilon_i \in H^R} \rho_\varepsilon(\varepsilon_i|\varphi, H_t, n) * \prod_{\varepsilon_i' \in H_t \setminus H^R} (1 - \rho_\varepsilon(\varepsilon_i'|\varphi, H_t, n))\right]}$$

Combining all of these elements, the probability that $\mathbf{H}^R(\varphi) = H^R$ is then

$$\Pr\{l(N) = n\} \frac{\prod_{\varepsilon_i \in H^R} \rho_\varepsilon(\varepsilon_i|\varphi, H_t, n) * \prod_{\varepsilon_i' \in H_N \backslash H^R} (1 - \rho_\varepsilon(\varepsilon_i'|\varphi, H_t, n))}{\sum_{H^R \in \mathbb{H}(H_t,n)} \left[\prod_{\varepsilon_i \in H^R} \rho_\varepsilon(\varepsilon_i|\varphi, H_t, n) * \prod_{\varepsilon_i' \in H_t \backslash H^R} (1 - \rho_\varepsilon(\varepsilon_i'|\varphi, H_t, n))\right]}$$

where $n = |H^R|$.

Of primary interest for the purposes of this study are the observable consequences of memory for decision outcomes (e.g. asset prices, portfolio choice). Let $q : \Theta \to \mathbb{R}$ denote an increasing function of θ. For example, $q(\theta)$ could represent the value of a firm's equity or the probability a firm will not default on its debt given an unknown, underlying firm profitability θ. Since the agent's judgment is based on what she recollects, the agent's expectation of $q(\theta)$ is a random variable that is a function of the recollection. Denote this random variable as

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

2.3 Bias in Prices

Recall our pricing equation

$$P_t(\varphi) = \frac{d_t + \gamma\varepsilon_t}{r} + \frac{1 + \gamma + r}{r(1 + r)} E_t[\varepsilon_{t+1}|\varphi]$$

Since the true mean of ε_{t+1} is 0, we can describe the bias as

$$\delta(\varphi) = \frac{1 + \gamma + r}{r(1 + r)} E_t[\varepsilon_{t+1}|\varphi]$$

We simplify our framework by focusing on situations where the representative agent recollects a large sample of data from memory, but this sample is both incomplete and biased. In effect the representative agent is confident in his beliefs regarding θ, but these beliefs exhibit a bias that depends on φ. The following comparative static is our principal result on the how affective state, or mnemonic cues more generally, influence market prices.

Theorem 1. *Suppose that as $N \to \infty$ we have $l(N) \overset{a.s.}{\to} \infty$ and $\frac{l(N)}{N} \overset{a.s.}{\to} 0$. Then for large N, $\varphi_1 \geq \varphi_2$ implies $\delta(\varphi_1) \geq \delta(\varphi_2)$.*

Since we have assumed that the signal distribution $f(\varepsilon_i|\theta)$ is log supermodular, high ε_i values are good news in the sense of providing unambiguous evidence of higher values of θ and $q(\theta)$. Higher cue states weight the recalled data towards higher values of ε_i for any given length of memory n due to the log supermodularity of the recall process. Combining these facts, for each fixed n and $\varphi_1 \geq \varphi_2$ 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, this relation holds when we consider randomizations over n. Taking aasymptotic limits as in Theorem 1 both provides tractability and reflects markets where prices are based on a large, but biased and incomplete, set of information.

Example 1. *Suppose $\varepsilon_i \sim N(\theta, 1)$ and $\rho(\varepsilon_i|\varphi) = \exp(2 * \varepsilon_i * \varphi)$. 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. The dashed line represents the distribution of these points recollect by an agent when $\varphi > 0$. The distribution of recalled data first order stochastically dominates the unbiased recollection. In the case of a normal distribution and under the conditions of Theorem 1, so we can write*

$$E[\varepsilon_i|\varphi] = \int s \frac{f(s|\theta)\rho(s|\varphi)}{\int f(x|\theta)\rho(x|\varphi)dx} ds$$

Using the fact that the true value is $\theta = 0$ and our choice of $\varepsilon_i \sim N(\theta, 1)$ and $\rho(\varepsilon_i|\varphi) =$

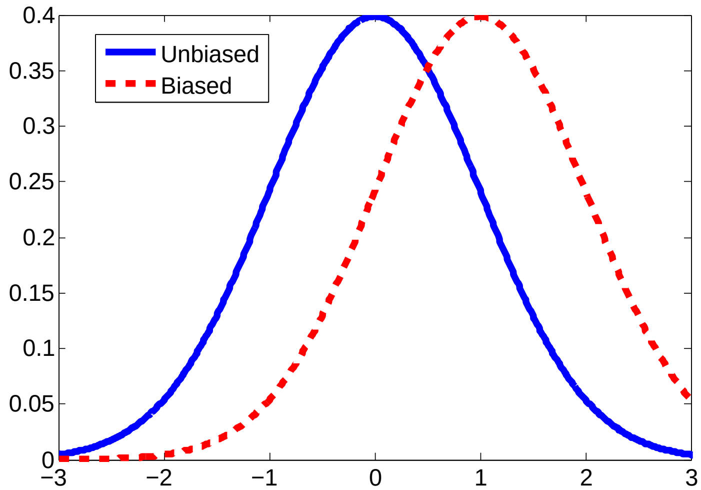

Figure 1: Biases in Recollection

$\exp(2 * \varepsilon_i * \varphi)$ *we have*

$$\begin{aligned}\delta(\varphi) &= \frac{1+\gamma+r}{r(1+r)}\int s e^{-0.5*(x-\varphi)^2} dx \\ &= \frac{1+\gamma+r}{r(1+r)}\varphi\end{aligned}$$

Although our results do not rely on this functional form, it is particularly tractable for embedding into other models.

2.4 Dynamic Model of Recall

In order to generate predictions regarding the time-series of asset prices, we must describe how the valence of the sentiment of the representative agent evolves over time. We assume that good news, $\eta_t \geq 0$, increases valence of agent affect, while bad news, $\eta_t \leq 0$, has the opposite effect.[21]

[21]The attribution of valence of the future can be indirect. For example, even if the agent does not hold the security, the associations would be strengthened if the good news was a signal of future good news

We caputure the dynamics of mood using a continuous transition function $g : [0, 1] \times \mathbb{R} \to [0, 1]$ with typical realization $g(\varphi_t, \eta_{t+1}) = \varphi_{t+1}$, which defines a mapping from the affective valence of period t and the dividend innovation in period $t + 1$ into the affective valence in period $t + 1$. We assume the following about the transition of affect

$$
\begin{aligned}
g(\varphi_t, 0) &= 0 && (2.1)\\
\frac{\partial}{\partial \eta_{t+1}} g(\varphi_t, \eta_{t+1}), \frac{\partial}{\partial \varphi_t} g(\varphi_t, \eta_{t+1}) &> 0 && (2.2)\\
\underset{\eta \to \underline{\eta}}{Lim} g(\varphi_t, \eta) &= 0, \underset{\eta \to \overline{\eta}}{Lim} g(\varphi_t, \eta) = 1 && (2.3)
\end{aligned}
$$

where $\underline{\eta}$ and $\overline{\eta}$ denote the upper and lower bounds of the support of η. Equation 2.1 imposes a normalization on the transition function. Equation 2.2 asserts that increases in dividend innovations improve the affective valence of the representative consumer and that a higher valence of affect in the current period implies a higher valence next period for any given dividend innovation. Equation 2.3 insures that the states of the model are accessible.

Together these assumptions imply that the process φ_t is ergodic, and we denote the invariant measure π. Importantly the ergodocity implies that shocks to affective state are transient, which differentiates the permanent, informative component of signals from the transient, affective component.

Lemma 1. *φ_t is ergodic.*

Proof. Assumption 2.1 implies that the process is strongly aperiodic. Assumption 2.3 implies the process is Harris recurrent, which when combined with strong aperiodicity implies that the affect process has an invariant measure and is irreducible (Theorem 10.4.2 of Meyn and Tweedie [43]). Together the properties of irreducibility, existence of an

regarding elements of his portfolio or the market generally.

invariant measure, and Harris recurrence imply the the process is ergodic and converges to the invariant measure π in the total variation norm (Theorem 13.3.1 of Meyn and Tweedie [43]) □

3 Market Effects

In this section we trace out linkages between affective state, beliefs, and market prices for financial assets. We show that the effects of dividend announcement on mood can explain over- and underreaction of securities prices to the release of news, predicts excess volatility of asset prices, implies asset prices exhibit excess correlation, and suggests that prices will respond to non-fundamental or repeated news events. Our model provides the unique predictions that excess volatility is maximized during market downturns and that knowledgeable agents may be more susceptible to these biases than agents who are relatively ignorant. Finally we provide of novel empirical predictions regarding asset price correlations and portfolio choice.

3.1 Dynamics of Over- and Underreaction

Let E_t^* refer to the expectations of an outside observer (e.g. an econometrician examining market price data). We assume that the information set of the outside observer does not include the affective valence unless otherwise stated. Recall that for an asset with a price P_t, in the model with perfect recall the future price increments are unpredictable and we would expect $E_t^*[P_{t+1} - P_t|\eta_t \geq 0] = E_t^*[P_{t+1} - P_t|\eta_t < 0]$ since all of the information from η_t would have been already incorporated into P_t.

Definition 1. ***Underreaction*** *is a short run phenomenon wherein securities for which good news has been recently reported tend to exhibit momentum over short time horizons.*

Formally given a security with current price P_t

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

In the case of underreaction, a short series of positive news events induces the market price to drift higher in the next period. The representative agent incorporates the news into the evaluation of asset value, but does not expect the price to drift upwards since he is naive regarding mnemonic biases.

Definition 2. ***Conditional Overreaction*** *is a long run phenomenon wherein securities for which good news has been repeatedly reported tend to exhibit mean reversion in the future. Formally for* j *sufficiently large*

$$E_t^*[P_{t+1} - P_t|\eta_t \geq 0, ..., \eta_{t-j} \geq 0] < 0 < E_t^*[P_{t+1} - P_t|\eta_t \leq 0, ..., \eta_{t-j} \leq 0]$$

A long series of positive news announcements pushes agent affective state to a high level, causing market participants to have extremely optimistically biased expectations of asset performance. Once this occurs good news events change beliefs only through their informational value, while bad news events influence beliefs through both informational and cuing effects. This asymmetry eventually pushes the agents' affective states towards a neutral valence and less biased beliefs, which causes asset prices to mean revert in the long run.

Finally, we provide a prediction regarding unconditional overreaction from a single news event, formally defined as follows

Definition 3. ***Unconditional Overreaction*** *is a long run phenomenon wherein securities for which good news has been reported once exhibit mean reversion in the future.*

Formally

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

$$\underset{\tau\to\infty}{Lim} E_t^*[P_{t+\tau 1} - P_t|\eta_t \leq 0] > 0$$

Recollect affect is an ergodic process. Therefore, if the representative agent's affect is shifted by a news event, it will eventually converge back to the ergodic distribution. The ergodic convergence implies the affect component of any price change fades with time, although the informational component remains.

3.2 Analysis of Market Effects

Empirical tests of the underreaction and overreaction phenomena examine panels of security prices. We study these phenomena by examining the drift of prices observed by an unbiased econometrician following valenced news events.

Theorem 2. *Consider an arbitrary $\varphi_t \in (0,1)$ and an asset with period t price P_t. If $f(\varepsilon_i|0)$ is log-concave in ε_i and differentiable, then asset prices exhibit underreaction in the short run*

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

Log-concavity of $f(\varepsilon_i|0)$ implies that good news this period raises the probability of good news next period, and the informational value of the anticipated good news is incorporated calculations of the asset value. If agent memory were unbiased, then market prices would form a random walk. The autocorrelation of changes in mood, rooted in the autocorrelation of η_t, generates predictable short run changes in asset prices. Good news next period further shifts the affective state of the market participants and increases next period's price further than the asset traders' estimates predict, which causes asset price underreaction.

We now show that a sufficiently long series of good (or bad) news events causes conditional overreaction. After a long series of good news the affect of the representative agent approaches the upper bound of $\varphi = 1$, and further positive news events have negligible effect on affect. However, a bad news event can significantly reduce agent affect and influence the memory biases. This asymmetry of cue evolution causes asset price conditional overreaction. We require an additional assumption on the effects of long sequences of good or bad news on affect to prove our result.

Assumption 5. *Consider an arbitrary initial state $\varphi_0 \in [0,1]$, $c > 0$, and sequence of $(\eta_1, \eta_2, ...)$ where $\eta_t > c$. Let φ_{k+1} be recursively defined as $\varphi_{k+1} = g(\varphi_k, \eta_{k+1})$. Then for any $\gamma > 0$ we can choose N such that $\varphi_n > 1 - \gamma$ for all $n > N$. Symmetrically, if $\eta_t < -c$ for all t, then we can choose N' such that $\varphi_n < \gamma$ for all $n > N'$.*

Assumption 5 implies that states of high affective valence do not require an extrordinairy event to be realized and can be reached through a lengthy sequence of arbitrarily positive events.

Theorem 3. *Consider an arbitrary $\varphi_t \in (0,1)$ and suppose assumption 5 holds. Asset prices exhibit conditional overreaction in the long run*

$$E_t^*[P_{t+1} - P_t | \eta_t \geq 0, ..., \eta_{t-j} \geq 0] < 0$$
$$< E_t^*[P_{t+1} - P_t | \eta_t \leq 0, ..., \eta_{t-j} \leq 0]$$

for sufficiently large j.

Our third prediction on asset price time-series behavior concerns unconditional mean reversion after a single news event. Our model predicts that an econometrician studying asset price time-series data would find that the cue effects fade over time as the distribution of cue states asymptotically returns to the ergodic distribution. Critical for this theorem, the unconditional overreaction must be taken as an average over the ergodic distribution. In effect, price changes regress when we look at the average effect over time.

Theorem 4. *Asset prices exhibit unconditional overreaction in the long run on average.*

$$\underset{\tau\to\infty}{Lim}\int E_t^*[P_{t+\tau}-P_t|\varphi_t,\eta_t]*\pi(d\varphi_{t-1})\leq 0 \; if \; \eta_t\geq 0$$
$$\underset{\tau\to\infty}{Lim}\int E_t^*[P_{t+\tau}-P_t|\varphi_t,\eta_t]*\pi(d\varphi_{t-1})\geq 0 \; if \; \eta_t\leq 0$$

Figure 2 describes the autocorrelation of changes in the representative agent's expectations of net present value of an asset. Without associative memory effects, there would

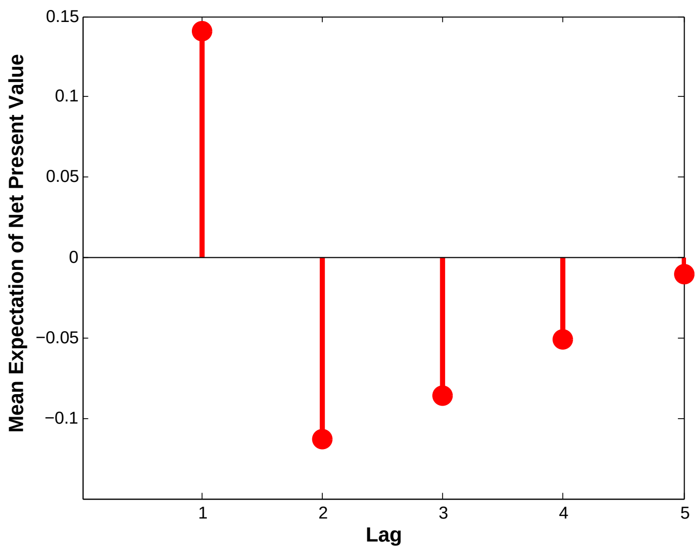

Figure 2: Biases in Recollection

be no autocorrelation. However, we see short-run underreaction in that changes in expectation about net present value are positively correlated over two periods. Long run overreaction is captured by the negative autocorrelation of changes in expectation over longer horizons. In cases where multiple successive news events of the same valence are realized, these correlations could extend over longer time horizons.

Finally, our theory suggests that agents may act as though dividends are more correlated across securities than the fundamentals suggest. Since affective cues have a general effect on cognitive processing, if an agent observes an affective cue regarding one element of his portfolio, then his affective state could bias his beliefs about unrelated segments of his portfolio. Positive or negative announcements by highly visible firms cause positive shifts in the representative agents affective state and make the representative agent optimistic about all assets in the industry (or the entire market). Our theory predicts that changing market affect will cause market prices to covary more than the correlation of the fundamentals would predict, but that these correlations will fade as the cue becomes less salient. However, our theory predicts that this excess covariance will be transient as the cue effects fade, which distinuishes our model from a purely informational theory.

3.3 Excess Volatility and Non-Fundamental News

In our model changes in asset prices are driven by two complementary effects. Information about future dividends is revealed by dividends in the current period. This information is incorporated into beliefs about the net present value of the asset, and the movements of price caused by this updating are referred to as *fundamental volatility.* Empirical studies find that asset prices exhibit volatility greater than can be explained by the fundamental contribution alone (see, for example, Shiller [59], [60]), and our model of asset pricing suggests that variations in affect can (partially) explain this excess volatility. If the arrival of new information also induces changes in agent affect, the movement of affective state will cause price volatility that we denote as *affective volatility.* Furthermore, information that drives fundamental changes in price also drives changes in affect, resulting in an *affective* *volatility of information.*

In our model the three effects can be separated allowing us to generate predictions about the relative magnitude of the affective and fundamental contributions to volatility in different market conditions. From period t to period $t+1$ the price increment is

$$P_{t+1} - P_t = \varepsilon_{t+1}\frac{1+\gamma}{r} + \delta(\varphi_{t+1}) - \delta(\varphi_t)$$

We interpret the variance of this price increment as volatility. We identify the fundamental contribution to volatility as

$$\left[\frac{1+\gamma}{r}\right]^2 * Var\left(\varepsilon_{t+1}\right) = \frac{1}{\gamma}\left[\frac{1+\gamma}{r}\right]^2 Cov(\eta_{t+1}, \eta_{t+2})$$

As the variance of ε_{t+1} increases, the fundamental contribution to volatility increases.

Without taking a stand on the particular parameterization of the recall probability function, ρ_ε, or the evolution of affective valence, g, it is difficult to make sharp predictions about the relationship of affective valence to volatility. Below we take two approaches. First, we describe conclusions that can be drawn about the relationship between affect and volatility without any parametric assumptions. Second, we describe particular regression formulas by assuming a linearized form of our model. In particular, we assume that

$$\delta(\varphi_t) = m * \varphi + c \text{ where } m, c \in \mathbb{R}$$

This relationship is exactly correct if $\varepsilon \sim N(\theta, \sigma^2)$ where σ^2 is known and for some $c, d > 0$ we take $\rho_\varepsilon(\varepsilon_i|\varphi) = c * \varepsilon_i * \varphi + d$.

Since the dynamic of affect drive many of our results, the functional form assumed for g is important. Our choice of g is designed to capture *immune neglect*, which denotes the empirical regularity that bad news has relatively little effect on positively valenced affects (Gilbert et al. [22]). The following functional form has the property that bad news ($\eta_{t+1} < 0$) causes smaller changes in φ_{t+1} than an equivalent[22] good news event.

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

Finally, to achieve closed forms we assume $\gamma = 0$ and note that all of our results are continuous and hold for small values of $\gamma > 0$. Sharp results in other cases can be generated using simulations.

The affective volatility is identified as

$$Var\left(\delta(\varphi_{t+1}) - \delta(\varphi_t)|\varphi_t\right)$$

The more mercurial affective state is with respect to news revelations, the greater the excess volatility will be. This can be written in terms of the current affective state and the innovation as

$$Var\left(\delta(g(\varphi_t, \eta_{t+1})) - \delta(\varphi_t)|\varphi_t\right)$$

This term is increasing in the variance of η_{t+1}, which implies that affective and fundamental volatility are positively related. Under our parametric assumptions we then have that affective volatility takes the form

$$Var\left(\delta(g(\varphi_t, \eta_{t+1})) - \delta(\varphi_t)|\varphi_t\right) = \frac{1}{2}m^2 * \left[(1-\varphi_t)^2 * Var(e^{-\eta_{t+1}}) + \varphi_t^2 Var(e^{\alpha\eta_{t+1}})\right]$$

Under our parametric assumptions,affective volatility is highest when pessimism is preva-

[22] Given a state φ_t and innovation $\eta_{t+1} \geq 0$, an equivalent bad news event is state $1-\varphi_t$ and innovation $-\eta_{t+1}$.

lent in the market (i.e. φ_t is close to 0) since $Var(e^{-\eta_{t+1}}) > Var(e^{-\alpha\eta_{t+1}})$.

The final component of volatility, the affective volatility effect of information, is the covariance of fundamental information with affective state

$$\frac{1+\gamma}{r} * Cov\left(\varepsilon_{t+1}, \delta(\varphi_{t+1}) - \delta(\varphi_t)|\varphi_t\right)$$

We can simplify this term to

$$\frac{1+\gamma}{r} * Cov\left(\varepsilon_{t+1}, \delta(g(\varphi_t, \eta_{t+1}))|\varphi_t\right)$$

Again, the affective volatility effect of information is correlated with our prior two sources of volatility.

Under our parametric assumptions we have

$$Cov\left(\varepsilon_{t+1}, \delta(g(\varphi_t, \eta_{t+1}))\right) = m * \varphi_t * E_t^*\left[\varepsilon_{t+1}e^{\alpha\varepsilon_{t+1}}\right] + m * (1-\varphi_t) * E_t^*\left[-\varepsilon_{t+1}e^{-\varepsilon_{t+1}}\right]$$

Again we find that the affective volatility effect of information is greatest during downturns (periods where φ_t) is low than during market rallies since $E_t^*\left[\varepsilon_{t+1}e^{\alpha\varepsilon_{t+1}}\right] < E_t^*\left[\varepsilon_{t+1}e^{\varepsilon_{t+1}}\right]$.

3.4 Levels of Knowledge

In this section we retreat from our use of the representative agent, and consider two classes of agents who possess different amounts of information stored in memory. One would expect such a bias to remain for (or even be exacerbated in) individuals who possess a large store of knowledge since these knowledgeable agents could, potentially, recollect an extremely biased set of data. Therefore equity analysts and professional traders could be more subject to these biases than semi-professional traders.

We provide an example to illustrate this effect. Let us assume, as above, that the agents are trying to estimate the mean of ε_t and that ε_t is a standard normal variable. There are two kinds of agents, inexperienced and experienced investors. Inexperienced investors have two realization of ε_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 all agents have the same relative sampling probability ρ_ε. Agents can be either in a neutral mood with unbiased recall ($\varphi = 0$ with $\rho_\varepsilon(\varepsilon_i|\varphi = 0) = 1$) or in a positive mood with optimistic recall ($\varphi = 1$ with $\rho_\varepsilon(\varepsilon_i|\varphi = 1) = \exp(\varepsilon_i)$). Note that since the agents recall a single datum, their expectation is equal to the value recalled.

Under a neutral affect, the estimates of both populations are standard normal distributions. Under optimistic recall, agents are biased towards recalling (a single) higher signal stored in memory. For participants with a great deal of knowledge, this places a very high likelihood on very high signals being recalled. Figure 3 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 inexperienced agents is biased by approximately $\frac{1}{3}$ of a standard deviation, while the experience agents are biased by more than 1 standard deviation. This effect is driven by the log-supermodularity of ρ_ε.

La Porta [38] shows that forecasts by stock market analysts reflect the pattern of long-run overreaction, which suggests that the beliefs of experienced market participants are subject to biases. We do not claim that the result of La Porta [38] 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 [38].

An alternative to our model of biased recall is that mood is treated by agents as an

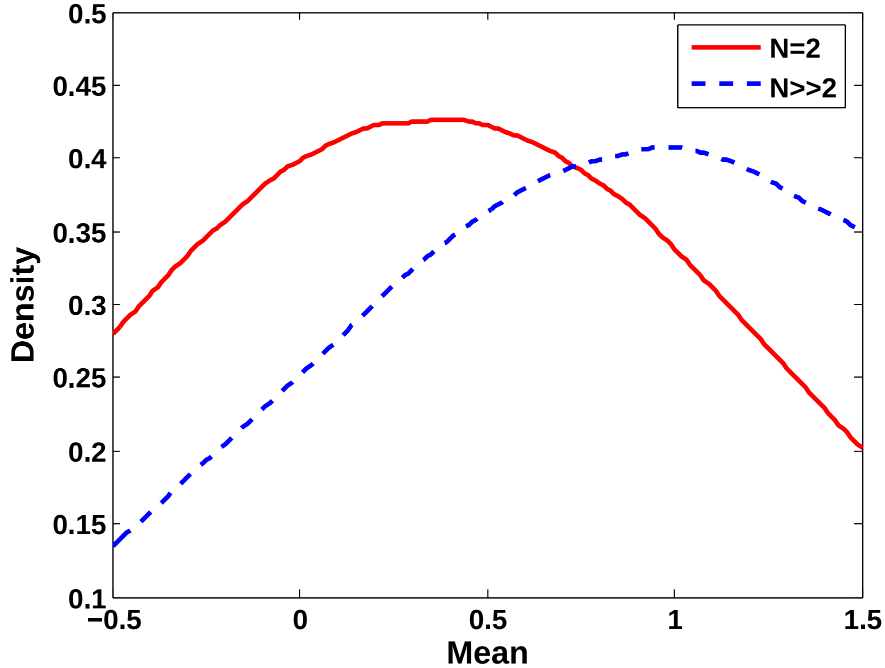

Figure 3: Biases in Recollection

informative signal (Schwarz and Clore [56] and [57]). 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. The predictions of this section dilenate our theory from this alternative.

3.5 Additional Empirical Tests

We divide our discussion of additional empirical predictions into three parts. The first part discusses novel predictions about the behavior of aggregate variables such as firm equity valuations in response to predictors of agent mood that serve as direct measures of market affect. The second part discusses the effect of mnemonic cues on how revenues from different markets are integrated to find the value of a firm. The third part discusses predictions in terms of portfolio choice that can be detected in cross-sectional data.

3.5.1 Exogenous Mood Predictors

A novel mood measure is the Gallup Daily: U.S. Mood Poll,[23] and our theory predicts that positive affect on trading days will have a positive effect on market prices. In the context of our theory, the advantage of examing mood as a predictor is that it highlights the exact psychological channel we are studying whereas weather is an indirect measure of affect. The great disadvantage of poll results relative to weather is that weather clearly has little informational content, but polls could reflect significant (if diffuse) amounts of information.

Two issues arise with interpreting a naive regression of asset prices on the mood expressed in the Gallup poll. First, the U.S. Mood Poll is presumably correlated with consumer confidence,[24] which plausibly contains fundamental information that ought to be correlated with asset prices. As suggested in Tetlock [63], 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 predictions of asset price changes (at least partially) reflects the influence of mood, then our theory predicts partial reversion of these changes. The partial reversion is the smoking gun phenomena for testing our theory. Second, reverse causality could be present if positive stock market moves induce poll respondents to express a more positive mood. One could use lagged poll reponses to eliminate effects of this nature.

The predictions our model predicts should also be correlated. Securities that exhibit a high degree of excess volatility are predicted to be subject to a higher degree of short-run price underreaction, long-run price overreaction, and response to non-fundamental news

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

[24] I thank a referee for the excellent example of how the mood poll could reflect fundamental information.

such as weather events or day length. If our effects are mediated through mood or public cues, one would expect the effects to hold on the industry- and market-level as well as for individual securities. These correlations, both across phenomena and across securities, have not been tested to my knowledge.

3.5.2 Generic Cues

We predict that mnemonic cues may affect how the different components of a firm's revenue influence market price. For example consider General Electric, a firm that sells capital equipment important for green-tech projects in addition to other divisions operating in a variety of other industries. When cues for future profits in the green-tech industry are salient (e.g. the bankruptcy of Solyndra), information regarding green-tech related events is over-represented in the recalled data and equity markets overweight the contribution of green-tech to GE's future profits. If GE has done well in the green-tech market GE's equity will exhibit an upward mispricing, and if GE has done poorly in the green-tech market GE's equity will be downwardly mispriced.

As in the case of the Gallup Poll, these cues may carry information on the fundamental profitability of GE's future profits. However, our theory predicts that these price effects will (partially) reverse as the cues become less salient with time. To test this prediction, the price of GE stock should be regressed on the valence of the cue (to control for momentum effects as documented above), the interaction of the presence of the cue and the performance of GE in the relevant market, and a set of traditional control variables. Our prediction is that the interaction variable is positive in the short run with reversion present in the long run.

3.5.3 Portfolio Choice

We can further test our model using panel data covering asset portfolio holdings in combination with data regarding cues the agents are facing. For example, we would expect asset holders in cloudy cities to divest from risky assets as their view of the net present value of future dividend flows decreases relative to the beliefs of asset holders in sunnier regions. If knowledgeable investors are more afflicted by this bias, then we would expect securities that are predominantly held by informed investors to exhibit a greater degree of these effects. Although it is well known that the equity of small firms is predomianntly held by individual investors, it is not immediate that these investors are well informed. This caveat aside, the results of Tetlock's analysis of the Fama-French small-minus-big factor provides tentative support for this prediction in the context of the WSJ "Abreast of the market" column (Tetlock [63]).

4 Conclusion

The principle goal of our paper is to model the effect of mood on associative memory and the resulting impact on asset prices. We are able to provide a unified explanation for the short-run underreaction and long-run overreaction of price to news, excess price volatility, and the response of prices to non-fundamental events. Our model provides the novel predictions that excess volatility will be highest in market downturns, knowledgable agents will be more prone to biases than less knowledgable agents, and asset prices will be more correlated than fundamentals indicate.

In addition to the empirical tests proposed, we hope to test our theory using an experimental asset market. A controlled experiment would allow us to study the response of experimental subjects to manipulations of affective state and test for interactions between demographic traits and response to mnemonic cues. In particular, we could monitor the data recalled by the subjects in each period of the experiment to insure that our affect manipulation has an effect on memory recall. Furthermore, we could elicit the subjects' beliefs to insure that our assumption of Bayesian updating provides a reasonable model of belief formation.

Although our model is founded on stylized facts of the psychology of memory, our theory could be reinterpreted to capture a number of other phenomena. For example, it could be that the agent has perfect recall, but the cues direct an agent's limited attention to particular pieces of information. If it is attention and not memory that is bounded, agents will require interventions that direct them to the most useful information rather than a memory aid that provides redundant information.[25]

Memory errors could also play a significant role in how shareholders evaluate executive competency and the timing of IPOs and other strategic market events. 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. 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 [11]. Modeling any of these effects requires incorporating our comparative statics into traditional finance models, which highlights an important advantage of the portability of Theorem 1. We hope to explore these applications in future work.

[25] An example of attention cuing is confirmatory bias (see Rabin and Schrag [52] for a review of the psychology literature). Suppose the agent entertains a theory about the relation between variables X and Y, and the agent's theory cues attention towards phenomena predicted by the theory. By focusing attention on data that is supportive of the theory, the attention bias amplifies the agent's belief in the theory's veracity. Furthermore agents holding diverse theories would become more certain in their different theories as evidence accumulated.

References

[1] Athey, S. (1998) "Monotone Comparative Statics Under Uncertainty," *Quarterly Journal of Economics*, 117, pp. 187 - 223.

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

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

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

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

[6] Benabou, R. and J. Tirole. (2006c) "Identity, Dignity, and Taboos: Beliefs as Assets," *mimeo.*

[7] Benartzi, S. and R.H. Thaler (1995) "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, Vol. 110, pp. 73-92.

[8] Bower, G.H. "Mood and Memory," *American Psychologist*, 36, pp. 129 - 148.

[9] Campbell, J. and R. Shiller (1988) "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, pp. 305-341.

[10] 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.

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

[12] Cowan, N. (1995) *Attention and Memory: An Integrated Framework.* Oxford University Press, New York City.

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

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

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

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

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

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

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

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

[21] Futia, C. (1982) "Invariant Distributions and the Limiting Behavior of Markovian Economic Models," *Econometrica*, Vol. 50, pp. 377-408.

[22] 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 Pyschology,* 75 (3), pp. 617 - 638.

[23] Gottlieb, D. (2009) "Imperfect Memory and Choice Under Risk," *mimeo.*

[24] 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.

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

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

[27] 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.

[28] 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.

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

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

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

[32] 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.

[33] 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.

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

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

[36] 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.

[37] Kyle, A.S. (1985) ”Continuous Auctions and Insider Trading,” *Econometrica*, 53, pp. 1315-1335.

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

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

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

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

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

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

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

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

[46] Miller, E. (1977) "Risk Uncertainty, and Divergence of Opinion," *The Journal of Finance*, 32, pp. 1151 - 1168.

[47] Miller, D. and K. Rozen (2009) "Need I Remind You? Monitoring with Collective Memory," *mimeo.*

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

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

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

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

[52] Rabin, M. and J. Schrag (1999) "First Impressions Matter: A Model of Confirmatory Bias," *The Quarterly Journal of Economics*, 114, pp. 37-82.

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

[54] 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.

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

[56] 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.

[57] 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.

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

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

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

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

[62] 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.

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

[64] Topkis, D. (1998) *Supermodularity and Complementarity.* Princeton, New Jersey: Princeton University Press.

[65] Wilson, A. (2003) "Bounded Memory and Biases in Information Processing," *mimeo.*

A Proofs

A.1 Notes on Stochastic Orders

The model and subsequent analysis make use of multivariate stochastic orders. Random variable $\mathbf{X}$ 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 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$,

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

where $\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), ..., \max(x_N, y_N))$ and $\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), ..., \min(x_N, y_N))$. This condition is identical to affiliation and log supermodularity (log-spm). 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. Muller and Stoyan [49] show that $\mathbf{X} \succcurlyeq_{tp2} \mathbf{Y}$ implies $\mathbf{X} \succcurlyeq \mathbf{Y}$.

A.2 Proofs from Section Two

Our first theorem shows that improved affect represses the recollection of bad news from memory.

Lemma 2. *If* $\rho_\varepsilon(\varepsilon|\varphi, H_N, n)$ *is log-spm in* (ε, φ)*, then* $1 - \rho_\varepsilon(\varepsilon|\varphi, H_N, n)$ *is log-spm in* $(-\varepsilon, \varphi)$

Proof. Note that $1 - \rho_\varepsilon(\varepsilon|\varphi, H_N, n)$ is log-spm in $(-\varepsilon, \varphi)$ is equivalent by definition to the claim that for all $\varepsilon \geq \varepsilon'$, $\frac{1-\rho_\varepsilon(\varepsilon|\varphi,H_N,n)}{1-\rho_\varepsilon(\varepsilon'|\varphi,H_N,n)}$ is decreasing in φ. Let $g(\varepsilon'; \varphi) = \frac{1}{\rho_\varepsilon(\varepsilon'|\varphi,H_N,n)}$ and

$$L(\varepsilon, \varepsilon'; \varphi) = \frac{\rho_\varepsilon(\varepsilon|\varphi, H_N, n)}{\rho_\varepsilon(\varepsilon'|\varphi, H_N, n)}$$

Then

$$\frac{1 - \rho_\varepsilon(\varepsilon|\varphi, H_N, n)}{1 - \rho_\varepsilon(\varepsilon'|\varphi, H_N, n)} = \frac{g(\varepsilon'; \varphi) - L(\varepsilon, \varepsilon'; \varphi)}{g(\varepsilon'; \varphi) - 1}$$

Noting that $L(\varepsilon, \varepsilon'; \varphi)$ is increasing in φ by assumption 3, we have that $\frac{1-\rho_\varepsilon(\varepsilon|\varphi,H_N,n)}{1-\rho_\varepsilon(\varepsilon'|\varphi,H_N,n)}$ is decreasing in φ. □

We now prove that the recall of sets of signals inherits the log-spm properties assumed for the function ρ_ε defining the relative recall probabilities of the individual signals.

Lemma 3. *Assumptions 3 and 4 imply that* $\rho\left(H^R|\varphi, H_N, n\right)$ *is log-spm in* (H^R, φ)*.*

Proof. (of Lemma 3) We will proceed by showing the log-spm relation holds between any pair of variables in (H^R, φ). See Karlin and Rinott [35] for a proof of the sufficiency of this argument.

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

$$\begin{aligned}\frac{\rho\left(H \cup\{\varepsilon_i, \varepsilon_j\}|\varphi, H_N, n\right)}{\rho\left(H \cup\{\varepsilon_i, \widehat{\varepsilon_j}\}|\varphi, H_N, n\right)} &= \frac{\rho_\varepsilon(\varepsilon_j|\varphi, H_N, n) * [1-\rho(\widehat{\varepsilon_j}|\varphi, H_N, n)]}{\rho_\varepsilon(\widehat{\varepsilon_j}|\varphi, H_N, n) * [1-\rho(\varepsilon_j|\varphi, H_N, n)]} \\ &= \frac{\rho\left(H \cup\{\widehat{\varepsilon_i}, \varepsilon_j\}|\varphi, H_N, n\right)}{\rho\left(H \cup\{\widehat{\varepsilon_i}, \widehat{\varepsilon_j}\}|\varphi, H_N, n\right)}\end{aligned}$$

Therefore we have log-spm in pairs $\{\varepsilon_i, \varepsilon_j\}$.

Suppose that $\varepsilon \geq \varepsilon'$,$\varphi \geq \varphi'$, and $H_1^R, H_2^R \in \mathbb{H}(H_N, n)$ differ only in that $\varepsilon \in H_1^R$, $\varepsilon' \notin H_1^R$ and $\varepsilon' \in H_2^R$, $\varepsilon \notin H_2^R$. Then we have $H_1^R \geq H_2^R$ (where this inequality refers to the majorizations of both recollected histories). Therefore

$$\frac{\rho\left(H_1^R|\varphi_1, H_N, n\right)}{\rho\left(H_2^R|\varphi_1, H_N, n\right)} = \frac{\rho(\omega|\varphi_1, H_N, n) * (1-\rho(\omega'|\varphi_1, H_N, n))}{\rho(\omega'|\varphi_1, H_N, n) * (1-\rho(\omega|\varphi_1, H_N, n))}$$

From Assumptions 3 and 4 it follows that

$$\begin{aligned}\frac{\rho\left(H_1^R|\varphi_1, H_N, n\right)}{\rho\left(H_2^R|\varphi_1, H_N, n\right)} &= \frac{\rho_\varepsilon(\varepsilon|\varphi_1, H_N, n) * (1-\rho_\varepsilon(\varepsilon'|\varphi_1, H_N, n))}{\rho_\varepsilon(\varepsilon'|\varphi_1, H_N, n) * (1-\rho_\varepsilon(\varepsilon|\varphi_1, H_N, n))} \\ &\geq \frac{\rho_\varepsilon(\varepsilon|\varphi_2, H_N, n) * (1-\rho_\varepsilon(\varepsilon'|\varphi_2, H_N, n))}{\rho_\varepsilon(\varepsilon'|\varphi_2, H_N, n) * (1-\rho_\varepsilon(\varepsilon|\varphi_2, H_N, n))} \\ &= \frac{\rho\left(H_1^R|\varphi_2, H_N, n\right)}{\rho\left(H_2^R|\varphi_2, H_N, n\right)}\end{aligned}$$

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

Log-supermodularity is preserved by multiplication, which is convenient in our context since the probability of observing a history H^R is the product of the relative sampling probabilities for each recollected datum, $\rho_\varepsilon(\varepsilon|\varphi, H_N, N)$, and each forgotten datum, $1 - \rho_\varepsilon(\varepsilon'|\varphi, H_N, N)$. Our assumptions insure that both $\rho_\varepsilon(\varepsilon|\varphi, H_N, N)$ and $1-\rho_\varepsilon(\varepsilon'|\varphi, H_N, N)$ have the requisite log supermodularity properties. Since log supermodularity is a property of the ratio of probabilities, the complex denominator of equation (3.5), which is common across all of the histories of length n, plays no role in our proof.

Lemma 4. *([35] or theorem 3.11.4, [49]) 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.*

The above 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 the measure can be taken to be the counting measure over the elements of H_N. Finally, we note that although we have established that $\rho\left(H^R|\varphi, H_N, n\right)$ is 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.

We now extend Milgrom's [44] representation result to the case of multiple signals. Assume that the agent has recollected a set of signals $H^R = (\varepsilon_1, ..., \varepsilon_n)$. 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 $G(\theta|H^R)$.

Lemma 5. *If the condition density of signals, $f(\varepsilon|\theta)$, is log-spm in (ε, θ), then $H_1^R = \overrightarrow{\varepsilon}_1 = (\varepsilon_1, ..., \varepsilon_n) \geq H_2^R = \overrightarrow{\varepsilon}_2 = (\widehat{\varepsilon_1}, ..., \widehat{\varepsilon_n})$ implies $G(\theta|\overrightarrow{\varepsilon}_1)$ first order stochastic dominates $G(\theta|\overrightarrow{\varepsilon}_2)$.*

Proof. Choose θ^* such that $0 < G(\theta^*) < 1$. Our assumption of the log-spm of $f(\varepsilon|\theta)$ implies that $f(\varepsilon_1, ..., \varepsilon_n|\theta) = \prod_{i=1}^{n} f(\varepsilon_i|\theta)$ has the log supermodularity property in $(\varepsilon_1, ..., \varepsilon_n; \theta)$ as products of log-spm functions are log-spm. This then implies

$$\frac{\int_{\overline{\theta}\geq\theta^*} f(\varepsilon_1,...,\varepsilon_n|\overline{\theta})G(d\overline{\theta})}{f(\varepsilon_1,...,\varepsilon_n|\theta)} \geq \frac{\int_{\overline{\theta}\geq\theta^*} f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\overline{\theta})G(d\overline{\theta})}{f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\theta)}$$

$$\begin{aligned}
&\implies \frac{f(\varepsilon_1,...,\varepsilon_n|\theta)}{\int_{\overline{\theta}\geq\theta^*} f(\varepsilon_1,...,\varepsilon_n|\overline{\theta})G(d\overline{\theta})} \leq \frac{f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\theta)}{\int_{\overline{\theta}\geq\theta^*} f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\overline{\theta})G(d\overline{\theta})} \\
&\implies \frac{\int_{\overline{\theta}\leq\theta^*} f(\varepsilon_1,...,\varepsilon_n|\theta)G(d\overline{\theta})}{\int_{\overline{\theta}\geq\theta^*} f(\varepsilon_1,...,\varepsilon_n|\overline{\theta})G(d\overline{\theta})} \leq \frac{\int_{\overline{\theta}\leq\theta^*} f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\theta)G(d\overline{\theta})}{\int_{\overline{\theta}\geq\theta^*} f(\widehat{\varepsilon_1},...,\widehat{\varepsilon_n}|\overline{\theta})G(d\overline{\theta})} \\
&\implies \frac{G(\theta^*|\varepsilon_1,...,\varepsilon_n)}{1-G(\theta^*|\varepsilon_1,...,\varepsilon_n)} \leq \frac{G(\theta^*|\widehat{\varepsilon_1},...,\widehat{\varepsilon_n})}{1-G(\theta^*|\widehat{\varepsilon_1},...,\widehat{\varepsilon_n})} \\
&\implies G(\theta^*|\varepsilon_1,...,\varepsilon_n) \leq G(\theta^*|\widehat{\varepsilon_1},...,\widehat{\varepsilon_n})
\end{aligned}$$

as required. □

The above proposition captures the notion of a multi-dimensional signal as either good or bad news in the sense of Milgrom [44]. 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.

The condition that $|H_1^R| = |H_2^R|$ is an important requirement of Lemma 5. As noted earlier, if the recollected histories are of different lengths, no stochastic ordering may be possible. To see this, consider the following simple example of updating a normal distributed prior with normally distributed signals.

Example 2. *Consider prior beliefs $g(\theta) = N(0,1)$ and signals distributed as $\varepsilon_i \sim N(\theta_0, 1)$. 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 $G(\circ|H_1^R) \sim N(0, \frac{1}{3})$ and $G(\circ|H_2^R) \sim N(\frac{1}{4}, \frac{1}{2})$. Although $G(\circ|H_2^R)$ has a higher mean than $G(\circ|H_1^R)$, the lower variance of $G(\circ|H_2^R)$ implies that neither a first nor a second order stochastic dominance relation can* *be established between these posteriors.*

Prior to proving Theorem 1, we use the theory of Monotone Comparative Statics to prove useful lemmas that may be of independent interest if our model is applied outside of our asset pricing context. 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 we will show that $g(\theta|H^R)$ possesses the log-spm property required to apply Athey's monotone comparative statics results.

Lemma 6. *$g(\theta|H^R, n)$ is log supermodular in (θ, H^R) where n is fixed.*

Proof. Note that

$$g(\theta|H^R, n) = g(\theta) * \frac{1}{\widetilde{f}(H^R)} * f(H^R|\theta) = \frac{g(\theta)}{\widetilde{f}(H^R)} \prod_{\varepsilon_i \in H^R} f(\varepsilon_i|\theta) \tag{A.1}$$

where $\widetilde{f}(H^R)$ and $f(H^R|\theta)$ denote the unconditional and conditional distributions of signals in a length $|H^R| = n$ series. As log supermodularity is preserved by multiplication, this implies that $\prod_{\varepsilon_i \in H^R} f(\varepsilon_i|\theta)$ 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, $\frac{g(\theta)}{\widetilde{f}(H^R)} \prod_{\varepsilon_i \in H^R} f(\varepsilon_i|\theta)$ is log supermodular in (θ, H^R). From the above equalities, we then have that $g(\theta|H^R)$ is log supermodular in (θ, H^R). □

As we 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) where n is fixed, 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 7. *Let $x^*(H^R) \in \arg\max_{x \in X} \int u(x; \theta) G(d\theta|H^R)$. If $|H^R| = |\widehat{H^R}|$ and $H^R \geq \widehat{H^R}$,* *then we have* $x^*(H^R) \geq x^*(\widehat{H^R})$[26]

Proof. This follows directly from [1] since we have, by assumption, $u(x;\theta)$ obeys the single crossing property and shown in Lemma 6 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.

□

Now we prove a monotone comparative static result regarding stochastic decision problems under the influence of affect. This would be of interest analyzing, for example, portfolio choice problems under the influence of affect.

Lemma 8. *Let* $x^*(H^R) \in \arg\max_{x\in X} \int u(x;\theta)G(d\theta|H^R)$, $x^*(H^R)$ *be a singleton, and* $u(x;\theta)$ *satisfy 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. (Proof of Theorem 8) From Lemma 7 we 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 $(\overrightarrow{\omega}, \varphi)$, we have that $\mathbf{H}^R(\varphi_1; n) \succcurlyeq \mathbf{H}^R(\varphi_2; n)$. Therefore, from Theorem 3.3.11 of [49], we 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, we have that the same result holds for the unconditional distribution of $x^*(\mathbf{H}^R(\varphi_1; n))$, so $x^*(\mathbf{H}^R(\varphi_1)) \succcurlyeq x^*(\mathbf{H}^R(\varphi_2))$. □

Now that our lemmas are complete, we proceed to prove Theorem 1. Since our claim is essentially a statement about Bayesian estimation, we employ our monotone comparative statics result to make statements regarding the maximizer of a log-likelihood function in the limit as $l(N) \overset{a.s.}{\to} \infty$

[26] As in Athey [1], the ordering on the choice correspondence refers to the strong set order.

Theorem 1. *Suppose that as $N \to \infty$ we have $l(N) \to \infty$ and $\frac{l(N)}{N} \to 0$. Then for large N, $\varphi_1 \geq \varphi_2$ implies $\delta(\varphi_1) \geq \delta(\varphi_2)$.*

Proof. The representative agent uses a set of recalled data $\{\varepsilon_1, ..., \varepsilon_{l(N)}\}$ to form beliefs regarding the true value of θ. Standard results on Bayesian estimation imply that the agent chooses the estimate

$$\widehat{\theta}_{l(N)}(\varphi) \in \arg\max_{\widehat{\theta}} \frac{g(\widehat{\theta})}{l(N)} + \frac{1}{l(N)} \sum_{i=1}^{l(N)} \ln f(\varepsilon_i | \widehat{\theta})$$

Let us consider the problem in the limit as $N \to \infty$

$$\theta(\varphi) \in \arg\max_{\widehat{\theta}} E\left[\ln f(\varepsilon_i|\widehat{\theta})|\varphi\right]$$

From assumption 2 and $l(N) \to \infty$ as $N \to \infty$ we have $\widehat{\theta}_{l(N)}(\varphi) \to \theta(\varphi)$ as $N \to \infty$. Note that assumption 1 implies that $\ln f(\varepsilon_i|\widehat{\theta})$ is supermodular. Our claim then follows from Lemma 8. □

A.2.1 Proofs of Price Time-Series Effects

Before proving our theorems, we first provide the following useful lemma on stochastic orderings of η_{t+1} conditional on η_t.

Lemma 9. *If $f(\cdot|\theta = 0)$ is log-concave and differentiable, then for any increasing function $\phi : \mathbb{R} \to \mathbb{R}$ that $E_t^* \left[\phi(\eta_{t+1})|\eta_t\right]$ is increasing in η_t.*

Proof. To see this it suffices to note that since $f(\cdot|\theta = 0)$ is log-concave we have that $f(x - b|\theta = 0)$ is log-supermodular in (x, b). To see this note, consider the following relation for $x > \widehat{x}$

$$\begin{aligned} \frac{d}{db}\left[\frac{f(x-b|0)}{f(\widehat{x}-b|0)}\right] &= -\frac{f'(x-b|0)}{f(\widehat{x}-b|0)} + \frac{f(x-b|0)f'(\widehat{x}-b|0)}{\left(f(\widehat{x}-b|0)\right)^2} \geq 0 \\ &\Longrightarrow \frac{f'(x-b|0)}{f(x-b|0)} \leq \frac{f'(\widehat{x}-b|0)}{f(\widehat{x}-b|0)} \end{aligned}$$

This relation holds only if

$$\frac{f'(x-b|0)}{f(x-b|0)}$$

is decreasing in x, which is given by the log-concavity of $f(\cdot|0)$. We then have that $f(x-b|0)$ is log-supermodular. Therefore the quantity $f(x-b|0)f(y-b|0)$ is log-supermodular in (x, y, b) since log-supermodularity is preserved by multiplication. Finally, log-supermodularity is preserved by integration (Karlin and Rinott [35]), so

$$\int_{-\infty}^{\infty} f(x-b|0)f(y-b|0)f(b|0)db$$

is log-supermodular in (x, y). But note that this implies that the following is increasing in y

$$\begin{aligned} \frac{\Pr\{\eta_{t+1} = x|\eta_t = y, \theta = 0\}}{\Pr\{\eta_{t+1} = \widehat{x}|\eta_t = y, \theta = 0\}} &= \frac{\Pr\{\varepsilon_{t+1} + \gamma\varepsilon_t = x|\varepsilon_t + \gamma\varepsilon_{t-1} = y, \theta = 0\}}{\Pr\{\varepsilon_{t+1} + \gamma\varepsilon_t = \widehat{x}|\varepsilon_t + \gamma\varepsilon_{t-1} = y, \theta = 0\}} \\ &= \frac{\int_{-\infty}^{\infty} f(x-\gamma\varepsilon_t|0)f(\frac{y-\varepsilon_t}{\gamma}|0)f(\varepsilon_{t-1}|0)d\varepsilon_{t-1}}{\int_{-\infty}^{\infty} f(\widehat{x}-\gamma\varepsilon_t|0)f(\frac{y-\varepsilon_t}{\gamma}|0)f(\varepsilon_{t-1}|0)d\varepsilon_{t-1}} \end{aligned}$$

Since $\Pr\{\eta_{t+1} = x|\eta_t = y, \theta = 0\}$ is log-supermodular in (x, y), we then have that for any increasing function ϕ that $E_t^*\left[\phi(\eta_{t+1})|\eta_t\right]$ is increasing in η_t. □

Theorem 2. *Consider an arbitrary $\varphi_t \in (0, 1)$ and an asset with period t price P_t. If $f(\varepsilon_i|0)$ is log-concave in ε_i and differentiable, then asset prices exhibit underreaction in the* *short run*

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

Proof. Note that

$$E_t^*[\varepsilon_{t+1}|\eta_t > 0] = E_t^*[\varepsilon_{t+1}|\eta_t < 0] = 0$$

so our claim can then be written as

$$E_t^*[\delta(\varphi_{t+1})|\eta_t \geq 0, \varphi_t] \geq E_t^*[\delta(\varphi_{t+1})|\eta_t < 0, \varphi_t]$$

Using our affect dynamic formula we have the equivalent statement

$$E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t \geq 0] \geq E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t < 0]$$

From Theorem 1 and the monotonicity of g with respect to η_{t+1}, we have that $\delta(g(\varphi_t, \eta_{t+1}))$ is increasing in η_{t+1}. From Lemma 9 we have that $E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t = x]$ is then increasing in x. We can therefore write

$$\begin{aligned} E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t &\geq 0] = \int_0^\infty E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t = x](2 * dK(x)) \\ &\geq \int_{-\infty}^0 E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t = x](2 * dK(x)) \\ &= E_t^*[\delta(g(\varphi_t, \eta_{t+1}))|\eta_t \leq 0] \end{aligned}$$

where the factor of 2 corrects for the fact that the distribution of η_t is symmetric about 0. This then implies

$$E_t^*[\delta(\varphi_{t+1})|\eta_t \geq 0] \geq E_t^*[\delta(\varphi_{t+1})|\eta_t < 0]$$

as required. □

Theorem 3. *Consider an arbitrary $\varphi_t \in (0,1)$. Asset prices exhibit conditional overreaction in the long run*

$$E_t^*[P_{t+j+1} - P_{t+j} | \eta_{t+j} \geq 0, ..., \eta_t \geq 0] < 0$$
$$< E_t^*[P_{t+j+1} - P_{t+j} | \eta_{t+j} \leq 0, ..., \eta_t \leq 0]$$

for sufficiently large j.

Proof. We need to derive the distribution of

$$P_{t+1} - P_t = \varepsilon_{t+1} \frac{1+\gamma}{r} + \big(\delta(\varphi_{t+1}) - \delta(\varphi_t)\big)$$

contingent on $\{\eta_t \geq 0, ..., \eta_{t-N} \geq 0\}$ and $\{\eta_t \leq 0, ..., \eta_{t-N} \leq 0\}$. We can state our overreaction result as

$$E_t^*[\delta(\varphi_{t+j+1}) - \delta(\varphi_{t+j}) | \eta_{t+j} \geq 0, ..., \eta_t \geq 0] <$$
$$E_t^*[\delta(\varphi_{t+j+1}) - \delta(\varphi_{t+j}) | \eta_{t+j} \geq 0, ..., \eta_t \leq 0]$$

Consider an arbitrary sequence $(\eta_t, \eta_{t+1}, \eta_{t+2}, ..)$ such that $\eta_{t+k} > 0$ From Assumption 5, for any initial state φ_t and $\gamma > 0$ we can choose a sequence sufficiently long (j sufficiently large) that $\varphi_{t+j} > 1 - \gamma$ where $\varphi_{t+k+1} = g(\varphi_{t+k}, \eta_{t+k+1})$. The continuity of g implies that for any $\rho < 1$ we can choose j^* sufficiently large that $\varphi_{t+j^*} > 1 - \gamma$ for a measure $1 - \rho$ of paths $(\eta_t, \eta_{t+1}, ..., \eta_{t+j^*})$ in the event $\{\eta_{t+j^*} \geq 0, ..., \eta_t \geq 0\}$.

From equation 2.3, there exists a probability $\varepsilon > 0$ that η_{t+j^*+1} sufficiently small is realized that $\varphi_{t+j^*+1} \leq \frac{1}{2}$. Therefore we can bound $E_t^*[\delta(\varphi_{t+j^*+1}) | \eta_{t+j} \geq 0, ..., \eta_t \geq 0]$ by

$$E_t^*[\delta(\varphi_{t+j^*+1}) | \eta_{t+j^*} \geq 0, ..., \eta_t \geq 0] \leq (1-\rho) * \left[(1-\varepsilon) * \delta(1) + \varepsilon * \delta(\frac{1}{2})\right] + \rho\delta(1)$$

From the boundedness of g and δ we can choose ρ sufficiently small that

$$E_t^*[\delta(\varphi_{t+j^*+1}) - \delta(\varphi_{t+j^*})|\eta_{t+j^*}] \geq 0, ..., \eta_t \geq 0] \leq (1-\rho) * \left[(1-\varepsilon) * \delta(1) + \varepsilon * \delta(\frac{1}{2})\right] +$$
$$\rho\delta(1) - \delta(1-\gamma)$$
$$= (1-\rho) * \left[(1-\varepsilon) * [\delta(1) - \delta(1-\gamma)] + \varepsilon * \left[\delta(\frac{1}{2}) - \delta(1-\gamma)\right]\right] +$$
$$\rho\,[\delta(1) - \delta(1-\gamma)]$$

For choices of $\gamma, \rho > 0$ sufficiently small we can make the terms $[\delta(1) - \delta(1-\gamma)]$ and $\rho\,[\delta(1) - \delta(1-\gamma)]$ arbitrarily close to 0, while $\left[\delta(\frac{1}{2}) - \delta(1-\gamma)\right]$ remains significant and negative. Therefore, there exists j^* such that

$$E_t^*[\delta(\varphi_{t+j^*+1}) - \delta(\varphi_{t+j^*})|\eta_{t+j^*} \geq 0, ..., \eta_t \geq 0] < 0$$

Symmetric arguments imply there exists a j^{**} sufficiently large that

$$E_t^*[\delta(\varphi_{t+j^{**}+1}) - \delta(\varphi_{t+j^{**}})|\eta_{t+j^{**}} \leq 0, ..., \eta_t \leq 0] > 0$$

□

Theorem 4. *Asset prices exhibit overreaction in the long run on average.*

$$\underset{\tau\to\infty}{Lim} \int E_t^*[P_{t+\tau} - P_t|\varphi_{t-1}, \eta_t] * \pi(d\varphi_{t-1}) \leq 0 \ \textit{if}\ \eta_t \geq 0$$
$$\underset{\tau\to\infty}{Lim} \int E_t^*[P_{t+\tau} - P_t|\varphi_{t-1}, \eta_t] * \pi(d\varphi_{t-1}) \geq 0 \ \textit{if}\ \eta_t \leq 0$$

Proof. Consider an arbitrary realization of $\eta_t > 0$. Let ν denote the distribution of φ_t

where for any $x \in \mathbb{R}$

$$\nu((-\infty, x]) = \pi((-\infty, g^{-1}(x, \eta_t)])$$

where $g^{-1}(\cdot, \eta_t)$ is the inverse of $g(\cdot, \eta_t)$.[27] Furthermore, note that ν first order stochastically dominates π since $g(x, \eta_t) > x$ for all $x < 1$ and $\eta_t > 0$.

As in our earlier proofs, we have

$$\begin{aligned} E_t^*[P_{t+\tau} - P_t | \eta_t &\geq 0, \varphi_{t-1}] = E_t^*[\delta(\varphi_{t+\tau}) - \delta(\varphi_t) | \eta_t \geq 0, \varphi_{t-1}] \\ &= E_t^*[\delta(\varphi_{t+\tau}) | \eta_t \geq 0, \varphi_{t-1}] - E_t^*[\delta(\varphi_t) | \eta_t \geq 0, \varphi_{t-1}] \end{aligned}$$

From the ergodicity of φ_t, in the limit as $\tau \to \infty$ the distribution of $\varphi_{t+\tau}$ approaches π in the total variation norm (and hence weakly). We then have for any φ_t

$$\underset{\tau \to \infty}{Lim} E_t^*[\delta(\varphi_{t+\tau}) | \eta_t \geq 0, \varphi_t] = \int \delta(\varphi) * \pi(d\varphi)$$

It is straightforward to note that

$$\int E_t^*[\delta(\varphi_t) | \eta_t \geq 0, \varphi_{t-1}] * \pi(d\varphi_{t-1}) = \int \delta(\varphi_t) * \nu(d\varphi_t)$$

Since ν first order stochastically dominates π and δ is increasing we have

$$\int \delta(\varphi_t) * \pi(d\varphi_t) \leq \int \delta(\varphi_t) * \nu(d\varphi_t)$$

which yields

$$\underset{\tau \to \infty}{Lim} \int E_t^*[\delta(\varphi_{t+\tau}) - \delta(\varphi_t) | \eta_t \geq 0, \varphi_t] * \pi(d\varphi_t) \leq 0$$

as required. Symmetric reasoning implies the result in the case $\eta_t \leq 0$. □

[27] The inverse is well defined from the strict monotonicity of $g(\cdot, \eta_t)$.