A retrieved-context theory of financial decisions

Jessica A. Wachter* Michael Jacob Kahana†

February 8, 2019

Work in progress

Please do not circulate without permission

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 the context framework from the memory literature 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.

*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 in the 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 in reality is impossibly large and ever-changing in way that could seems impossible to capture by low-dimensional models in use in economics and elsewhere. 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 association between experiential features of the environment, and underlying contextual states. 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.

In this paper, we use these concepts to address puzzles in asset pricing and portfolio choice that defy the standard Bayesian paradigm. Chief among these are the result that

[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).

life experience has near-permanent effects on financial decision making (Malmendier and 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 show that the framework can be applied to a broader set of phenomena, such as the appearance, ex ante out of nowhere (though of course ex post there are many explanations) of the 2008 financial crisis and the phenomenon in the cross-section of stock returns known as momentum.

We present examples of how principles of memory might inform financial puzzles, rather than a general framework of decision making and memory. Nonetheless, certain concepts and notation are common to all of the examples. These are a space of *features vectors* and a space of *context vectors*.

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. A features vector, then is an element of $\mathcal{B}^n \subset R^n$, where $\mathcal{B}^n$ is a set of basis vectors that spans n-dimensional space. One could have, e.g. non-exclusive event states n. If there were two possible outcomes, the features vector would be in a space of size 2^n. 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 persistent (usually) and endows the agent with an understanding of possibly latent aspects of the environment that are relevant for the decision at hand. Context may also be subjective. Specifically, define the *context space* as the standard simplex $\mathcal{A}_c \subset R^m$, for $m \leq n$. That is $\mathcal{A}_c = \{c_t = [c_{1t}, \ldots, c_{mt}]^\top \in R^m \,|\, \iota^\top c_{it} = 1\}$, 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.

Our work is related to recent research incorporating memory laws into economic decisions. Bordalo et al. (2017) develop a model based on a concept known as geometric similarity. They focus on the ability of geometric similarity to account for the strange 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, the notion of context is embedded in the environment, and thus is static; the feature layer of the environment and the context layer are the same.

Whereas many applications of psychological principles to economic decision making have focused on cognitive biases such as loss aversion and narrow framing, 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.

2 Retrieved context and the persistence of beliefs

This section builds a simple model of portfolio choice under context reinstatement. It shows how events in the distant past can influence decisions, even in the presence of a potentially infinite amount of data. The logic follows well-established principles from psychology (Lohnas et al., 2015). The motivating example is that of Malmendier and Nagel (2011), who show that those who experienced the Great Depression are less likely to hold equities, years into the future.

2.1 The economic problem

Consider an agent solving a one-period portfolio choice problem. The agent receives risky labor income $\tilde{y}$ and chooses between a bond, earning return of zero, and a risky stock, earning net return of $\tilde{r}$. Labor income $\tilde{y}$ can take on two possible values, $y(1) > y(2)$. We will interpret labor-income-state 1 as a normal state, and labor-income-state 2 as a depression state. The stock return can take on two possible values $r(1) > r(2)$. We will think of $r(1)$, as an increase in stock prices relative to the mean and $r(2)$ as a decrease. Thus $r(1)$ and $r(2)$ represent normal fluctuations in the stock market. The agent also has access to a riskless asset, which we will assume has net return zero.

There are thus 4 states. Consider a 2×2 matrix P of joint probabilities. That is

$$P = \begin{bmatrix} \text{Prob(up \& normal)} & \text{Prob(down \& normal)} \\ \text{Prob(up \& depression)} & \text{Prob(down \& depression)} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} - p \\ 0 & p \end{bmatrix}.$$

where $p \in [0, \frac{1}{2}]$. We make the plausible assumption that the occurrence of depression would be reflected in stock prices, and thus an up market and a depression cannot possibly coincide.

The laws of probability imply that p is the unconditional probability of a depression, whereas (unconditionally), up and down market states occur with equal probability

(this assumption is made for simplicity and is in no way essential). Note that the requirement of $p < \frac{1}{2}$ ensures that the depression is the unusual, not the normal state. An increase in p (up to the limit of 1/2), increases uncertainty.

To best illustrate the mechanism, we assume the simplest possible setting in which the agent trades of risk and return: quadratic utility. Let $\tilde{X}$ denote the agent's wealth. The agent solves

$$\max_{\theta \geq 0} E^*[\tilde{X}] - \frac{1}{2}\text{Var}^*(\tilde{X})$$

subject to

$$\tilde{X} = 1 + \theta\tilde{r} + \tilde{y},$$

where $*$ is used to denote moments taken under the agent's subjective probability measure. The problem has solution

$$\theta = \frac{E^*\tilde{r} - \text{Cov}^*(\tilde{r}, \tilde{y})}{\text{Var}^*(\tilde{r})}. \tag{1}$$

2.2 Memory

Given an integer point in time t, let c_t be a 2 dimensional column vector denoting the "context" and f_t a two dimensional column vector denoting "features." We will identify f_t with realizations of the stock market, so that $f_t = [1, 0]^\top$ represents the "up" state, and $f_t = [0, 1]^\top$ represents the "down" state. We will assume c_t represents the unobserved labor market state. $c_t = [1, 0]^\top$ implies 100% probability on no depression in the labor market.

Define a "features to context" matrix as

$$W_t^{f \to c} = \frac{1}{t + \tau} \sum_{s=-\tau}^{t} c_s f_s^\top, \tag{2}$$

where τ represents the length of the prior sample. The formulation (2) implies that

the features-to-context matrix captures the agents' experienced joint contingencies of labor-market and stock-market states. We can think of (2) therefore as the prior beliefs of the agent.

For τ sufficiently large, it is clear that

$$W_0^{f\to c} = P \tag{3}$$

However, consider a set of distorted beliefs about the probability of the Great Depression.

$$P^* = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} - p^* \\ 0 & p^* \end{bmatrix},$$

where, as with p, $p* \in [0, \frac{1}{2}]$. We will assume the agent's prior is represented by these distorted beliefs:

$$W_0^{f\to c} = P^*. \tag{4}$$

Under Bayesian learning,

$$\text{plim}_{t\to\infty} W_t^{f\to c} = P.$$

With a large sample size, the agent behaves like one who began with the correct priors.

However, now suppose the agent retrieves the context from the features. Let

$$c_t = \frac{1}{\iota^\top W_{t-1}^{f\to c} f_t} W_{t-1}^{f\to c} f_t, \tag{5}$$

be the *retrieved context*, where ι is a column vector of ones. The term $\frac{1}{\iota^\top W_{t-1}^{f\to c} f_t}$ ensures that the elements of c_t sum to one.

Continue to assume (4). Note that

$$P^* \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix},$$

so that, at time 1, if an "up" move happen, the retrieved context will put 100% weight on the normal labor income state. Furthermore,

$$P^* \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - p^* \\ p^* \end{bmatrix} \propto \begin{bmatrix} 1 - 2p^* \\ 2p^* \end{bmatrix} \tag{6}$$

so that if a "down" move happens, the retrieved context (5) places weight $2p^*$ on a depression.

We now calculate the effect on the $W^{f \to c}$ matrix. Note that

$$W_t^{f \to c} = \frac{\tau}{t + \tau} W_0^{f \to c} + \frac{t}{t + \tau} \left(\frac{1}{t} \sum_{s=1}^{t} c_s f_s^\top \right). \tag{7}$$

That is, posterior probabilities are a weighted average of probabilities from the prior sample and the current sample. If, at time s, an "up" move occurs in the stock market,

$$c_s f_s^\top = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1, 0] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{8}$$

When a "down" move occurs, the depression context is reinstated (see Equation 6), and

$$c_s f_s^\top = \begin{bmatrix} 1 - 2p^* \\ 2p^* \end{bmatrix} [0, 1] = \begin{bmatrix} 0 & 1 - 2p^* \\ 0 & 2p^* \end{bmatrix}. \tag{9}$$

Let k denote the number of "up" moves and $t - k$ the number of "down" moves since time 0. To $W_0^{f \to c}$, we add k elements of the form (8) and $t - k$ elements of the form (9), appropriately scaled. Thus (7) implies

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

Then

$$\text{plim}_{t\to\infty} W_t^{f\to c} = P^*.$$

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

The original matrix over-associates a downturn with a depression due to prior beliefs. What (5) does is recall the depression with every down realization. The act of recalling the depression reinstates the depression context, or rather (5) recalls a context in which there is a high probability of depression. Thus high probability of depression remains associated with a down market in the mind of the agent.

Interestingly, if the agent happened to arrive at the correct probabilities at the beginning, the updating rule (5) would have produced the correct probabilities P.

2.3 Implications for decision making

Let $\tilde{r}(1) = 1 + \sigma$, $\tilde{r}(2) = 1 - \sigma$, for $\sigma > 0$, $\tilde{y}(1) = y > 0$ and $\tilde{y}(2) = 0$. Then. $E^*\tilde{r} = 1$, $\text{Var}^*(\tilde{r}) = \sigma^2$. Note that $\tilde{y}$ is a Bernoulli process multiplied by a constant y, so that

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

Direct calculation implies

$$\text{Cov}^*(\tilde{r},\tilde{y}) = E^*\left[(\tilde{r} - E^*\tilde{r})\tilde{y}\right] = \frac{1}{2}\sigma y - \left(\frac{1}{2} - p^*\right)\sigma y = p^*\sigma y,$$

so that (1) implies

$$\theta = \frac{1 - p^* y\sigma}{\sigma^2}. \tag{10}$$

For $p^* \geq p$, the agent's allocation is strictly lower than is optimal. That is, the mechanism outlined here prevents the agent from learning that stocks are relatively

safe investments. Their remains a wedge between the subjective distribution and the physical one, leading to persistently low allocation to equities.

Figure 1 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 $y = 2$. The Bayesian investor has an uninformative, but still well-defined, prior that assumes a sample size of 2.[2] Bayesian analysis gives a distribution for the probability of a depression; Figure 1 shows the mean of the posterior distribution, which is all that is required to compute (1), 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 knows the probability for certain; in this economic problem, parameter uncertainty has no effect.

Specifically, Figure 1 shows that 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. Bayesian updating is therefore inconsistent with the finding in Malmendier and Nagel (2011) that lifetime experience has significant effects on portfolio choice. Even in the case where it is a rare event that is learned about, 20 years of common data suffice to render beliefs nearly the same. On the other hand, the agent for whom context is retrieved does not learn, and can maintain an incorrect probability even in the face of many years of evidence.

[2]Specifically, the prior is Beta$(p^*\tau, \tau)$, where τ is the prior sample size. Given the likelihood implied by Bernoulli observations on $\tilde{y}$, the posterior is also Beta.

3 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.[3]

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 intermediation in the economy. Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2013) develop models in which the balance sheets of intermediaries are responsible for business cycle fluctuations. However, while it may be necessary to have specialized institutions trade certain complicated investments, it is not clear why common stocks are necessarily intermediated assets: trading costs for common stocks, already quite low for the past half-century, have only gotten lower. Another possibility is that Lehman represented a sunspot that caused a run on other intermediaries, and in fact on all debt. However, why should this cause a broad-based stock market decline, when most companies have very low leverage and can fund themselves through retained earnings?[4].

A third possibility, introduced by Gennaioli and Shleifer (2018), is that individuals and banks took on too much debt because they incorrectly extrapolated that the environment was risky. This debt created unstable conditions. They model the idea that that 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 the model below, the risk is wholly illusory. While this is extreme, in fact, the Great Recession was nothing like the Great Depression.[5] The

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

[4]Kahle and Stulz (2013) argue that firms dependent on bank-lending were not unduly affected by the crisis. Gomes et al. (2018) argue that fluctuations in borrowing conditions are more likely to be affected by investment opportunities than the other way around.

[5]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.

realized outcome does not seem commensurate with the panic in the fall of 2008.

Our hypothesis is that the financial crisis was a psychological event caused by the failure of Lehman Brothers. The specter of a important financial institution failing in the absence of insurance reminded investors of the Great Depression. [6] 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., 2011). Once this feeling entered the discourse, it proved hard to shake. Only when subsequent events showed that there was no Great Depression, and this caused the stock market to rise. The model below formalizes this intuition.

As before, we consider a 2×2 matrix P of joint probabilities. That is

$$P = \begin{bmatrix} \text{Prob(no crisis \& no depression)} & \text{Prob(crisis \& no depression)} \\ \text{Prob(no crisis \& depression)} & \text{Prob(crisis \& depression)} \end{bmatrix}$$

We identify features f_t with a financial crisis (signified, e.g. by the failure of a major financial institution). We identify the context c_t with whether or not the economy is in a disaster state. Having seen 100 or so years of data, with roughly one disaster, the agents' estimate of P, through the features-to-context matrix, looks like

$$W_{-1}^{f \to c} = \frac{1}{\tau} \sum_{s=-\tau}^{-1} c_s f_s^\top \approx P \tag{11}$$

and

$$P = \begin{bmatrix} (1-p^c)(1-\epsilon) & p^c(1-q) \\ (1-p^c)\epsilon & p^c q \end{bmatrix},$$

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

where

$$\begin{aligned} p^c &= \text{probability of a financial crisis} \\ q &= \text{probability of an economic disaster, given a financial crisis} \\ \epsilon &= \text{probability of an economic disaster, given no financial crisis.} \end{aligned}$$

Roughly speaking, we think of $p^c \approx 1/20$, $q \approx 1/2$, $\epsilon \approx 1/500$.

Note that the sum in (11) contains mostly elements of the form of

$$c_s f_s^\top = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1, 0]$$

(representing neutral times) and some elements of the form (representing the Great Depression)

$$c_s f_s^\top = \begin{bmatrix} 0 \\ 1 \end{bmatrix} [0, 1],$$

and sufficiently few mixed elements that we can ignore them. Another view is that agents come to associate the Great Depression with bank failures based on historical and fictionalized accounts that emphasized the role of banks.[7] Thus, in what follows, we will simplify the analysis and consider $\epsilon = 0$ so that

$$W_0^{f \to c} = \begin{bmatrix} 1 - p^c & p^c(1-q) \\ 0 & p^c q \end{bmatrix}. \tag{12}$$

We assume the a generalized rule for updating context found in Lohnas et al. (2015):

$$c_t = \rho c_{t-1} + \zeta c_t^{\text{in}}, \tag{13}$$

[7]For example, "Its a Wonderful Life" among others.

where $\zeta = 1 - \rho$, and where c_t^{in} is retrieved from the features to context matrix:

$$c_t^{\text{in}} = \frac{1}{\iota^\top W_{t-1}^{f\to c} f_t} W_{t-1}^{f\to c} f_t. \tag{14}$$

we assume that $\rho \in [0, 1)$, so that context gradually decays. Note that context is a weighted-average of past context and what can be interpreted as a "shock" to context, c_t^{in}.[8] We assume that mostly neutral features have been observed, namely $f_t = [1, 0]^\top$. Neutral features imply the neutral context

$$c_t^{\text{in}} = \frac{1}{\iota^\top W_{t-1}^{f\to c} \begin{bmatrix} 1 \\ 0 \end{bmatrix}} W_{t-1}^{f\to c} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Given sufficiently many observations of neutral features, context reaches a steady state value of $[1, 0]^\top$. This can be seen by setting $c_t = c_{t-1}$ in (13).[9]

As in the neglected risk model (Gennaioli and Shleifer, 2018), the depression state has no weight in the context vector. However, this does not imply that the depression state is forgotten. Representing the failure of Lehman brothers is $f_0 = [0, 1]^\top$, the well-publicized failure of a major financial institution. Given this value of f_0, we have

$$c_0^{\text{in}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

[8] Appendix B discusses the precise relation between the evolution implied by (13) and (14) and the assumptions in Lohnas et al. (2015).

[9] There are two points of contrast between this model and the previous one. Here, features change rarely, whereas in the previous example, they change frequently. Adding context dynamics of the form (13) to the model in the previous section would greatly complicate the analysis with little ultimate effect.

The agent's context is then

$$c_0 = \rho \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \zeta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \rho \\ \zeta \end{bmatrix}.$$

The investor's weight on the depression context has changed from 0 to ζ.

We assume, for simplicity, that all investors are identical and all price assets treating c_t as a vector of probabilities. We assume that the agent views the probabilities in c_t as permanent. That is, agents are not forward-looking in that they do not understand that their beliefs given a probability of disaster evolve according to (13).

Let D_t be the aggregate dividend and S_t the aggregate stock price. Then the ratio of the price to the dividend is given by

$$\frac{S_t}{D_t} = \frac{\Phi(p_t)}{1 - \Phi(p_t)}$$

where

$$p_t = c_t^\top \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

namely, the probability of a depression. The function $\Phi(p_t)$ is a decreasing function of p_t under reasonable parameter values (see Appendix A).

We calibrate the model assuming $p_C = .05$, $q = 0.5$, a decline in aggregate consumption of 40%, risk aversion of 1 (log utility) and leverage of 3. Here, and in the applications that follow, we assume $\rho = 0.8$, in the range of values typically found in the memory literature (Lohnas et al., 2015).

See Appendix A for more detail. Panel A of Figure 2 shows the time path of the price-dividend ratio, assuming a jump-back-in-time (JBIT) event. Prior to the event, prices are higher than in the full-information case, due to neglected risk. The introduction of the crisis features leads to an immediate decline followed by recovery

over the next decade, as only neutral features are observed. Panel B shows the real riskfree rate. The JBIT causes an immediate decline in the riskfree rate, which then rises slowly over the subsequent ten years. Both features roughly match the behavior of stock prices and interest rates during and after the financial crisis.

The memory of the crisis event decays approximately exponentially, as given by the context evolution equation (13). The decay in the memory of the JBIT is not exactly exponential because of the updating of the features-to-context matrix. Indeed, beliefs converge after about 10 periods to a new steady-state that are above the full-information case – again exhibiting neglected risk. The new steady-state is below the initial beliefs, however. The agent remembers, permanently, that a depression context can be accompanied by neutral features, simply because of the act of recall.

4 Sticky context as an explanation for price momentum

The model presented in Section 3 provides a natural explanation of the momentum effect in stock returns. Momentum, noted by Jegadeesh and Titman (1993), is the finding that stocks with the best past price performance outperform, over the next year, stocks with the worst performance, by a wide margin. This stocks with the best performance are known as "winners" with the others as "losers." When sorted into deciles, the WML (winners-minus-losers) portfolio, formed on the extreme deciles, generates an annual return of about 11%.

We assume a simple neoclassical production model in which productivity, earnings, and dividends all covary up to a scale factor. Firms can experience high or low productivity growth. They can also experience high or low observed earnings. The joint contingency matrix, in which observed earnings growth varies along the columns, and

the true growth state varies along the rows is:

$$P = \begin{bmatrix} \Pr\left(\frac{E_{t+1}}{E_t} = g_H \ \& \ g = g_H\right) & \Pr\left(\frac{E_{t+1}}{E_t} = g_L \ \& \ g = g_H\right) \\ \Pr\left(\frac{E_{t+1}}{E_t} = g_H \ \& \ g = g_L\right) & \Pr\left(\frac{E_{t+1}}{E_t} = g_L \ \& \ g = g_L\right) \end{bmatrix},$$

where we use the shorthand E_{t+1}/E_t to denote growth in earnings from one period to the next.

These growth rates are permanent, and, for simplicity, there is no noise, so high observed growth is uniquely associated with high true growth.[10] There is a measure of p of high growth firms. Therefore

$$P = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \tag{15}$$

We associate context with the long-run productivity growth rate of the firm, and features with high or low observed earnings growth. Investors observe a large sample of firms over a long time period. They store the relation between features and context in a stable features-to-context matrix, in which the high growth context is overwhelmingly associated with high observed growth, and the low growth context with low observed growth. This is assumption will greatly simplify our analysis below.[11] That is

$$W^{f\to c} = \sum_{t=-\tau}^{0} \sum_{i=1}^{N} c_{it} f_{it}^{\top} \approx P. \tag{16}$$

where i indexes firms and t indexes time, and where (as before), without loss of generality, we assume we are currently at time 0.

Now consider a set of firms indexed by j ($j = 1, \ldots, N$) for which investors do not

[10] Useful generalizations would be to assume a Markov switching model, and/or have an iid shock (or potentially measurement error) create a wedge between observed and true earnings. For our basic qualitative result, this model is sufficient however.

[11] Accounting for short-term context dynamics in the $W^{f\to c}$ matrix would complicate the analysis without changing the conclusions.

yet know the growth rate. At time t, some of these firms will be assigned the high growth rate, while the complement will be assigned the low growth rate. Those that are assigned the high growth rate are "winners" (in the sense of Jegadeesh and Titman (1993)) whereas those assigned the low growth rate are "losers." This terminology will be justified shortly. The total set of firms is sufficiently small so that the estimation of the $W^{f\to c}$ matrix is unaffected.

Assume that, as in the model in Section 3, investors do not accurately forecast the time path of context. Rather, they price assets using the context vector to assign probabilities to permanent states. Here, we also assume risk-neutral investors and an exogenous discount rate r, such that $r > g_H > g_L$. If investors assign subjective probability $\tilde{p}_{jt}$ at time t to firm j having high growth, then the price-dividend ratio for firm j is given by

$$\begin{aligned}
\frac{P_{jt}}{D_{jt}} &= \tilde{p}_{jt} E\left[\sum_{s=1}^{\infty}(1+r)^{-s}\frac{D_{j,t+s}}{D_{jt}} \,|\, g = g_H\right] + (1-\tilde{p}_{jt}) E\left[\sum_{s=1}^{\infty}(1+r)^{-s}\frac{D_{j,t+s}}{D_{jt}} \,|\, g = g_L\right] \\
&= \tilde{p}_{jt}\sum_{s=1}^{\infty}\left(\frac{1+g_H}{1+r}\right)^s + (1-\tilde{p}_{jt})\sum_{s=1}^{\infty}\left(\frac{1+g_L}{1+r}\right)^s \\
&= \tilde{p}_{jt}\frac{1+g_H}{r-g_H} + (1-\tilde{p}_{jt})\frac{1+g_L}{r-g_L}. \qquad (17)
\end{aligned}$$

We assume, for all firms j, that

$$c_{j0} = \begin{bmatrix} p \\ 1-p \end{bmatrix}$$

Winning firms have high productivity growth rates, and so experience $f_{j1} = [1,0]^\top$. Losing firms have low productivity growth rate, and so experience $f_{j1} = [0,1]^\top$. It follows from the features-to-context matrix (16) that $c_{j1}^{\text{in}} = f_{j1}$ for all j. Applying (13),

winners have time-1 context of

$$c_{j1} = \begin{bmatrix} \rho p + 1 - \rho \\ \rho(1-p) \end{bmatrix}.$$

Because $\rho p + 1 - \rho > p$, the subjective probability of high growth has increased. Losers have time-1 context of

$$c_{j1} = \begin{bmatrix} \rho p \\ \rho(1-p) + 1 - \rho \end{bmatrix}.$$

Because $\rho p < p$, the subjective probability of high growth has decreased.

Finally note that for stock j, the price change is given by

$$\frac{P_{j1}}{P_{j0}} = \frac{D_{j1}}{D_{j0}} \frac{P_{j1}/D_{j1}}{P_{j0}/D_{j0}}. \tag{18}$$

It follows from (17), that the price-dividend ratio is an increasing function of $\tilde{p}$. Moreover, high growth firms also have higher realized values of dividend growth $\frac{D_{j1}}{D_{j0}}$. The price change (18) is therefore higher for winners than for losers, justifying the terminology.

The high-growth features of time-1 is what makes the winning stocks "winners" and the losing stocks "losers." It is not surprising that positive earnings surprises should lead investors to update their probabilities of a stock having high long-run earnings growth. The key implication of context dynamics (13) is that this updating is incomplete.

Indeed, in subsequent periods, context puts ever-increasing weight on the high-growth state for winning firms and ever-decreasing weight for losing firms. Eventually the context vectors converge, with convergence being faster for lower values of ρ. When convergence has been reached, returns on the two firms are the same. Before this limit, however, returns on winning firms will always be above returns on losing firms.

Figure 3 shows the prices and returns on winners and losers and constrasts the implications of the retrieved context model with Bayesian updating (in this case, Bayesian updating is identical to full information). The key assumption in this model that differentiates it from Bayesian updating is the slow evolution of context implied by (13). In the simple model outlined above, a Bayesian investor should update immediately to 100% probability in the high growth state upon observing one observation. More generally, beliefs under Bayesian updating should follow a martingale. One cannot have positive "surprises" systematically following positive events. Yet that is what the data appear to show (Chan et al., 1996).

From a psychological perspective, the reason updating is slow is that c_0, representing a low probability of a high growth state is still in investors' context when they evaluate the new information. Repeated observations are required before c_0 disappears from context.

5 Fear and asset allocation

Guiso et al. (2018) observe that watching a horror movie influences the risk premium investors require to hold risky assets. This is also been observed in Cohn et al. (2015). What is striking about this experiment is that fear alone, as opposed to new information, has a substantial effect on risk taking. We hypothesize that fear operates through the memory channel.

Note that this model allows for both short-run effects (because of the autoregressive structure) and highly persistent effects (through the features to context matrix) of negative associations.

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 the 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.

In the model in Section 3, agents update their probabilities of a Great Depression. Because stocks are sensitive to such an event, this results in a decline in stock valuations. In this section, as in Section 3, we wish to explain an updating of probabilities in response to a stimulus. However, unlike the setting in Section 3, the agents know that the underlying probability distribution has not changed.

As in Section 2, we assume that the agent is endowed with one unit of human capital, and one unit of financial wealth. The agent invests financial wealth in a risky and a riskless asset. Again, we assume the riskless asset return to be zero. The agent therefore solves

$$\max_{\theta} Eu(1 + \tilde{y} + \theta\tilde{r}), \tag{19}$$

where

$$\tilde{y} = \begin{cases} 1 & \text{with probability } 1 - p \\ 1 - \tilde{\xi} & \text{with probability } p, \end{cases}$$

where $\tilde{\xi}$ is a random variable with support on $[0, 1]$.

We assume that most of the mass of $\tilde{\xi}$ is centered on values large in magnitude relative to realizations of $\tilde{r}$. Thus a nonzero outcome of $\tilde{x}$ represents a disaster state, similar to the Great Depression. For example, $\tilde{\xi}$ might equal 80% while $\tilde{r}$ might have

an upper bound of 10%.[12]

We further assume the form

$$u(x) = \frac{x^{1-\gamma}}{1-\gamma} \tag{20}$$

in which it will be helpful for the calibration to have $\gamma \geq 1$, so that risk aversion is greater than or equal to what one finds for log utility, and also that the agent endogenously avoids outcomes associated with negative consumption. Note that this utility function implies decreasing absolute risk aversion, a helpful property in explaining the Guiso et al. (2008) results.

A direct interpretation of (19) is that $\tilde{y}$ represents background labor market risk as in Section 2. Other interpretations of $\tilde{y}$ are possible. Under the specification (20), $\tilde{y}$ shifts the curvature of the utility function. Viewing (19) as a utility over $\theta\tilde{r}$, the agent exhibits decreasing *relative* risk aversion, and thus implies time-varying risk aversion. One interpretation is that emotions associated with context enter the utility function (as in Loewenstein (2000)). Another, related, interpretation is that the agent recalls a different risk aversion depending on the context.

The probability p is determined by the context vector. Assume $c_t = [1, 0]^\top$ corresponds to zero risk probability $c_t = [0, 1]^\top$ is probability 1 of risk. The matrix P representing the joint contingencies equals

$$P = \begin{bmatrix} \Pr(nr, nd, nc) & \Pr(nr, d, nc) & \Pr(nr, d, c) \\ \Pr(r, nd, nc) & \Pr(r, d, nc) & \Pr(r, c, d) \end{bmatrix}, \tag{21}$$

where r stands for risk nr stands for no risk, etc. We let p be the unconditional probability of danger, and assume, for simplicity, that risk and danger always co-occur,

[12] Consider that the 10,000 euros offered to participants by Guiso et al. (2008) is likely to be small relative to the fear of the negative outcome evoked by the movie.

so that

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

We assume context follows (13) and (14). Without loss of generality, label the time that the movie occurs as $t = 0$, so $t = -1$ refers to the context prior to the experiment. We assume that the agent has accurately observed and recalled a sufficiently long sample, so that

$$W_{-1}^{f \to c} = P$$

As in the previous example, this example features neglected risk.

As in Section 3, we assume that only neutral features, namely $f_t = e_1$ have been observed for sufficiently long for context to converge to a stationary value:[13]

$$c_{-1} = [1, 0]^\top.$$

Because neutral features are incompatible with danger, the agent puts zero weight on the danger state.

The horror movie represents $f_0 = e_2$. Namely, the horror movie evokes a sense of danger without crisis. We have

$$c_0^{\text{in}} \propto W_{-1}^{f \to c} f_0 = W_{-1}^{f \to c} e_2$$

which implies

$$c_0^{\text{in}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

[13]This convergence is implied by (13) given a long string of identical features vectors f_t. See Section 3 for further discussion.

Therefore, the new context is:

$$c_0 = \rho c_0 + \begin{bmatrix} 0 \\ \\ \zeta \end{bmatrix}$$

While the agent previously had put zero weight on the high-risk state, now the probability rises to 20%.

Figure 4 shows (19) as a function of θ, prior to the stimulus, when the agent exhibits neglected risk, after the stimulus, when the agent is transported into a context that has a 20% weight on the risk state, and the case when the agent has the correct beliefs. The agent with neglected risk has an optimum outside the [0,1] range. After the stimulus, this optimum shifts into the interior, below that for the full information case. This shift accounts for the finding in Guiso et al. (2018).

6 Conclusion

To be completed.

A 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

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

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

$$\log D_{t+1} = \log D_t + \mu + u_{t+1} + \lambda v_{t+1} \tag{22}$$

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

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

so that

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

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

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

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

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

It is convenient to define the notation

$$\Phi(p) = \beta e^{(1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2} \left(e^{-p} + (1-e^{-p}) E\left[(1-\xi)^{\lambda-\gamma}\right]\right) \tag{23}$$

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

Taking the log of both sides of (23) gives a convenient approximation

$$\begin{aligned}\log \Phi(p) &= \log \beta + (1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2 + \log\left(e^{-p} + (1-e^{-p})E\left[(1-\xi)^{\lambda-\gamma}\right]\right) \\ &\approx -\log \beta + (1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2 - pE\left[1-(1-\xi)^{\lambda-\gamma}\right]. \qquad (24)\end{aligned}$$

Note that $1-\xi \in (0,1]$. Thus the term inside the expectation in (24) 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 (22). Let P_t denote the price of this claim. The condition for equilibrium, applied to this claim implies

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

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

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

This equation is solved by

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

The model is calibrated using $\mu = 0$, $\sigma = 2\%$, $\beta = e^{-.03}$, $\gamma = 1$, $\lambda = 3$.

B Two approaches to scaling

To be completed...

References

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. (2017). 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.

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 Maréchal, M. A. (2015). Evidence for countercyclical risk aversion: An experiment with financial professionals. *The American Economic Review*, 105(2):860–885.

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.

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.

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.

Gomes, J. F., Grotteria, M., and Wachter, J. A. (2018). Cyclical dispersion in expected defaults. *The Review of Financial Studies*, page hhy085.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the stock market. *Journal of Finance*, 63(6):2557–2600.

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.

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

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

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

Lohnas, L. J., Polyn, S. M., and Kahana, M. J. (2015). Expanding the scope of memory search: Modeling 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.

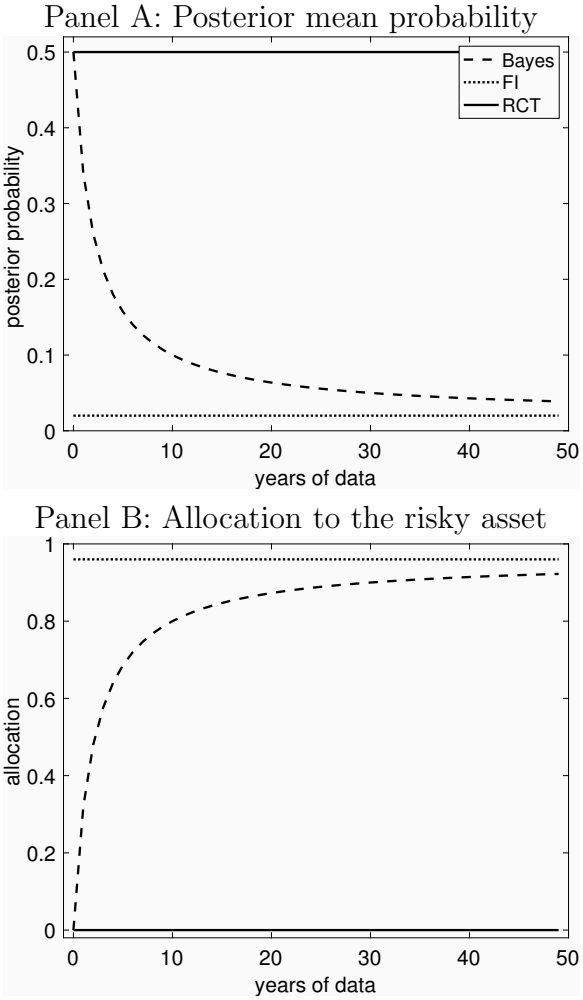
Malmendier, U. and Shen, L. S. (2018). Scarred consumption. NBER Working Paper No. 24696.

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

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

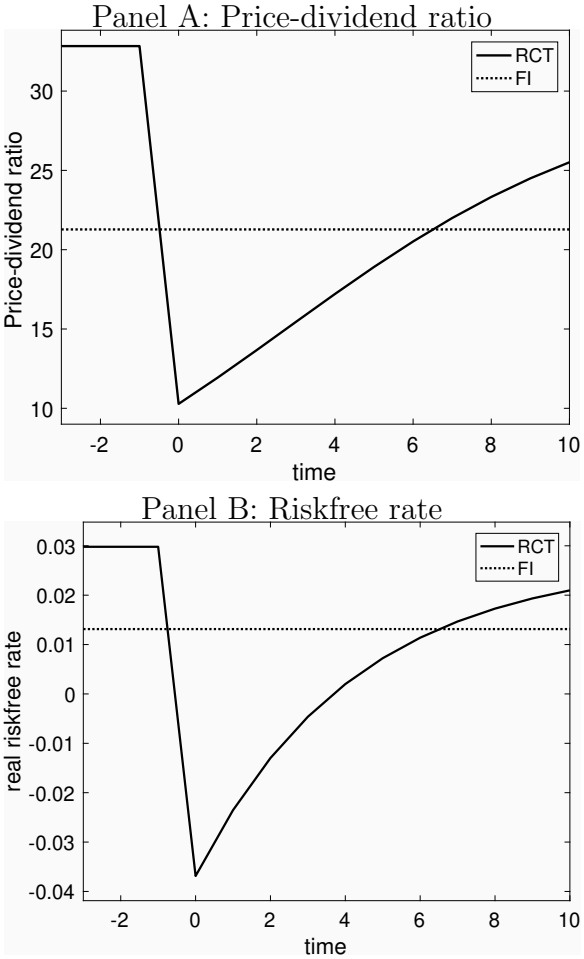
Savage, L. (1954). *The Foundations of Statistics*. Wiley, New York.

Figure 1: Posterior probability and asset allocation as a function of sample length



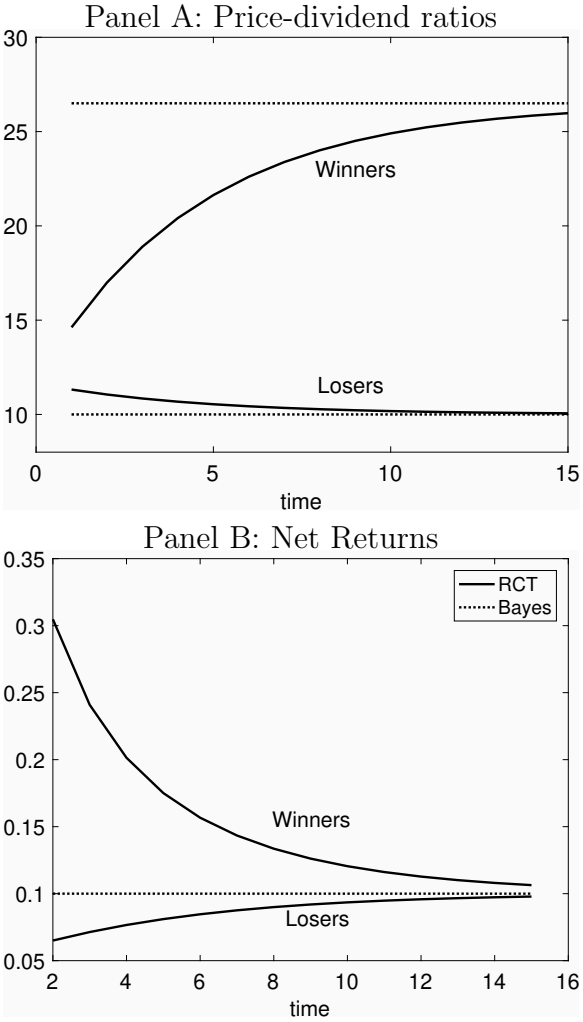
Notes: The figure shows posterior mean of the probability of a depression (Panel A) and the resulting asset allocation (Panel B) for the model presented in Section 2. FI refers to full-information. RCT refers to Retrieved context theory.

Figure 2: A jump back in time



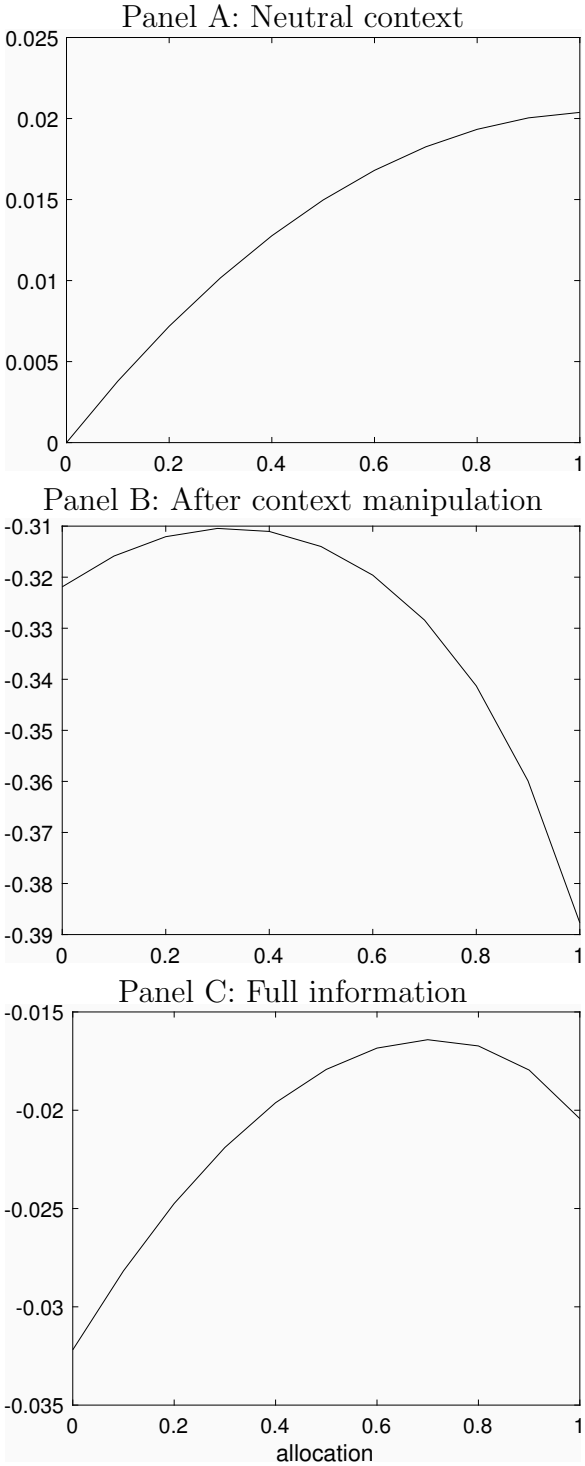
Notes: Panel A shows the aggregate price-dividend ratio pre-and post-banking crisis as a function of time since the crisis (the crisis is assumed to occur at time 0) for the model in Section 3. Panel B shows the real riskfree rate.

Figure 3: Prices and returns in the model for momentum



Notes: The figure shows price-dividend ratios (Panel A) and returns (Panel B) for the model presented in Section 4. RCT refers to prices and returns under retrieved context theory. Bayes refers to prices and returns under Bayesian updating (returns on winners and losers are identical).

Figure 4: Expected utility under context manipulation



Notes:Expected utility as a function of allocation to the risky asset in the model of Section 5. Panel A shows utility under the neutral context. Panel B shows utility after context has been manipulated by introducing a feature suggestive of danger (specifically, a horror movie). Panel C shows expected utility under beliefs indicative of full information.