NBER WORKING PAPER SERIES

INFORMATION AVERSION

Marianne Andries
Valentin Haddad

ABSTRACT

The main features of households' attention to savings are rationalized by a model of information aversion, a preference-based fear of receiving flows of news. In line with the empirical evidence, information averse investors observe the value of their portfolios infrequently; inattention is more pronounced for more risk averse investors and in periods of low or volatile stock prices. The model also explains how changes in information frequencies affect risk-taking decisions, as observed in the field and the lab. Further, we find that receiving state-dependent alerts following sharp downturns improves welfare, suggesting a role for financial intermediaries as information managers.

Marianne Andries
Toulouse School of Economics
21 allee de Brienne
31000 Toulouse
France
marianne.andries@gmail.com

Valentin Haddad
University of California at Los Angeles
Anderson School of Management
Office C4.19
110 Westwood Plaza
Los Angeles, CA 90024
and NBER
vhaddad@ad.ucla.edu

1 Introduction

Experiencing the roller coaster of life can be stressful. A natural way to avoid it is to close your eyes for the ride. We show that a model of *information aversion* building on this idea rationalizes the main stylized facts about households' attention to savings. Namely, it explains why households observe the value of their portfolios infrequently, more so when they are more risk averse and in periods of low or volatile stock prices. In addition, we show the model is consistent with a large body of evidence in the field and the lab, on how changes in the information environment affect decisions. As such, our approach provides an empirically appealing preference-based model of inattention.

An aversion to information arises naturally from our choice of preferences: a recursive implementation of disappointment aversion (Gul, 1991). In this model, agents inflate the probabilities of disappointing outcomes, making them even more costly. Whenever information arrives, the agent readjusts her expectations, and runs the risk of being disappointed. Each piece of news is evaluated separately, under the recursive setup, so receiving fragmented information multiplies the scope for disappointment. To avoid this, the agent prefers observing information infrequently.

In a classical consumption and savings problem, we analyze how an information averse investor balances the endogenous utility cost of paying attention to her wealth with its benefits, mainly the opportunity to invest in high returns risky assets. We show that she optimally chooses to observe her risky portfolio infrequently. How disappointment averse she is governs both her aversion to risk and to information, so more risk averse agents are also more inattentive. Similarly, more risk makes information more costly, so households lower their observation frequencies in risky periods. This can happen even when expected returns, and the incentives to take risk, are high – a feature specific to our our model that allows it to rationalize the main stylized facts on investors' inattention to their savings.

Preference-based utility costs of information – information aversion – are key to the ability of our model to rationalize all these facts. They are consistent with observations of information avoidance in various economic settings, and with a large body of experimental evidence showing that, when given more fragmented information,

subjects are often less willing to participate in risky lotteries.

As an extension, we consider whether better flows of informations can be obtained with the help of a third party. We find simple state-dependent rules can help: receiving distress signals following sharp market downturns increases welfare. This indicates financial institutions can foster more investment by providing such signals, and proposes a new rationale for the role of intermediaries, as information stress managers; a view supported by the empirical evidence in Foerster et al. (2014).[1]

Our approach differs fundamentally from technological or cognitive limitations as a motivation for inattention: in our model agents desire to stay away from information. Understanding the sources of investors' information decisions is not only interesting on its own, but also because of their consequences. Investors' inattention creates portfolio inertia, as documented following fluctuations in income (Brunnermeier and Nagel, 2008), stock prices (Bilias et al., 2010), and wealth (Calvet et al., 2009); and leads to missed financial opportunities for households (Andersen et al., 2014). At the aggregate level, these slow-moving portfolio adjustments affect the demand for financial assets, and thus their pricing. Duffie et al. (1996, 1997) analyze how bond prices, in levels and in the term-structure, react to changes in information filtrations. Gabaix and Laibson (2002) show inattention has a profound impact on equilibrium asset prices.[2] By proposing a novel rationale for inattention and its dynamics, our model provides a foundation on which to re-evaluate these questions. More broadly, our approach could also naturally be applied to any economic context where agents choose how attentive they want to be.

After a literature review, Section 2 introduces our model and solves the investor's problem. We discuss the ability of the model to rationalize the empirical evidence on inattention in Section 3. Section 4 considers richer information structures and delegation. Mathematical proofs, and some extensions, are in the Appendix.

Related literature. Previous work in the literature studies inattention to savings as arising from exogenously specified costs of information. Duffie and Sun (1990) consider a portfolio problem similar to ours, assuming observations and transactions

---

[1]They find wealth accounts managed by advisors do not outperform passive investments, but have higher risky asset holdings and trading activity than those managed directly by investors.

[2]Lynch (1996) also studies the equilibrium implications of infrequent transactions, but without explicitly attributing them to inattention.

must be synchronized, and come at a fixed cost. Abel et al. (2007) derive optimal inattention periods and portfolio decisions under monetary costs of information. Abel et al. (2013) add transaction costs; Alvarez et al. (2012) durable consumption. In these models, the benefit of information is similar to our setting and optimal policies exhibit some similarities. However, our endogenous information costs have a different impact on the determination of inattention, which helps rationalize how it varies with market conditions, in particular why it can increase when index prices are low.[3]

Another strand of literature focuses on the implications of preference models on the role information plays. Benartzi and Thaler (1995) point out the frequency of observation matters for risk premia under loss aversion. In contrast to our model where it is an optimal choice, they consider this low frequency to be the result of another behavioral trait, myopia. Pagel (2017), in a contemporaneous paper, considers a consumption and savings problem under the news-utility theory of Kőszegi and Rabin (2009), who explicitly model flows of information as costly in the utility function. In contrast to ours, her framework is time inconsistent. There are many other ways in which information, irrespective of how it affects decisions, can enter preferences: e.g. curiosity, anticipatory feelings (Caplin and Leahy, 2001, 2004), optimal expectations (Brunnermeier and Parker, 2005), preferences for early or late resolution of uncertainty (Epstein and Zin, 1989), commitment (Carrillo and Mariotti, 2000; Benabou and Tirole, 2002). Our choice of preferences builds on the axiomatic characterizations of Dillenberger (2010) and Artstein-Avidan and Dillenberger (2011) to derive information aversion, and target inattention, by making only a simple deviation from expected utility, disappointment aversion

Gul (1991)'s disappointment aversion features a well documented (see Ericson and Fuster (2014)) behavioral trait concerning risk taking, first-order risk aversion – similar to loss aversion. It has proven useful in understanding risk-taking in financial markets: portfolio choices (Ang et al., 2005), equilibrium aggregate prices (Routledge and Zin, 2010; Bonomo et al., 2011), and the cross-section of expected returns (Ang et al., 2006; Lettau et al., 2013). However, none of this work accounts for the implications of disappointment aversion for information choices, our focus.

---

[3]Our finding that state-dependent rules can help also contrasts with the findings Abel et al. (2013).

2 Consumption-Savings under Information Aversion

We define dynamic disappointment aversion, and show it captures a notion of information aversion. We then analyze how investors trade off their fear of information with the benefits they can derive from it, in a consumption-savings framework.

2.1 Investor Preferences

Dynamic disappointment aversion. The first building block of our model is disappointment aversion (Gul, 1991). For a static lottery with payoff X, distributed according to F, the certainty equivalent is defined as:

$$\mu_\theta\left(X|F\right) = \frac{\int x dF\left(x\right) + \theta \int_{x < \mu_\theta(X|F)} x dF\left(x\right)}{1 + \theta \int_{x < \mu_\theta(X|F)} dF\left(x\right)},\tag{1}$$

where $\theta \geq 0$ is the coefficient of disappointment aversion. $\mu_\theta(.)$ is a weighted average of potential payoffs, where disappointing ones receive a higher weight $(1 + \theta)$, and create discontinuously more disutility than comparable gains.[4] What defines an outcome as disappointing is wether it is below the certainty equivalent, or reference point, $\mu_\theta(.)$ itself.[5] In this simple piecewise linear specification, the only source of risk aversion comes from the kink at the reference point, and the concavity it entails.

To derive optimal decisions over consumption plans and information structures, we extend these static preferences to a dynamic setting in which consumption happens continuously and information is revealed over time. We use the recursive model of Epstein and Zin (1989) and Duffie and Epstein (1992) – our second building block.

Definition 1. *For an increasing information filtration $\{\mathcal{F}_t\}_{t \in [0,\infty)}$ and an adapted consumption process $\{C_t\}_{t \in [0,\infty)}$, the value function is the solution to the recursion*

$$\mathcal{V}_t = \left[C_t^{1-\alpha} dt + e^{-\rho dt}\left(\mu_\theta\left[\mathcal{V}_{t+dt}|\mathcal{F}_t\right]\right)^{1-\alpha}\right]^{\frac{1}{1-\alpha}}.\tag{2}$$

At each point in time, the agent evaluates the utility of her wealth by aggregating

---

[4]This asymmetry is reminiscent of loss aversion introduced by Kahneman and Tversky (1979).

[5]Equation (1) is a fixed point problem in $\mu_\theta\left(X|F\right)$, which always admits a unique solution.

her consumption today, C_t, and her future utility, $\mathcal{V}_{t+dt}$. The rate of time discount is $\rho > 0$. The parameter $\alpha > 0$ controls the elasticity of intertemporal substitution between present and future consumption. The agent adjusts for the risk to her future utility, according to her current information, by using the disappointment aversion certainty equivalent $\mu_\theta(.)$ of Equation (1).[6]

We favor this recursive formulation for several reasons. First, the preferences of Epstein and Zin (1989) are broadly used in macroeconomics and finance.[7] Second, this formulation ensures time consistency. Third, and most important for our purpose, under these dynamics, the revelation of information over time creates compound lotteries that are evaluated successively. We now explain why this feature, combined with disappointment aversion, gives rise to information aversion.

Information aversion. Receiving information in a fragmented way multiplies the chances of receiving negative news. The idea of information aversion is that this is a source of discomfort. The preferences of Definition 1 formalize this intuition. To illustrate, consider the following simple situation: a payoff X is revealed and consumed at a final date, and there is no intermediate consumption. We compare the initial value of this lottery under two different information structures: when an intermediate signal about X will be received, $\mathcal{V}^{info}$, or not, $\mathcal{V}^{no\ info}$. Assume the initial date is $t = 0$, the intermediate signal is received at $t = 1$, and the payoff occurs at $t = 2$. Under Definition 1, the value without information is: $\mathcal{V}_0^{no\ info} = e^{-2\rho}\mu_\theta\left[X|\mathcal{F}_0\right]$. With information at $t = 1$, the value is: $\mathcal{V}_0^{info} = e^{-2\rho}\mu_\theta\left[\mu_\theta\left(X|\mathcal{F}_1\right)|\mathcal{F}_0\right]$, formed by compounding the certainty equivalents when information arrives at date 1, and at the final date 2. We show (Proposition 1 below) that $\mathcal{V}_0^{no\ info} \geq \mathcal{V}_0^{info}$ for any payoff X and $\mathcal{F}_0 \subset \mathcal{F}_1$, capturing the notion of information aversion.

To understand why, observe that the news revealed at the two dates, $t = 1$ and $t = 2$, is not considered jointly but successively, a byproduct of the recursive model of Epstein and Zin (1989), and where our choice of dynamics matters. Receiving an uncertain intermediate signal thus corresponds to shifting some of the risk from date 2 — a date with uncertainty (the payoff X), to date 1 — a date with no prior uncer-

---

[6]Formally, $\mathcal{V}_{t+dt}$ is $\mathcal{F}_{t+dt}$-measurable, and the expectation is formed conditional on $\mathcal{F}_t$.

[7]Routledge and Zin (2010) and Bonomo et al. (2011) combine them with disappointment aversion to explain equilibrium asset prices.

tainty. This is where disappointment aversion plays a role: under Gul (1991)'s model, adding risk to a certain lottery is more hurtful than to an uncertain lottery of same value – capturing the Allais (1979) paradox.[8] Cutting information in smaller pieces corresponds to spreading risk across previously safe times, and is painful, compared to revealing all at a single date. The following proposition formalizes this intuition:

Proposition 1. *Across all information structures $\mathcal{F}$ for a final payoff X, with same initial distribution $\mathcal{F}_0$, a dynamically disappointment averse agent prefers those revealing the realization of X at a single date.*

To clarify, the agent is indifferent as to when information is revealed, as long as it is at a single date. In particular, information aversion does not entail a preference for early or late resolution of uncertainty, and both full-information and no-information intermediate signals have zero implied (utility) cost, making information aversion irreconcilable with exogenous information costs or cognitive constraints.[9]

2.2 Consumption and Savings Problem

An application of Proposition 1 is the situation of an investor who owns a stock that she will sell for sure at the end of the year, and chooses how often to observe its price during the year. If she exhibits dynamic disappointment aversion, she finds it valuable not to observe it until the final date, as information is painful. Of course, in practice, information is also useful: she may need to know her portfolio value to manage her consumption and her trades. In this section, we model such information trade-offs, and make concrete predictions on how information averse households allocate and pay attention to their savings.

Investment opportunity set. An agent, with the preferences of Definition 1, uses her wealth W_t to consume and to save, at any time t, as in the classic setup of Merton

---

[8]The "Allais Ratio Paradox" is that people choose a) $200 with probability 1 over $300 with probability 0.8 and $0 with probability 0.2, b) $200 with probability 0.5 and $0 with probability 0.5 over $300 with probability 0.4 and $0 with probability 0.6. These choices are incompatible with the independence axiom, but compatible with disappointment aversion. The link between information aversion and the Allais (1979) paradox is not specific to disappointment aversion: see Dillenberger (2010).

[9]Any level of mutual entropy between the signal and the outcome can be attained at no utility cost, with an intermediate signal which reveals either nothing, with probability p, or the final outcome.

(1969). She has access to two investment accounts, which she can rebalance at no transaction cost: a risk-free asset with constant continuously compounded interest rate r, and a risky asset, which price X_t follows a geometric brownian motion with drift g and volatility σ.[10] The risky asset has higher expected returns than the safe asset, $g > r$, and, to ensure finite utility, $g < \frac{\rho}{1-\alpha}$, when $\alpha < 1$. Given an initial wealth $W_0 \geq 0$, the agent's sequence budget constraint is

$$dW_t = -C_t dt + S_t \frac{dX_t}{X_t} + r(W_t - S_t) dt \tag{3}$$

$$W_t \geq 0.$$

where S_t is the wealth invested in the risky asset at date t. Because the asset price can drop to 0, the agent cannot lever up and $S_t \leq W_t$ at all time (the natural borrowing limit in the absence of exogenous sources of income).

Information choice. At any time, the agent can choose to close her eyes, and receive no information, or observe, at no direct cost, the current value of her risky portfolio. In between observations, she makes decisions based on the last information she collected. Noting $\{\bar{\mathcal{F}}_t\}$ the full information filtration generated by the process $\{X_t\}$ appropriately completed, and $\{\mathcal{F}_t\}$ that of the agent, the constraint on information is

$$\forall t, \mathcal{F}_t = \bar{\mathcal{F}}_{\tau(t)}, \tag{4}$$

where $\tau(t) \leq t$ is the last observation time, at time t.[11]

Optimization problem. Given her initial wealth, the agent chooses her information filtration $\{\mathcal{F}_t\}$, and her consumption and savings $\{(C_t, S_t)\}$ $\mathcal{F}_t$-measurable, to maximize the value function of Equation (2) under the budget constraint of Equation (3) and the information constraint of Equation (4). At any observation time t, the utility function is homogenous of degree 1. The opportunity set is linear in current wealth

---

[10]Returns are i.i.d. and log-normally distributed: $\log(X_{t+\tau}/X_t) \sim \mathcal{N}([g - \frac{1}{2}\sigma^2]\tau, \sigma^2 \tau)$, $\forall \tau > 0$. The constant risk-free rate and Brownian assumption capture the salient features of asset returns, and is the workhorse of the finance literature. It is common to models of inattention, e.g. Duffie and Sun (1990), Abel et al. (2007), Alvarez et al. (2012), and Abel et al. (2013). We show in section 2.4 our results extend to more general risk processes.

[11]$\tau(t)$ is increasing, right continuous and admits a left limit.

W_t and identical across time (returns are i.i.d). This implies the consumption-savings policies and the value function are proportional to wealth and information acquisition optimally happens at constant time intervals.[12] In between observations, consumption is deterministic, to ensure the budget constraint $W_t \geq 0$ is always satisfied. We note $\mathcal{V}$, the utility per unit of wealth, which is constant: $\mathcal{V}_t = W_t \mathcal{V}, \forall t$.

At observation time t, the agent's problem reduces to choosing, for unit wealth, i) T, the length of time until next observation, ii) $\mathcal{C} = \int_0^T e^{-r\tau} C_{t+\tau} d\tau$, how much to set aside in the risk-free account to finance consumption during the interval $[t, t+T]$, and iii) S, how much to invest in the risky asset. She solves the optimization problem:

$$\mathcal{V} = \max_{T, \{C_\tau\}_{\tau \in [0,T]}, S} \left[\int_0^T e^{-\rho\tau} C_\tau^{1-\alpha} d\tau + e^{-\rho T} \left(S\mu_\theta \left[\frac{X_{t+T}}{X_t} \mid \mathcal{F}_t \right] + (1 - \mathcal{C} - S) e^{rT} \right)^{1-\alpha} \mathcal{V}^{1-\alpha} \right]^{\frac{1}{1-\alpha}}$$

(5)

$$\text{s.t. } \mathcal{C} = \int_0^T e^{-r\tau} C_\tau d\tau \text{ , and } S + \mathcal{C} \leq 1.$$

2.3 Optimal Risk-taking and Information Decisions

In Problem (5), the return of the risky asset X_{t+T}/X_t enters through its certainty equivalent. We derive from it a risk-adjusted rate per unit of time, a unit comparable to the safe rate r and useful to simplify our analysis.

Definition 2. Certainty equivalent rate. *Given a time interval T, the certainty equivalent rate $v(T)$ is defined as*

$$\exp(v(T)T) = \mu_\theta \left[\frac{X_{t+T}}{X_t} \mid \mathcal{F}_t \right], \ \forall t.$$

(6)

Lemma 1. *The certainty equivalent rate is the sum of two elements: the expected growth rate g, and a risk adjustment which is i) independent of g, ii) negative, iii) decreasing in the volatility σ and the coefficient of disappointment aversion θ, and iv) increasing in the observation interval T.*

---

[12]This feature arises naturally in our setting, whereas the literature on exogenous cost typically needs to assume information costs scale with wealth, e.g. Duffie and Sun (1990).

Investors are risk averse, so $v(.)$ is lower than the expected growth rate, the more so the greater the risk or the risk aversion (determined by θ). They are also information averse, so $v(.)$ varies with the length of time between observations. Going back to the illustrative example of Proposition 1, preferring to observe the price at date 2 only (interval of length $T = 2$) rather than at date 1 and 2 (intervals of length $T = 1$) corresponds to the comparison $v(2) > v(1)$. More generally, $v(.)$ increasing in T accounts for the fact that more frequent observations are more painful for the agent. The certainty equivalent rate fully encodes how information affects the valuation of risk, and, as we show below, allows us to determine investors' optimal decisions.

Optimal strategy. Using Definition 2 for the valuation of risk, Problem (5) becomes a deterministic dynamic program. We first characterize the optimal allocations, then the optimal length of the observation interval.

Proposition 2. *Given an observation interval T, the optimal consumption and savings strategies are:*

$$
\begin{cases}
\mathcal{C} = 1 - \exp\left[\left(-\frac{\rho}{\alpha} + \frac{1-\alpha}{\alpha}v(T)\right)T\right], S = 1 - \mathcal{C} \ \textit{if} \ v(T) > r \\
\mathcal{C} = 1 - \exp\left[\left(-\frac{\rho}{\alpha} + \frac{1-\alpha}{\alpha}r\right)T\right], S = 0 \ \textit{if} \ v(T) \le r.
\end{cases}
\tag{7}
$$

$\mathcal{C}$, the wealth set aside for consumption in between observations, is increasing in T.

Across observations, the optimal consumption is that of a standard isoelastic utility, where the savings instrument accrues at the highest rate between r and $v(T)$. With too frequent observations, the risk becomes so glaring, $v(T) \le r$, that the investor chooses to exit the risky asset altogether, $S = 0$. At lower information frequencies, the agent does want to save in the higher returns risky market, $S > 0$ when $v(T) > r$. But she has less wealth to invest, as she must set more aside for her consumption over longer inattention intervals, $\mathcal{C}$ increasing in T. This tension between information and risk-taking determines the optimal attention choice:

Proposition 3. *The optimal time interval between observations T^* is the unique solu-*

tion of

$$\frac{dv(T^*)}{d\log(T)} = \left(\frac{\rho}{1-\alpha} - v\left(T^*\right)\right)\left[1 - \frac{f\left(\frac{\rho}{1-\alpha} - r, T^*\right)}{f\left(\frac{\rho}{1-\alpha} - v\left(T^*\right), T^*\right)}\right] \tag{8}$$

where $f\left(x, T\right) = x/\left(\exp\left(\frac{1-\alpha}{\alpha}xT\right) - 1\right)$. **It satisfies** $T^* > 0$ *and* $v(T^*) > r$.

The right-hand side of Equation (8) is increasing in $v(.)$ and in T, and represents the opportunity cost incurred when setting wealth aside for consumption at the risk-free rate, rather than in the high returns risky asset ($v(T^*) > r$): it formalizes the benefits of information, and is common to models with infrequent transactions à la Baumol-Tobin.[13] The left-hand side of Equation (8) encodes the marginal cost of information, and arises endogenously from the preferences – the novelty of our approach. Instead of an ad-hoc exogenous fixed cost, the downside to receiving more frequent information is that it makes risky savings less appealing – as given by the elasticity of the certainty equivalent rate $v\,(.)$, with respect to the observation interval T. As the investor balances her desire to invest her wealth in the risky asset to take advantage of its superior returns, with her fear of frequent disappointments, some inattention is optimal ($T^* > 0$), at a level where risk remains attractive ($v(T^*) > r$).

Figure 1 depicts how the value function varies with the observation interval T. At high frequencies (small T), her fear of information is such that the investor cannot cope with taking any risk: she exits the market. As T increases from zero, her information "stress" decreases, and the perceived value of her wealth increases. But, the more inattentive she becomes, the more wealth she has to divert into her consumption account, incurring greater investment opportunity costs. Those dominate for long inattention intervals: utility is hump-shaped, with a unique maximum.

The determinants of inattention. Both the benefits of information and its endogenous costs vary with the economic environment, affecting optimal attention decisions. The following proposition summarizes how.

Proposition 4. *The optimal inattention interval* T^* *is*

    1. increasing in disappointment aversion θ,

---

[13]Another benefit of information, the ability to rebalance the portfolio and adjust the rate of consumption, does not appear in Equation (8) because it is of second order around the optimum.
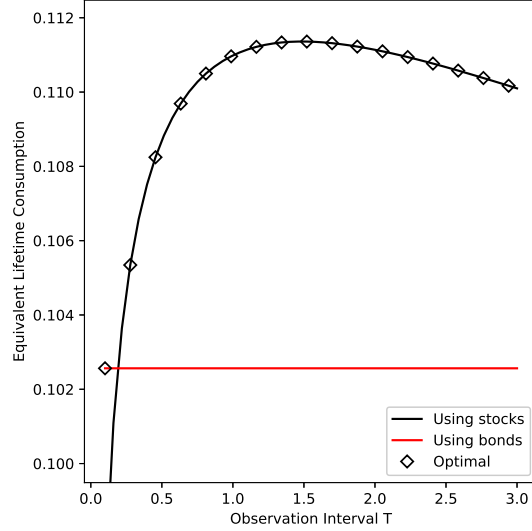
Figure 1: **Utility in function of the observation interval**. The figure reports the equivalent lifetime consumption — constant annual consumption — corresponding to the optimal policy at each observation interval T. $r = 1\%, g = 7.3\%, \sigma = 17\%$, matching properties of stock market and short-term Treasury returns (1926-2016). $\theta = 0.5, \alpha = 0.5, \rho = 0.2$, standard preference parameters. The optimal T^*, around 1 year, is consistent with existing surveys (Section 3).

2. *decreasing in expected returns g,*

3. *increasing in volatility σ,*

4. *increasing in volatility even when v is unchanged.*

To understand these results, consider how the various parameters affect the certainty equivalent rate $v(.)$, both in slope, controlling the cost of information, and level, controlling the benefit of information. Decreases in expected returns, increases in risk, increases in disappointment aversion: all imply a lower risk-adjusted rate $v(.)$, i.e. a decrease in the benefits of information. More risk or greater disappointment aversion also make observations more "stressful" – they increase information costs, while variations in expected returns leave them unchanged: $dv(.)/d\log(T)$ is increasing in σ and θ, and independent of g. As they increase information costs and/or lower its benefits, decreases in g, increases in σ, and in θ, imply more inattention – points 1,

2 and 3 in Proposition 4. Consider now an increase in risk compensated by an increase in expected returns, such that the certainty equivalent rate remains constant. Why this particular case? Because it leaves the benefits of information unchanged (like v). But the preference-based costs of information still go up, so inattention increases, the final result in Proposition 4. By continuity, even when $v(.)$, the risky asset's investment opportunity, increases, a higher volatility can lead to more inattention.

2.4 Generalizations

In the remainder of the paper, we discuss the ability of our model to help understand the observed information decisions made by households managing their savings. In order to do so in a meaningful way, we assess the robustness of our results, by extending the model in various realistic directions. First, we let the stochastic process differ from Brownian: we derive general results and illustrate them with two other processes, useful when considering financial risks. Second, we assume more standard forms of risk aversion, beyond that implied by disappointment aversion, and add concavity in the preferences on both sides of the reference point. This allows us to obtain more realistic savings allocations than the bang-bang optimal choice of Proposition 2. Third, we introduce transaction costs, to decouple information and transaction decisions, as is the case in practice.

Return dynamics. Because our results are characterized by the certainty equivalent rate of Definition 2, the model accommodates non-normal returns in a straightforward way. In particular, Equations (7) and (8) are valid for any stochastic process with i.i.d. increments. General properties of $v(.)$ capture the intuitions behind the results of Propositions 2, 3 and 4. From Proposition 1, the value of *any* uncertain payoff decreases when risk is revealed in multiple pieces: for any process $\{X_t\}$ with i.i.d growth, and $t_1, t_2 \geq 0$, $(t_1 + t_2)v(t_1 + t_2) \geq t_1 v(t_1) + t_2 v(t_2)$. This inequality formalizes the idea that a lower frequency of observations reduces the stress induced by risk. At the limit, when the time intervals become infinitely large, the preference-based information costs disappear altogether, and $v(.)$ converges to its expected rate of return: $\lim_{T \to \infty} v(T) = \log(\mathbb{E}[X_{t+1}/X_t])$. Therefore, for large T, information has positive benefits and zero implied costs: to never observe her wealth is suboptimal, and the agent

chooses a finite interval T^*. Some inattention is optimal and $T^* > 0$ for a general stochastic process with i.i.d. increments if observing it continuously makes the risky investment less attractive than the risk-free rate ($v(.)$ less than r at the limit $T = 0$), or if it is increasing at the continuous limit ($v(.)$ increasing around $T = 0$).[14] Finally, the risk adjusted rate $v(.)$ is decreasing in the coefficient of disappointment aversion θ and in risk, for instance captured by a mean-preserving spread in the returns distribution – and therefore so are the benefits of information.

We illustrate these properties for two processes: a pure Poisson jump process, and the empirical distribution of stock market returns.[15] Figure 2 depicts their two value functions – illustrating how remarkably similar they are to the Brownian case.[16]

(a) Jump Risk   (b) Empirical Stock Returns
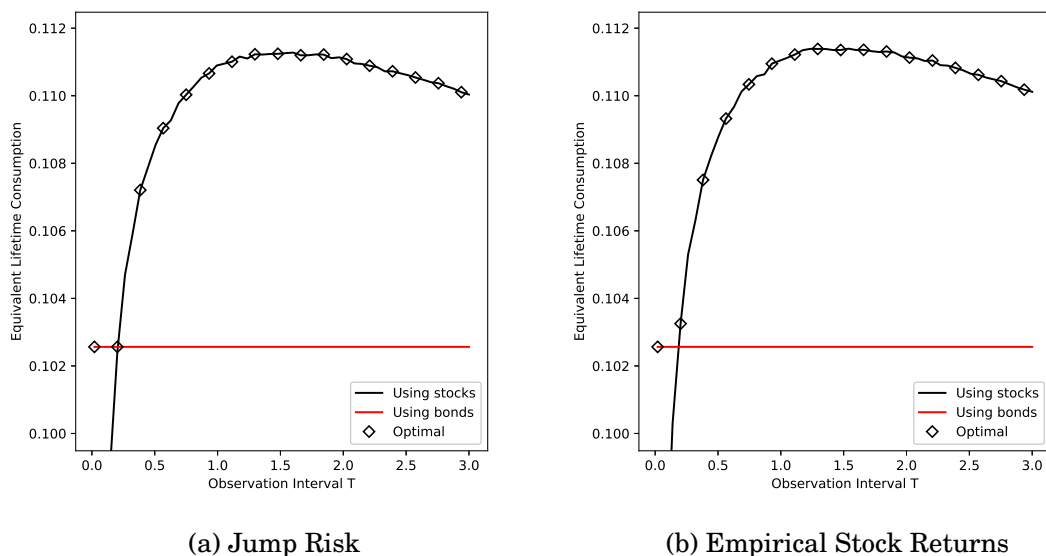
Figure 2: **Other Risk Distributions**. The left panel corresponds to a Poisson jump process with jump intensity $\lambda = 10$ and size $\sigma = 17\%/\sqrt{10}$. The right panel corresponds to the historical distribution of daily stock returns. Other parameters are as in Figure 1.

Concave utility. Gul (1991)'s disappointment aversion model allows for more standard forms of risk aversion on both sides of the reference point. Adding CRRA-type

---

[14]These are sufficient but not necessary conditions.

[15]Daily value-weighted returns of all NYSE, AMEX, and NASDAQ stocks, from CRSP, 1926 to 2017.

[16]The value functions are jagged, due to the discreteness of the distributions, but maintain the same overall behavior as in the Brownian case – see Appendix B.4 and B.5.

risk aversion $\gamma \geq 0$, for instance, modifies Equation (2) in Definition 1 to:

$$\mathcal{V}_t = \left[C_t^{1-\alpha} dt + (1 - \rho dt) \left((1-\gamma) \mu_\theta \left[\frac{1}{1-\gamma} \mathcal{V}_{t+dt}^{1-\gamma} \right]^{\frac{1-\alpha}{1-\gamma}} \right) \right]^{\frac{1}{1-\alpha}}. \tag{9}$$

Under the preferences of Equation (9), Proposition 1 – and the notion of information aversion, still holds. The homotheticity is maintained, so the optimal consumption and observation policies take a similar form to our baseline model: the length of time between observations is constant, and the consumption and portfolio policies are proportional to wealth. Where the models differ is in the choice of the optimal savings portfolio. With standard risk aversion, the investor may prefer to invest a fraction $\phi^*(T) \in [0,1]$, rather than all, in the risky asset, with $\phi^*(T)$ solution to the static problem, for a given time interval T:

$$\phi^*(T) = \arg\max_{\phi \in [0,1]} \left((1-\gamma) \mu_\theta \left[\frac{1}{1-\gamma} \left(\phi \frac{X_{t+T}}{X_t} + (1-\phi)r \right)^{1-\gamma} \right] \right)^{\frac{1-\alpha}{1-\gamma}}. \tag{10}$$

The returns of the optimal saving portfolio over a time interval of length T become $R_T^* = \phi^*(T)\frac{X_{t+T}}{X_t} + (1 - \phi^*(T))r$, and we re-define the investment portfolio certainty equivalent rate $v(.)$:

$$e^{v(T)T} = \left((1-\gamma) \mu_\theta \left[\frac{1}{1-\gamma} (R_T^*)^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}. \tag{11}$$

Replacing the certainty equivalent rate of Definition 2, Equation (6), with that of Equation (11), the information trade-offs of Proposition 3 are unchanged, and inattention is determined as in Section 2.3.

Transaction costs. To add transaction costs in a tractable way, we follow the approach of Alvarez et al. (2012). The saving technologies are the same as our baseline model but consumption now arises from a durable good which depreciates at a given rate. To change the stock of durable, a fraction of it is paid, each time.[17]

Under the investor's optimal choice, at each observation, she considers the ratio of

---

[17]We formally set the problem up, and derive its qualitative implications, in Appendix C.4.

her stock of durable to her financial wealth. She chooses a finite, non-zero time until her next observation, and, only when the ratio falls outside of two time-independent bounds, she transacts, and brings it back to a time-independent value. Our inattention results go through; in addition transactions happen less often than observations, a realistic feature, as we discuss below.

3 Information choices: Empirical evidence

We review the evidence on households' inattention and assess the ability of our information aversion model to rationalize it. We find, and explain why, our theory is well suited to capture the main features of inattention to savings, in comparison in particular to models with fixed-costs of information. We then turn to other sources of evidence, including experiments, on the links between information and risk-taking decisions, at the heart of our mechanism.

3.1 Inattention to Savings

Empirical evidence. Several recent papers study direct evidence of investors' inattention to their savings.[18] Alvarez et al. (2012) analyze the frequency of portfolio observations, frequency of active trades, risk preferences and demographics, from two surveys of Italian Households. The first focuses on 1,800 clients of the bank Unicredit, matched with a panel of administrative records for their accounts; the second is the Survey of Households Income and Wealth, done by the Bank of Italy on a representative sample of 8,000 households. Karlsson et al. (2009) and Sicherman et al. (2016) study the online activities of over a million Vanguard's investors (mostly 401(k) accounts and money purchase pension plans).[19] For each account, and day, they observe the number of logins, and trades, as well as monthly balance information, and some demographics. Karlsson et al. (2009) also use aggregate data on daily logins from the Swedish premium pension system.

---

[18]The low correlation between perceived equity holdings and the stock index – Kennickell et al. (2000) and Dynan and Maki (2001) – provides indirect evidence.

[19]Vanguard is one of the world's largest investment companies, with more than 20 million investors, and about $4 trillion in global assets under management, as of December 31, 2016.

Several stylized facts emerge from these various sources of evidence:

1. *Investors observe their accounts infrequently.* Alvarez et al. (2012) reports that the average Italian household observes the value of their portfolio three to four times a year; direct stockholders a little more. In Sicherman et al. (2016), the median number of logins is between once and twice per month.

2. *Inattention varies across investors.* Investors with a greater appetite for risk observe their portfolios more frequently. In Alvarez et al. (2012), to the question: "Which of the following statements comes closest to the amount of financial risk that you are willing to take when you make your financial investment?", the investors responding "a very high return, with a very high risk of loosing the money" observe their portfolios 3.9 times per year, those responding "high return and high risk" 1.9 times per year, and those responding "moderate return and moderate risk" 1.6 times per year, on average.[20] They find equity holdings increase with the appetite for risk, and thus the observation frequency. Sicherman et al. (2016) document the same positive relation between the portion of stock holdings and the number of logins.

3. *Inattention varies with market conditions.* It varies negatively with stock prices. Karlsson et al. (2009) document and name this "ostrich behavior", whereby investors stop checking their portfolios when the market goes down. Sicherman et al. (2016) find this effect is stronger for investors with larger bond positions. Inattention varies positively with the market expected volatility. Sicherman et al. (2016) note investors logins decrease when the VIX index increases.[21]

4. *Investors actively trade assets more rarely than they observe their account positions.* Alvarez et al. (2012) find an average ratio of three observations per trade. In Sicherman et al. (2016), active trading occurs once a year on average.

Information aversion. The model of Section 2 rationalizes the evidence above. From Proposition 3, information averse investors optimally observe their portfolios

---

[20]Under the preferences of Definition 1, our model, this question elicits differences in the coefficient of disappointment aversion θ.

[21]Constructed from S&P 500 index options trading prices by the Chicago Board Options Exchange (CBOE), the VIX index measures the expected volatility of the S&P 500 index.

at discrete points in time – stylized fact 1. A high disappointment aversion corresponds to a high risk aversion *and* information aversion, leading to lower observation frequencies (Proposition 4), and lower risk investments (Proposition 2) – stylized fact 2. To understand stylized fact 3, note the empirical and theoretical asset pricing literature shows that periods of low stock prices and high volatility coincide with higher expected returns and Sharpe ratios.[22] This has two opposite effects, in our model. The high value of the risky asset pushes the agent to be more attentive, and invest more. But the utility cost of information also increases with volatility, encouraging more inattention. The latter effect dominates if increases in the valuation of risk are not too large (Proposition 4), as is the case empirically – stylized fact 3.

Information aversion does not directly cause the lower frequency of trades than observations – stylized fact 4, but does not contradict it, as discussed in Section 2.4, in the extension with transaction costs.

Alternative explanations. Within the set-up of Section 2, investors with expected utility are not inattentive, as information entails no inherent cost, and can help make better decisions. In the CRRA model with risk aversion γ, for instance, the certainty equivalent rate of Definition 2 becomes: $v(T) = \bar{v}$, a constant – e.g. $\bar{v} = g - \frac{1}{2}\gamma\sigma^2$ under Brownian risk. The characterization of optimal attention from Equation (8) still holds, with a left-hand side, the preference-based costs of information, equals to $d\bar{v}/d\log(T) = 0$; and a right-hand side, the benefits of information, increasing in its frequency: it is optimal to observe one's portfolio continuously.

To derive inattention, an additional assumption must be made: that households face fixed costs of information – the main alternative to our model, and the benchmark in the literature on investors' inattention.[23] In line with this framework, and as a basis for comparison, we modify our model and replace disappointment aversion with CRRA, with risk aversion γ, for the risk-adjustment in Definition 1; and assume that each observation costs a fraction ϕ_o of total wealth. We note R_T the returns of the

---

[22]Shiller (1981) and Fama and French (1988) first document the predictability of returns from index levels. Ghysels et al. (2005) and Lundblad (2007) study the effect of volatility on expected returns. Lettau and Ludvigson (2010) surveys the joint dynamics of risk and returns.

[23]Duffie and Sun (1990), Abel et al. (2007), Abel et al. (2013), and Alvarez et al. (2012) study such models. In contrast Luo (2010), who assumes agents have cognitive constraints as in Sims (1998, 2006) to explain portfolio inertia, finds investors should continuously observe their wealth.

optimal saving portfolio over a time interval of length T, and define the certainty equivalent rate $v(.)$, which now explicitly includes the observation costs:

$$e^{v(T)T} = (1 - \phi_o) \, \mathbb{E} \left[R_T^{1-\gamma} \right]^{\frac{1}{1-\gamma}} . \tag{12}$$

As per Equation (8), the marginal cost of attention is:

$$\frac{dv}{d \log T} (T) = -\frac{1}{T} \log (1 - \phi_o) , \tag{13}$$

exactly proportional to the number of observations (given by $\frac{1}{T}$ for an interval of length 1). Continuously observing one's wealth requires paying fixed information costs infinitely many times, so investors are optimally inattentive. But Equation (13) shows these costs are the same across preference parameters and return processes, or market conditions. All variations in inattention must thus be driven by variations in the benefits of information – determined by the perceived value of risk investments relative to the risk-free asset.[24] Because empirically the value of risk tends to be high when the index is low and vice versa, the exogenous cost framework implies inattention should decrease in such times, in direct contradiction with the evidence in Karlsson et al. (2009) and Sicherman et al. (2016) – stylized fact 3 above.[25]

Under information aversion, variations in inattention across investors and market conditions derive from variations in both the benefits *and* the intrinsic costs of information, and changes to both sides are needed to explain all the stylized facts.

3.2 Preference-based information costs

The preference-based, instead of exogenous, information cost assumption is central to the success of our model at explaining the evidence on inattention to savings. We argue below that such an approach is independently justified by observed behaviors, across fields, with respect to information and risk decisions. We review alternative preference models with implied information costs, and discuss why we favor ours.

Information avoidance. Under the expected utility model, agents care about infor-

---

[24]e.g. Abel et al. (2007), where inattention is function of the risky asset's Sharpe ratio.
[25]The other stylized facts – 1,2 and 4 – are compatible with exogenous information costs.

mation strictly insomuch as it allows them to make better decisions. Freely available information, because potentially useful, would never be declined. However, examples abound of agents actively choosing to avoid receiving signals free of monetary or time cost.[26] This attitude – information avoidance – disputes the traditional view (see Golman et al. (2017) for a review), and suggests information is not simply a tool to make better decisions; it directly affects agents' well-being – as further supported by neuroscience evidence (Knutson and Peterson, 2005).

To account for the impact of information on agents' utility, we choose a rational and time consistent framework where information always enters as an endogenous cost – *information aversion* – the most natural counterpart to the benchmark exogenous cost model. Segal (1987, 1990) discusses sufficient conditions for systematic costs of information in two-stage binary lotteries. Palacios-Huerta (1999) extends their framework and introduces the timing of resolution of uncertainty as an economic variable. Dillenberger (2010) provides an axiomatic structure for preferences with an intrinsic disutility of information in compound lotteries. Among this general characterization, only disappointment aversion (Gul, 1991) belongs to the extended set of rational agents' utility functions proposed and studied in applied theory (Dillenberger, 2010; Artstein-Avidan and Dillenberger, 2011). In contrast, rank-dependent utility (Segal, 1987, 1990), quadratic utility, any twice-differentiable utility function, and loss aversion, as in Benartzi and Thaler (1995) or Kőszegi and Rabin (2009), do not imply unambiguous information costs.

Information and risk decisions. Numerous papers study, in the lab or in the field, how changes in the information environment affect risk decisions. Their evidence is consistent with our framework.

Starting with Gneezy and Potters (1997) and Thaler et al. (1997), lab experiments have shown that subjects' valuations of risky outcomes diminish when they are given more frequent information.[27] Galai and Sade (2006) find similar outcomes using field data on Israeli treasury bills and commercial banks deposits. Bellemare et al. (2005) show that disentangling the frequency of information from decisions flexibility does

---

[26]e.g. subjects refusing to be tested for diseases, or to receive the results of such tests (Ganguly and Tasoff, 2016; Sullivan and Drake, 2004; Lerman et al., 1999; Lyter et al., 1987; Oster et al., 2013).

[27]Also Benartzi and Thaler (1999), Barron and Erev (2003), Gneezy et al. (2003), Haigh and List (2005), Fellner and Sutter (2009) and Anagol and Gamble (2011).

not change these results. This suggests it's not the fear of making "bad", irrational, choices when information is received – inattention as a commitment device – that drives them, but rather intrinsic costs of information within a rational framework, our approach (Proposition 1).

Two recent papers focus on the taking of risk, rather than the valuation of risk, with respect to information. In a framed field experiment, Beshears et al. (2012) document that more frequent signals (weekly versus bi-annual) do not reduce the average risk-taking of subjects, to whom relatively large sums of money are given to invest. Larson et al. (2016) find the opposite result in a natural field experiment involving professional traders: receiving market information every four hours, rather than second-by-second – keeping unchanged the flexibility to change their positions minute-by-minute – results in a 33% higher risk-investment (and a 53% higher profit over the experiment).[28]

These results may appear contradictory with each other, and with the evidence above of a monotone relation between information and the value of risk. But they actually are all consistent with the interplay of attention and risk-taking under information aversion characterized in Proposition 2. If information frequency is such (second-by-second in Larson et al. (2016)) that some agents no longer value the risky over the risk-free asset ($v(T) \leq r$) and exit the market, the average risk taking can go down. On the other hand, as long as information is rare enough (as in Beshears et al. (2012)) that the risky asset remains attractive ($v(T) > r$) to some investors, they can, and will, allocate more wealth to the risky asset, even though its value goes down, when the information frequency goes up. Even though additional information is painful, if the agent receives it (as an optimal choice, or through external forces), she may as well take advantage of it, by engaging in more risk-taking.[29]

Other information models. Other preference models take information into account. In Brunnermeier and Parker (2005), agents distort their expectations to be optimistic about future outcomes.[30] Caplin and Leahy (2001, 2004) assume anxiety,

---

[28]See Harrison and List (2004) for the taxonomy of field experiments.

[29]In the model this nonmonotonicity is sharp due to the bang-bang nature of optimal portfolio choice and would be smoothed out with strictly concave rather than piecewise linear preferences.

[30]Oster et al. (2013) show it can explain the observed information choices of at-risk patients, in the testing for Hutchinson disease.

or anticipatory feelings, in a time-inconsistent model. When anxiety follows a U-shape over time, it can explain the lack of savings for retirement of the baby boomers generation – a form of inattention.[31] In Epstein and Zin (1989), agents display preferences for early or late resolution of uncertainty: they care about "when" information is received, even if they cannot use it to amend their decisions. Carrillo and Mariotti (2000) and Benabou and Tirole (2002) justify inattention as a commitment device when agents are time inconsistent, or react sub-optimally to news and information. These are but a few examples: Wakker (1988) shows forms of information avoidance can arise in any preference framework that departs from expected utility.

Our model of information aversion is, by comparison, a minimal deviation from the classical framework: it does not appeal to behavioral traits specific to information such as learning constraints, non-Bayesian updating, time inconsistency or non-rational responses to information. Instead, it features first-order risk aversion – a well documented behavioral trait concerning risk taking – and, applied to finance, can explain equilibrium asset prices, market participation and portfolio allocations. The model of Section 2 can explain and reconcile the evidence concerning the impact of information frequency on the valuation of risk and on risk-taking decisions; and yields a cost-benefit analysis similar to the classical framework, directly interpretable, and grounded in the data on investors' inattention to their wealth. Its success across applications and fields attests to the versatility of, and provide strong support for, our preference-based information cost framework.

4 Richer Information Choices

So far, our theoretical analysis has focused on an investor with access to the simplest tool to manage information: choosing when to observe or not the value of her portfolio. We extend the model and study how more subtle information choices could be made, with the help of computers or other agents. The insights we develop have both positive and a normative implications: providing a rationale for information systems used in

---

[31]Averill and Rosenn (1972) find more information causes sometimes more, sometimes less, anxiety, depending on subjects and contexts; making the relation between anxiety and information choices ambiguous.

practice, or suggesting ways to help agents deal with information aversion.[32]

In Appendix E, we discuss and derive solutions for a few other natural extensions: the role of diversification; the link between leverage and information decisions; the relation between learning, information and risk-taking dynamics; and the demand for opaque versus transparent signals.

Information Delegation. Our framework makes possible the comparison of arbitrary structures of information flows under the preferences of Definition 1; a contrast to fixed cost models of inattention, where each information choice must be ascribed an ad hoc exogenous cost.[33] Rather than studying all the feasible information structures, we consider wether simple signals can help the investor.[34]

In addition to the time-dependent observations the agent can obtain on her own, we include state-dependent alerts she receives when her wealth falls to a fraction $\eta \geq 0$ of its previous value. At any time t she does not receive a signal, the investor knows her wealth from her last observation at time $\tau(t) \leq t$, *and* that it has remained above the threshold $\eta W_{\tau(t)}$ ever since. Let $\bar{\mathcal{F}}_t^\eta$ be the information set given by $\{W_s > \eta W_{\tau(t)}, \forall s \in [\tau(t), t]\}$. The investor's information filtration follows:

$$\forall t, \mathcal{F}_t = \bar{\mathcal{F}}_{\tau(t)} \cap \bar{\mathcal{F}}_t^\eta. \tag{14}$$

At observation time $\tau(t)$, the investor chooses an interval of time T after which to observe her wealth, if she has received no signal, where T constant is the optimal rule. Her next observation time will either be $\tau(t) + T$ or the earliest $s(t)$ such that $\tau(t) \leq s(t) \leq \tau(t) + T$ and $W_{s(t)} \leq \eta W_{\tau(t)}$.[35]

The agent optimizes her information choices – the observation interval T and the signal threshold $\eta \geq 0$ – and her consumption and savings $\{(C_t, S_t)\}$ $\mathcal{F}_t$-measurable, to maximize the value function of Equation (2) under the budget constraint of Equation (3), as in Section 2.2, and the information constraint of Equation (14). Our base-model Problem (5) corresponds to imposing $\eta = 0$.

---

[32]Providing these signals can also entail some costs of its own, for instance due to agency issues.

[33]To respect Bayesian updating and no-forgetting, information policies must simply satisfy: the filtration $\{\mathcal{F}_t\}$ is increasing, and in the maximum information available at date t, $\bar{\mathcal{F}}_t$.

[34]Obtaining a general characterization of the optimal information structure is not readily tractable and beyond the scope of this paper.

[35]The limit $\eta \to +\infty$ corresponds to $s(t) = \tau(t), \forall t$, and continuous observations.

Proposition 5. *The optimal threshold is such that $\eta^* > 0$: the agent is strictly better off if she can add, to her regular portfolio observations, a state dependent signal she receives following poor portfolio performances.*

Adding state-dependent alerts entails a more frequent information, a utility cost. On the other hand, the agent can adjust her consumption and savings following an alert, a benefit. More subtle, the agent also modifies her ex-ante investment strategy. Because the poor-performance alerts provide a backstop for her investment account before it gets depleted, she can now finance some of her deterministic consumption between regular observations from the risky rather than risk-free asset. For low signal thresholds, the utility gains from the increased risk-taking are first-order, whereas the additional information costs, because they occur rarely, are negligible.[36]

In practice, the alarms we consider are natural to implement: a broker or bank could contact their clients, or send automatic messages, following poor portfolio performances. Our result suggests such delegation enhances risk-taking, a result supported by the evidence in Foerster et al. (2014). More generally, becoming aware of extreme events is useful, a result consistent with the media representation of news. The insight of Proposition 5 extends beyond our particular model of consumption and savings: obtaining information precisely when it is needed most is worth the additional utility cost.

Decision Delegation. Our investor might want to go further, and delegate all decisions to a portfolio manager. It seems the perfect tool against information aversion: pay somebody to be fully informed and make all decisions, and you no longer have to look for information. Though intuitive it is partly false: delegated management cannot always shield the agent from receiving information. In our model for instance, the value of her portfolio can be hidden from her, but the agent must always observe her consumption, thereby inferring her wealth, and limiting the scope for delegation.[37]

Consider, in contrast, the case of a dynamic disappointment aversion investor who does not consume until a final date, and where expected returns are time-varying.

---

[36]This result generalizes to non-Brownian distributions where the probability of reaching a close to 0 threshold decreases fast enough. A similar analysis for upper thresholds signals shows the additional information costs and benefits are of same order of magnitude, so adding them is *not* always desirable.

[37]Formally, S_t is $\bar{\mathcal{F}}_t$-measurable but C_t is $\mathcal{F}_t$-measurable, in the notations of Section 2.

If she cannot delegate, she will trade off the benefits of active portfolio management with her utility costs of information. She would prefer, however, to delegate all portfolio management decisions, and only observe the final value of her wealth, when she consumes – and be fully shielded from all information costs.

This example provides some guidance as to when delegation is most useful: for decisions with long-term implications. Consistent with this argument, the delegation of portfolio management of retirement accounts is common in practice, while short-term consumption and saving decisions are typically made by the agents themselves.

5   Conclusion

Under dynamic disappointment aversion, agents are intrinsically information averse. We propose a theory of inattention solely based on these preferences, absent any cognitive limitations, or external costs of acquiring information. We show that, as information averse investors balance the utility cost of paying attention to the economic environment with the benefits of making informed decisions, their information decisions echo the empirical evidence on attention to savings, and rationalizes all its stylized facts. In particular, their attention can decrease in turbulent times – when risk increases, information is more stressful – even when market opportunities are greater, a result unique to our model of inattention. We explore how to better tailor information flows to the specific needs of information averse investors, and show state-dependent poor-performance alerts make them strictly better off.

Taking the point of view that people *want* to stay away from information draws a very different picture of inattention than in standard models: facilitating the access to information might entail a waste of resources, or even be detrimental. This insight has pervasive implications for decision-making under uncertainty. For instance, the desire to remain uninformed naturally gives rise to agency problems with suppliers of information.[38] We believe the simplicity of our approach suggests an avenue for future research, exploring the theoretical predictions and empirical implications of information aversion in other economic contexts.

---

[38]Cohen et al. (2013) find evidence of firms manipulating information to their advantage.

References

ABEL, A. B., J. C. EBERLY, AND S. PANAGEAS (2007): "Optimal inattention to the stock market," *The American economic review*, 97, 244–249.

——— (2013): "Optimal Inattention to the Stock Market With Information Costs and Transactions Costs," *Econometrica*, 81, 1455–1481.

ACHARYA, V. V. AND B. M. LAMBRECHT (2014): "A theory of income smoothing when insiders know more than outsiders," Tech. rep., National Bureau of Economic Research.

ALLAIS, M. (1979): "The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American School (1952)," in *Expected utility hypotheses and the Allais paradox*, Springer, 27–145.

ALVAREZ, F., L. GUISO, AND F. LIPPI (2012): "Durable Consumption and Asset Management with Transaction and Observation Costs," *The American Economic Review*, 102, 2272–2300.

ANAGOL, S. AND K. GAMBLE (2011): "Does Presenting Investment Results Asset by Asset Lower Risk Taking?" *University of Pennsylvania working paper*.

ANDERSEN, S., J. Y. CAMPBELL, K. M. NIELSEN, AND T. RAMADORAI (2014): "Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market," .

ANG, A., G. BEKAERT, AND J. LIU (2005): "Why stocks may disappoint," *Journal of Financial Economics*, 76, 471–508.

ANG, A., J. CHEN, AND Y. XING (2006): "Downside risk," *Review of Financial Studies*, 19, 1191–1239.

ARTSTEIN-AVIDAN, S. AND D. DILLENBERGER (2011): "Dynamic disappointment aversion," .

AVERILL, J. R. AND M. ROSENN (1972): "Vigilant and nonvigilant coping strategies and psychophysiological stress reactions during the anticipation of electric shock," *Journal of Personality and Social Psychology*, 23, 128 – 141.

BARRON, G. AND I. EREV (2003): "Small feedback-based decisions and their limited correspondence to description-based decisions," *Journal of Behavioral Decision Making*, 16, 215–233.

BELLEMARE, C., M. KRAUSE, S. KRGER, AND C. ZHANG (2005): "Myopic loss aversion: Information flexibility vs. investment flexibility," *Economics Letters*, 87, 319–324.

BENABOU, R. AND J. TIROLE (2002): "Self-Confidence and Personal Motivation*," *The Quarterly Journal of Economics*, 117, 871.

BENARTZI, S. AND R. THALER (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73–92.

——— (1999): "Risk aversion or myopia? Choices in repeated gambles and retirement investments," *Management Science*, 45, 364–381.

BESHEARS, J., J. CHOI, D. LAIBSON, AND B. MADRIAN (2012): "Can Psychological Aggregation Manipulations Affect Portfolio Risk-Taking? Evidence from a Framed Field Experiment," *Working paper*.

BILIAS, Y., D. GEORGARAKOS, AND M. HALIASSOS (2010): "Portfolio Inertia and Stock Market Fluctuations," *Journal of Money, Credit and Banking*, 42, 715–742.

BONOMO, M., R. GARCIA, N. MEDDAHI, AND R. TÉDONGAP (2011): "Generalized disappointment aversion, long-run volatility risk, and asset prices," *Review of Financial Studies*, 24, 82–122.

BRUNNERMEIER, M. K. AND S. NAGEL (2008): "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-evidence on Individuals," *American Economic Review*, 98, 713–36.

BRUNNERMEIER, M. K. AND J. A. PARKER (2005): "Optimal Expectations," *The American Economic Review*, 95, 1092–1118.

CALVET, L. E., J. Y. CAMPBELL, AND P. SODINI (2009): "Fight or Flight? Portfolio Rebalancing by Individual Investors*," *The Quarterly Journal of Economics*, 124, 301.

CAPLIN, A. AND J. LEAHY (2001): "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 116, 55–79.

——— (2004): "The Supply of Information by a Concerned Expert," *Economic Journal*, 114, 487–505.

CARRILLO, J. D. AND T. MARIOTTI (2000): "Strategic Ignorance as a Self-Disciplining Device," *The Review of Economic Studies*, 67, 529.

COHEN, L., D. LOU, AND C. MALLOY (2013): "Playing Favorites: How Firms Prevent the Revelation of Bad News," Working Paper 19429, National Bureau of Economic Research.

DILLENBERGER, D. (2010): "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 78, 1973–2004.

DUFFIE, D. AND L. G. EPSTEIN (1992): "Stochastic Differential Utility," *Econometrica*, 60, 353–394.

DUFFIE, D., M. SCHRODER, AND C. SKIADAS (1996): "Recursive valuation of defaultable securities and the timing of resolution of uncertainty," *The Annals of Applied Probability*, 1075–1090.

——— (1997): "A term structure model with preferences for the timing of resolution of uncertainty," *Economic Theory*, 9, 3–22.

DUFFIE, D. AND T.-S. SUN (1990): "Transactions costs and portfolio choice in a discrete-continuous-time setting," *Journal of Economic Dynamics and Control*, 14, 35–51.

DYNAN, K. E. AND D. M. MAKI (2001): "Does stock market wealth matter for consumption?" *FEDS Discussion Paper*, 23.

EPSTEIN, L. G. AND S. E. ZIN (1989): "Substitution, risk aversion, and the intertemporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 57, 937–969.

ERICSON, K. M. M. AND A. FUSTER (2014): "The endowment effect," .

FAMA, E. F. AND K. R. FRENCH (1988): "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96, 246–273.

FELLNER, G. AND M. SUTTER (2009): "Causes, consequences, and cures of myopic loss aversion — an experimental investigation," *Economic Journal*, 119, 900–916.

FOERSTER, S., J. LINNAINMAA, B. MELZER, AND A. PREVITERO (2014): "The Costs and Benefits of Financial Advice," .

GABAIX, X. AND D. LAIBSON (2002): "The 6D bias and the equity-premium puzzle," in *NBER Macroeconomics Annual 2001, Volume 16*, MIT Press, 257–330.

GALAI, D. AND O. SADE (2006): "The Ostrich Effect" and the Relationship between the Liquidity and the Yields of Financial Assets," *Journal of Business*, 79, 2741–2759.

GANGULY, A. R. AND J. TASOFF (2016): "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future," *Management Science, Forthcoming*.

GHYSELS, E., P. SANTA-CLARA, AND R. VALKANOV (2005): "There is a risk-return trade-off after all," *Journal of Financial Economics*, 76, 509 – 548.

GNEEZY, U., A. KAPTEYN, AND J. POTTERS (2003): "Evaluation periods and asset prices in a market experiment," *Journal of Finance*, 58, 821–837.

GNEEZY, U. AND J. POTTERS (1997): "An experiment on risk taking and evaluation periods," *Quarterly Journal of Economics*, 112, 631–645.

GOLMAN, R., D. HAGMANN, AND G. LOEWENSTEIN (2017): "Information Avoidance," *Journal of Economic Literature*, 55, 96–135.

GUL, F. (1991): "A theory of disappointment aversion," *Econometrica: Journal of the Econometric Society*, 667–686.

HAIGH, M. AND J. LIST (2005): "Do professional traders exhibit myopic loss aversion? An experimental analysis," *Journal of Finance*, 60, 523–534.

HARRISON, G. W. AND J. A. LIST (2004): "Field experiments," *Journal of Economic literature*, 42, 1009–1055.

KAHNEMAN, D. AND A. TVERSKY (1979): "Prospect theory: an analysis of decision under risk," *Econometrica*, 47, 263–291.

KARLSSON, N., G. LOEWENSTEIN, AND D. SEPPI (2009): "The ostrich effect: Selective attention to information," *Journal of Risk and Uncertainty*, 38(2), 95–115.

KENNICKELL, A. B., M. STARR-MCCLUER, AND B. J. SURETTE (2000): "Recent changes in US family finances: Results from the 1998 Survey of Consumer Finances," *Fed. Res. Bull.*, 86, 1.

KNUTSON, B. AND R. PETERSON (2005): "Neurally reconstructing expected utility," *Games and Economic Behavior*, 52, 305 – 315, special Issue on Neuroeconomics.

KŐSZEGI, B. AND M. RABIN (2009): "Reference-Dependent Consumption Plans," *American Economic Review*, 99, 909–936.

LARSON, F., J. A. LIST, AND R. D. METCALFE (2016): "Can Myopic Loss Aversion Explain the Equity Premium Puzzle? Evidence from a Natural Field Experiment with Professional Traders," Tech. rep., National Bureau of Economic Research.

LERMAN, HUGHES, TROCK, AND ET AL (1999): "Genetic testing in families with hereditary nonpolyposis colon cancer," *JAMA*, 281, 1618–1622.

LETTAU, M. AND S. LUDVIGSON (2010): "Measuring and modeling the evolution of risk-return tradeoffs," *Handbook of Financial Econometrics*.

LETTAU, M., M. MAGGIORI, AND M. WEBER (2013): "Conditional Risk Premia in Currency Markets and Other Asset Classes," Tech. rep., National Bureau of Economic Research.

LUNDBLAD, C. (2007): "The risk return tradeoff in the long run: 18362003," *Journal of Financial Economics*, 85, 123 – 150.

LUO, Y. (2010): "Rational Inattention, Long-run Consumption Risk, and Portfolio Choice," *Review of Economic Dynamics*, 13, 843–860.

LYNCH, A. W. (1996): "Decision frequency and synchronization across agents: Implications for aggregate consumption and equity return," *The Journal of Finance*, 51, 1479–1497.

LYTER, D. W., R. O. VALDISERRI, L. A. KINGSLEY, W. P. AMOROSO, AND J. CHARLES R. RINALDO (1987): "The HIV Antibody Test: Why Gay and Bisexual Men Want or Do Not Want to Know Their Results," *Public Health Reports (1974-)*, 102, 468–474.

MERTON, R. C. (1969): "Lifetime portfolio selection under uncertainty: The continuous-time case," *The review of Economics and Statistics*, 247–257.

MITCHELL, M., T. PULVINO, AND E. STAFFORD (2002): "Limited arbitrage in equity markets," *The Journal of Finance*, 57, 551–584.

OSTER, E., I. SHOULSON, AND E. R. DORSEY (2013): "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, 103, 804–30.

PAGEL, M. (2017): "A News-Utility Theory for Inattention and Rebalancing in Portfolio Choice," *Working paper*.

PALACIOS-HUERTA, I. (1999): "The aversion to the sequential resolution of uncertainty," *Journal of Risk and uncertainty*, 18, 249–269.

ROUTLEDGE, B. R. AND S. E. ZIN (2010): "Generalized disappointment aversion and asset prices," *The Journal of Finance*, 65, 1303–1332.

SEGAL, U. (1987): "Some remarks on Quiggin's anticipated utility," *Journal of Economic Behavior & Organization*, 8, 145–154.

——— (1990): "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 58, 349–77.

SHILLER, R. J. (1981): "Do stock prices move too much to be justified by subsequent changes in dividends?" *American Economic Review*, 71, 421–436.

SICHERMAN, N., G. LOEWENSTEIN, D. J. SEPPI, AND S. P. UTKUS (2016): "Financial Attention," *The Review of Financial Studies*, 29, 863–897.

SIMS, C. A. (1998): "Stickiness," in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 49, 317–356.

——— (2006): "Rational inattention: Beyond the linear-quadratic case," *The American economic review*, 158–163.

SULLIVAN, PATRICK S., A. L. AND A. DRAKE (2004): "Failure to Return for HIV Test Results among Persons at High Risk for HIV Infection: Results from a Multistate Interview Project," *JAIDS Journal of Acquired Immune Deficiency Syndromes*.

THALER, R., A. TVERSKY, D. KAHNEMAN, AND A. SCHWARTZ (1997): "The effect of myopia and loss aversion on risk taking: an experimental test," *Quarterly Journal of Economics*, 112, 647–661.

WAKKER, P. (1988): "Nonexpected utility as aversion of information," *Journal of Behavioral Decision Making*, 1, 169–175.

A   Information Aversion

Proof of Proposition 1 Focus on a setup with three dates: 0, 1, and 2. At date 2, the agent receives a final outcome X with cumulative distribution function F. The certainty equivalent under the static disappointment aversion model with linear realized utility and coefficient θ is $\mu_\theta(X \mid F)$ – Equation (1). Define:

$$h(\mu) = \int_{x \geq \mu} (x - \mu) \, dF(x) + (1 + \theta) \int_{x < \mu} (x - \mu) \, dF(x)$$

The function h is continuous, decreasing in μ, and converges to $+\infty$ when μ tends to $-\infty$ and $-\infty$ when μ tends to $+\infty$: there exists as unique zero, the certainty equivalent $\mu_\theta(X \mid F)$.

At date 1, the agent can receive signals $i \in \mathcal{I}$, with cumulative distribution function S, and update her belief on the distribution of X from F to F_i, where $F = \int_i F_i dS(i)$. We are interested in comparing the certainty equivalent at date $t = 0$ of the compound lottery with date $t = 1$ signals, $\mu_\theta[\mu_\theta(X \mid F_i) \mid S]$, as per dynamic disappointment aversion – Definition 1, with that of a lottery without intermediate signal, $\mu_\theta(X \mid F)$.

For all $i \in \mathcal{I}$ the function

$$h_i(\mu) = \int_{x \geq \mu} (x - \mu) \, dF_i(x) + (1 + \theta) \int_{x < \mu} (x - \mu) \, dF_i(x)$$

admits $\mu_\theta(X \mid F_i)$ as a unique zero. To simplify notations, we write $\mu_\theta(X \mid F_i) = \mu_i$.

Under these notations, $\mu_\theta[\mu_\theta(X \mid F_i) \mid S] = \mu_\theta[\mu_i \mid S]$ is the unique zero of

$$h_{info}(\mu) = \int_{\{i \text{ s.t. } \mu_i \geq \mu\}} (\mu_i - \mu) \, dS(i) + (1 + \theta) \int_{\{i \text{ s.t. } \mu_i < \mu\}} (\mu_i - \mu) \, dS(i)$$

To simplify the calculations below, we write $\mu_\theta[\mu_i \mid S] = \mu_{\text{info}}$, and take the notational liberty $\{\mu_i \geq \mu\} = \{i \text{ s.t. } \mu_i \geq \mu\}$, $\{\mu_i < \mu\} = \{i \text{ s.t. } \mu_i < \mu\}$.

Let us compute $h(\mu_{\text{info}})$:

$$
\begin{aligned}
h(\mu_{\text{info}}) &= \int (x - \mu_{\text{info}}) \, dF(x) + \theta \int_{x < \mu_{\text{info}}} (x - \mu_{\text{info}}) \, dF(x) \\
&= \int_i \left[\int (x - \mu_{\text{info}}) \, dF_i(x) + \theta \int_{x < \mu_{\text{info}}} (x - \mu_{\text{info}}) \, dF_i(x) \right] dS(i) \\
&= \int_i \left[(\mu_i - \mu_{\text{info}}) + \theta \int_{x < \mu_i} (\mu_i - x) \, dF_i(x) + \theta \int_{x < \mu_{\text{info}}} (x - \mu_{\text{info}}) \, dF_i(x) \right] dS(i)
\end{aligned}
$$

where we use $F = \int_i F_i dS(i)$ first, and $h_i(\mu_i) = 0, \forall i$ second. Re-organizing the terms yields:

$$h\left(\mu_{\text{info}}\right) = \theta \int_{\mu_i < \mu_{\text{info}}} \left[\left(\mu_{\text{info}} - \mu_i\right) \int_{x \geq \mu_{\text{info}}} dF_i\left(x\right) + \int_{\mu_i \leq x < \mu_{\text{info}}} \left(x - \mu_i\right) dF_i\left(x\right)\right] dS(i) \qquad (15)$$

$$+ \theta \int_{\mu_i \geq \mu_{\text{info}}} \left[\left(\mu_i - \mu_{\text{info}}\right) \int_{x < \mu_{\text{info}}} dF_i\left(x\right) + \int_{\mu_{\text{info}} \leq x < \mu_i} \left(\mu_i - x\right) dF_i\left(x\right)\right] dS(i)$$

Observe all the terms on the right-hand side are positive, so that

$$h\left(\mu_{\text{info}}\right) \geq 0$$

Remember h is decreasing with $\mu_\theta\left(X \mid F\right)$ as its unique zero. Therefore we can conclude

$$\mu_\theta\left[\mu_\theta(X \mid F_i) \mid S\right] \leq \mu_\theta\left(X \mid F\right)$$

Q.E.D

To compare information aversion to models with exogenous information costs, and with models with cognitive constraints, we analyze under which conditions intermediate signals entail no utility costs: $\mu_\theta\left[\mu_\theta(X \mid F_i) \mid S\right] = \mu_\theta\left(X \mid F\right)$ and $h\left(\mu_\theta\left[\mu_\theta(X \mid F_i) \mid S\right]\right) = 0$. From Equation (15), it is straightforward that all $\mu_i = \mu_{\text{info}}$ terms are equal to zero. Suppose $\mu_j \neq \mu_{\text{info}}$. If $\mu_j < \mu_{\text{info}}$, the positive contribution to $h\left(\mu_{\text{info}}\right)$ of the j term is:

$$\left(\mu_{\text{info}} - \mu_j\right) \int_{x \geq \mu_{\text{info}}} dF_j\left(x\right) + \int_{\mu_j \leq x < \mu_{\text{info}}} \left(x - \mu_j\right) dF_j\left(x\right)$$

The first term is zero iff $\forall x \geq \mu_{\text{info}}, F_j\left(x\right) = 0$, i.e. if all outcomes are below μ_{info} in the F_j distribution. Supposing that is the case, let us analyze the second term. From $\mu_j \leq \int x dF_j\left(x\right)$, we know the interval $\mu_j \leq x < \mu_{\text{info}}$ has a non-zero probability weight. Under these conditions, the second term $\int_{\mu_j \leq x < \mu_{\text{info}}} \left(x - \mu_j\right) dF_j\left(x\right)$ is null if and only if $x = \mu_j$ under F_j, and the signal j is degenerate. A similar result obtains if $\mu_j > \mu_{\text{info}}$.

We find:

$$\mu_\theta\left[\mu_\theta(X \mid F_i) \mid S\right] = \mu_\theta\left(X \mid F\right)$$

$$\Leftrightarrow \forall i, \begin{cases} \mu_\theta\left(X \mid F_i\right) = \mu_\theta\left(X \mid F\right) \text{ or} \\ F_i \text{ is degenerate} \end{cases}$$

Information has no utility cost if the intermediate signals are either fully revealing (F_i degenerate), or non-informative, i.e. they leave unchanged the value of the risky payoff ($\mu_\theta\left(X \mid F_i\right) = \mu_\theta\left(X \mid F\right)$). This result proves exogenous information costs cannot be made model-equivalent to information aversion. It also implies that for any level of mutual entropy at the first stage of the lottery, there are intermediate signals that come at no utility costs. Indeed, consider signals that reveals the final out-

come with probability p or nothing with probability $1 - p$. Clearly, they satisfy the conditions above. One can choose p to attain any level of mutual entropy between the first stage outcome and the final outcome.

B Certainty Equivalent Rate

Take the time interval T as given. By Definition 2, valid under any i.i.d process for risk:

$$\exp\left(v\left(T\right)T\right) = \mu_\theta\left(\frac{X_{t+T}}{X_t} \mid \mathcal{F}_t\right), \forall t.$$

We rewrite, for any t

$$\frac{X_{t+T}}{X_t} = \exp\left(gT\right)Y_T,$$

where

$$\log(\mathbb{E}[X_{t+1}/X_t]) = g$$

and

$$\mathbb{E}\left(Y_T\right) = 1.$$

Then

$$\exp\left(\left(v\left(T\right) - g\right)T\right) = \mu_\theta\left(Y_T \mid \mathcal{F}\right),$$

and the certainty equivalent rate is the sum of two elements: the expected growth rate g and a risk-adjustment $\tilde{v}(.)$ where

$$\exp\left(\tilde{v}\left(T\right)T\right) = \mu_\theta\left(Y_T \mid \mathcal{F}\right).$$

To simplify the notations, we assume $g = 0$ and study $\tilde{v}(.) = v(.)$. We often use v for $v(.)$.

B.1 Brownian case - Proof of Lemma 1

Let $\{X_t\}$ follow a Brownian with expected growth $g = 0$ and volatility σ:

$$\frac{dX_t}{X_t} = \sigma dZ_t,$$

and $\log\left(\frac{X_{t+T}}{X_t}\right) = \log Y_T = -\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\epsilon$, where ϵ is distributed $\mathcal{N}\left(0, 1\right)$, in the notations above.

The certainty equivalent of payoff Y_T is:

$$\mu_\theta\left(Y_T \mid \mathcal{F}\right) = \frac{1 + \theta \int_{Y_T < \mu_\theta(Y_T \mid \mathcal{F})} Y_T dF\left(Y_T\right)}{1 + \theta \int_{Y_T < \mu_\theta(Y_T \mid \mathcal{F})} dF\left(Y_T\right)}.$$

Expanding, we get

$$
\exp\left(v\left(T\right)T\right) = \frac{1 + \theta \int_{\epsilon < \left(v(T) + \frac{1}{2}\sigma^2\right)\frac{\sqrt{T}}{\sigma}} \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\epsilon\right) \frac{\exp\left(-\frac{1}{2}\epsilon^2\right)}{\sqrt{2\pi}} d\epsilon}{1 + \theta \int_{\epsilon < \left(v(T) + \frac{1}{2}\sigma^2\right)\frac{\sqrt{T}}{\sigma}} \frac{\exp\left(-\frac{1}{2}\epsilon^2\right)}{\sqrt{2\pi}} d\epsilon}
$$

$$
= \frac{1 + \theta \int_{\epsilon - \sigma\sqrt{T} < \left(v(T) - \frac{1}{2}\sigma^2\right)\frac{\sqrt{T}}{\sigma}} \frac{\exp\left(-\frac{1}{2}\left(\epsilon - \sigma\sqrt{T}\right)^2\right)}{\sqrt{2\pi}} d\epsilon}{1 + \theta \int_{\epsilon < \left(v(T) + \frac{1}{2}\sigma^2\right)\frac{\sqrt{T}}{\sigma}} \frac{\exp\left(-\frac{1}{2}\epsilon^2\right)}{\sqrt{2\pi}} d\epsilon}
$$

and, finally,

$$
\exp\left(v\left(T\right)T\right) = \frac{1 + \theta\Phi\left(\frac{\sqrt{T}}{\sigma}\left(v\left(T\right) - \frac{1}{2}\sigma^2\right)\right)}{1 + \theta\Phi\left(\frac{\sqrt{T}}{\sigma}\left(v\left(T\right) + \frac{1}{2}\sigma^2\right)\right)}, \tag{16}
$$

where Φ is the cumulative distribution function of the normal $\mathcal{N}(0,1)$.

Risk adjustment. From Equation (16), it is immediate that $\exp\left(v\left(T\right)T\right) < 1 \implies v(.) < 0$.

Role of observation interval T**.** Write $\frac{1}{\sigma}\sqrt{T}v\left(T\right) = g\left(T, \sigma\right) < 0$, then Equation (16) becomes

$$
\exp\left(\sigma\sqrt{T}g\left(T,\sigma\right)\right) = \frac{1 + \theta\Phi\left(g\left(T,\sigma\right) - \frac{1}{2}\sigma\sqrt{T}\right)}{1 + \theta\Phi\left(g\left(T,\sigma\right) + \frac{1}{2}\sigma\sqrt{T}\right)}.
$$

Let $z = \frac{1}{2}\sigma\sqrt{T}$, then $g(T, \sigma) = g(z) < 0$ with

$$
\exp(2zg\left(z\right)) = \frac{1 + \theta\Phi\left(g\left(z\right) - z\right)}{1 + \theta\Phi\left(g\left(z\right) + z\right)} \tag{17}
$$

Differentiating, we obtain

$$
2g\left(z\right) + 2zg'\left(z\right) = \theta\left(\frac{\left(g'\left(z\right) - 1\right)\Phi'\left(g\left(z\right) - z\right)}{1 + \theta\Phi\left(g\left(z\right) - z\right)} - \frac{\left(g'\left(z\right) + 1\right)\Phi'\left(g\left(z\right) + z\right)}{1 + \theta\Phi\left(g\left(z\right) + z\right)}\right)
$$

$$
= -2\theta\frac{\Phi'\left(g\left(z\right) - z\right)}{1 + \theta\Phi\left(g\left(z\right) - z\right)}. \tag{18}
$$

Let $u\left(x\right) = \log\left(1 + \theta\Phi\left(x\right)\right)$, $u'\left(x\right) = \frac{\theta\Phi'\left(x\right)}{1 + \theta\Phi\left(x\right)} > 0$ and $u''\left(x\right) = \frac{-\theta\Phi'\left(x\right)\left[x\left(1 + \theta\Phi\left(x\right)\right) + \theta\Phi'\left(x\right)\right]}{\left(1 + \theta\Phi\left(x\right)\right)^2}$ with $\left[x\left(1 + \theta\Phi\left(x\right)\right) + \theta\Phi'\left(x\right)\right]' = \left(1 + \theta\Phi\left(x\right)\right) > 0$. u'' goes from $+\infty$ to $-\infty$ with a unique zero $\implies u$ is increasing convex then concave with a unique inflection point x^*.

From Equation (17), it is relatively straightforward to show $\frac{\theta\Phi'(g(z) + z)}{1 + \theta\Phi(g(z) + z)} = \frac{\theta\Phi'(g(z) - z)}{1 + \theta\Phi(g(z) - z)}$ and $u'\left(g\left(z\right) + z\right) = u'\left(g\left(z\right) - z\right) \implies \forall z, g\left(z\right) - z \leq x^* \leq g\left(z\right) + z$.

Because u is convex between $g(z) - z$ and x^*, and concave between $g(z) + z$ and x^*,

$$u(x^*) - u(g(z) - z) \geq (x^* - (g(z) - z)) u'(g(z) - z)$$
$$u(g(z) + z) - u(x^*) \geq ((g(z) + z) - x^*) u'(g(z) + z).$$

Putting these results together, $u(g(z) + z) - u(g(z) - z) \geq 2zu'(g(z) + z)$ and, using Equation (17), the definition of $u(.)$, and $u'(g(z) + z) = u'(g(z) - z)$:

$$-2zg(z) \geq 2zu'(g(z) - z)$$

As per Equation (18), this proves $g'(z) \geq 0$ for all z.

From the definition of g, $v(T) = \frac{\sigma}{\sqrt{T}} g\left(\frac{1}{2}\sigma\sqrt{T}\right)$, and $v'(T) = \frac{\sigma}{2T\sqrt{T}}(zg'(z) - g(z)) \implies v'(T) \geq 0$ for all T.

Role of the volatility σ. From the definition of g, and for T given, let $v(\sigma) = \frac{\sigma}{\sqrt{T}} g\left(\frac{1}{2}\sigma\sqrt{T}\right)$. We find $\sqrt{T} v'(\sigma) = g(z) + zg'(z) \leq 0$, from Equation (18) $\implies v(.)$ in always decreasing in σ.

Role of the disappointment aversion θ. Let T and σ be given, and differentiate the right-hand side of Equation (17) with respect to θ. We obtain: $\frac{\Phi(g(z) - z) - \Phi(g(z) + z)}{(1 + \theta\Phi(g(z) + z))^2} \leq 0 \implies v(.)$ in always decreasing in θ.

Q.E.D

B.2 Brownian case - Properties of $\frac{dv(T)}{d\log(T)}$ useful for Propositions 2, 3 and 4

- $Tv(T)$ **is decreasing with finite limit in** $+\infty$.

  From $z = \frac{1}{2}\sigma\sqrt{T}$, $v(T) = \frac{\sigma}{\sqrt{T}} g(\frac{1}{2}\sigma\sqrt{T})$ and $v'(T) = \frac{\sigma}{2T\sqrt{T}}(zg'(z) - g(z))$ above, obtain $Tv'(T) + v = \frac{\sigma}{2\sqrt{T}}(zg'(z) + g(z)) \leq 0$ from Equation (18) $\implies Tv(T)$ is decreasing everywhere.

  Its limit in $+\infty$ directly derives from Equation (16), and $v(.)$ converging to zero in $+\infty$.

- **Continuous information limit:** $\sqrt{T} v(T) \underset{T \to 0}{\longrightarrow} -\kappa\sigma$ **with** κ **finite function of** θ.

  From $z = \frac{1}{2}\sigma\sqrt{T}$, $v(T) = \frac{\sigma}{\sqrt{T}} g(\frac{1}{2}\sigma\sqrt{T})$ and $v'(T) = \frac{\sigma}{2T\sqrt{T}}(zg'(z) - g(z))$ above, obtain $\sqrt{T} v'(T) + \frac{1}{2\sqrt{T}} v = \frac{\sigma}{2T} zg'(z) \geq 0$ from our previous analysis $\implies \sqrt{T} v(T)$ is increasing everywhere, and admits a limit in zero (either finite or $-\infty$).

  Suppose $\sqrt{T} v(T) \to -\infty$, and $Tv(T) \to 0$ from above. From Equation (16):

$$1 + Tv(T) \approx \left(1 + \theta\Phi\left(\frac{1}{\sigma}\sqrt{T}\left(v(T) - \frac{1}{2}\sigma^2\right)\right)\right)\left(1 - \theta\Phi\left(\frac{1}{\sigma}\sqrt{T}\left(v(T) + \frac{1}{2}\sigma^2\right)\right)\right)$$

and $\sqrt{T} v(T) \approx -\theta\frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} T(v(T))^2\right)$, a contradiction in orders of magnitude if $\sqrt{T} v(T) \to -\infty \implies \sqrt{T} v(T) \to -\kappa\sigma$, where $\kappa \geq 0$ is finite.

37

Reorganizing Equation (16) at the limit $T \to 0$ yields $g(\kappa) = \kappa + \theta\kappa\Phi(-\kappa) - \theta\Phi'(-\kappa) = 0$. Note $g(0) < 0$, $g \xrightarrow[+\infty]{} +\infty$ and $g'(x) = 1 + \theta\Phi(-x) > 0$ for all $x \geq 0$. These guarantee the existence and uniqueness of $\kappa > 0$ solution to $g(\kappa) = 0$, and function of only the parameter θ.

- $\frac{dv(T)}{d\log(T)}$ **increasing in σ.**

  We differentiate $Tv'(T) = \frac{\sigma}{2\sqrt{T}}(zg'(z) - g(z))$ with respect to σ, and use $z = \frac{1}{2}\sigma\sqrt{T}$, to obtain:

  $$2\sqrt{T}\frac{d(Tv'(T))}{d\sigma} = \left(zg'(z) - g(z) + z\sigma\frac{dz}{d\sigma}g''(z)\right) = zg'(z) - g(z) + z^2 g''(z).$$

  From Equation (18), $g(z) + zg'(z) = -u'(z + g(z)) = -u'(z - g(z))$, so $2g'(z) + zg''(z) = -(1 + g'(z))u''(z + g(z))$, and $zg'(z) - g(z) + z^2 g''(z) = -z(1 + g'(z))u''(z + g(z)) - (g(z) + zg'(z))$.

  Equation (18) yields $g(z) + zg'(z) \leq 0$, and, from above, we also have $g'(z) \geq 0$ and $u''(z + g(z)) < 0 \implies zg'(z) - g(z) + z^2 g''(z) \geq 0$ and $\frac{d}{d\sigma}\left(\frac{dv(T)}{d\log(T)}\right) \geq 0$.

- $\frac{dv(T)}{d\log(T)}$ **increasing in θ.**

  From $Tv'(T) = \frac{\sigma}{2\sqrt{T}}(zg'(z) - g(z))$ and Equation (18), we derive $\frac{2\sqrt{T}}{\sigma}Tv'(T)) = (-u'(g(z) + z) - 2g(z))$ and so

  $$\frac{2\sqrt{T}}{\sigma}\frac{d(Tv'(T))}{d\theta} = -\frac{dg(z)}{d\theta}\left(\frac{du'(g(z) + z)}{d\theta} + 2\right).$$

  $\frac{du'(.)}{d\theta} \geq 0$ is trivial, and as shown above $\frac{dg(z)}{d\theta} \leq 0 \implies \frac{d}{d\theta}\left(\frac{dv(T)}{d\log(T)}\right) \geq 0$.

- $\frac{dv(T)}{d\log(T)}$ **decreasing in T.**

  Differentiate $Tv'(T) = \frac{\sigma}{2\sqrt{T}}(zg'(z) - g(z))$ with respect to T to obtain $\frac{2}{\sigma}\frac{d(Tv'(T))}{dT} = -\frac{1}{2T\sqrt{T}}(zg'(z) - g(z)) + \frac{1}{\sqrt{T}}z\frac{dz}{dT}g''(z) = \frac{2}{\sigma}\frac{d(Tv'(T))}{dT} = -\frac{1}{2T\sqrt{T}}(zg'(z) - g(z) - z^2 g''(z))$.

  Assume $g''(z) \leq 0$, then $zg'(z) - g(z) - z^2 g''(z) \geq 0$ is immediate, and $\frac{d(Tv'(T))}{dT} \leq 0$ obtains.

  Now assume $g''(z) \geq 0$. Reorganizing $2g'(z) + zg''(z) = -(1 + g'(z))u''(z + g(z))$, we derive $zg'(z) - g(z) - z^2 g''(z) = 3zg'(z) - g(z) - z^2(1 - (g'(z))^2)u'(z + g(z))$.

  Remember $g(z) + u'(z + g(z)) \leq 0$, so, if $z^2(1 - g'^2) \leq 1$, we find again $zg'(z) - g(z) - z^2 g''(z) \geq 0$ and $\frac{d(Tv'(T))}{dT} \leq 0$ obtains.

  Finally, if $z^2(1 - g'^2) \geq 1$, then we have $(zg''(z))' = (z(1 - g'^2)u'(z + g) - 2g')' = (1 - z^2(1 - g'^2))(1 - g'^2)u'(z + g) - 2g'' \leq 0$, and, since $g''(z) \geq 0$, $zg'(z) - g(z) - z^2 g''(z)$ is increasing.

  In $T = 0$, $v(T) \approx -\kappa\sigma/\sqrt{T}$, and both $g''(z)$ and $zg'(z) - g(z) - z^2 g''(z)$ are positive at the limit $z = 0$. As long as g'' stays positive, $zg'(z) - g(z) - z^2 g''(z)$ remains positive. If, at any further point, g'' turns from negative to positive, since g'' is smooth, it is equal to zero at the turning point, and $zg'(z) - g(z) - z^2 g''(z)$ increases from $zg'(z) - g(z) \geq 0$ thereafter (for as long as g'' is positive). We find $zg'(z) - g(z) - z^2 g''(z) \geq 0$ for all $z \implies \frac{d}{dT}\left(\frac{dv(T)}{d\log(T)}\right) \leq 0$ for all T.

38

B.3   General Properties - Section 2.4

Using the definition of Appendix A, and noting $x_T T \sim \log(\frac{X_{t+T}}{X_t})$, $v(T)$ is the unique zero of the decreasing function h:

$$h(v) = \int_{x_T \geq v} (\exp(x_T T) - \exp(vT)) \, dF(x_T) + (1+\theta) \int_{x_T < v} (\exp(x_T T) - \exp(vT)) \, dF(x_T),$$

$$h(v) = 1 - \exp(vT) + \theta \int_{x_T < v} (\exp(x_T T) - \exp(vT)) \, dF(x_T).$$

- h is decreasing in θ, so its zero is also decreasing in $\theta \to v(T)$ is decreasing in θ, the coefficient of disappointment aversion.

- For a given v, increasing the support for $x_T < v$ and the probability $F(x_T < v)$ implies a lower $h(v) \implies v(.)$ is decreasing in the the quantity of risk, e.g. the mean-preserving spread in the returns distribution.

- The first term $1 - \exp(vT)$ is decreasing in v with a zero at $v = 0$. $\int_{x_T < v} (\exp(x_T T) - \exp(vT)) \, dF(x_T)$ is also decreasing in v and strictly negative when x_T is not degenerate $\to v(T) < 0$.

- **Limit in** $T \to +\infty$**.** By the central limit theorem, as T becomes large, x_T approaches a normal, and $v(T) \underset{+\infty}{\longrightarrow} 0$, and $v(T)T$ bounded, as in the Brownian case.

- **Subadditivity** From Proposition 1, $v(x+y) > \frac{x}{x+y} v(x) + \frac{y}{x+y} v(y), \forall (x, y) > 0$, and $-Tv(T)$ is subadditive.

B.4   Jump distribution - Section 2.4

We consider a pure jump process. Write N_t the counting variable for a Poisson jump process with intensity λ. Define the process $\{X_t\}$ by the stochastic differential equation:

$$\frac{dX_t}{X_{t-}} = \lambda \sigma dt - \sigma dN_t,$$

where $\sigma < 1$, and without loss of generality $X_0 = 1$. The value of X_t decreases geometrically at each jump. The drift term compensates for the average decrease, so that $\{X_t\}$ is a martingale. Solving this S.D.E. with initial condition $X_0 = 1$, $X_t = \exp(\lambda \sigma t + \log(1 - \sigma) N_t)$. We are interested in the certainty equivalent of a lottery paying X_T for various values of T.

A few standard results on Poisson jump processes will be useful:

$$P[N_t = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}; P[N_t = 0] = e^{-\lambda t}; P[N_t \leq k] = \frac{\Gamma(k+1, \lambda t)}{k!}$$

where $\Gamma(.,.)$ is the incomplete gamma function.

If the certainty equivalent is between the points of the distribution corresponding to k and $k+1$ jumps, we can compute it exactly. This corresponds to the condition:

$$(1-\sigma)^{k+1} \leq \exp(v(T)T)\exp(-\lambda\sigma T) \leq (1-\sigma)^{k}.$$

Then, we get immediately

$$\exp\left(-\lambda\sigma T\right)V = \frac{(1+\theta)\exp\left(-\lambda\sigma T\right) - \theta\mathbb{E}\left[(1-\sigma)^{N_t}1_{N_T \leq k}\right]}{(1+\theta) - \theta P\left[N_T \leq k\right]}.$$

where we use $\frac{\mathbb{E}[y]+\theta\mathbb{E}[y1_{y<V}]}{1+\theta\mathbb{P}[y<V]} = \frac{(1+\theta)\mathbb{E}[y]-\theta\mathbb{E}[y1_{y\geq V}]}{(1+\theta)-\theta\mathbb{P}[y\geq V]}$.

Note that

$$\mathbb{E}\left[(1-\sigma)^{N_t}1_{N_T \leq k}\right] = e^{-\lambda T}\sum_{i=0}^{k}(1-\sigma)^i\frac{(\lambda T)^i}{i!}$$

$$= e^{-\lambda T+(1-\sigma)\lambda T}e^{-(1-\sigma)\lambda T}\sum_{i=0}^{k}\frac{((1-\sigma)\lambda T)^i}{i!}$$

$$= e^{-\lambda\sigma T}\frac{\Gamma\left(k+1,(1-\sigma)\lambda T\right)}{k!}.$$

Re-organizing, we obtain

$$\exp(v(T)T) = \frac{1 - \frac{\theta}{1+\theta}\frac{\Gamma(k+1,(1-\sigma)\lambda T)}{k!}}{1 - \frac{\theta}{1+\theta}\frac{\Gamma(k+1,\lambda T)}{k!}}.$$

As the certainty equivalent is unique, there is a unique k so that the corresponding $\exp(v(T)T)$ falls in the right interval.

Remark 1. In matlab, the incomplete gamma function is defined such that $\Gamma(k+1,x)/k! = \mathbf{gammainc}(x,k+1)$.

Remark 2. At the points where we go from one k to the next, we have $\exp(v(T)T) = (1-\sigma)^k\exp\left(\lambda\sigma T\right)$.

Some properties:

- **Continuous information limit.** We prove $v(T) \to -\theta\sigma\lambda$ and $v(.)$ increasing around $T=0$. In the limit where T gets close to 0, the certainty equivalent falls in the region between 0 and 1 jumps. We guess and verify this result and obtain the limiting behavior of V as T converges to 0. In this case we have $\exp(v(T)T) = \frac{1-\frac{\theta}{1+\theta}\exp(-(1-\sigma)\lambda T)}{1-\frac{\theta}{1+\theta}\exp(-\lambda T)}$, which clearly converges to 1 as T converges to 0 so the guess is verified. In the limit $T \to 0$, we get $v(T) \approx -\theta\sigma\lambda+\theta\lambda^2\sigma(1+\theta-\frac{\sigma}{2})T$, and the stated results.

- **Role of the shock size σ.** We show v is decreasing in σ.

if $\frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} \notin \mathbb{N}$, then for any $\sigma > 0$, $k(\sigma) = k(\sigma + \epsilon)$ for $|\epsilon|$ sufficiently small, and v is decreasing in σ simply because $\Gamma(x, y)$ is decreasing in y.

if $\frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} \in \mathbb{N}$, observe we can equivalently use $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$, or $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} - 1$. Because $\frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$ is increasing in σ, $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$ remains valid for $\sigma + \epsilon$ and $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} - 1$ remains valid for $\sigma - \epsilon$, for any $\epsilon > 0$ sufficiently small. Once more, we find v is decreasing in σ simply because $\Gamma(x, y)$ is decreasing in y.

- **Role of the shock intensity** λ. We show v is decreasing in λ.

  if $\frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} \in \mathbb{N}$, observe we can equivalently use $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$, or $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} - 1$. Because $\frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$ is increasing in λ, $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)}$ remains valid for $\lambda + \epsilon$ and $k = \frac{(v(T)-\lambda\sigma)T}{\log(1-\sigma)} - 1$ remains valid for $\lambda - \epsilon$, for any $\epsilon > 0$ sufficiently small. We can thus simply look at the derivative of v with respect to λ, keeping k as exogenous. We find $dv/d\lambda$ of same sign as $(1-\sigma)^{k+1} - \exp(v(T)T)\exp(-\lambda\sigma T)$, which, by definition of k proves v decreasing in λ.

B.5   Empirical distribution of returns - Section 2.4

We obtain daily returns on the CRSP value-weighted market index of all stocks listed on the NYSE, AMEX and NASDAQ from 1926 to 2016. We compute the certainty equivalent rate at horizon T by simulating 1,000,000 realization of returns. To obtain one such realization, we draw $\lfloor 252 \times T \rfloor$ independent equiprobable draws of historical daily returns. The factor 252 reflects the average number of trading days in a given year.

C   Problem (5): optimal information, consumption and savings decisions

C.1   Proof of Proposition 2

For a given optimization time t, we let the time interval T be fixed, and, because all the values scale up with wealth, we assume initial wealth is $W_t = 1$. Given $\mathcal{C}$ set aside to finance consumption over $[t, t+T]$, the deterministic consumption stream $\{C_\tau\}_{\tau \in [0,T]}$ optimizes:

$$\max_{\{C_\tau\}_{\tau \in [0,T]}} \left[\int_0^T e^{-\rho\tau} C_\tau^{1-\alpha} d\tau \right]^{\frac{1}{1-\alpha}} \text{ s.t. } \int_0^T e^{-r\tau} C_{t+\tau} d\tau = \mathcal{C},$$

with solution: $\forall \tau, C_\tau = \mathcal{C} \left(\frac{1-\exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{\frac{\rho+(\alpha-1)r}{\alpha}} \right)^{-1} e^{\frac{r-\rho}{\alpha}\tau}$ s.t. $\int_0^T e^{-\rho\tau} C_{t+\tau}^{1-\alpha} d\tau = \mathcal{C}^{1-\alpha} \left(\frac{1-\exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{\frac{\rho+(\alpha-1)r}{\alpha}} \right)^{\alpha}$.

From $1 - \mathcal{C}$, the wealth remaining, S is invested in the risky asset. The continuation value after T is:

$$\mathcal{V}_{t+T} = \mathcal{V}\frac{W_{t+T}}{W_t} = \mathcal{V}\left(S\frac{X_{t+T}}{X_t} + (1 - \mathcal{C} - S)e^{rT}\right).$$

Because the savings invested in the risk-free asset grow deterministically, $(1 - \mathcal{C} - S)e^{rT}$ contains no risk: $\mu_\theta(S\frac{X_{t+T}}{X_t} + (1 - \mathcal{C} - S)e^{rT} \mid \mathcal{F}_t) = S\mu_\theta(\frac{X_{t+T}}{X_t} \mid \mathcal{F}_t) + (1 - \mathcal{C} - S)e^{rT}$, and we derive the optimization problem in Equation (5).

Using Definition 2, we replace $\mu_\theta(\frac{X_{t+T}}{X_t} \mid \mathcal{F}_t)$ with $\exp(v(T)T)$ and S optimizes: $\max_{S \leq 1-\mathcal{C}} Se^{v(T)T} + (1 - \mathcal{C} - S)e^{rT})$ which yields the corner solution: $S = 0$, if $v(T) < r$, $S = 1 - \mathcal{C}$, if $v(T) > r$.

Assume $v(T) > r$ (we'll verify it is the case at the optimum), and insert into the value function the optimal consumption over the time interval, and the optimal risk savings, given $\mathcal{C}$. Combining Definition 1 and Definition 2:

$$\mathcal{V}^{1-\alpha} = \left[\mathcal{C}^{1-\alpha}\left(\frac{1 - \exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{\frac{\rho+(\alpha-1)r}{\alpha}}\right)^\alpha + \exp(-\rho T)(1-\mathcal{C})^{1-\alpha}\mathcal{V}^{1-\alpha}\exp((1-\alpha)v(T)T)\right].$$

and the FOC condition in $\mathcal{C}$ yields:

$$(\mathcal{C})^{-1}\left(\frac{1 - \exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{\frac{\rho+(\alpha-1)r}{\alpha}}\right) = \left(\frac{1 - \exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{\frac{\rho+(\alpha-1)r}{\alpha}}\right) + \left(\exp(-\rho T)(\mathcal{V}\exp(v(T)T))^{1-\alpha}\right)^{\frac{1}{\alpha}}$$

Through a series of (uninteresting) re-arranging steps, we derive:

$$\mathcal{V}^{\frac{1-\alpha}{\alpha}}\left(\frac{\rho + (\alpha - 1)r}{\alpha}\right) = \frac{1 - \exp\left(-\frac{\rho+(\alpha-1)r}{\alpha}T\right)}{1 - \exp\left(-\frac{\rho+(\alpha-1)v(T)}{\alpha}T\right)}, \tag{19}$$

and

$$\mathcal{C} = 1 - \exp\left[\left(-\frac{\rho}{\alpha} + \frac{1-\alpha}{\alpha}v(T)\right)T\right],$$

the optimal wealth set aside in the consumption account given T.

Note that all of the analysis above holds for any distribution of returns with i.i.d increments: Equation (7) is a general result.

To complete the proof of Proposition 2, in the Brownian case with growth rate g, derive $\frac{d\mathcal{C}}{dT}$. It is of same sign as $\rho - (1 - \alpha)(v(T) + Tv'(T))$.

If $\alpha \leq 1$, from result $T(v(T) - g)$ decreasing in Appendix B.2, and $\rho > (1-\alpha)g$ by assumption, we get $v(T) + Tv'(T) \leq g \leq \frac{\rho}{1-\alpha}$ and $\frac{d\mathcal{C}}{dT} \geq 0$.

If $\alpha \geq 1$, from $v(T) > r \geq 0$ (assumed for now, proven later to hold at the optimal time interval T^*) and $v'(T) \geq 0$, we trivially find $\rho + (\alpha - 1)(v(T) + Tv'(T)) \geq 0$ and $\frac{d\mathcal{C}}{dT} \geq 0$.

Q.E.D

C.2 Proof of Proposition 3

From Equation (19), the optimal value $\mathcal{V}$ varies with $\tilde{\mathcal{V}} = \left[\frac{1 - \exp\left(-\frac{\rho + (\alpha - 1)r}{\alpha}T\right)}{1 - \exp\left(-\frac{\rho + (\alpha - 1)v(T)}{\alpha}T\right)} \right]^{\frac{\alpha}{1-\alpha}}$.

- If $v(T) = r$, then $\tilde{\mathcal{V}} = 1$.

- at $T \to +\infty$, $v(T) - g \to 0$ and $(v(T) - g)T$ bounded from the general properties of $v - g$ derived in Appendix B.3. So $\tilde{\mathcal{V}} \approx \left[\frac{1 - \exp\left(-\frac{\rho + (\alpha - 1)r}{\alpha}T\right)}{1 - \exp\left(-\frac{\rho + (\alpha - 1)g}{\alpha}T\right)} \right]^{\frac{\alpha}{1-\alpha}} > 1$ since $g > r$, for large $T \implies$

- if $v(0) \leq r$, the value when $T \to +\infty$ is greater than the value the investor gets from continuous observations (and investing in the risk-free rate) $\implies$ some inattention is optimal, $T^* \neq 0$, and $v(T^*) > r$.

- if $g \geq v(0) > r$ bounded, close to $T = 0$, $\tilde{\mathcal{V}} \approx \left[\frac{\rho + (\alpha - 1)r}{\rho + (\alpha - 1)v(T)} \right]^{\frac{\alpha}{1-\alpha}}$ increasing in $v(T) \implies$

- if $v(0) > r$ and $v(.)$ increasing around $T = 0 \implies$ some inattention is optimal, $T^* \neq 0$, and $v(T^*) > r$.

Note these results are general, and true for any distribution of returns with i.i.d increments. They guarantee some inattention is optimal in the Brownian case, and in the Jump case, since $v(.)$ is increasing from $T = 0$. Also $v(T^*) > r$.

We derive the fist order condition for the optimal observation interval, by deriving $\frac{d\mathcal{V}_0}{dT} = 0$ from Equation (19). Some cumbersome calculations yield:

$$\frac{d\mathcal{V}}{dT} = 0 \Leftrightarrow \begin{cases} \frac{\left(r - \frac{\rho}{1-\alpha}\right)\exp\left(\frac{1-\alpha}{\alpha}\left(r - \frac{\rho}{1-\alpha}\right)T\right)}{1 - \exp\left(\frac{1-\alpha}{\alpha}\left(r - \frac{\rho}{1-\alpha}\right)T\right)} \\ -\frac{\left(v - \frac{\rho}{1-\alpha}\right)\exp\left(\frac{1-\alpha}{\alpha}\left(v - \frac{\rho}{1-\alpha}\right)T\right)}{1 - \exp\left(\frac{1-\alpha}{\alpha}\left(v - \frac{\rho}{1-\alpha}\right)T\right)} \end{cases} = v'(T)\,T\frac{\exp\left(\frac{1-\alpha}{\alpha}\left(v - \frac{\rho}{1-\alpha}\right)T\right)}{1 - \exp\left(\frac{1-\alpha}{\alpha}\left(v - \frac{\rho}{1-\alpha}\right)T\right)}.$$

Re-organizing the terms yields, at the optimum T^*:

$$\frac{dv(T^*)}{d\log(T)} = \left(\frac{\rho}{1-\alpha} - v(T^*) \right) \left[1 - \frac{f\left(\frac{\rho}{1-\alpha} - r, T^*\right)}{f\left(\frac{\rho}{1-\alpha} - v(T^*), T^*\right)} \right]$$

where $f(x, T) = x / \left(\exp\left(\frac{1-\alpha}{\alpha}xT\right) - 1 \right)$ – Equation (8), which holds for any distribution of returns with i.i.d increments..

Q.E.D

43

C.3 Proof of Proposition 4

We now assume a Brownian risk. From Appendix B.2, the left-hand side of Equation (8) – the implied utility costs of information – is increasing in σ, in θ and decreasing in T. It is unchanged by variations in g.

From these results, if the right-hand side of Equation (8) is kept constant – the case when $v(.)$ is constant – then increases in σ (and in θ) make the optimal T^* greater, i.e. increase inattention: Point 4 in Proposition 4.

Let's turn to the right-hand side of Equation (8). To simplify the calculations, note $x = \frac{\rho}{1-\alpha} - v(T)$, $y = \frac{\rho}{1-\alpha} - r$, and our interest is in the right-hand side: $RHS = x - y\frac{e^{\frac{1-\alpha}{\alpha}xT}-1}{e^{\frac{1-\alpha}{\alpha}yT}-1}$.

Since $v(T) > r$ around the optimum, $\frac{\rho}{1-\alpha} - v(T) < \frac{\rho}{1-\alpha} - r$, i.e. $x < y$ and we can show $\frac{\partial RHS}{\partial T} > 0$, i.e. the right-hand side of Equation (8) – the implied benefits of information – is increasing in T. (A first-order approximation at $T = 0$ yields $RHS \approx \frac{1}{2}\frac{1-\alpha}{\alpha}x(y-x)T$ which is always positive, so $RHS \geq 0$ for all T).

Now, compute $\frac{\partial RHS}{\partial x}$. We can easily show that $\frac{\partial^2 RHS}{\partial x^2}$ is of same sign as $-y$.

If $\alpha \geq 1$, the second derivative is positive (since $y < 0$, and $\frac{\partial RHS}{\partial x}$ is increasing in x. Its greatest value is for the highest possible x, i.e. $x = y$, where we find $\frac{\partial RHS}{\partial x}|_{x=y} < 0 \Longrightarrow RHS$ is decreasing in x, i.e. increasing in $v(T)$.

If $\alpha \leq 1$, the second derivative is negative (since $y > 0$, and $\frac{\partial RHS}{\partial x}$ is decreasing in x. Its greatest value is for the lowest possible $x = \frac{\rho}{1-\alpha} - g$, where we find $\frac{\partial RHS}{\partial x}|_{v(T)=g} < 0 \Longrightarrow RHS$ is decreasing in x, i.e. increasing in $v(T)$.

Combining these results we obtain:

- an increase in g causes an increase in $v(.)$, and thus on the right-hand side (RHS) of Equation (8); it leaves the left-hand side (LHS) unchanged. To bring the RHS back down, T decreases $\Longrightarrow$ inattention decreases.

- an increase in σ causes a decrease in $v(.)$, and thus on the right-hand side (RHS) of Equation (8); it increases the left-hand side (LHS). To bring the RHS back up, and the LHS back down, T must increase $\Longrightarrow$ inattention increases.

- an increase in θ causes a decrease in $v(.)$, and thus on the right-hand side (RHS) of Equation (8); it increases the left-hand side (LHS). To bring the RHS back up, and the LHS back down, T must increase $\Longrightarrow$ inattention increases.

- an increase in σ compensated by an increase in g that leaves $v(.)$ unchanged increases the left-hand side (LHS) of Equation (8); it leaves the right-hand side (RHS) unchanged. To bring the LHS back down, T increases $\Longrightarrow$ inattention increases.

Q.E.D.

C.4 Extension to transaction costs

We consider a model with transaction costs in addition to information aversion, along the lines of Alvarez et al. (2012). We still assume the agent has recursive disappointment aversion but now utility is defined over durable rather than non-durable consumption. The price of a unit of durable good is 1 and it depreciates at rate δ. The agent can also invest in the same financial assets as before. Finally, at each adjustment of the stock of durable, a fraction ϕ_T of the existing stock is lost. The current position of the agent can be summarized by three state variables: her position in durables C_t, her wealth invested in financial assets A_t, and the fraction of these financial assets invested in the risky asset θ_t. For simplicity of notations, we consider the case where $\theta_t = 1$, and note R_T the random return over an horizon T. Allowing optimization over the portfolio is straightforward.

First, note that the value function is homogenous of degree 1 with respect to (C_t, A_t) because the consumption opportunity set is linear in those, and the utility function is homogenous of degree 1 in consumption. Therefore it is enough to solve for values $(1, A_t)$, and we drop the first argument in the notations hereafter. We are interested in the value right after an observation. We write $\bar{\mathcal{V}}$ the value function conditional on choosing to not transact. We write $\hat{\mathcal{V}}$ the value function conditional on choosing to transact. Then clearly,

$$\mathcal{V} = \max\left(\bar{\mathcal{V}}, \hat{\mathcal{V}}\right)$$

Conditional on no transaction, the optimization problem is

$$\bar{\mathcal{V}}(A)^{1-\alpha} = \max_{T} \int_0^T e^{-\rho\tau} e^{-\delta\tau(1-\alpha)} dt + e^{-(\rho+(1-\alpha)\delta)\tau} \mu_\theta \left[\mathcal{V}\left(A e^{\delta\tau} R_T\right)\right]^{1-\alpha}.$$

Conditional on a transaction, the agent solves

$$\hat{\mathcal{V}}(A)^{1-\alpha} = \max_{T, 0 \leq A' \leq A+1-\phi_T} \int_0^\tau e^{-\rho t} \left([A+(1-\phi_T)-A']e^{-\delta t}\right)^{1-\alpha} dt$$
$$+ e^{-(\rho+(1-\alpha)\delta)\tau} [A+(1-\phi_T)-A']^{1-\alpha} \mu_\theta \left[\mathcal{V}\left(\frac{A'}{A+(1-\phi_T)-A'} e^{\delta\tau} R_T\right)\right]^{1-\alpha}.$$

Note that this second equation implies

$$\hat{\mathcal{V}}(A) = [A+(1-\phi_T)]\hat{v},$$

and that the optimal fraction of wealth invested in financial assets $A^* = A'/(A+(1-\phi_T)$ does not depend on the curent level of wealth.

Applying the same arguments as Alvarez et al. (2012), we can see that the optimal policy takes the following form. Within an interval $[\underline{A}, \bar{A}]$, it is optimal to not transact, choose a time until the next observation which depends of the current wealth, $\bar{T}(A)$. If the position in financial assets falls outside of this interval, it is optimal to transact, bring the fraction of risky assets back to a fixed level A^*, and

45

choosing a time until the next observation T^* that does not depend of current wealth.

There are a few notable features of this policy. First, continuous observation, even if there are no transactions likely in the next instant, is not optimal. Even without transaction, fluctuations in wealth still generate variation in utility — $\bar{\mathcal{V}}$ is strictly increasing in A — and create a scope for repeated disappointment, which the agent can avoid using infrequent observation. Second, there are less observations than transaction. For instance, if the agent learns that there was no change in her financial position relative to her durable position, there is no point in paying the transaction cost.

D   Richer information choices - Proof of Proposition 5

We take the set-up of the optimal frequency problem, and set $T = T^*$ the optimal frequency. We want to analyze if the investor would be better off if she received an intermediate signal when her wealth fell below a given threshold. This is, naturally, a sufficient condition for her to be strictly better off when she can also re-adjust the time interval T to a $T \neq T^*$ of her choice. We can represent this problem by comparing two two-period set-ups, where, for simplicity, the rate of time discount and the risk-free rate are both set to 0. The agent has preferences as in Equation 2. To simplify the notations, we write $\mu_\theta(.)$ for $\mu_\theta(. \mid \mathcal{F})$, when the distribution $\mathcal{F}$ is self evident.

- At time 0: the agent's initial wealth is normalized to 1. She invests a fraction of her wealth in a risky account, at price P_0, and the rest in the risk-free asset.

- At time 1: the risky asset trades at price $P_1 = P_0 \exp(g - \frac{1}{2}\sigma^2 + \sigma\epsilon_1)$, where $\epsilon_1 \sim \mathcal{N}(0,1)$. We assume g and σ such that $\mu_\theta(\frac{P_1}{P_0}) > 1$ (the agent is better off investing in the risky asset rather than the risk-less one, even over half the total horizon). The agent consumes some of her wealth. She can opt to receive an alert signal if $\frac{P_1}{P_0}$ falls below a threshold δ.

- At time 2: the risky asset is worth $P_2 = P_1 \exp(g - \frac{1}{2}\sigma^2 + \sigma\epsilon_2)$, where $\epsilon_2 \sim \mathcal{N}(0,1)$ independent of ϵ_1. Naturally, $\mu_\theta(\frac{P_2}{P_1}) = \mu_\theta(\frac{P_1}{P_0}) > 1$, and $\mu_\theta(\frac{P_2}{P_0}) > (\mu_\theta(\frac{P_2}{P_1}))^2 > 1$. The agent consumes all her remaining wealth.

Some useful results:

1. For X and Y independent, $\mu_\theta(XY) \geq \mu_\theta(X)\mu_\theta(Y)$ (Direct from Proposition 1)

2. For all X and Y, $\mu_\theta(X + Y) \geq \mu_\theta(X) + \mu_\theta(Y)$

Proof: Note $-\mu_\theta(X) \geq -E(X) = E(-X) \geq \mu_\theta(-X)$. From there, if $Y = \alpha X$, we get $\mu_\theta(X + Y) \geq \mu_\theta(X) + \mu_\theta(Y)$, $\forall \alpha$. We now just need to prove the inequality is also true for X and Y independent. Since $\mu_\theta(X) + \mu_\theta(Y) = \mu_\theta(\mu_\theta(X) + Y)$, the inequality is a direct consequence of Proposition 1, when X and Y independent.

"No signal" set-up The agent does not observe her wealth until $t = 2$. At $t = 0$, she allocates her wealth, of value V_0, between her intermediate consumption C_1, and her savings S_0. She optimally

invests all S_0 in the risky asset, i.e. whe does not set some wealth aside in the risk-free rate between time 0 and time 1 to invest in the risky asset at time 1. Indeed, from "useful result" number 2, $\forall x \in [0,1]$, $\mu_\theta(x\frac{P_2}{P_0} + (1-x)\frac{P_2}{P_1}) + (1-x)\mu_\theta(\frac{P_2}{P_1}(\frac{P_1}{P_0} - 1)) \leq \mu_\theta(\frac{P_2}{P_0})$. Her optimization problem becomes:

$$V_0 = \max_{0 \leq S_0 \leq 1} \left[(1-S_0)^{1-\alpha} + S_0^{1-\alpha}\mu_\theta(\frac{P_2}{P_0})^{1-\alpha} \right]^{\frac{1}{1-\alpha}},$$

whith solutions $S_0 = \frac{\mu_\theta(\frac{P_2}{P_0})^{\frac{1-\alpha}{\alpha}}}{1+\mu_\theta(\frac{P_2}{P_0})^{\frac{1-\alpha}{\alpha}}}$, $C_1 = 1 - S_0 = \frac{1}{1+\mu_\theta(\frac{P_2}{P_0})^{\frac{1-\alpha}{\alpha}}}$, and $V_0^{1-\alpha} = (1+\mu_\theta(\frac{P_2}{P_0})^{\frac{1-\alpha}{\alpha}})^\alpha$.

"Intermediate information" set-up:

- $t = 0$: the agent allocates her wealth, of value $\tilde{V}_0$, between her risky savings $\tilde{S}_0$ and the risk-free rate to optimize: $\tilde{V}_0 = \max_{0 \leq \tilde{S}_0 \leq 1} \mu_\theta(\tilde{V}_1)$.

- $t = 1$, **if $\frac{P_1}{P_0} \leq \delta$:** the agent receives a "bad" signal, she observes her wealth $\tilde{W}_1 = \tilde{S}_0\frac{P_1}{P_0} + (1 - \tilde{S}_0)$, chooses $\tilde{C}_1$, and allocates the rest to the risky asset because $\mu_\theta(\frac{P_2}{P_1}) > 1$. Her optimization problem is:

$$\tilde{V}_1 \mid_{\text{signal}} = \max_{0 \leq \tilde{C}_1 \leq \tilde{S}_0\frac{P_1}{P_0} + (1-\tilde{S}_0)} \left[\tilde{C}_1^{1-\alpha} + (\tilde{S}_0\frac{P_1}{P_0} + (1-\tilde{S}_0) - \tilde{C}_1)^{1-\alpha}\mu_\theta(\frac{P_2}{P_1})^{1-\alpha} \right]^{\frac{1}{1-\alpha}},$$

whith solutions $\tilde{C}_1 \mid_{\text{signal}} = \frac{\tilde{S}_0\frac{P_1}{P_0} + (1-\tilde{S}_0)}{1+\mu_\theta(\frac{P_2}{P_1})^{\frac{1-\alpha}{\alpha}}}$ and $(\tilde{V}_1 \mid_{\text{signal}})^{1-\alpha} = (\tilde{S}_0\frac{P_1}{P_0} + (1-\tilde{S}_0))^{1-\alpha}(1+\mu_\theta(\frac{P_2}{P_1})^{\frac{1-\alpha}{\alpha}})^\alpha$.

- $t = 1$, **if $\frac{P_1}{P_0} > \delta$:** the agent receives no signal, and knows her wealth $\tilde{W}_1 > \tilde{S}_0\delta + (1 - \tilde{S}_0)$. Her optimization problem is:

$$\tilde{V}_1 \mid_{\text{no signal}} = \max_{0 \leq \tilde{C}_1 \leq \tilde{S}_0\delta + (1-\tilde{S}_0)} \left[\tilde{C}_1^{1-\alpha} + \mu_\theta[(\tilde{S}_0\frac{P_1}{P_0} + (1-\tilde{S}_0) - \tilde{C}_1)(\frac{P_2}{P_1})]^{1-\alpha} \right]^{\frac{1}{1-\alpha}},$$

where all the wealth set aside in the risk-free rate between $t = 0$ and $t = 1$ is consumed at $t = 1$ (see proof above).

Proof of Proposition 5

Because $\tilde{V}_1 \mid_{\text{no signal}} \geq \tilde{V}_1 \mid_{\text{signal}} \geq 0$, $\tilde{V}_0 \geq \frac{P(\text{no signal})}{1+\theta P(\text{signal})} \max_{0 \leq \tilde{S}_0 \leq 1} \tilde{V}_1 \mid_{\text{no signal}}$, where $P(\text{signal}) = 1 - P(\text{no signal})$ is the probability of the intermediate signal.

We assume $\delta << 1$, so $\frac{S_0}{1-\delta} \leq 1$, where $S_0 < 1$ is the optimal risky investment in the framework without intermediate information. Then, $\max_{0 \leq \tilde{S}_0 \leq 1} \tilde{V}_1 \mid_{\text{no signal}} \geq \tilde{V}_1 \mid_{\text{no signal}} (\tilde{S}_0 = \frac{S_0}{1-\delta})$. $\tilde{S}_0 = \frac{S_0}{1-\delta} \implies \tilde{S}_0\delta + (1-\tilde{S}_0) = 1 - S_0$, so $C_1 = 1 - S_0$, the optimal intermediate consumption in the framework without intermediate information, is in the feasible set of $\tilde{V}_1 \mid_{\text{no signal}} (\tilde{S}_0 = \frac{S_0}{1-\delta})$, and:

$$\tilde{V}_1 \mid_{\text{no signal}} \geq \left[C_1^{1-\alpha} + \mu_\theta[(\tilde{S}_0\frac{P_1}{P_0} - \delta\tilde{S}_0)(\frac{P_2}{P_1})]^{1-\alpha} \right]^{\frac{1}{1-\alpha}}.$$

Using $\tilde{S}_0 \frac{P_2}{P_0} = S_0 \frac{P_2}{P_0} - \delta \tilde{S}_0 \frac{P_2}{P_0}$ when $\tilde{S}_0 = \frac{S_0}{1-\delta}$,

$$\tilde{V}_1 \mid_{\text{no signal}} \geq \left[C_1^{1-\alpha} + \mu_\theta [S_0 \frac{P_2}{P_0} + \delta \tilde{S}_0 \frac{P_2}{P_1} (\frac{P_1}{P_0} - 1)]^{1-\alpha} \right]^{\frac{1}{1-\alpha}} .$$

And from "useful results" number 1 and 2,

$$\tilde{V}_1 \mid_{\text{no signal}} \geq \left[V_0^{1-\alpha} + (\delta \tilde{S}_0)^{1-\alpha} [\mu_\theta (\frac{P_1}{P_0})(\mu_\theta (\frac{P_1}{P_0}) - 1)]^{1-\alpha} \right]^{\frac{1}{1-\alpha}} .$$

So

$$\tilde{V}_0 \geq \frac{P(\text{no signal})}{1 + \theta P(\text{signal})} [V_0^{1-\alpha} + (\delta \tilde{S}_0)^{1-\alpha} [\mu_\theta (\frac{P_1}{P_0})(\mu_\theta (\frac{P_1}{P_0}) - 1)]^{1-\alpha}])^{\frac{1}{1-\alpha}} . \tag{20}$$

$\frac{P(\text{no signal})}{1+\theta P(\text{signal})} < 1$ represents the utility cost of information, from the risk of receiving a signal at $t = 1$. It tends to push $\tilde{V}_0$ below V_0, the first term in Equation (20), and depends on $P(\text{signal}) \approx \frac{1}{-\log(\delta)} \exp(-(\log(\delta))^2$ when $\delta << 1$. The term $\delta \tilde{S}_0 [\mu_\theta (\frac{P_1}{P_0})(\mu_\theta (\frac{P_1}{P_0}) - 1)]$ in Equation (20) represents the extra savings, $\delta \tilde{S}_0 = \tilde{S}_0 - S_0$, the agent can make when she knows she will be alerted if the market goes down – the information gains of such signals. When $\delta << 1$, $\frac{1}{-\log(\delta)} \exp(-(\log(\delta))^2 << \delta \implies \tilde{V}_0 > V_0$.

This argument generalizes to the continuous-time case. We can obtain a similar bound on the cumulative probability of the Brownian motion hitting the threshold over a finite time interval, making the information cost negligible compared to the first-order gains from increases in the risky savings.

Q.E.D

E   Extensions

E.1   Diversification and the multiplication of information flows

A robust insight of portfolio theory is that diversification is valuable. When presented with two assets with imperfectly correlated returns, it is optimal to invest in both. Because our disappointment averse agents are risk-averse, the rationale for diversification obtains. However, in our framework, not only does the distribution of the final payoffs matter, but also the structure of the information flow. It is plausible investing in a larger number of assets corresponds to more frequent arrivals of information, which might diminish and even overcome the benefit of diversification.

To characterize the tradeoff between the costs and benefits of diversification in our model, we study a simple example. Suppose the agent receives at date τ the final value $\lambda X_\tau^{(1)} + (1 - \lambda) X_\tau^{(2)}$, where $X^{(1)}$ and $X^{(2)}$ are two arithmetic Brownian motions with volatility σ and correlation ρ, and $\lambda \in [0, 1]$ can be thought of as a portfolio share. Let the agent's observations, at intervals of length T, alternate systematically between the two processes.
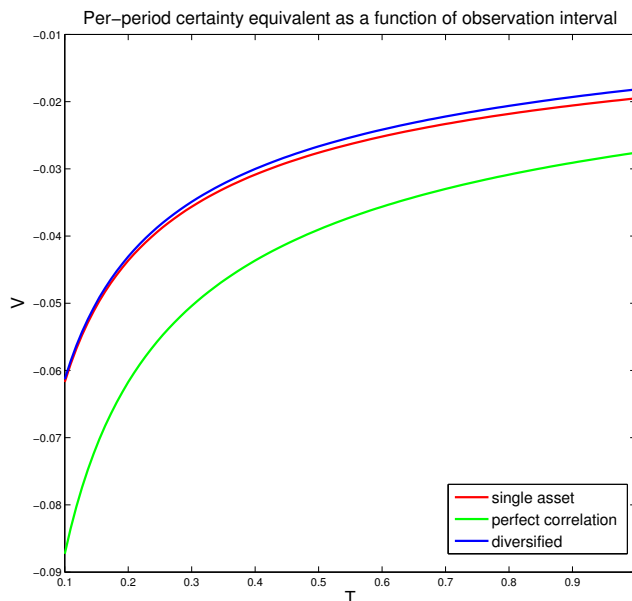
Figure 3: **Role of diversification**. Certainty equivalent as a function of the observation interval for (a) a single asset portfolio (red), (b) an equal-weight perfectly correlated portfolio (green), (c) and an equal weight independent payoff portfolio (blue). For arithmetic Brownian motions with volatility $\sigma = 0.1$, and for $\theta = 1$.

In Figure 3, we consider (a) investing in only one asset ($\lambda = 0$), (b) investing in two perfectly correlated assets ($\lambda = 50\%$, $\rho = 1$), and (c) investing equally in two fully independent assets ($\lambda = 50\%$, $\rho = 0$). Investments (a) and (b) share the exact same payoff structure, but have different information flows, with higher signal frequency for case (b), and thus a lower valuation. Investments (b) and (c) have same information flows, but different risk exposures: the diversified investment (c) yields a higher value. The most interesting comparison is between (a) and (c). In this particular example, when assets are perfectly uncorrelated, the diversification motive dominates over the information costs: the agent always prefers the diversified portfolio with frequent information (c) over the portfolio with one asset and infrequent information (a).[39] The benefits of diversification, however, are greatly diminished by the multiplication of information flows. At higher frequency, in particular, as information costs increase, the agent is close to indifferent between holding a diversified portfolio with more information versus holding a single asset.

---

[39]This result remains true for arbitrary values of λ.

E.2 Leverage and information decisions

In Section 2, the investor cannot lever up, the natural borrowing limit under time-dependent rules, since an agent with levered positions cannot avoid bankruptcy, unless she observed the Brownian risk continuously: levering up requires being and keeping informed at all times. Since continuous observations are infinitely costly, as long as there is a finite (arbitrarily high) borrowing limit, an agent with access only to time dependent rules would not take any leverage.

In the richer information filtrations set of Section 4 however, downturn alerts allow the agent to lever up into the attractive risky asset, without incurring the dissuasive cost of continuous information flows. Even then, though, the agent would not choose to lever up above a given level: the more levered her portfolio, the higher the threshold she needs to set for her downturn alerts, and the more frequent, and thus costly, they become.

Information aversion yields non-trivial interconnections between leverage decisions and attention decisions. The information costs due to the monitoring of levered positions are a likely complement to the margin costs emphasized by Mitchell et al. (2002).

E.3 Learning, information and risk-taking dynamics

For a given signal about a lottery, the information cost is hump-shaped in the quantity of information (see Appendix A). This non-monotonicity has potentially rich implications for learning decisions.

As an illustration, take the framework of Section 2, and assume the agent does not know the growth rate g. She can learn about g by observing her wealth (provided she invests a given portion of her wealth in the risky asset), and from independent signals she automatically receives when she observes her wealth. If these signals are very informative, they are comparatively not very costly, and she would optimally observe them very fast in order to fully learn the value of g; and then adjust her information frequency down to the optimal one of Section 2. If they are not very informative, they are still not extremely costly, but observing her wealth is, so she might maintain a very low frequency dynamics, where only a small part of her wealth is invested in the risky asset, up to the point she learns $g > r$, where she switches her whole portfolio to the risky asset, and increases her observation frequency to the optimal one at steady-state. Finally, if the signals are in a middle range of informativeness, they might be too costly to be worth observing, and in this intermediate region, the investor might fully abstain from investing in the risky asset.

This simple example illustrates the complex interconnections between learning, information and risk-taking, and their dynamics, when agents are information averse. The existence of no-learning regions would not only affect investment decisions, but could have broad implications in other areas of decision making where learning is key, e.g. the decision to engage in medical testing, to collect information and invest in a new technology etc.

E.4 The supply of information

In practice, information is produced and disseminated by agents or institutions. In an economy populated by information averse agents, suppliers of information need to adapt and appropriately tailor the flow of information. As such our model provides the basis for a theory of optimal opaqueness. While providing an in-depth treatment of this question is left for future research, we briefly outline a few simple implications.

One way to help information averse agents is to lump news together in bundles delivered at precise points in time. Such a behavior is consistent with firm's disclosure policies organized around scheduled earnings announcement.[40] Similarly, monetary policy is disclosed at precise points in time, at 2:00p.m. following FOMC meetings, most of which are scheduled in advance. Other macroeconomic announcements, such as employment numbers, quarterly growth etc., also follow discretely spaced releases. A more detailed examination of our framework could provide further guidance for the design of such information release policies.

Agents do not want to receive information too often. However, when they do observe information (either of their own choice or due to external forces), Appendix A indicates they want it to be as precise and "transparent" as possible. Ours is a framework in which suppliers sometimes refraining from releasing any information can be beneficial; but suppliers releasing partial or distorted information is not. Further, even if this form of opaqueness is intrinsically desired by economic agents, it generates asymmetric information, and agency problems are likely to arise in those situations, e.g. between an investor and her wealth manager. These potentially counteract the motive for information withholding. To account for information aversion, optimal compensation contracts need to provide the necessary incentives, while minimizing the information needed to enforce them.

---

[40]Acharya and Lambrecht (2014) provides an alternative theory of earnings target set to manage investors' expectations.