A retrieved-context theory of financial decisions*

Jessica A. Wachter† Michael Jacob Kahana‡

December 31, 2019

Abstract

Studies of human memory indicate that features of an event evoke memories of prior associated contextual states, which in turn become associated with the current event's features. This mechanism allows the remote past to influence the present, even as agents gradually update their beliefs about their environment. We apply a version of retrieved context theory, drawn from the literature on human memory, to four problems in asset pricing and portfolio choice: over-persistence of beliefs, providence of financial crises, price momentum, and the impact of fear on asset allocation. These examples suggest a recasting of neoclassical rational expectations in terms of beliefs as governed by principles of human memory.

JEL codes: D91, E71, G11, G12, G41

*We are grateful to Ernst Fehr, Cary Frydman, Nicola Gennaioli, Robin Greenwood, Valentin Haddad, Ulrike Malmendier, Nikolai Roussanov, Andrei Shleifer, and seminar participants at the California Institute of Technology, Northwestern University, Wharton, the Miami Behavioral Finance Conference, the NBER Behavioral Finance Meeting, the NBER Summer Institute, the Red Rock Finance Conference, the Sloan/Nomis Workshop on the Cognitive Foundations for Human Behavior for helpful comments.

†Department of Finance, The Wharton School, University of Pennsylvania and NBER; Email: jwachter@wharton.upenn.edu

‡Department of Psychology, University of Pennsylvania; Email: kahana@psych.upenn.edu

1 Introduction

Standard decision-making under uncertainty starts with a probability space and an information structure. The information structure implies that the agent associates a value with every subset of the space and then maximizes expected utility. This is the approach of Savage (1954). The difficulty that agents have in forming beliefs over an entire state space has been formulated in the Ellsberg paradox (Ellsberg, 1961), ambiguity aversion formalized by Gilboa and Schmeidler (1989) and in the alternative representations of choice as a probabilistic selection among a small set of alternatives, due to Luce (1959) and McFadden (2001).[1] The set of possible states of nature is impossibly large and ever-changing. Nonetheless, we as individuals do manage to make decisions under uncertainty.

In this paper, we propose a memory-based model of decision-making under uncertainty. A wealth of data support the idea of a human memory system that maintains a record of associations between experiential features of the environment, and underlying contextual states (Kahana, 2012). This record of associations, together with inference about the current contextual state, constitutes a belief system that could potentially affect any kind of choice under uncertainty. This belief system responds to the current environment through retrieved context. The mechanism of retrieved context is how memory "knows" what information is most relevant to bring forward to our attention at any given time. At the same time, any new experience, and the context itself, is then stored again in the memory system (Howard and Kahana, 2002).

This paper applies these concepts to puzzles in asset pricing and portfolio choice that defy the standard Bayesian paradigm. Chief among these are the result that life experience has near-permanent effects on financial decision making (Malmendier and

[1]The problem of determining the underlying state space continues to be a point of contention in recent literature on ambiguity aversion: see, for example, the debate concerning rectangularity of the model set (Epstein and Schneider, 2003; Hansen and Sargent, 2018).

Nagel, 2011, 2016; Malmendier et al., 2017; Malmendier and Shen, 2018), and that an exogenous cue, such as a horror movie, can influence financial decisions (Guiso et al., 2018). More speculatively, we then apply the framework to a broader set of phenomena, such as the sudden onset of a financial crisis, and the momentum effect in the cross-section of asset returns.

When making a decision, an agent is confronted by certain *features* of the environment. For simplicity, we assume that features are perceived as discrete, and that there are a finite number of possible features pertaining to a particular decision. In the benchmark model, a feature vector is an element of $\mathcal{B}^n \subset \mathbb{R}^n$, where $\mathcal{B}^n$ is a set of basis vectors that spans n-dimensional space. For convenience, we assume the standard basis. That is, the time$-t$ features vector f_t has ith element equal to 1 if the ith instance of the feature is realized at time t and all other elements equal to zero. One can think of the features vector as a mathematical representation of objective, verifiable, and most likely transitory, aspects of the environment.

Features are connected over time through *context*. Context is an internal mental state that endows the agent with an understanding of possibly latent aspects of the environment that are relevant for the decision at hand. Specifically, define the *context space* as the standard simplex $\mathcal{A}_c \subset \mathbb{R}^m$, for $m \leq n$. That is $\mathcal{A}_c = \{c_t = [c_{1t}, \ldots, c_{mt}]^\top \in \mathbb{R}^m \,|\, \iota^\top c_t = 1\}$, where ι denotes a conforming vector of ones. We will think of c_t as assigning probabilities to the underlying states of nature, and at that point proceed in a manner similar to the standard economic approach. Indeed, a special case of our framework will be the Bayesian problem under which the agent learns about an unobserved state (context) from observed data (features). Principles of memory, however, can lead context to evolve in ways that are distinctly non-Bayesian.

Whereas many applications of psychological principles to economic decision making have focused on cognitive biases such as loss aversion and narrow framing (see Barberis (2013)), or on limited attention (see Gabaix (2019)) the literature on human learning and memory offers a different perspective. Three major laws govern the human memory system: similarity, contiguity, and recency: Similarity refers to the priority accorded to information that is similar to the presently active features, contiguity refers to the priority given to features that share a history of co-occurrence with the presently active features, and recency refers to priority given to recently experienced features. All three "laws" exhibit universality across agents, feature types, and memory tasks and thus provide a strong basis for a theory of economic decision-making.

While few economic models explicitly incorporate these laws, there are exceptions. Gilboa and Schmeidler (1995) replace axiomatic expected utility with utility computed using probabilities that incorporate the similarity of the current situation to past situations. Mullainathan (2002) proposes a model in which agents tend to remember those past events which resemble current events, and where a previous recollection increases the likelihood of future recollection. He applies the model to the consumption-savings decision. Nagel and Xu (2018) show that a constant-gain learning rule about growth in dividends can explain a number of asset pricing puzzles; they motivate this learning rule using the memory principle of recency. Recency-bias is present also in models of extrapolative expectations (Barberis et al., 2015) and in natural expectations (Fuster et al., 2010). These models do not employ context-based retrieval, which is the focus of our paper. Bordalo et al. (2019) develop a model based on the geometric similarity of representations in memory. They focus on the the role that similarity in memory representations plays in accounting for the propensity of agents to make large expenditures on housing or durable goods when lower expenditures would appear optimal by standard theory. Their work differs from ours in that we focus on the retrieval of prior contextual states, and we directly model contextual evolution. In their model, as in psychological studies such as Godden and Baddeley (1975), context is embedded in the environment, and thus is static; the feature layer of the environment and the context layer are the same.

2 Integrating Memory into Decision Making

Unless agents have full access to all decision-relevant information at the moment of choice, they must use their memory of past experiences to guide their decisions. The question of how past experiences influence present behavior has occupied the attention of experimental psychologists for more than a century (Ebbinghaus, 1913; Müller and Pilzecker, 1900; Jost, 1897; Müller and Schumann, 1894; Ladd and Woodworth, 1911; Carr, 1931). Because memories of recent experiences readily come to mind, early scholars sought to uncover the factors that lead to forgetting. Their experiments quickly challenged the folk assertion that memories decay over time, eventually becoming completely erased. Rather, they found that removing a source of interference, or reinstating the "context" of original learning, readily restored these seemingly forgotten memories (McGeoch, 1932; Underwood, 1948; Estes, 1955). In these early papers, context represents the set of latent (or background) information not specifically related to the present stimulus. Such contextual information could include internal states of the agent, such as thoughts, emotions, goals, and concerns that form the cognitive milieu in which new learning occurs (Kahana, 2012).

According to modern memory theories, the set of psychological (or neural) features that represent a stimulus enter into association with a mental representation of context, and the database of such associations form the basis for performance in subsequent recall, recognition, and categorization tasks (Howard and Kahana, 2002).

In this paper, we use a dynamic model of contextual coding based closely on that developed by Howard and Kahana (2002) to account for data on the dynamics of memory search. The experimenter presents the subject with a list of items, often words, to remember. We consider this as analogous to the data in which nature presents the agent with features of the environment. Howard and Kahana (2002) model these features of the environment as basis vectors in an n-dimensional space f_t. The key innovation of the Howard and Kahana model over prior models is the existence of internal mental context that links these features through time (below, we discuss specific evidence in support of this hypothesis). Following this and subsequent papers we model context as a norm-1 vector of length m that evolves based on past context and current features:

$$c_t = (1-\zeta)c_{t-1} + \zeta c_t^{\text{in}}, \tag{1}$$

where c^{in} (the "in" stands for input) is the aspect of context that arises (is retrieved) from the current environment, and where $\zeta \in [0,1]$. Specifically, the agent forms a record of associations between features and the internal context. This record is stored in memory in an $m \times n$ matrix $W^{f\to c}$. Then

$$c_t^{\text{in}} = \frac{W_{t-1}^{f\to c} f_t}{||W_{t-1}^{f\to c} f_t||}, \tag{2}$$

where the denominator in (2) implies that c_i^{in} is scaled so that its length equals one for a given norm $||\cdot||$.[2]

The matrix $W^{f\to c}$ is called the "features-to-context matrix" for reasons that (2) makes apparent. Following the memory literature, e.g. Lohnas et al. (2015), we close the model by initializing $W_0^{f\to c}$ using long-standing associations (semantic memory) and recursively defining

$$W_t^{f\to c} = W_{t-1}^{f\to c} + c_t f_t^\top, \tag{3}$$

where $c_t f_t^\top$ is an outer product. In what follows, we simplify notation by assuming that time has run for sufficiently long so that this initial matrix is not relevant, and that all associations take the form of $c_t f_t^\top$. This is equivalent to assuming that $W_0^{f\to c}$

[2] In what follows, we depart from the memory literature in using the L^1-norm. Namely context is scaled by the sum of the absolute values of its elements. The memory literature, e.g. Polyn et al. (2009), typically uses the L^2-norm, with $c_t = \rho_t c_{t-1} + \zeta c_t^{\text{in}}$ and $\rho_t \approx 1-\zeta$, to maintain c_t on the unit circle. Because elements of context sum to 1, (1) is exact.

is formed by viewing a prior sample.

The model is associative in the sense that "cueing" with an item (a feature) recovers the contexts previously associated with that item:

$$\begin{aligned} c_t^{\text{in}} &= \frac{W_{t-1}^{f\to c} f_t}{||W_{t-1}^{f\to c} f_t||} \\ &\propto \sum_{s=0}^{t} (c_s f_s^\top) f_t \\ &= \sum_{s=0}^{t} c_s (f_s^\top f_t). \end{aligned}$$

Note that for orthonormal basis vectors, $f_s^\top f_t$ is either zero or 1, depending on $f_s = f_t$ or $f_s \neq f_t$. If the agent experiences a feature under multiple contexts, then the feature recalls a weighted average of the contexts. In a more general model, features need not be basis vectors, and the interpretation still holds.

In retrieved context theory, context determines what an agent is most likely to remember. Figure 2 illustrates the mechanism. The current state of context contains both an autocorrelated component that overlaps with the contexts of recent experiences, and a retrieved context component that overlaps with items experienced close in time to the just-recalled item(s). The figure illustrates these two effects as spotlights shining down on memories arrayed on the stage of life. Memories are not truly forgotten, but just obscured when they fall outside of the spotlights.

The theory of Howard and Kahana (2002) and in numerous follow-up papers treats memory as an outcome of a mechanistic process. There is no explicit decision-maker facing an objective function, nor is there an underlying probability space on which the agent forms beliefs. To map the above framework into financial decisions, we assume a link between memory and the subjective probability the agent assigns to future events. The equations themselves suggest such a link. For example, (2) reflects storage of co-occurrences of features with contexts, while (1) reflects (partial) updating of beliefs based on new information.

To define behavior, all we require is an agent with an objective function and constraints. However, to map this language into the more standard language of decision-making under uncertainty, as well as to provide a role for empirical evidence uncovered by an econometrician, it is useful to define an underlying economic environment.

We use the standard economic setting in models of macroeconomics and finance. There is a persistent underlying state of nature, represented by a random variable $\tilde{X}$. For simplicity, we assume throughout that this random variable takes on finitely many states, $\{x_1, \ldots, x_m\}$. We allow the state to be persistent over time, and define a probability of transitioning from one state to another (the state is Markov in the sense that this probability depends only on the current state). Let X_t be the value $\tilde{X}$ takes on at time t. Then we can define notation for these transition probabilities as $p^X_{ij} \equiv \text{Prob}(X_{t+1} = x_j \,|\, X_t = x_i)$. An important special case will be one in which $\tilde{X}$ is independently and identically distributed (iid) across time, in which case we use the notation $p^X_i = \text{Prob}(\tilde{X} = x_i)$.

As in much of the recent literature, the agent may not perfectly observe outcomes $\tilde{X}$. Rather, the agent observes a related random variable $\tilde{Y}$ that takes on finitely many values $\{y_1, \ldots, y_m\}$. Nature provides a mapping from $\tilde{X}$ to $\tilde{Y}$: let $p^Y(y_k|x)$ denote the probability that $Y_t = y_k$ conditional on $\tilde{X} = x$ for $k = 1, \ldots, n$. This set-up is analogous to that used in the rational inattention framework of Sims (2003). Assuming the distributions p^X and p^Y were known, inference on $\tilde{X}$ conditional on $\tilde{Y}$-observations forms a discrete-time Kalman filter. The more realistic case is one where the agents must also infer these distributions.

The following is our benchmark definition of a features vector:

Definition. *Given a time t, features f_t is an n-dimensional vector such that*

$$f_t(j) = \begin{cases} 1 & \textit{if} \quad Y_t = y_j \\ 0 & \textit{otherwise} \end{cases} \tag{4}$$

That is, $f_t = e_j$, the jth basis vector if the jth outcome of the random variable $\tilde{Y}$ obtains. Then c_t and $W_t^{f \to c}$ are defined as in (1) and (3) respectively.

Here, and elsewhere in the paper, we use $W^{f \to c}(i, j)$ to denote the ijth element of the $W^{f \to c}$ matrix, $c(i)$ to denote the ith element of a context vector, and $f(j)$ the jth element of a features vector.

Assumption 1. *The agent has an initial record of associations $W_0^{f \to c}$, such that $W_0^{f \to c}(i, j)$ represents the agent's prior belief regarding the joint occurrence of $(\tilde{X}, \tilde{Y}) = (x_i, y_j)$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$. The context vector $c_0(i)$ represents the probability that $X_0 = x_i$.*

Given initial conditions as in Assumption 1 and a time series of features, context then evolves according to (1), (2), and (3). That is, the time series of features together with the initial conditions, fully define context. As discussed previously, our notational convention is that time has run for sufficiently long that prior beliefs can be captured as associations $c_t f_t^\top$, so we can ignore the effects of $W_0^{f \to c}$.

The next assumptions describe how associations and context enters into decision-making.

Assumption 2. *At time t, the agent assigns the unconditional probability of $(\tilde{X}, \tilde{Y}) = (x_i, y_j)$ to be proportional to $W_{t-1}^{f \to c}(i, j)$.*[3]

Assumption 3. *At time t, the agent assigns the conditional probability of $\tilde{X} = x_i$ using the elements of the context vector. That is, $c_t(i)$ is the probability of state i*

[3]The constant of proportionality is such that the entries of $W_t^{f \to c}$ sum to 1.

at time t conditional on features observed up to time t, and on the agent's record of associations $W_{t-1}^{f\to c}$.

As an initial sanity check on these assumptions, we consider their implications for a simpler alternative world in which both X and Y are observed, and context is directly analogous to features.

Theorem 1. *Assume features are defined by (4). Assume context c_t is such that $c_t(i) = 1$ if $X_t = x_i$ and 0 otherwise. Consider a Bayesian agent with an uninformative prior (so that all elements of $W_0^{f\to c}$ are equal). Then the elements of $W_t^{f\to c}$ are proportional to the means of the posterior probabilities of each state.*

Intuitively, when the agent observes X_t and Y_t then (3) implies that $W_t^{f\to c}$ "counts" the joint contingencies. Thus there exists a simple rational expectations benchmark in which the memory matrix $W_t^{f\to c}$ yields the correct Bayesian inference. This a remarkable property given that the theory was developed with no notion of decision-making under uncertainty, and entirely as a means of matching data in memory experiments.

Analogously, (2) implies, in this simple alternative world, that context indeed reflects the conditional distribution.

Theorem 2. *Assume $W_{t-1}^{f\to c}$ is proportional to the agent's beliefs about the joint states of $(\tilde{X}, \tilde{Y})$. Then retrieved context (2) is the conditional probability of X_t given Y_t.*

Appendix A gives the proofs. To summarize: these theorems motivate the notion of the $W^{f\to c}$ matrix and of the context vector as probabilities. Clearly the scenario outlined above can only serve as a starting point: it assumes symmetric contexts and features and leaves no role for context dynamics (1). It is nonetheless remarkable that there is a straightforward mapping between the memory model, developed independently of any notion of uncertainty or optimization and Bayesian decision-making.

In what follows, we take as given a time series f_t, an initial condition on context, and on the features-to-context matrix $W_0^{f \to c}$. Context then evolves fully endogenously based on (1). That is, features are observed and context is endogenous.

Assumptions 2 and 3 imply that inference, in general, will not be Bayesian for several reasons. First, a direct implication is that the agent consults his or her memory when evaluating probabilities. If the agent has experienced a biased sample, beliefs, too, will be biased. Second, the agent "views" the internal mental state is a representation of an objective, though latent, external state. To the extent, then that context evolution does not match with Bayesian updating regarding a latent state, conditional probabilities will also be biased. The link between an external and an internal state is potentially controversial. However, that such internal states exist is a fundamental tenant of modern theories of memory, with extensive support in the data. The existence of an external latent state is not obvious other than in a hypothesized Savage (1954)-world. Should it exist, it is not clear what it is *other* than an internal mental representation. Finally, as we describe below, we assume that the agent can be manipulated in that features need not be generated by nature; they can be constructed to resemble nature (unlike context, they are nonetheless external).

We now consider general properties of the model. On one level, (1) implies that context is persistent, but eventually decays. It is a weighted average (with exponentially decaying weights) of previous contexts retrieved from experiences. On another level, the retrieval process itself is endogenous, and affected by experience, as (3) shows.

We focus first on a set of results for which the persistence and eventual decay of context is paramount. These results, based on the notion of beliefs as arising from recent experience, are reminiscent of the literature on extrapolative expectations. See Appendix A for formal definitions of terms and proofs.

Theorem 3 (Neglected risk). *Suppose that the agent has not observed any features associated with state $\tilde{X} = x_i$ for a sufficiently long amount of time. Then the agent's* *probability that $\tilde{X} = x_i$ becomes arbitrarily small. That is, the agent forgets (temporarily) that state $\tilde{X} = x_i$ could occur.*

Theorem 4 (Short-run momentum; long-term reversal)**.** *Consider an outcome $\tilde{X} = x_i$ such that a subset of outcomes of $\tilde{Y}$ can only occur in that state.*[4] *Assume this is reflected in the agents' associations, so that x_i is associated only with features vectors that can occur under state i. Assume that x_i is sufficiently rare and persistent.*[5] *Then*

1. *If the features are novel, beliefs exhibit momentum. That is, positive changes in beliefs are followed by still more positive changes:*

$$E[c_{t+1}(i) - c_t(i)|c_t(i) > c_{t-1}(i)] > E[c_{t+1}(i) - c_t(i)|c_t(i) \leq c_{t-1}(i)]$$

2. *If the features have been observed for a sufficiently long amount of time, then beliefs reverse: $E[c_{t+1}(i) - c_t(i)] < 0$.*

These theorems, focusing on exponential decay of context, shows how the model generates effects that resemble extrapolative expectations. It implies behavior as if the agent has forgotten elements in the past.

The following theorems focus on aspects of the model that go beyond exponential decay of recent experience. Specifically, the model generates predictions for how agents form associations, and how these associations affect decision-making. To focus on the associative mechanism, we temporarily turn off the exponential decay mechanism (setting $\zeta = 1$), and assume that $\tilde{X}$ is iid, just as Theorems 3 and 4 simplified the associative mechanism, by assuming associations that took a simple form. First, if a context i and features k become too greatly associated in the mind of an agent, then over-reaction can result.

[4]Formally: $y_k \in \Omega_i$ if $p^Y(y_k|x_i) > 0$ and $p^Y(y_k|x_j) = 0$.
[5]That is, $p^X_{ii} - \zeta > p_{ij}$ for all j.

Theorem 5 (Over-reaction)**.** *Assume $\tilde{X}$ is iid and $\zeta = 1$. Suppose that there is some $i \in \{1, \ldots, n\}$ such that*

$$\frac{W_{t-1}^{f \to c}(i,k)}{W_{t-1}^{f \to c}(j,k)} > \frac{p_i^X p^Y(y_k \mid x_i)}{p_j^X p^Y(y_k \mid x_j)}, \quad j \neq i,$$

Then retrieved context in response to features $f_t = e_k$ will overstate the conditional probability $\text{Prob}(X_t = x_i | Y_t = y_k)$.

Of course, this theorem is mainly interesting under circumstances in which $W^{f \to c}(i,k)$ may become over-stated relative to $W^{f \to c}(j,k)$. One possibility is that features and context become over-associated through outsized emphasis in the media or in an agent's particular experience. However, the present theory, combined with repetitive retrieval of features from context, can endogenously generate this result (see Appendix D).

A related implication of the model is that it matters how similar events appear, even if the underlying data generating process could be different.

Theorem 6 (Similarity)**.** *Consider the case in which the agent experiences features f_j such that $f_j^\top f_t \approx 1$. That is, features f_j closely resemble, but are not exactly the same as f_t. Then conditional probabilities at time t are as if f_t occurs. That is, f_j "reminds" the agent of f_t and of the associated context.*

Finally, the model generates the surprising implication that beliefs can remain stable over time.

Theorem 7. *Assume $\zeta = 1$. Contextual retrieval remains stable, assuming features are basis vectors. That is, suppose $s > t$, and $f_s = f_t$. Then $c_s = c_t$, despite the fact that $W_s^{f \to c} \neq W_t^{f \to c}$.*

It is of interest to contrast Theorem 7 and Theorem 1. In the case of Theorem 1, both $\tilde{X}$ and $\tilde{Y}$ are observed. The agent forms a record of associations using (3) that matches the data. Inference is the same as for a Bayesian agent who views $\tilde{X}$ and $\tilde{Y}$, and updates beliefs on the density functions p^X and p^Y.

In the case of Theorem 7, the agent also forms a record of associations using (3). Here, the record of associations is different than in the case where $\tilde{X}$ is observed. The agent associates features f_t with retrieved context; in other words the agent behaves as if the imagined context has been observed. In effect, the agent forms associations by thinking.

The consequences for behavior are significant. In the case of Theorem 1, the effect of prior beliefs quickly disappears. On the other hand, in Theorem 7, prior beliefs can persist indefinitely, because they become self-reinforcing.

3 Psychological and Neural Basis for Contextual Retrieval

Before turning to the application of the memory model to financial decision-making, we briefly summarize the psychological and neural evidence for context as an internal state.

In the memory laboratory, researchers create experiences by presenting subjects with lists of easily identifiable items, such as common words or recognizable pictures. Subjects attempt to remember the previously experienced items under varying retrieval conditions. These conditions include *free recall*, in which subjects recall as many items as they can in any order, *cued recall*, in which subjects attempt to recall a particular target item in response to a cue, and *recognition* in which subjects judge whether or not they encountered a test item on a study list. In each of these experimental paradigms, memory obeys the classic "Laws of Association" which appear first in the work of Aristotle, and later in Hume (1748). The first of these is *recency*: human subjects exhibit better memory for recent experiences, *semantic similarity*: we remember experiences that are most similar in meaning to those we are currently experiencing, and finally, *temporal contiguity*: we remember items that occurred contiguously in time to recently-recalled items. Although quantified in the memory laboratory, each of these phenomena appears robustly in real-world settings, as described below.

A longstanding and persistently active research agenda in experimental psychology seeks to uncover the cognitive and neural mechanisms that could give rise to these regularities. Experimental psychologists have proposed many hypotheses and have (accordingly) refined the tasks above in a number of ways. Some striking findings include the fact that recency and contiguity have similar magnitudes at short and long time scales.[6]

Several classic explanations, though successful in many ways, struggled to explain this scale invariance. One highly influential class of explanations posits the existence of a specialized retrieval process for recently-experienced items (short-term memory).[7] A related idea is that associations chain together in the mind of the subject.[8] In con-

[6]To measure the effect of contiguity on memory retrieval, researchers examine subjects' tendency to successively recall items experienced in proximate list positions. In free recall, this tendency appears as decreasing probability of successively recalling items f_i and f_{i+lag} as a function of lag, conditional on the availability of that transition (Kahana, 1996). This function reaches its maximum at $lag = \pm 1$, but also exhibits a forward asymmetry in the form of higher probability for positive as compared with negative lags. Equations 1–3 generate a forward asymmetry in the contiguity effect because recalling an item reinstates both its associated study-list context and its associated pre-experimental context. Whereas the study-list context became associated, symmetrically, with both prior and subsequent list items, the pre-experimental context became associated only with subsequently encoded list items, leading to a forward asymmetric contiguity effect, as seen in the data.

[7]The view that recency arises from specialized retrieval processes associated with short-term memory rose to prominence in the 1960s. According to these dual-store models, separate short-term and long-term memory stores support retrieval of information experienced at short and long time scales, with short-term (or "working") memory holding a small number of information units through an active rehearsal process and supporting the rapid and accurate retrieval and manipulation of that information. According to these models, retrieval from long-term memory involved a search process guided by interitem associations and context-to-item associations, and subject to interference from similar memories (Kahana, 2012).

[8]Continental philosophers saw contiguity as the result of chained associations (Herbart, 1834) that could be easily disrupted by interfering mental activity (Thorndike, 1932). This idea took form in cognitive models that conceived of associations as being forged in a limited-capacity short-term memory store (Atkinson and Shiffrin, 1968; Raaijmakers and Shiffrin, 1980), perhaps arising as the

trast, retrieved context theory does not derive contiguity and similarity through direct interitem associations. Rather, they arise because the contextual information retrieved during the recall of an item overlaps with the contextual information associated with similar and neighboring items. Underlying context naturally generates scale invariance.

Figure 1 summarizes some of the evidence supportive of context retrieval. Figure 1A shows that interitem distraction and test delay do not disrupt the temporal contiguity effect (TCE) seen in the relation between transition probability and lag, which speaks directly to time-scale invariance. Figure 1B-D shows that the TCE appears robustly for both younger and older adults, for subjects of varying intellectual ability, and for both naïve and highly practiced subjects. Figure 1E shows that the TCE appears even for transitions between items studied on completely distinct lists, despite these items being separated by many other item presentations. Figure 1F-H shows that the TCE also predicts confusions between different study pairs in a cued recall task, in errors made during probed recall of serial lists, and in tasks that do not depend on inter-item associations at all, such as picture recognition (see caption for details). Finally, long-range contiguity appears in many real-life memory tasks, such as recalling autobiographical memories(Moreton and Ward, 2010) and remembering news events (Uitvlugt and Healey, 2019).

A second source of data in favor of retrieved context arises from neurobiology. At a neurobiological level, this implies that the brain states representing the context of an original experience reactivate or replay during the subsequent remembering of that experience. Several studies tested this idea using neural recordings. These studies found that in free recall (Manning et al., 2011a), cued recall (Yaffe et al., 2014) and recognition memory (Howard et al., 2012a; Folkerts et al., 2018) brain activity during memory retrieval resembles not only the activity of the original studied item, but also the brain states associated with neighboring items in the study list. Thus, one

result of imagery or linguistic mediation (Murdock, 1974).

observes contiguity both at the behavioral and at the neural level, with these effects being strongly correlated (Manning et al., 2011a). Finally, this recursive nature of the contextual retrieval process offers a unified account of many other psychological phenomena including the spacing effect (Lohnas and Kahana, 2014b), the compound cueing effect (Lohnas and Kahana, 2014a), and the phenomena of memory consolidation and reconsolidation (Sederberg et al., 2011).

Memory theory thus indicates that remembering an item involves a jump-back-in-time to the state of mind that obtained when the item was previously experienced. This neural reinstatement, in turn, becomes re-encoded with the new experience and also persists to flavor the encoding of subsequently experienced items. The persistence of the previously retrieved contextual states enables memory to carry the distant past into the future, allowing the contextual states associated with an old memory to re-enter one's life following a salient cue and associate with subsequent "neutral" memories. While the original memory is retained in association with its encoding context, the retrieval and re-experiencing of that memory forms a new memory in association with the mixture of the prior and retrieved context. Memory theory thus also predicts that multiple recalls of an item will largely appear to the agent as if there were multiple experiences, when in fact there was only perhaps a single experience (this is also what the data show: see Rubin and Berntsen (2009) and Rubin et al. (2008)). The well-known existence of post-traumatic stress disorder attests to the power of continually recurring danger that is wholly in the mind of an agent.

4 Applications

4.1 Retrieved-context theory and the persistence of beliefs

A classic problem in asset allocation is that of an investor allocating wealth between a risky asset (with unknown return) and a riskless asset with known return (Arrow, 1971; Pratt, 1976). This deceptively simple problem is the subject of a large and sophisticated literature (Wachter, 2010). In a new take on this classic problem, Malmendier and Nagel (2011) report an intriguing pattern in the portfolio choice of investors in the Survey of Consumer Finances. Investors whose lifetime experience includes periods with lower stock returns invested a lower percentage of their wealth in stocks as compared with investors whose lifetime experience includes periods with higher returns. While, on one level this may seem intuitive, it is a puzzle from the point of view of standard asset allocation theories. For one thing, experience should not matter, only objective data on returns. For another, even if investors over-weight their own experience, and under-weight returns outside of their experience, investors in the sample had experiences of sufficiently long length (and the return distribution is sufficiently ergodic) such that their beliefs should quickly converge.

Here, we abstract from many interesting features of the Malmendier and Nagel (2011) study. For example, investors exhibit a recency effect (their portfolio choice depends more on recent observations than on past observations) which we do not emphasize here, but which is very much in the spirit of a memory model. We focus on a qualitative implication of their results, namely, that personal experiences can continue to influence investors' beliefs, even though there are sufficient data (if investors were Bayesian) to over-ride a specific time path of experience. Thus, in this section, we focus simply on the question of persistence of beliefs, abstracting both from many features of memory models, and many features of the portfolio choice data. We do so to highlight the novel implications of the theory.

4.1.1 The portfolio choice problem

We consider the problem of an investor choosing to allocate wealth between a risky stock, with net return $\tilde{r}$, and a riskfree bond. For simplicity, we assume the bond has a net return of zero. The agent also receives risky labor income $\tilde{\ell}$. We assume the agent begins with wealth of one. In the language of Section 2, stock market gains and losses correspond to $\tilde{Y}$ and labor market states (which are iid) correspond to $\tilde{X}$. The agent prefers more wealth to less, and is risk averse. We tractably capture these preferences by assuming utility is an increasing function of the mean of wealth and a decreasing function of the variance (Markowitz, 1952):

$$\max_{\pi} E[(1+\pi\tilde{r}+\tilde{\ell})] - \frac{1}{2}\mathrm{Var}(1+\pi\tilde{r}+\tilde{\ell}) \tag{5}$$

where π is the percent allocation to the risky asset. The expectation and the variance in (5) are with respect to the agents' subjective preferences. Setting the derivative of the objective function with respect to π equal to zero leads to

$$\pi = \frac{E\tilde{r} - \mathrm{Cov}(\tilde{r},\tilde{\ell})}{\mathrm{Var}(\tilde{r})}. \tag{6}$$

If stocks deliver a low return in a negative labor income state, that makes them unattractive.

Assume that the risky return takes on two possible values $r(gain) > r(loss)$. Assume labor income $\tilde{\ell}$ takes on two possible values $\ell(normal) > \ell(depression)$. We consider beliefs that take the following form: gain and loss states each occur with probability 1/2, that a gain and a depression cannot co-occur, and that a depression has (unconditional) probability p, for $p \leq 1/2$. The following matrix captures the state space and the probabilities:

$$P = \begin{bmatrix} \text{Prob}(\mathit{gain}\ \&\ \mathit{normal}) & \text{Prob}(\mathit{loss}\ \&\ \mathit{normal}) \\ \text{Prob}(\mathit{gain}\ \&\ \mathit{depression}) & \text{Prob}(\mathit{loss}\ \&\ \mathit{depression}) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} - p \\ 0 & p \end{bmatrix},$$

where $p \in [0, \frac{1}{2}]$.

We assume $\tilde{r}$ has mean 1, which implies $\tilde{r}(\mathit{gain}) = 1 + \sigma$, $\tilde{r}(\mathit{loss}) = 1 - \sigma$, where σ is the standard deviation of $\tilde{r}$. Let $\tilde{\ell}(\mathit{normal}) = \ell > 0$ and $\tilde{\ell}(\mathit{depression}) = 0$. Then. $E\tilde{r} = 1$, $\text{Var}(\tilde{r}) = \sigma^2$. Note that $\tilde{\ell}$ is a Bernoulli random variable multiplied by a constant y, so that:

$$\begin{aligned} E\tilde{\ell} &= (1-p)\ell \\ \text{Var}(\tilde{\ell}) &= p(1-p)\ell^2. \end{aligned}$$

Direct calculation implies

$$\text{Cov}(\tilde{r}, \tilde{\ell}) = E\big[(\tilde{r} - E\tilde{r})\tilde{\ell}\big] = \frac{1}{2}\sigma\ell - \left(\frac{1}{2} - p\right)\sigma\ell = p\sigma\ell.$$

Then the optimal allocation (6) equals

$$\pi(p) = \frac{1 - p\sigma\ell}{\sigma^2}. \tag{7}$$

The greater the probability that the agent assigns to the depression, the less he or she allocates to the risky asset. A useful benchmark is Theorem 1. When the agent observes labor income and stock return realizations, the agent updates her estimates of p. We discuss this Bayesian benchmark in what follows.

4.1.2 Memory for stock market gains and losses

In contrast to the benchmark in which the agent observes $\tilde{r}$ and $\tilde{\ell}$, we assume the agent observes only $\tilde{r}$ and retrieves observations on $\tilde{\ell}$ from memory. Thus, in the language of Section 2, $f_t = [1, 0]^\top$ represents stock market gain, and $f_t = [0, 1]^\top$ represents loss. The agent has a prior record of associations such that

$$W_t^{f \to c} \propto P^* \equiv \begin{bmatrix} \frac{1}{2} & \frac{1}{2} - p^* \\ 0 & p^* \end{bmatrix}. \tag{8}$$

The special case of $P^* = P$ implies that the agent's associations correspond to the correct joint population frequencies of labor market and stock market states. Instead, we assume the agent may have seen a biased sample. We focus on the case of $p^* > p$, so that a depression is over-represented.

We apply the setting of Theorem 7, which implies that $\zeta = 1$, and we have the following recursion for context:

$$c_{t+1} \propto W_t^{f \to c} f_{t+1}. \tag{9}$$

A stock market gain retrieves 100% probability on the normal labor income state:

$$c_{t+1} \propto W_t^{f \to c} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \propto \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

A stock market loss, on the other hand, retrieves a positive probability of a depression, *even if one has not occurred*:

$$c_{t+1} \propto W_t^{f \to c} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - p^* \\ p^* \end{bmatrix} \propto \begin{bmatrix} 1 - 2p^* \\ 2p^* \end{bmatrix}$$

That is, the agent recalls the depression. The key difference between this model and the Bayesian benchmark (with observed $\tilde{\ell}$) described in Theorem 1 is that this act of recollecting implies that there is a new "depression" observation in the agent's mental database.

Consider then what happens to the $W^{f\to c}$ matrix:

$$W_{t+1}^{f\to c} = W_t^{f\to c} + c_{t+1} f_{t+1}^{\top},$$

where

$$c_{t+1} f_{t+1}^{\top} = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1,0] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } gain \\ \begin{bmatrix} 1-2p^* \\ 2p^* \end{bmatrix} [0,1] = \begin{bmatrix} 0 & 1-2p^* \\ 0 & 2p^* \end{bmatrix} & \text{if } loss \end{cases} \tag{10}$$

Regardless of whether a gain or loss occurs, the columns of $W_{t+1}^{f\to c}$ relate to those of $W_t^{f\to c}$ by a constant of proportionality. Thus, at time 2, a stock market gain retrieves $[1,0]^{\top}$, whereas a stock market loss retrieves $[1-2p^*, 2p^*]^{\top}$. The agent's probability distribution is the same as before.

Applying Theorem 7, it follows, that after τ periods of which k are gains:

$$W_{t+\tau}^{f\to c} \propto \begin{bmatrix} \frac{1}{2}\frac{t}{\tau+t} + \frac{k}{\tau+t} & (\frac{1}{2} - p^*)\frac{t}{\tau+t} + (1-2p^*)\frac{\tau-k}{\tau+t} \\ 0 & p^* \frac{t}{\tau+t} + 2p^* \frac{\tau-k}{\tau+t} \end{bmatrix}. \tag{11}$$

It follows that, in the limit at τ approaches infinity, $W_{t+\tau}^{f\to c}$ approaches P^*:[9]

$$\text{plim}_{\tau\to\infty} \frac{1}{t+\tau} W_{t+\tau}^{f\to c} = P^*.$$

It does not matter how much data the agent observes: probabilities remain distorted.

[9]Note that $k/(\tau+t) \to 1/2$, $(\tau-k)/(\tau+t) \to 1/2$, and $t/(\tau+t) \to 0$.

Why, intuitively, does the agent fail to update his or her probabilities? The reason is that the agent's memory over-associates a stock market loss with a depression. The appearance of a stock market loss, then reinstates the depression context. This act of recalling the depression context is similar to the experiencing the depression. Thus, a high probability of depression remains associated with losses in the mind of the agent. Interestingly, if the agent happened to arrive at the correct probabilities at the beginning, the updating rule (9) would have produced the correct probabilities P.

Figure 3 contrasts implications for three types of agents: the agent who knows the true probability, the agent who starts with an incorrect prior and learns the true probability according to Bayesian updating, and the agent who starts with the same incorrect prior and whose learning is subject to context retrieval. For the purposes of the figure, $p = 0.02$, $p^* = 0.50$, $\sigma = 1$, and $\ell = 2$.[10] Figure 3 shows the mean of the posterior distribution for p. The Bayesian agent's beliefs converge quickly to something close to the truth. Thus, while precise convergence to a 2% probability of Depression takes many years, updating is very fast for values of the probability that are far from the truth. Twenty years of data suffice to bring the probability sufficiently close so that the resulting portfolio allocation is virtually indistinguishable from that of the full-information agent. On the other hand, the agent who relies on memory does not learn, and can maintain an incorrect probability even in the face of many years of evidence.[11] Memory itself produces a distorted database because the agent relives his worst fears when a stock market downturn occurs.

[10] The Bayesian investor has an uninformative prior. Given the likelihood implied by Bernoulli observations on $\tilde{\ell}$, the posterior Beta (see Appendix B). We report the mean of this distribution, which is all that is required to (6), since the covariance is linear in the depression probability. Note that the Bayesian agent who infers the correct probability thus behaves the same as the agent who is certain about the probability; we abstract from the effect of parameter uncertainty.

[11] Recent survey evidence (Goetzmann et al., 2017) indicates irrationally high levels of fear of stock market crashes, and that exogenous events can trigger such fears. The latter point is specifically addressed in the model below.

4.2 Context and the jump back in time: Application to the financial crisis

The failure of Lehman Brothers is widely recognized as a point of inflection in the 2008 financial crisis.[12]

An open question is: why was the failure of Lehman Brothers so pivotal? A growing line of research answers this question by focusing on the importance of financial intermediation to the overall the economy. Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2013) develop models in which the balance sheets of intermediaries contribute to business cycle fluctuations. However, while it may be necessary to have specialized institutions trade certain complicated investments, it is not clear why the failure of a financial institution should be followed by a broad-based stock market decline. Common stocks are not intermediated assets: trading costs for common stocks, already quite low for the past half-century, have only gotten lower (Jones, 2002). Another possibility is that Lehman represented a sunspot that caused a run on other intermediaries, and other forms of debt (Allen and Gale, 2009; Gorton and Metrick, 2012). Unanswered is why this should cause the stock market to crash, as it did in the fall of 2008, when most companies have very low leverage and can fund themselves through retained earnings?[13].

A third possibility, which Gennaioli and Shleifer (2018) emphasize, is that individuals and banks took on too much debt because they incorrectly extrapolated that the environment was riskfree. This debt created unstable conditions. The Lehman bankruptcy reminded agents of the risk that they faced. This possibility is most in the spirit of the model here. However, while related, the two explanations are distinct. In our model, the risk is illusory, whereas is the neglected risk hypothesis (as usually

[12]See, for example, French et al. (2010).

[13]Kahle and Stulz (2013) argue that firms dependent on bank-lending were not unduly affected by the crisis. Gomes et al. (2019) argue that fluctuations in borrowing conditions are more likely to be affected by investment opportunities than the other way around.

formulated), the risk caused by excessive debt is real. While our view may be extreme, in fact, the Great Recession was nothing like the Great Depression.[14] The realized outcome does not seem commensurate with the panic in the fall of 2008.[15]

Our hypothesis is that the financial crisis was a psychological event caused by the failure of Lehman Brothers. The actual realization of a important financial institution failing in the absence of insurance reminded investors of the Great Depression.[16] Some felt that they had – literally – returned to the Great Depression. Investors experienced what the memory literature refers to as a *jump back in time* (Manning et al., 2011b; Howard et al., 2012b). Once this feeling entered the discourse, it proved hard to shake. Subsequent events showed that in fact there was no Great Depression. This was only revealed, though, over time. Somehow, what emerged from the crisis and recession was not a feeling of relief but rather a renewed emphasis on the fragility of the financial sector and the possibility that a Great Depression might in fact occur. The model below formalizes this intuition.

4.2.1 Asset prices

We consider an economy in which there is a single, representative agent. This agent faces a consumption and investment choice. Following Lucas (1978), we assume an endowment economy, in which there is no technology for moving resources across periods. Thus prices equilibrate to make consumption optimal.

In this economy, stock prices are expected discounted values of future cash flows. When cash flows occur at times that are risky for the agent, they receive a higher

[14]The effect of the financial crisis on aggregate consumption was relatively minor: from the start of 2008 to the end of 2009, aggregate consumption fell by 3%, and consumption began to recover by 2010. In contrast, consumption fell by 16% in the Great Depression.

[15]This outcome was endogenous to the policy response, which may have prevented further declines. Note however that policy makers may be subject to the same context dynamics discussed here; they may also be responsive to stock market outcomes. The resultant multiple equilibria stemming from belief dynamics are beyond the scope of this article.

[16]See, for example, the reporting of *The Guardian* on the day's events: https://www.theguardian.com/business/2008/sep/15/marketturmoil.stockmarkets.

discount rate: namely a premium. Thus, if the economy suddenly becomes riskier, stock prices may suddenly fall, even if very little has changed in terms of observable cash flows. The disaster risk framework (Tsai and Wachter, 2015) offers a way to think about how prices can change suddenly even if observables do not.

Specifically, the agent faces a consumption process that has normal risk, and rare-event risk. The rare events occur with probability approximately equal to p, which is small (the use of exponentials below implies convenient analytical expressions as the time interval shrinks). We use the model of Barro (2006) for the consumption process.

$$\log C_{t+1} = \log C_t + \mu + u_{t+1} + v_{t+1}, \tag{12}$$

where u_{t+1} and v_{t+1} are independent, $u_{t+1} \sim N(0, \sigma^2)$ and

$$v_{t+1} = \begin{cases} 0 & \text{with prob. } e^{-p} \\ \log(1-b) & \text{with prob. } 1 - e^{-p} \end{cases} \tag{13}$$

where b is a random variable with support on $[0, 1)$. The aggregate market is a claim to cash flows satisfying

$$\log D_{t+1} = \log D_t + \mu + u_{t+1} + \lambda v_{t+1} \tag{14}$$

with $\lambda > 1$. This assumption captures the fact that dividends fall by more than consumption during financial disasters (Longstaff and Piazzesi, 2004).[17]

[17]The specification (13–14) implies dividends and consumption are perfectly conditionally correlated. It also implies that dividends have equal volatility to consumption in normal times. In the data, dividends have greater normal-times volatility than consumption, and they are imperfectly correlated with consumption. Both facts could be introduced into the model by adding independent shocks to (14). Because these shocks are unpriced, they would have a negligible effect on the results of interest.

We assume that at every period, the agent maximizes utility

$$E_t \sum_{s=1}^{\infty} \beta^s \log C_s$$

Let S_t equal the value of the aggregate stock market, namely the claim to cash flows (14). The first-order conditions of the agent imply

$$S_t = E_t \left[M_{t+1}(S_{t+1} + D_{t+1}) \right], \tag{15}$$

where the intertemporal marginal rate of substitution equals

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-1}.$$

Equilibrium requires that optimal consumption satisfy (12) and (13) and that cash flows equal (14). Asset prices adjust to satisfy the first-order conditions for the representative agent.

In Appendix C, we show (15) has solution

$$S_t = D_t \sum_{n=1}^{\infty} \Phi(p)^n = \frac{\Phi(p)}{1 - \Phi(p)} \tag{16}$$

for

$$\Phi(p) = \beta \left(e^{-p} + (1 - e^{-p}) E \left[(1-b)^{\lambda - \gamma} \right] \right) \tag{17}$$

When $\lambda > 1$, an increase in p (in a comparative statics sense), lowers the price.

The riskfree rate also solves first-order condition, and equals

$$\begin{aligned} 1 + r_f &= E \left[M_{t+1} \right]^{-1} \\ &= \beta^{-1} e^{\mu - \frac{1}{2}\sigma^2} \left(e^{-p} + (1 - e^{-p}) E \left[\frac{1}{1-b} \right] \right)^{-1}, \end{aligned}$$

From (18), we can see that an increase in the disaster probability p lowers the riskfree rate. This is intuitive: an increase in the disaster probability leads the investor to want to save to protect against the disaster realization. Bond prices rise, and riskfree rates fall.

4.2.2 Memory for rare events

We identify features with the state of the financial system, so $f_t = [1, 0]^\top$ represents normal times and $f_t = [0, 1]^\top$ represents a crisis. In the language of Section 2, $\tilde{Y}$ is the state of the financial system, whereas $\tilde{X}$ is the state of the economy. The agent has a prior record of associations such that

$$W_t^{f \to c} \propto \begin{bmatrix} \text{Prob}(\textit{no crisis \& no depression}) & \text{Prob}(\textit{crisis \& no depression}) \\ \text{Prob}(\textit{no crisis \& depression}) & \text{Prob}(\textit{crisis \& depression}) \end{bmatrix} \tag{18}$$

$$\propto \begin{bmatrix} 1 - p^c & p^c(1-q) \\ 0 & p^c q \end{bmatrix}, \tag{19}$$

where

$$p^c = \text{probability of a financial crisis}$$
$$q = \text{probability of depression, given a financial crisis,}$$

namely, in investors' minds, depression is always accompanied by crisis. Equation 18 implies that, in the benchmark case of (1), $c_t = [1, 0]^\top$ would represent normal times, and $c_t = [0, 1]^\top$ a depression state.

We connect context to asset prices by assuming, for simplicity, homogeneous investors aggregating to the representative agent of the previous section. Agents extract probabilities of a disaster from c_t (the probability is the second element of c_t), and view these probabilities (again, for simplicity) as permanent.

We assumed the generalized context evolution (1), with context retrieval (2). We assume that in the recent past, the agent observes mainly neutral features: $f_t = [1, 0]^\top$. Assuming that the features-to-context matrix is in the steady state given by (18), neutral features imply the neutral context:

$$c_t^{\text{in}} \propto W_t^{f \to c} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \propto \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{20}$$

Theorem 3 implies that, given sufficiently many observations of neutral features, context reaches a steady state value $c_t = [1, 0]^\top$. Intuitively, this follows from setting $c_t = c_{t-1}$ in (1), treating the $W^{f \to c}$ as fixed.[18] The model thus implies neglected risk (Gennaioli and Shleifer, 2018).

Though agents neglect the depression state, they have not forgotten it. Representing the failure of Lehman brothers is $f_1 = [0, 1]^\top$, the well-publicized failure of a major financial institution. It follows that

$$c_t^{\text{in}} \propto W_t^{f \to c} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \propto \begin{bmatrix} 1 - q \\ q \end{bmatrix}. \tag{21}$$

Equation 21 represents reinstatement of the depression context. Even though a depression has not occurred, the agent is reminded strongly of a depression because of the financial crisis. The stronger the association between depression and crisis (the higher is q), the greater this reinstatement.[19]

Because context is autoregressive, the agent is still partially in the non-crisis con-

[18]Because crisis features are orthogonal to the no-depression context, this is correct, as Theorem 3 shows.

[19]Whether or not this is an over-reaction (Theorem 5), or a correct Bayesian response depends on whether the relation between crisis and depression is over-stated. See Appendix D for why it might be over-stated.

text. The retrieved depression context mixes with the prior neutral context to form

$$c_{t+1} = (1-\zeta)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \zeta \begin{bmatrix} 1-q \\ q \end{bmatrix} \tag{22}$$

The probability of a depression changes from zero to ζq, causing an immediate decline in stock prices (16) and in the riskfree rate (18).

What does the model say about the time path of context, and hence that of prices and the riskfree rate, following the event? We discuss in detail one such possible path. Consistent with events in late 2008, we assume several (specifically, three) observations of crisis features, and then neutral features. First, consider the effect of the crisis on memory.

$$\begin{aligned} W_{t+1}^{f\to c} &= W_t^{f\to c} + c_{t+1} f_{t+1}^\top \\ &= t\begin{bmatrix} 1-p^c & p^c(1-q) \\ 0 & p^c q \end{bmatrix} + \begin{bmatrix} 0 & (1-\zeta)+\zeta(1-q) \\ 0 & \zeta q \end{bmatrix} \end{aligned} \tag{23}$$

Memory updates with the term $c_1 f_1^\top$. The appearance of f_1 states that a crisis has occurred – it says nothing about a depression. However, c_1 does contain some probability of a depression, specifically, ζq. Thus, regardless of whether or not a depression actually occurs, the agent updates memory, represented by $W_1^{f\to c}$, with a partial observation of a depression, co-occurring with crisis.

Now suppose that the agent again observes crisis features, retrieving, again, the depression context:

$$c_{t+2}^{\text{in}} \propto W_{t+1}^{f\to c} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{24}$$

Retrieved context mixes with the prior context to form

$$\begin{aligned} c_{t+2} &= (1-\zeta)c_{t+1} + \zeta c_{t+2}^{\text{in}} \\ &= (1-\zeta)^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1-\zeta)\zeta \begin{bmatrix} 1-q \\ q \end{bmatrix} + \zeta c_{t+2}^{\text{in}} \end{aligned}$$

The agent starts to forget that normal features were part of the environment, as an ever-decreasing weight is placed on the original neglected risk context $[1, 0]^\top$. Similarly:

$$c_{t+3} = (1-\zeta)^3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1-\zeta)^2\zeta \begin{bmatrix} 1-q \\ q \end{bmatrix} + (1-\zeta)\zeta c_{t+2}^{\text{in}} + \zeta c_{t+3}^{\text{in}}$$

As long as the agent continues to observe crisis features, the weight on the depression state increases and the weight on the normal state decreases. Memory continues to be updated, as crisis features, and the depression state combine.

Something interesting happens when the agent finally observes neutral features again. Suppose for concreteness that this occurs at time 4. First, the agent will retrieve the neutral context, with no probability on depression:

$$c_{t+4}^{\text{in}} \propto W_{t+3}^{f\to c} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

this is because only the (1,2) and (2,2) elements of $W_{t+3}^{f\to c}$ have changed relative to $W_{t+0}^{f\to c}$. However, the depression is still in context. Thus the agent associates the depression not only with crisis features, but also with neutral features. That is, *even if the depression did not occur, and regardless of the number of observed neutral features*, the agent will continue to remember the depression.

As an example calibration, we assume $p^c = .025$, $q = 0.5$, a decline in aggregate consumption of 20%, $\lambda = 2$. Here, and in the applications that follow, we assume

$\zeta = 0.35$, found in Polyn et al. (2009) and Healey and Kahana (2016). We assume the agent observes a prior sample of 20 years. The agent begins in a neutral context, observes three periods worth of crisis features, and then 10 periods of neutral features. Figure 4 shows the time path of the price-dividend ratio. The real riskfree rate, which begins at 3%, and falls as low as -1%, follows a very similar path.

First note that the crisis leads to a jump back in time, namely an immediate decline in the price-dividend ratio and in the interest rate. Both continue to decline, as the agent continues to observe crisis features. Both occurred during the 2008 financial crisis. While the economy recovers, following observation of neutral features, recovery is incomplete. Because the agent continues to remember a depression (that did not occur), neither stock prices nor interest rates return to pre-crisis levels, even after 15 years.

4.3 Fear and asset allocation

Guiso et al. (2018) find that professional investors in Italy required a double the premium to accept a lottery following the financial crisis than before. To rule out the explanation that the apparent shift in risk aversion followed from a decrease in wealth (as it would if, say, agents exhibited decreasing relative risk aversion), they conducted an experiment with undergraduates students as subjects. The students were assigned at random to watch, or not watch, a scene from a horror movie. The authors found that the students who watched the scene required a 50% greater premium to accept the lottery. Similarly, Cohn et al. (2015) report results from an experiment on financial professionals, in which some viewed a fictive chart of a booming stock market, while others viewed a chart with a market crash. In both cases, professionals answered questions about their trading strategies during the event in question. They then performed an investment task. Investors in the boom condition invested 17 percentage points more in the risky asset that those in the bust condition.

The results from these experiments are striking in that fear alone, as opposed to new information, has a substantial effect on risk taking. Here, we apply retrieved-context theory to explain how an emotional experience can change portfolio holdings.

We will need to modify the simple model required to explain the effect of the Lehman Brothers bankruptcy on stock market valuations of the previous section. In Section 4.2, a cue triggered a jump back in time, namely a sudden jump in beliefs regarding the probability of a Great Depression. This stimulus had previously been directly associated with a Great Depression. However, in that case, agents could have believed that the true probability of a Great Depression had changed. In this section, where we seek to explain experimental evidence, subjects were told explicitly that risks had not changed, and yet there was a change in portfolio choice.

We hypothesize that fear operates through the memory channel. As we have shown, the context-retrieval mechanism allows negative associations to have both a short-lived effect (through the autoregressive structure) and a highly persistent effect (through the features to context matrix). It will be the first that that is the focus of this section.

We assume that the feature state can consist of the presence of danger, which may or may not be associated with a financial crisis. Danger is evoked by the kind of movie that Guiso et al. (2018) showed in their experiment. The feature space consists of:

$$f_t = \begin{cases} e_1 & \text{if} \quad \text{no danger \& no crisis} \\ e_2 & \text{if} \quad \text{danger \& no crisis} \\ e_3 & \text{if} \quad \text{danger \& crisis} \end{cases}$$

where e_j is the jth basis vector. We assume a two-dimensional context vector, depending on whether the underlying state represents a high level of risk or a low level of risk. We refer to $f_t = e_1$ as neutral features.

As in Section 4.1, we consider the portfolio choice problem of an agent investing in a risky asset and a riskless asset. Let $\tilde{r}$ denote the risky asset return, and π the percent allocation to the risky asset. Without loss of generality, we assume the agent starts the period with financial wealth equal to one, so that end-of-period financial wealth equals $1+\pi\tilde{r}$. Similarly to Section 4.1, the agent also faces the possibility of a negative labor market outcome, which we denote by $\tilde{\ell}$. We can think of $\tilde{\ell}$ as health expenditures or other financial obligations (such as a mortgage), net of labor income. To summarize, the agent solves

$$\max_{\pi} Eu(1+\pi\tilde{r}+\tilde{\ell}). \tag{25}$$

We model $\tilde{\ell}$ as a Bernoulli random variable:

$$\tilde{\ell} = \begin{cases} 0 & \text{with probability } 1-p \\ -b & \text{with probability } p, \end{cases}$$

with $b \in [0,1]$. We assume $\tilde{r}$ also takes on two possible outcomes (each with equal probability), and has mean μ and standard deviation σ. Unlike the model in Section 4.1, $\tilde{\ell}$ and $\tilde{r}$ are uncorrelated.

In Section 4.1, agents invested less in stocks in response to fear about a depression state. The mechanism in that section was a covariance between labor in come and the stock return. In this section, we hypothesize that agents experience fear of physical danger after watching the horror movie. However, we cannot rely on covariance between physical danger and the risky asset return (which would be implausible) to generate the decreased investment in the risky asset.

To allow $\tilde{\ell}$ to affect the agent's portfolio choice, we assume log utility, as in Section 4.2. Fixing an outcome $\tilde{\ell}=\ell$, (25), given $u(x)=\log x$, implies decreasing relative risk aversion; the agent is very averse to declines in wealth that are close to b (the decline in labor income). Now allowing $\tilde{\ell}$ to be variable, the greater the probability, the more weight the agent places on this possible outcome in her decisions. This formulation, together with the context dynamics below, endogenizes time-varying risk aversion. It also endogenously produces a role for emotional state in the utility function, as suggested by Loewenstein (2000).

As in previous examples, assume the context vector determines the agent's subjective risk probability p. In the language of Section 2, the random variable $\tilde{Y}$ consists of three states $\{y_1, y_2, y_3\}$, where y_1 represents no danger and no crisis, y_2 represents danger and no crisis, and y_3 represents both danger and crisis. The random variable $\tilde{X}$ equals the outcome of $\tilde{\ell}$. True joint contingencies are given by:

$$P = \begin{bmatrix} \Pr(x_1, y_1) & \Pr(x_1, y_2) & \Pr(x_1, y_3) \\ \Pr(x_2, y_1) & \Pr(x_2, y_2) & \Pr(x_2, y_3) \end{bmatrix}, \tag{26}$$

Let p be the unconditional probability of danger, and q be the probability of crisis given danger (we make the reasonable assumption that crisis is always accompanied by danger, which can include dangers inherent in losing one's money).[20] We assume, for simplicity, that risk and danger always co-occur, so that

$$P = \begin{bmatrix} 1-p & 0 & 0 \\ 0 & p(1-q) & pq \end{bmatrix}. \tag{27}$$

Our results do not depend on this assumption.

Assume that context follows (1) and (2) and that the agent has accurately observed and recalled a sufficiently long sample, so that:

$$W_t^{f\to c} \propto P,$$

[20]This structure does have the implication that stock returns are uncorrelated with the crisis outcome. A richer model might have two types of risk which share a common component of $\tilde{\ell}$, but one with an additional component that correlates with stock returns.

and

$$c_t = \begin{bmatrix} 1-p \\ p \end{bmatrix}.$$

That is, the agent begins with the correct probabilities. However, our results do not depend on these assumption. Our results are robust to variation in this assumption, as all that is required is that the stimulus introduced in the experiment drives context sufficiently far away from its pre-experimental state.

The stimulus represents $f_{\text{movie}} \approx e_2$, namely, an experience that shares some aspects of danger, but is not the same as danger. Following Guiso et al. (2018), who choose a movie that is both intense and obscure, agents do not have prior associations with f_{movie}, except that the movie is in some respects similar (by design) to danger in real life.

We have

$$c_{t+1}^{\text{in}} \propto W_t^{f\to c} f_{\text{movie}}$$

which implies, by Theorem 6,

$$c_{t+1}^{\text{in}} \approx \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Therefore, the new context is:

$$c_{t+1} = (1-\zeta)c_t + \begin{bmatrix} 0 \\ \zeta \end{bmatrix},$$

so that the subjective probability of the risky state rises from p to $(1-\zeta)p+\zeta$.[21]

Figure 5 shows expected utility (25) as a function of portfolio allocation π, prior to and after the stimulus, with p equal to the second element of context. We assume an excess return $\mu = 4\%$, a standard deviation $\sigma = 20\%$, a prior probability of the

[21] Note that $(1-\zeta)p+\zeta > p$ for $p < 1$, because $(1-\zeta)p+\zeta$ is a weighted average of p and 1.

negative labor market outcome $p = 2\%$, and a percent decline $b = -0.8$, should the outcome occur. As elsewhere, $\zeta = 0.35$. When the agent has the correct probabilities, the portfolio allocation equals 70%, falling to 30% after the stimulus. Note that the model would imply the same shift for a financial crisis. This accounts for the finding of Guiso et al. (2018) that (a) viewing a horror movie and (b) exposure to a financial crisis increases effective risk aversion.

The horror movie changes the beliefs of the agent about the risk the agent might face. It is as if the movie reminds the agent that the world is a risky place, and one thus should not take risks with one's financial wealth. Our set-up could either be interpreted as the one in which the agent is reminded of why wealth is necessary (that is the literal interpretation above), or is reminded of how painful (through time-varying risk aversion), low-wealth states are.

The response of the agent to the experiment cannot be Bayesian: a movie has not changed anything about the outside world. In that sense, the response of risk-taking to viewing a horror movie is a good test of our theory. The experiment shows that financial decisions in one context do not resemble financial decisions in another, even though the financial decision in both cases is materially the same. Context "should" be irrelevant, and yet it is not. The agent may know, intellectually, that nothing has changed, and yet the powerful pull of context implies that choices change anyway.

5 Conclusion

Our past experiences, and our knowledge about the world, constitute a vast database of information that potentially informs every decision we make. Does the human memory system discard most of this information to abstract a small, and possibly biased, subset? Modern research on human memory supports an alternative view in which much of our past information remains in storage, to be retrieved based on a latent dynamic context (Kahana, 2012). According to this view, context updates recursively; features of the environment evoke past contextual states via associative memory. These associations then are permanently stored to be themselves evoked at later times. Thus past contextual states drive the evolution of context itself.

Here we introduce memory into the decision problem of an economic agent, through a formal model of retrieved context theory. Features represent observed stock prices or exceptionally salient news such as a large bank failure. The associative matrices linking context to features draw out the agent's beliefs given these observations. Our model allows for important deviations from Bayesian updating, such as the influence of events in the distant past, the influence of irrelevant events, and slow updating to new information. We apply retrieved-context theory to four illustrative problems in financial economics: the effects of life experience on choices, the sudden onset of a financial crisis, the appearance of momentum in stock returns, and time-variation in risk aversion due to exogenous factors.

A Proofs of Theorems in Section 2

.

Notation: Throughout the Appendix, we use e_j to denote the conforming standard jth basis vector. For a vector v, $v(i)$ denotes the ith element of v.

Proof of Theorem 1:

Proof. The agent has a flat prior and a likelihood function that $(\tilde{X}, \tilde{Y})$ is multinomial, with mn states. A standard result is that the posterior distribution of the probability $\text{Prob}(\tilde{X} = x_i, \tilde{Y} = y_j) = p_i^X p^Y(y_j \,|\, x_i)$ is Dirichlet, with posterior mean equal to the relative joint occurrence of x_i and x_j in the sample. See, e.g., Gelman et al. (2004). $\square$

Proof of Theorem 2:

Proof. With some abuse of notation, define $p(x_i, y_j) = p_i^X p^Y(y_j \,|\, x_i)$ to be the joint probability of state i and state j, and define a matrix P such that $P(i,j) = p(x_i, y_j)$. Then, under the assumptions of the theorem, $P(i,j) \propto W_{t-1}^{f\to c}$, with the constant of proportionality equal to the sum of the elements of $W_{t-1}^{f\to c}$. Suppose $Y_t = y_j$. Then f_t is the jth basis vector and

$$\begin{aligned} c_t^{\text{in}} &= \frac{W_{t-1}^{f\to c} f_t}{||W_{t-1}^{f\to c} f_t||} \\ &= \frac{Pe_j}{||Pe_j||} \\ &= \left(\sum_i p(x_i, y_j)\right)^{-1} \begin{bmatrix} p(x_1, y_j) \\ \vdots \\ p(x_m, y_j) \end{bmatrix}. \end{aligned}$$

Note that $\sum_i p(x_i, y_j)$ is simply the unconditional probability of y_j. Thus

$$c_t^{\text{in}}(i) = p(x_i, y_j) \left(\sum_i p(x_i, y_j) \right)^{-1} = p(x_i \,|\, y_j).$$

$\square$

Definition (Orthogonal features)**.** *A features vector $\bar{f}$ and a state $\tilde{X} = x_i$ are* orthogonal at time t *if for all $s \leq t$ such that $f_s^\top \bar{f} \neq 0$, $c_s(i) = 0$.*

We will simply state that $\bar{f}$ is orthogonal to state i at t (or state i is orthogonal to features $\bar{f}$ at t) when there is no potential for confusion.

Definition (Associated features)**.** *We say that state $\tilde{X} = x_i$ and features $\bar{f}$ are*

1. associated *at time t if they are not orthogonal at time t.*
2. uniquely associated *at time t if $\bar{f}$ is orthogonal to all states other than i at time t.*

Lemma 8. *If $\bar{f}$ is orthogonal to state i at time t, then, if $f_{t+1} = \bar{f}$, $c_{t+1}^{in}(i) = 0$.*

Proof. From (2),

$$\begin{aligned} c_{t+1}^{\text{in}} &\propto W_t^{f \to c} \bar{f} \\ &\propto \sum_{s=0}^{t} c_s (f_s^\top \bar{f}), \end{aligned}$$

and therefore

$$c_{t+1}^{\text{in}}(i) \propto \sum_{s=0}^{t} c_s(i)(f_s^\top \bar{f}) = 0,$$

because $c_s(i) \neq 0$ only when $f_s^\top \bar{f} = 0$, $\square$

Lemma 9. *If $\bar{f}$ is uniquely associated with state i at time t, then if $f_{t+1} = \bar{f}$, $c_{t+1}^{in}(i) =$ 1.*

Proof. Note that if $\bar{f}$ is uniquely associated with state i, then $\bar{f}$ is orthogonal to all other states j. That implies $c^{\text{in}}_{t+1}(j) = 0$ for $j \neq i$ (by Lemma 8), and $c^{\text{in}}_{t+1}(i) > 0$. Because the elements of the context vector sum to 1, after scaling, $c^{\text{in}}_{t+1}(i) = 1$. □

These definitions, and the subsequent lemmas, have psychological interpretations: If $\bar{f}$ is orthogonal to state i, then the agent has never experienced features $\bar{f}$ at times when context placed some weight on i. That means that if $\bar{f}$ comes along, retrieved context does not include state i. On the other hand, if $\bar{f}$ is uniquely associated with state i, it can only recall state i.

Note that features can start out being orthogonal and then become associated with a context. Suppose that $\bar{f}$ is orthogonal at time t, and context at time t is such that $c_t(i) \neq 0$. Then, whereas $c^{\text{in}}_t(i) = 0$, by Lemma 8, the context evolution equation (1) implies that $c_{t+1}(i) \neq 0$, though it will be smaller than $c_t(i)$. Then the updating of associations through (3) implies that $\bar{f}$ becomes associated with state i.

Strictly speaking, once associated, features and a state can never become orthogonal. However, if features are very common, and the context is very rare, then the association is negligible and the model behaves as if they are orthogonal. Consider the expression for the ith element of retrieved context:

$$c^{\text{in}}(i) = \frac{1}{\| \sum_{s=0}^{t} (c_s f_s)^\top \bar{f} \|} \sum_{s=0}^{t} (e_i^\top c_s)(f_s^\top \bar{f}).$$

Common features $\bar{f}$ are ones in which $f_s^\top \bar{f}$ is often equal (or close) to 1. If state i is rare, then $e_i^\top c_s$ will usually be zero, whereas (given how often $\bar{f}$ occurs) $\| \sum_{s=0}^{t} (c_s f_s)^\top \bar{f} \|$, will be large. Intuitively, a common feature cannot be used to recall a rare event because of interference with many other common features.

To avoid dealing with very weakly associated features, we employ a technique to reset context, which involves introducing a sufficient number of orthogonal features:

Lemma 10 (Context reset)**.** *Assume a sequence of features $f_{t+1}, \ldots, f_{t+\tau}$, orthogonal to one another, such that each is orthogonal to state i at time t. Then, as $\tau \to \infty$, $c_{t+\tau}(i) \to 0$.*

Proof. Fix $c_t(i) \neq 0$ (if $c_t(i) = 0$, then the lemma holds trivially). Choose an arbitrary $\epsilon > 0$. It suffices to show that

$$c_{t+\tau}(i) = (1 - \zeta)^{\tau} c_t(i), \tag{A.1}$$

because we can then choose τ sufficiently large so that $(1-\zeta)^{\tau} < \epsilon / c_t(i)$ and $c_{t+\tau}(i) < \epsilon$.

We now prove (A.1) by induction on τ. It holds trivially for $\tau = 0$. Assume it is true for $\tau - 1$. Introduce features $f_{t+\tau}$ orthogonal to i at t and orthogonal to $f_{t+1}, \ldots, f_{t+\tau-1}$. Note that

$$W^{f\to c}_{t+\tau-1} = W^{f\to c}_{t} + c_{t+1} f^{\top}_{t+1} + \cdots + c_{t+1} f^{\top}_{t+t+\tau-1}$$

Then

$$\begin{aligned} c^{\text{in}}_{t+\tau} &\propto W^{f\to c}_{t+\tau-1} f_{t+\tau} \\ &\propto W^{f\to c}_{t} f_{t+\tau} + c_{t+1} f^{\top}_{t+1} f_{t+\tau} + \cdots + c_{t+1} f^{\top}_{t+\tau-1} f_{t+\tau} \\ &\propto W^{f\to c}_{t} f_{t+\tau} \end{aligned}$$

Thus, by orthogonality at time t,

$$c^{\text{in}}_{t+\tau}(i) = 0.$$

Assume by induction that

$$c_{t+\tau-1}(i) = (1 - \zeta)^{\tau-1} c_t(i). \tag{A.2}$$

Then $c_{t+\tau}(i) = 0$ follows from (1). $\square$

Intuitively, exponential decay of context, combined with a sufficiently large number of orthogonal features, implies that context "reset" is possible. This is not simply a mathematical construction: novel features are used in the memory laboratory to "reset" context. In the memory laboratory, these orthogonal features are called "distractor tasks" and often involve solving arithmetic problems quickly. Whether or not there exist such a large variety of features, or whether the updated weight in $W^{f\to c}$ is negligible, or if the weights ζ are varying are less important than the result which is that it is possible, through distraction, to change context. This is the main result we need for the theorems that follow.

The proof of Theorem 3 is then a direct application of Lemma 10.

Proof. Fix a time T, and a prior time $T-\tau+1$. The assumption of the theorem implies that the agent has only observed features orthogonal to i for $t = T-\tau+1, \ldots T$. To be precise, these features can be orthogonal to i at time t, and orthogonal to each other. Then, by Lemma 10, the weight $c_{t+\tau}(i)$ can be made arbitrarily small. $\square$

Proof of Theorem 4

Proof. By (1), Lemma 9, and the assumption that y^k is uniquely associated with i,

$$c_t(i) - c_{t-1}(i) = \begin{cases} \zeta(1 - c_{t-1}(i)) & \text{if } x_t = x_i \\ -\zeta c_{t-1}(i) & \text{otherwise} \end{cases}$$

That is, if features $x_t = x_i$, then features f_k, for $k \in \Omega_i$ occur, and the agent's context updates weakly positively, and strictly positively if $c_{t-1}(i) \neq 1$. On the other hand, if $x_t \neq x_i$, then features inconsistent with x_i are observed, and the weight on i falls (or remains at zero).

We first consider the case such that beliefs are novel. Then $c_t(i) > c_{t-1}(i)$ implies $X_t = x_i$. We then have:

$$E[c_{t+1}(i) - c_t(i)|c_{t-1}] = \zeta \left(p_{ii}^X - (1-\zeta)c_{t-1}(i) - \zeta\right). \tag{A.3}$$

On the other hand, if $c_t(i) \leq c_{t-1}(i)$, then it must be that $X_t \neq x_i$, and

$$E[c_{t+1}(i) - c_t(i)|c_{t-1}] = \zeta\Big(\sum_{j\neq i} p_{ji}^X p(j) - (1-\zeta)c_{t-1}(i)\Big). \tag{A.4}$$

where $p(j)$ is the conditional probability of being in state j. Under the assumptions on persistence, (A.3) exceeds (A.4). Namely, if $X_t = x_i$ it is more likely that the state will remain, and beliefs will update further, than if $X_t \neq x_i$ and the state will switch to x_i and beliefs will update further. Because (A.3) exceeds (A.4) for any level of $c_{t-1}(i) < 1$, it must hold unconditionally.

We now suppose that $c_t(i) = c_{t-1}(i) = 1$, as would be the case if f_t has been observed for a sufficiently long amount of time. Then if we see $c_t(i) = c_{t-1}(i)$, it must be that $X_t = x_i$. Equation A.3 still holds; however, the value is strictly negative (except in the limiting case where i is an absorbing state and $p_{ii} = 1$). □

Proof of Theorem 5

Proof. Under the assumptions of the theorem

$$c_t = c_t^{\text{in}} \propto W_{t-1}^{f\to c} f_t.$$

Suppose $f_t = e_k$. Then

$$c_t = \left(\sum_{i=1}^{m} W_{t-1}^{f\to c}(i,k)\right)^{-1} \begin{bmatrix} W_{t-1}^{f\to c}(1,k) \\ \vdots \\ W_{t-1}^{f\to c}(m,k) \end{bmatrix}.$$

According to Assumption 3, the agent behaves as if the probability of state i upon viewing $f_t = e_k$ equals

$$c_t(i) = \left(\sum_{i=1}^{m} W_{t-1}^{f\to c}(i,k)\right)^{-1} W_{t-1}^{f\to c}(i,k).$$

The conditional probability, in population, however, is

$$\text{Prob}(\tilde{X} = x_i \,|\, \tilde{Y} = y_k) = \left(\sum_{i=1}^{m} p(i,k)\right)^{-1} p(x_i, y_k),$$

where we have followed the notational convention in the proof of Theorem 2. Dividing through by $W_{t-1}^{f\to c}(i,k)$ and by $p(x_i, y_k)$ proves the result. □

Proof of Theorem 6

Proof. Consider two features vectors f and f', and some previous features vector f_s. By the Cauchy-Schwarz inequality:

$$f_s^\top(f - f') \leq \|f_s\|\|f - f'\|.$$

for the standard distance (L^2) norm $\|\|$. Thus retrieved context (2) can be made arbitrarily close for f and f', provided that f and f' are arbitrarily close. The result then follows from (1). □

Proof of Theorem 7

Proof. We prove the statement by induction on the number of periods τ. When $\tau = 0$, the theorem holds trivially. Assume that for any features vector at time $t+\tau$, $f_{t+\tau}$,

$$W^{f\to c}_{t+\tau-1} f_{t+\tau} \propto W^{f\to c}_{t} f_{t+\tau}.$$

Note that

$$W^{f\to c}_{t+\tau} = W^{f\to c}_{t+\tau-1} + c_{t+\tau} f^{\top}_{t+\tau}.$$

where, under the assumptions of the theorem

$$c_{t+\tau} = (\|W^{f\to c}_{t+\tau-1} f_{t+\tau}\|)^{-1} W^{f\to c}_{t+\tau-1} f_{t+\tau}.$$

Therefore,

$$W^{f\to c}_{t+\tau} = W^{f\to c}_{t+\tau-1}(1 + (\|W^{f\to c}_{t+\tau-1} f_{t+\tau}\|)^{-1} f_{t+\tau} f^{\top}_{t+\tau})$$

The theorem follows from the induction step by separately considering $f_{t+\tau+1} = f_{t+\tau}$ in which case the inner product $f^{\top}_{t+\tau} f_{t+\tau_1} = 1$, and $f_{t+\tau+1} \neq f_{t+\tau}$, in which case the inner product is zero.

□

B Bayesian updating from rare events

Suppose a depression occurs with probability p. A Bayesian agent seeks to learn about p from observations of whether or not a depression occurs. The agent has prior

$$p \sim \mathrm{Beta}(p^{*}\tau + 1, (1 - p^{*})\tau + 1), \tag{B.1}$$

which has the interpretation of a pseudo-sample of length τ, during which there are $p^*\tau$ occurrences of a depression. The mean equals

$$E[p] = \frac{p^*\tau + 1}{\tau + 2}$$

The case of $\tau = 0$ implies corresponds to a uniform prior on $[0, 1]$ with a mean of 1/2. This is an uninformative prior. The greater is τ, the more informative the prior. As τ approaches infinity, the prior mean approaches p^*. The density function corresponding to (B.1) equals

$$f(p) \propto p^{p^*\tau}(1-p)^{(1-p^*)\tau},$$

which approaches the uniform prior as τ approaches zero.

Assume T years of data. Conditional on knowing the probability p, the likelihood of exactly N occurrences of a depression out of a total of T periods equals

$$\mathcal{L}(N \text{ disasters} \,|\, p) = \binom{T}{N} p^N (1-p)^{T-N}. \tag{B.2}$$

The assumption of a Beta$(p^*\tau+1, (1-p^*)\tau+1)$ prior implies the prior density function

$$f(p) \propto p^{p^*\tau}(1-p)^{(1-p^*)\tau}$$

where the constant of proportionality does not depend on p and can therefore be disregarded in what follows. Therefore the posterior distribution equals

$$\begin{aligned} f(p \,|\, N \text{ disasters}) &\propto \mathcal{L}(N \text{ disasters}|p) f(p) \\ &\propto p^{N+p^*\tau}(1-p)^{T+\tau-(N+p^*\tau)} \end{aligned}$$

where once again we have ignored terms that do not depend on p. This is proportional to the Beta density, so

$$p \,|\, N \text{ disasters} \sim \text{Beta}(N + p^*\tau + 1, T + \tau - (N + p^*\tau) + 1).$$

It follows from properties of the Beta distribution that the posterior mean equals

$$E[p \,|\, N \text{ disasters}] = \frac{N + p^*\tau + 1}{T + \tau + 2}.$$

The posterior mean depends on the sample path. Figure 3 shows the average posterior mean, assuming the likelihood (B.2):

$$\begin{aligned} E_{\#\text{disasters}}\left[E[p \,|\, N \text{disasters}]\right] &= \int \frac{N + p^*\tau + 1}{T + 2} \mathcal{L}(N \text{ disasters} \,|\, p)\, dN \\ &= \frac{pT + p^*\tau + 1}{T + \tau + 2} \end{aligned}$$

where we have used the fact that, conditional on p, N has a binomial distribution, and therefore $E[N \,|\, p] = pT$. The figure corresponds to the case of $\tau = 0$, however the results are very similar for $\tau > 0$, provided that the actual sample is large relative to the prior sample.

C Asset pricing with rare events

The model in this section follows that of Barro (2006). We assume a complete-markets endowment economy similar to Lucas (1978). Assume

$$\log C_{t+1} = \log C_t + \mu + u_{t+1} + v_{t+1},$$

where u_{t+1} and v_{t+1} are independent, $u_{t+1} \sim N(0, \sigma^2)$ and

$$v_{t+1} = \begin{cases} 0 & \text{with prob. } e^{-p} \\ \log(1-b) & \text{with prob. } 1 - e^{-p} \end{cases}$$

where b is a random variable with support on $[0, 1)$. We further assume that the dividend satisfies

$$\log D_{t+1} = \log D_t + \mu + u_{t+1} + \lambda v_{t+1} \tag{C.1}$$

with $\lambda > 1$. This assumption captures the fact that dividends fall by more than consumption during crisis (Longstaff and Piazzesi, 2004).

We assume that at every period, the agent maximizes utility

$$E_t \sum_{s=1}^{\infty} \beta^s \frac{C_s^{1-\gamma}}{1-\gamma}.$$

Note that p scales with the time interval. Thus we can make p arbitrarily small (without changing the underlying economics) by considering smaller and smaller time intervals (in effect approximating a Poisson process in discrete time). Note, however, that b is a fixed quantity as the time interval shrinks. Besides the closed-form expressions, we will give simpler formulas using

$$1 - e^{-p} \approx p,$$

and for x close to zero,

$$\log(1 + x) \approx x.$$

Define the stochastic discount factor as

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$$

It follows from the first-order condition of the representative agent that

$$
\begin{aligned}
1 + r_f &= E\left[M_{t+1}\right]^{-1} \\
&= \beta^{-1} E\left[\exp\left\{-\gamma(\mu + u_{t+1} + v_{t+1})\right\}\right]^{-1} \\
&= \beta^{-1} e^{\gamma\mu - \frac{1}{2}\gamma^2\sigma^2} \left(e^{-p} + (1 - e^{-p}) E\left[(1-b)^{-\gamma}\right]\right)^{-1},
\end{aligned}
$$

so that

$$
\log(1 + r_f) \approx -\log\beta + \gamma\mu - \frac{1}{2}\gamma^2\sigma^2 - pE\left[(1-b)^{-\gamma} - 1\right].
$$

The price S_{1t} of a claim to a one-period equity strip satisfies the equation

$$
S_{1t} = E_t\left[M_{t+1} D_{t+1}\right].
$$

It is straightforward to solve for this price by using the normalization

$$
\begin{aligned}
\frac{S_{1t}}{D_t} &= E_t\left[M_{t+1}\frac{D_{t+1}}{D_t}\right] \\
&= \beta e^{(1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2} \left(e^{-p} + (1 - e^{-p}) E\left[(1-b)^{\lambda-\gamma}\right]\right).
\end{aligned}
$$

Define

$$
\Phi(p) \equiv \beta e^{(1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2} \left(e^{-p} + (1 - e^{-p}) E\left[(1-b)^{\lambda-\gamma}\right]\right) \tag{C.2}
$$

as the price-dividend ratio for the one-period claim.

Taking the log of both sides of (C.2) gives a convenient approximation

$$
\begin{aligned}
\log\Phi(p) &= \log\beta + (1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2 + \log\left(e^{-p} + (1 - e^{-p}) E\left[(1-b)^{\lambda-\gamma}\right]\right) \\
&\approx -\log\beta + (1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2 - pE\left[1 - (1-b)^{\lambda-\gamma}\right]. \qquad \text{(C.3)}
\end{aligned}
$$

Note that $1 - b \in (0, 1]$. Thus the term inside the expectation in (C.3) is positive if and only if $\lambda > \gamma$. Under these circumstances, an increase in p lowers prices.

Now consider the claim to a stream of dividends following process (C.1). Let S_t denote the price of this claim. The condition for equilibrium, applied to this claim implies

$$S_t = E_t \left[M_{t+1}(S_{t+1} + D_{t+1}) \right]$$

which in turn implies a recursion for the price-dividend ratio

$$\frac{S_t}{D_t} = E_t \left[M_{t+1} \frac{S_{t+1}/D_{t+1} + 1}{S_t/D_t} \frac{D_{t+1}}{D_t} \right].$$

The solution equals:

$$\frac{S_t}{D_t} = \sum_{n=1}^{\infty} \Phi(p)^n = \frac{\Phi(p)}{1 - \Phi(p)}.$$

We calibrate the model using with $\mu = 0$, $\sigma = 2\%$, $\beta = e^{-.03}$, $\gamma = 1$, $\lambda = 3$, $b = 0.40$.

D Retrieval of Features

To be completed

References

Allen, F. and Gale, D. (2009). *Understanding Financial Crises.* Clarendon lectures in finance. Oxford University Press.

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing.* Markham Publishing Co, Chicago.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In Spence, K. W. and Spence, J. T., editors, *The Psychology of Learning and Motivation*, volume 2, pages 89–105. Academic Press, New York.

Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115(1):1–24.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, page 37.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1):173–196.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2019). Memory, attention, and choice. NBER Working Paper No. 23256.

Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, 104(2):379–421.

Carr, H. A. (1931). The laws of association. *Psychological Review*, 38:212–228.

Chan, L. K. C., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5):1681–1713.

Cohn, A., Engelmann, J., Fehr, E., and Marechal, M. A. (2015). Evidence for countercyclical risk aversion: An experiment with financial professionals. *The American Economic Review*, 105(2):860–885.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1839–1885.

Ebbinghaus, H. (1885/1913). *Memory: a contribution to experimental psychology*. Teachers College, Columbia University, New York.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Epstein, L. G. and Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, 62:145–154.

Folkerts, S., Rutishauser, U., and Howard, M. (2018). Human episodic memory retrieval is accompanied by a neural contiguity effect. *Journal of Neuroscience*, 38(17):4200–4211.

French, K. R., Baily, M. N., Campbell, J. Y., Cochrane, J. H., Diamond, D. W., Duffie, D., Kashyap, A. K., Mishkin, F. S., Rajan, R. G., Scharfstein, D. S., Shiller, R. J., Shin, H. S., Slaughter, M. J., Stein, J. C., and Stulz, R. M. (2010). *The Squam Lake Report: Fixing the Financial System*. Princeton University Press.

Fuster, A., Laibson, D., and Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives*, 24(4):67–84.

Gabaix, X. (2019). Behavioral inattention. In Bernheim, D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics*, volume 2. Elsevier.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL.

Gennaioli, N. and Shleifer, A. (2018). *A Crisis of Beliefs: Investor Psychology and Financial Fragility*. Princeton University Press, Princeton, NJ.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Gilboa, I. and Schmeidler, D. (1995). Case-based decision theory. *The Quarterly Journal of Economics*, 110(3):605–639.

Godden, D. R. and Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, 66:325–331.

Goetzmann, W. N., Kim, D., and Shiller, R. J. (2017). Crash beliefs from investor surveys. NBER Working Paper No. 22143.

Gomes, J. F., Grotteria, M., and Wachter, J. A. (2019). Cyclical dispersion in expected defaults. *Review of Financial Studies*, 32(4):1275–1308.

Gorton, G. and Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, 104(3):425–451.

Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.

Hansen, L. P. and Sargent, T. J. (2018). Structured uncertainty and model misspecification. Working paper, New York University and University of Chicago.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, 103(2):732–70.

Healey, M. K. and Kahana, M. J. (2016). A four–component model of age–related memory change. *Psychological Review*, 123(1):23–69.

Healey, M. K., Long, N. M., and Kahana, M. J. (2019). Contiguity in episodic memory. *Psychonomic Bulletin & Review*, 26(3):699—720.

Herbart, J. F. (1834). Lehrbuch zur Psychologie. In Kehrbach, K., editor, *Sämtliche Werke*, volume 4, pages 376–377. August Wilhelm Unzer, Königsberg, 2nd edition.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6):2143–2184.

Howard, M. W. and Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46(3):269–299.

Howard, M. W., Viskontas, I. V., Shankar, K. H., and Fried, I. (2012a). Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, 22:1833–1847.

Howard, M. W., Viskontas, I. V., Shankar, K. H., and Fried, I. (2012b). Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, 22:1833–1847.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Retrieved from Project Gutenberg: http://www.gutenberg.org/files/9662/9662-h/9662-h.htm.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48:65–91.

Jones, C. M. (2002). A century of stock market liquidity and trading costs. Working paper, Columbia University.

Jost, A. (1897). Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 16:436–472.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, 24(1):103–109.

Kahana, M. J. (2012). *Foundations of Human Memory*. Oxford University Press, New York, NY.

Kahle, K. M. and Stulz, R. (2013). Access to capital, investment, and the financial crisis. *Journal of Financial Economics*, 110(2):280–299.

Ladd, G. T. and Woodworth, R. S. (1911). *Elements of physiological psychology: A treatise of the activities and nature of the mind from the physical and experimental point of view*. Charles Scribner's Sons, New York, NY.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *The American Economic Review*, 90(2):426–432.

Lohnas, L. J. and Kahana, M. J. (2014a). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 40(1):12–24.

Lohnas, L. J. and Kahana, M. J. (2014b). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, 40(3):755–764.

Lohnas, L. J., Polyn, S. M., and Kahana, M. (2015). Expanding the scope of memory search: Intralist and interlist effects in free recall. *Psychological Review*, 122(2):337–363.

Longstaff, F. A. and Piazzesi, M. (2004). Corporate earnings and the equity premium. *Journal of Financial Economics*, 74:401–421.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica*, 46:1429–1445.

Luce, R. D. (1959). *Individual choice behavior*. Wiley, New York.

Malmendier, U. and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, 126(1):373–416.

Malmendier, U. and Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87.

Malmendier, U., Nagel, S., and Yan, Z. (2017). The making of hawks and doves: Inflation experiences on the FOMC. NBER Working Paper No. 23228.

Malmendier, U. and Shen, L. S. (2018). Scarred consumption. NBER Working Paper No. 24696.

Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., and Kahana, M. J. (2011a). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences, USA*, 108(31):12893–12897.

Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., and Kahana, M. J. (2011b). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences, USA*, 108(31):12893–12897.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

McFadden, D. (2001). Economic choices. *American Economic Review*, 91(3):351–378.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, 39:352–70.

Moreton, B. J. and Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review*, 17(4):510–515.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, 117(3):735–774.

Müller, G. E. and Pilzecker, A. (1900). Experimental contributions to memory theory. *Zeitschrift für Psychologie Eganzungsband*, 1:1–300.

Müller, G. E. and Schumann, F. (1894). Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 6(80-191):257–339.

Murdock, B. B. (1974). *Human memory: Theory and data.* Lawrence Erlbaum and Associates, Potomac, MD.

Nagel, S. and Xu, Z. (2018). Asset pricing with fading memory. Working paper, University of Chicago and University of Michigan.

Polyn, S. M., Norman, K. A., and Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, 116(1):129–156.

Pratt, J. W. (1976). Risk aversion in the small and in the large. *Econometrica*, 44(2):420–420.

Raaijmakers, J. G. W. and Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In Bower, G. H., editor, *The Psychology of Learning and Motivation: Advances in Research and Theory*, volume 14, pages 207–262. Academic Press, New York.

Rubin, D. C. and Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the lifespan. *Memory & Cognition*, 37(5):679–688.

Rubin, D. C., Berntsen, D., and Johansen, M. K. (2008). A memory-based model of posttraumatic stress disorder: Evaluating basic assumptions underlying the PTSD diagnosis. *Psychological Review*, 115(4):985–1011.

Savage, L. (1954). *The Foundations of Statistics*. Wiley, New York.

Sederberg, P. B., Gershman, S. J., Polyn, S. M., and Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, 18(3):455–468.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Thorndike, E. L. (1932). *The fundamentals of learning*. Bureau of Publications, Teachers College, New York.

Tsai, J. and Wachter, J. A. (2015). Disaster risk and its implications for asset pricing. *Annual Review of Financial Economics*, 7(1):219–252.

Uitvlugt, M. G. and Healey, M. K. (2019). Temporal proximity links unrelated news events in memory. *Psychological Science*, 30(1):92–104.

Underwood, B. J. (1948). 'Spontaneous recovery' of verbal associations. *Journal of Experimental Psychology*, 38:429–439.

Wachter, J. A. (2010). Asset allocation. *Annual Review of Financial Economics*, 2(1):175–206.

Yaffe, R. B., Kerr, M. S., Damera, S., Sarma, S. V., Inati, S. K., and Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences*, 111(52):18727–8732.

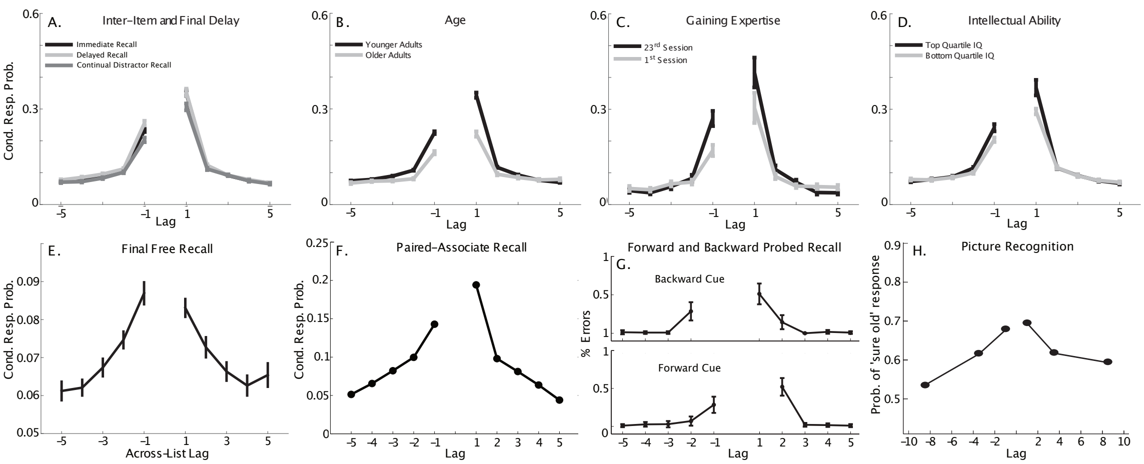


Figure 1: **Universality of Temporal Contiguity. A.** When freely recalling a list of studied items, people tend to successively recall items that appeared in neighboring positions. This temporal contiguity effect (TCE) appears as an in increase in the conditional-response probability as a function of the lag, or distance, between studied items (the lag-CRP). The TCE appears invariant across conditions of immediate recall, delayed recall, and continual-distractor recall, where subjects perform a demanding distractor task between each of the studied items. **B.** Older adults exhibit reduced temporal contiguity, indicating impaired contextual retrieval **C.** Massive practice increases the TCE, as seen in the comparison of 1st and 23rd hour of recall practice. **D.** Higher-IQ subjects exhibit a stronger TCE than individuals with average IQ. **E.** The TCE is not due to inter-item associations as it appears in transitions across different lists, separated by minutes, in a delayed final test given to subjects who studied and recalled many lists. **F.** The TCE appears in conditional error gradients in cued recall, where subjects tend to mistakenly recall items from pairs studied in nearby list positions. **G.** When probed to recall the item that either followed or preceded a cue item, subjects occasionally commit recall errors whose distribution exhibits a TCE both for forward and backward probes. **H.** The TCE also appears when subjects are asked to recognize previously seen travel photos. When successive test items come from nearby positions on the study list, subjects tendency to make high confidence "old" responses exhibits a TCE when the previously tested item was also judged old with high confidence. This effect is not observed for responses made with low confidence. Healey et al. (2019) provide references and descriptions of each experiment.

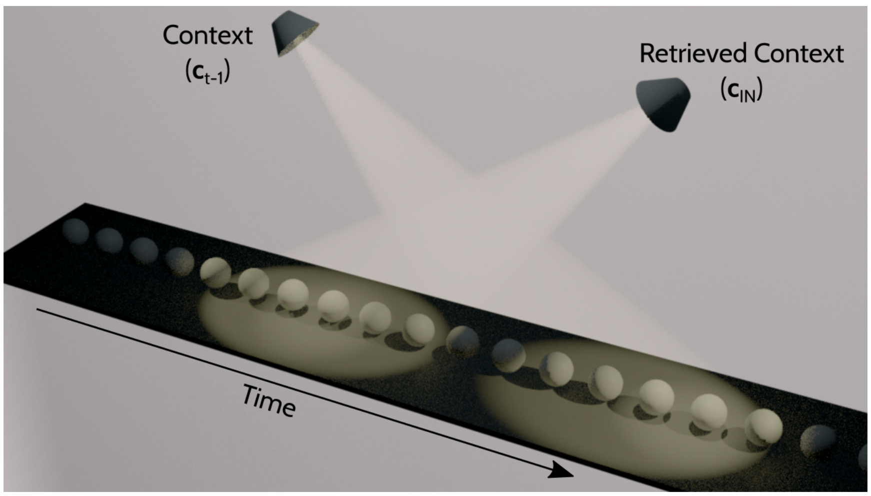
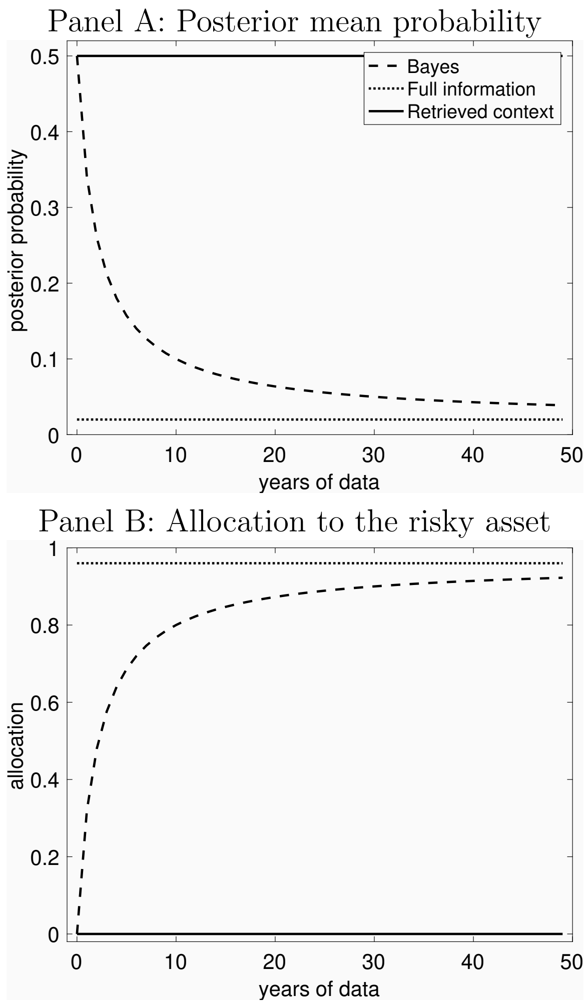


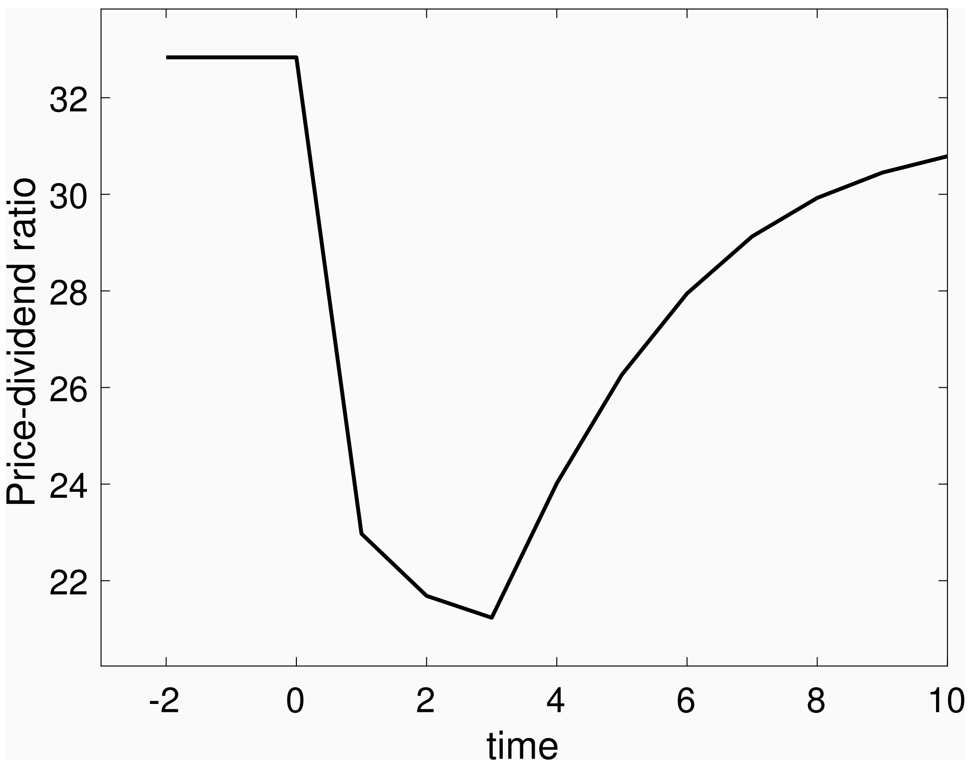
Figure 2: Retrieved Context and the spotlights of memory. In this illustration, memories appear as circles on the stage of life. All experiences that enter memory, as gated by perception and attention, take their place upon the stage. Context serves as a set of spotlights, each shining into memory and illuminating its associated features. The prior state of context c_{t-1} illuminates recent memories, whereas the context retrieved by the preceding experience, c^{IN}, illuminates temporally and semantically contiguous memories. Due to the recursive nature of context and the stochastic nature of retrieval, the lamps can swing over time and illuminate different sets of prior features.

Figure 3: Posterior probability and asset allocation as a function of sample length



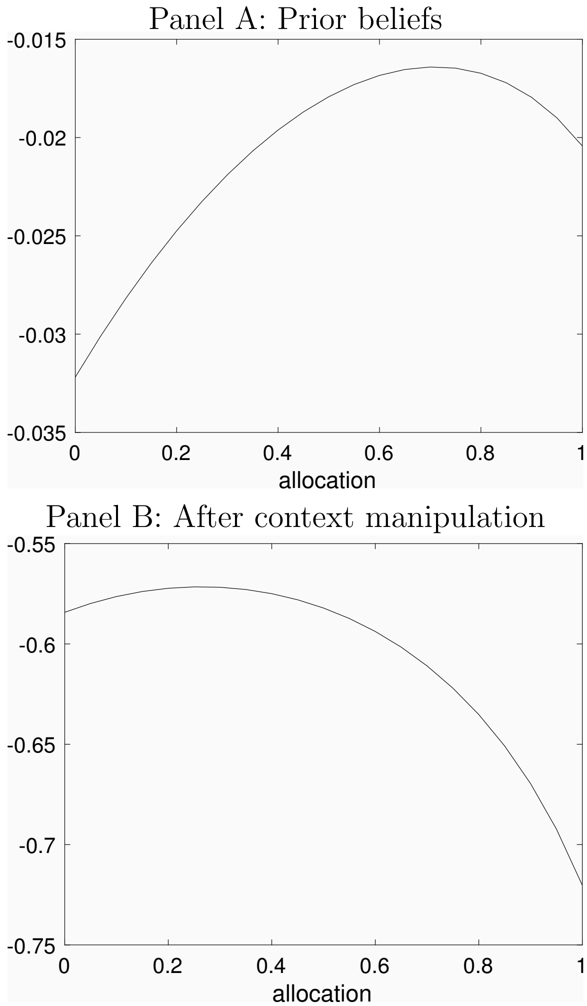
Notes: The figure shows posterior mean of the probability of a depression (Panel A) and the resulting asset allocation (Panel B) for the model presented in Section 4.1.

Figure 4: A jump back in time: the price-dividend ratio



Notes: The figure shows the equilibrium ratio of prices to dividends on the aggregate market (in the model of Section 4.2), assuming the agent starts in the fully neutral context and observes three periods of crisis features, followed by neutral features.

Figure 5: Expected utility under context manipulation



Notes:Expected utility as a function of allocation to the risky asset in the model of Section 4.3. Panel A shows utility prior to treatment by viewing a horror movie. Panel B shows utility after context has been manipulated by introducing a feature suggestive of danger (specifically, a horror movie).