Emotional Inattention

Lukas Bolte* Collin Raymond*

November 15, 2021

Most Recent Version:

https://lukasbolte.github.io/papers/emotionalInattention.pdf

Abstract

We introduce a model of emotional inattention. An agent decides how much attention to allocate to a decision. Attention has two consequences: 1) it increases material payoffs, and 2) leads to an emotional response, increasing the weight of that decision in her total payoff. The cost of attention is thus endogenous and depends on the relative payoff of the decision. Emotionally inattentive agents exhibit the ostrich effect and avoid thinking about low-payoff decisions. They also exhibit excessively volatile levels of attention and can be caught in attention traps, where cognitive scarcity along with dynamic inconsistency leads individuals to be inattentive to welfare-improving actions. Standard interventions to improve decision-making, such as the provision of free information or penalties for mistakes, can in fact worsen decision-making. In a consumption-savings application, the agent is shown to react asymmetrically to income shocks immediately adjusting to good and sometimes ignoring bad ones. In a portfolio choice problem, the agent avoids assets that require re-adjustments to information (an attention premium), which can lead to an excessive premium for risky assets.

JEL Classification Codes: D81, D83

Keywords: Attention, emotional response, information

*Lukas Bolte; Department of Economics, Stanford University, Stanford, CA; lbolte@stanford.edu. Collin Raymond; Department of Economics, Purdue University, West Lafayette, IN; collinbraymond@gmail.com. We thank Douglas Bernheim, Gabriel Carroll and Muriel Niederle, and seminar participants at Stanford University for helpful comments. All errors are our own.

[...] we have at least two distinct roles for our minds at play, that of the information processing and reasoning machine by which we choose what to consume out of the array of things that our resources can be exchanged for, and that of the pleasure machine or consuming organ, the generator of direct consumer satisfaction.

Schelling (1988)

1 Introduction

Researchers within economics have increasingly documented the fact that many individuals avoid information about negative outcomes even in situations where it would help them make better decisions, including not checking their investment accounts when markets are down (Karlsson et al., 2009; Sicherman et al., 2015) or refusing to take medical tests (Oster et al., 2013; Ganguly and Tasoff, 2017). Conversely, individuals often pay excessive attention to good situations, e.g., checking their investments repeatedly on weekends when no change can possibly have occurred when markets are up (Karlsson et al., 2009; Sicherman et al., 2015), or daydreaming about high-payoff future events (e.g., a vacation, winning the lottery, or getting a top publication) at the expense of relatively unpleasant, but important, tasks at hand. Evidence also suggests that individuals direct attention in order to minimize the impact of negative information, e.g., Falk and Zimmermann (2016) find information about an electrical shock is viewed differently depending on whether individuals have a distracting task available.

Two broad frameworks have recently been used to try and explain seemingly anomalous attitudes towards learning and information. The first model's individuals as directly caring about their beliefs regarding future outcomes (independent of the actual realizations). These models, including Kreps and Porteus (1978); Caplin and Leahy (2001); Kőszegi and Rabin (2009) and Dillenberger and Raymond (2020), implicitly or explicitly capture the emotional reaction that beliefs induce, such as anxiety about potential bad outcomes or the savoring of good ones. The second framework models individuals as having to pay costs to acquire information, which may be monetary and explicit, or psychological and implicit (i.e., the literature on rational inattention; see Sims (2003) for a seminal example). Here information avoidance captures a rational desire to avoid paying acquisition costs, even if it induces worse decision-making.

However, both approaches fail to fully account for much of the evidence. Rational avoidance of information is hard to square with the fact that we see individuals avoiding information that is extremely instrumentally valuable, such as important medical information.[1] Anticipatory models

[1] For example, individuals are typically aware of their risk of Huntington's disease, a hereditary and deadly disorder, which is around 50% when a parent carries the gene, and a test costs no more than \$300. As Oster et al. (2013)

(as do models of costly information acquisition) fail to explain why checking an investment account repeatedly can be beneficial, since observing the same information does not cause beliefs to change, or why the presence of a distracting task should change informational preferences towards negative outcomes.

Our paper's key contribution is to develop a model which links information acquisition and anticipatory utility using the key intuition that they both are driven by attention.[2] Attention thus serves a dual role in our model: Directing attention to a decision allows the decision-maker to better make that decision (through acquiring information but also reducing noisy actions, or moving away from a default), but it also generates emotional focus on that decision, which generates current utility from thinking about future outcomes (which we refer to as anticipatory utility). Thus, attending to a decision changes utility in two ways; one by improving the future material payoffs from that decision, and second by generating contemporaneous anticipatory utility about those future payoffs. Our approach can be seen as modeling costly attention, but where the costs, or benefits, are endogenous: they are the anticipatory emotions attached to the decision being paid attention to. For example, our model suggests that individuals avoid medical tests because, although they value the potential gain from learning their medical status, thinking about the potential negative outcomes imposes too much of an emotional cost. Conversely, investors may want to recheck their high-performing portfolios not to actually learn anything new, but rather simply to direct their emotional focus towards pleasant outcomes. A few other papers have considered this linkage as well, e.g., Karlsson et al. (2009), Golman and Loewenstein (2018), and Tasoff and Madarasz (2009). Although similar in spirit, and also designed to match ostrich-effect behavior (described in more detail below), they are very different in assumptions and focus on very different behaviors. We carefully compare our approach and results to the existing literature in Section 5.

Our model, which we call a model of emotional inattention, explicitly captures the dual role of attention and delivers a rich set of predictions. Some match well-studied behaviors but provide new perspectives, and others represent novel empirical tests of our theory for situations in which our predictions differ from those of existing explanations. In Section 2, we begin by sketching out the basic structure of our model in a one-shot setting. Our setting supposes that the individual faces some decision-problem, as well as an outside "trivial" task, which generates the same material payoff

points out, the value of knowing whether a person will live until retirement seems quite large in terms of making life choices about children, savings, etc., and so hard to justify with any model of rational inattention.

[2]The linkage between attention as a locus of control and emotional reactions has been documented in the psychology and neuroscience literature, e.g., Mrkva et al. (2020).

no matter what the decision-maker does. The decision-maker decides how much attention to devote to the decision versus the trivial task. We suppose that the payoff from the decision is increasing in the amount of attention paid, but make no further substantive assumptions. Thus, we take no substantive stand on why attention increases payoffs. We believe it could happen for a variety of reasons and may vary by environment. It could be that attention allows an individual to think about the decision and acquire information about the optimal action (e.g., via mental simulations of potential outcomes). Or it could be that the decision-maker chooses actions with noise (as in random choice models), and increased attention reduces this noise. Or it could be that in order to take any-non default action a decision-maker must pay at least a minimal amount of attention. The second key ingredient is that attention leads to an emotional response: we operationalize this by assuming that attention to either the trivial task or the decision task generates current flow utility of anticipated payoffs that is proportional to future expected payoffs from that task.

In Section 2.2, we show that our model (not surprisingly) generates the well-known asymmetric ostrich effect: Individuals avoid thinking about decisions that have low payoffs (they "bury their heads in the sand") but are happy to engage with high-payoff decisions. Our model has important implications for policies that are designed to increase information uptake (Section 2.3). We focus on two policies. First, we consider what happens when the decision-maker is provided with free information — i.e., information that is costlessly (in terms of attention) provided to the decision-maker before they decide on how much information to acquire endogenously. Second, increasing the instrumental return of information by increasing the difference in payoffs between good and bad actions. Both should improve material outcomes when inattention is not emotionally driven. Due to the fact that with emotional inattention the cost of attention is endogenously-determined anticipatory utility, in our model, this is not necessarily true. Emotional inattention makes nuanced predictions about when informational nudges and heightened incentives can backfire, consistent with the fact that these interventions have sometimes, but not always, been found to fail (e.g., the literature on choking Ariely et al., 2009).

In Section 2.4, we show how emotional attention can capture notions of cognitive scarcity (similar to Mullainathan and Shafir, 2013). Because attentional costs are endogenous, the model predicts escape thresholds, where individuals with low enough attentional capacity actually refuse to engage with unpleasant decisions, while those with high enough capacity devote full attention to them. Thus, small changes in the amount of attention a decision-maker has available can lead to drastic shifts in decision-making quality. The driving feature here is the complementarity between the two roles attention plays in the utility function. Higher attention can lead to higher material payoffs, which leads to a lower cost of attention. In particular, even if the material value function is concave in attention, the overall utility may be convex in attention. Consistent with this Olafsson and Pagel (2017) document a discrete jump in investors' attention to their bank accounts depending on whether the balance is negative or positive, and the Vanguard data of Karlsson et al. (2009) is also suggestive of bi-modal levels of attention (see their Figure 4).

In Section 3, we extend the model to multiple periods to study important dynamic implications of emotional inattention. Section 3.1 extends our analysis of attentional scarcity and shows that in conjunction with the dynamic inconsistency inherent in emotional attention it can lead to *attentional traps*. These capture situations where tasks that require sustained attention over many periods may never even be attempted. This is because an agents' future self devotes too little attention to potentially unpleasant tasks relative to the perspective of the current self (because the future self does not internalize the beneficial effect of future attention on the current self's anticipatory utility). In the presence of commitment, or the ability to devote sufficient attention to the decision in any given period, the agent would make a relatively good decision. But, when forced to spread attention out over both periods without commitment, the agent (rationally) anticipates that her future self will acquire little information (because the problem is unpleasant to think about), leading her current self to acquire little information (because without the future self willing to think, the current self faces low anticipatory utility). Such effects can leave selves at all time periods worse off. Thus limited cognitive resources can be especially pernicious because they force an individual to coordinate across different temporal selves, who have different preferences. This is consistent with underinvestment in long-term, potentially unpleasant, but high-return activities like schooling or retraining.

Section 3.2 discusses how emotional inattention generates intrinsic preferences for the timing of learning. In particular, emotionally inattentive agents have an intrinsic preference for paying attention (and so learning) earlier rather than later, even when there are no observable actions. This is because the agent anticipates in the future exhibiting an asymmetric ostrich effect — she will devote different amounts of attention depending on whether she learns the decision will generate a high or low payoff. This generates a hidden action on the part of the agent, and so a preference for early learning. Such a preference for early learning is consistent with experimental evidence on intrinsic preferences for information (e.g., Masatlioglu et al., 2017 and Nielsen, 2020).

We study the implications of our model in particular applications in Section 4. We first apply the model to study the consumption patterns in a consumption-savings model. The agent has to decide what to consume in the first of two periods before she knows her income. Once she receives her income, she can either consume at the initially planned level or devote attention to the problem and re-optimize her consumption. In the second period, she consumes whatever income remains. The agent's response to income is asymmetric: she always reacts to high-income realizations and increases her consumption while she may not to low-income realizations. Further, she anticipates such asymmetry and chooses her initially planned consumption level conservatively, i.e., suited for exactly these low-income realizations for which she does not devote attention.

We next study the portfolio choice of an emotionally inattentive agent. We find three distinct consequences of emotional inattention. First, not surprisingly, our model generates an ostrich effect, with emotionally inattentive agents selectively re-optimizing in later periods. Second, because risky assets may require more attention than safe assets due to a need to re-optimize after learning about potential returns, agents, rationally anticipating their later ostrich behavior, tend to exhibit an excessively high risk premia (as measured by initial investment decisions) relative to standard agents. However, this attentional premia is conceptually distinct from risk preferences, and may not even be attached to the risky asset if, for example, re-optimization is automatic. Third, even though the agent appears too risk-averse when making their initial portfolio choice, they actually exhibit a stronger preference for some increases in risk (relative to a standard agent). The reason is that a varied payoff allows her to focus on good realizations and ignore bad ones, which makes her payoff more convex than that of a standard agent — the same mechanism that leads the agent to prefer early resolution of uncertainty.

Section 5 discusses how emotional inattention differs from several other classes of models to which it is closely related. We consider in turn models of rational inattention, anticipatory emotions, and reference-dependence preferences. We then discuss how our approach compares to a small set of other papers which try to compare attention as well as utility and attention. Section 6 concludes.

2 Static Model

2.1 Setup

As discussed, our starting point is two recognize two roles for attention. First, attention allows for better decisions. There are many reasons why devoting attention to a task may improve payoffs in that task. For example, paying attention to a task allows one to collect information about the relevant states of the world, or payoffs of actions. Although we often focus on this information interpretation, there are also other reasons. For example, attention could induce individuals to "control" their actions more, and reduce random errors. More generally, to even make a decision, or take an action, one must focus attention at some point.

Our jumping-off point is to note that attention also leads to emotional focus and emotional responses (in this we mirror the point made by Sicherman et al., 2015 who note that attention in financial markets plays both a role as an input for decision-making and a way in which investors experience utility). This idea is well documented in neuroscience studies.[3] More abstract, non-immediate payoffs, those outside the "here and now," by definition, require attention to enter consideration in a decision problem. Thus, we assume the act of paying attention to a decision generates emotional/anticipatory utility for the agent that depends on the payoff from that decision.

There is evidence of people allocating attention taking into account its emotional focus: from laboratory settings (Falk and Zimmermann, 2016; Chen et al., 2018), medical testing (Oster et al., 2013), and financial investing (Karlsson et al., 2009; Olafsson and Pagel, 2017). The idea of attention willfully guided to satisfy some objective is referred to as "top-down" attention and central to our model: the agent freely chooses how to allocate her attention among the tasks driven by her goal to maximize total payoff.[4] Of course, there could also be some external stimuli capturing the agent's attention and posing additional constraints on the agent's problem (which we explore in e.g., in Sections 2.4 and 3.1).

Of course, the key feature in our model is that attention takes on both a role in assisting decisions as well as in generating anticipatory utility via emotional focus. In principle, this need not be the case; the mental process of processing information could be independent of the one assessing the material payoff of the decision. The tight linkage present in our model (and which the agent cannot control) is referred to as "bottom-up" attention.[5] There may be other linkages forcing certain consequences of attention.

Formally, we suppose an agent (she) faces two tasks. We refer to the first as *the decision* and the second as *the outside option*. The agent allocates a unit of attention among them with α denoting attention devoted to the decision. Attention has two effects.

First, a standard feature of attention in economic models, attention improves decision-making. Each task has a *material value* associated with it. Devoting attention to the decision increases its

[3]See Chen et al. (2018); Brosch et al. (2013); Yamaguchi and Onoda (2012); Wiens and Syrjänen (2013).

[4]See also Benedek et al. (2016) on a neurological account of how attention can be directed.

[5]See Katsuki and Constantinidis (2014) for a neurological account of "top-down" and "bottom-up" attention.

material value: attention may fuel information acquisition or processing, or it may be required to make a choice other than the default option. We denote the material value of the decision when the agent devotes attention α by $V^D(\alpha)$ and assume V^D is increasing, bounded and right-continuous. We provide micro foundations for V^D when attention increases information or when it increases available choices in Sections B.1 and B.2 in the Appendix, respectively. The material value of the outside option is constant and denoted by V^O.

Second, attention to a task leads to a *emotional response*, an emotional focus on that task. Specifically, it affects the weights given to the different tasks in the total payoff of the agent.

We assume that an individual receives flow utility not only from the realization of both tasks but also from their expectation of what they will receive — i.e., anticipatory utility. The anticipatory utility that an individual receives from a task is proportional to the amount of attention they pay to it.[6] The agent places a weight τ on anticipatory utility (with the weight on actual consumption utility being $1-\tau$). Thus, the total payoff of the agent is given by

$$\begin{aligned} V(\alpha) &= \overbrace{(1-\tau)[V^D(\alpha)+V^O]}^{\text{Consumption Utility}} + \underbrace{\alpha\tau V^D(\alpha)}_{\text{Decision Anticipatory Utility}} + \overbrace{(1-\alpha)\tau V^O}^{\text{Outside Option Anticipatory Utility}} \\ &= V^D(\alpha)[(1-\tau)+\alpha\tau] + V^O[(1-\tau)+(1-\alpha)\tau] \end{aligned} \tag{1}$$

When $\tau = 0$, the model collapses to that with a "standard" agent who optimally devotes attention to maximize the change in material values, here, to the decision; and when $\tau = 1$, the agent is able to completely ignore a task by not devoting any attention to it.

We frequently return to a particular parameterization which can be interpreted as resulting from an underlying decision problem: the agent has a normally distributed prior about a state of the world that she tries to match. The agent receives a signal about the state the precision of which is increasing in attention. The material value also includes a constant term, H^D (the "height"), and is thus given by

$$V^D(\alpha) = -\gamma\frac{1}{\beta+\psi\alpha} + H^D. \tag{2}$$

A full derivation of the expression above is provided in Section B.1 in the appendix.

[6] This is largely a normalization. If the anticipatory utility is proportional to $f(\alpha), 1-f(\alpha)$ for the decision and the outside option, respectively, and f is strictly increasing and continuous, then letting the agent choose $f(\alpha)$ directly reproduces the objective in (1).

We begin by noting that an optimal level of attention devoted to the decision exists.

Lemma 1. *There exists α that maximizes V.*

All proofs are omitted from the main text and instead given in Appendix A.

The optimal level of attention need not be unique. When we formulate results on the effects of parameter changes on attention, phrases such as "attention is increasing" are to mean comparative statics in the strong set order unless noted otherwise.

2.2 Ostrich Effect

The ostrich effect refers to individuals avoiding thinking about unpleasant news or situations by burying their figurative heads in the sand.[7] Evidence consistent with ostrich effects comes from several different domains.

One key domain is in the realm of medical testing. It is widely documented that individuals often avoid medical tests, and the assumption is that this is because individuals want to avoid thinking about the potential downsides associated with positive diagnoses. For example, individuals at risk of Huntington's disease, often avoid cheap and accurate predictive genetic testing (Oster et al., 2013; Shouldson and Young, 2011). This seems at odds with rational information acquisition, as the instrumental value of predictive testing would allow individuals to make better decisions with regard to investments such as in education, retirement choice, and family planning. Ganguly and Tasoff (2017) document similar avoidance of STD tests. Avoidance of unpleasant thoughts also are seen as motivating factors when individuals (often) fail to follow medical recommendations, e.g., taking medicine or self-screening for symptoms (Becker and Mainman, 1975; Sherbourne et al., 1992; Lindberg and Wellisch, 2001; DiMattero et al., 2007).

A second domain that has provided robust robust evidence of the ostrich effect is in individuals' attention to personal finances (Olafsson and Pagel, 2017; Karlsson et al., 2009; Sicherman et al., 2015). Karlsson et al. (2009) and Sicherman et al. (2015) study individuals' logins to their investments account and find that logins increase when the market is up. Olafsson and Pagel (2017) look at individuals' attention to their financial accounts and find increases after individuals are paid, and goes down when individuals are overdrawn, or have low account holding.

[7] Although the idea of ostriches burying their heads in the sand to avoid predators dates back to Roman times, they actually do not display this behavior. Instead, they put their heads into their nests (which are built on the ground) in order to check temperatures and rotate eggs.

Our model sheds light on such and other behaviors.[8] In it, agents attend to tasks whose material value is high relative to the outside option. In other words, they do not devote attention, or ignore, tasks with very low material value — the ostrich effect. The proposition below documents this effect in our model. Furthermore, agents with a strong emotional focus devote more attention to tasks with high material value (irrespective of the curvature).

Fix some increasing function $\tilde{V}^D$ and let $V^D(\alpha) = H^D + \tilde{V}^D(\alpha)$, where H^D is the *height* of the decision.

Proposition 1 (Ostrich effect)**.** *Attention to the decision*

1. *increases in the height of the decision; and*

2. *decreases in the material value of the outside option.*

In particular, for every $\tilde{V}^D$ *that is differentiable at* 0*, there exists* H^D *such that the agent devotes no attention to the decision.*

Furthermore, attention to the decision decreases in the importance of emotional focus.

Individuals at risk of a disease may thus avoid getting tested as doing so draws attention to their potential health problems leading to a negative emotional response. Similarly, low adherence to medical recommendations is explained by individuals avoiding constant reminders of health issues or threats. Attention patterns to investment portfolios or financial accounts are also consistent with the model: Attention to the market if it is up, or an account with a high balance, leads to a positive emotional response, whereas attention to the market when it is down, or to a negative balance, is avoided.

Lindberg and Wellisch (2001) finds a positive correlation between feelings of anxiety and ostrich-type behavior. This could be interpreted as evidence for higher weight on anticipatory utility, τ larger, leading to avoidance. While we do not discuss where an agent's τ comes from, Sicherman et al. (2015) find that investors show stable behavioral traits pertaining to asymmetric attention depending on market conditions, thus suggesting that τ may be a stable, individual-specific parameter.

[8] Our model also sheds light on the low-take up of social benefits (Currie, 2006). Eligible individuals may not want to complete the perhaps complex bureaucratic procedures as doing so draws attention to their dire financial situation.

2.3 Improving Information

As the previous proposition makes clear, emotionally inattentive individuals may often ignore information that could be materially beneficial. This raises the obvious question of how policymakers who want to improve the material payoffs of individuals may do so.[9] Our model of emotional inattention provides some insights into when and why different interventions work. As we show, some natural interventions may, perhaps unexpectedly, not work, i.e., the treated population may end up with less total information.

To understand the effects of different policies on an emotionally inattentive agent, it is important to study their interactions with the ostrich effect and the complementarity between the two roles of attention (increasing material payoffs and causing an emotional response). To illustrate the importance of the former, while the response of a standard agent to some intervention is a function of its effect on the curvature of the material value function only, an emotionally inattentive agent also reacts to changes in the level of material value. To understand the latter, consider an information intervention that allows the agent to have the same total information while paying less attention. The lower level of attention then reduces the emotional response and, due to the complementarity, may in fact decrease total information.

We consider two types of interventions, which when inattention is not linked to anticipatory utility, should improve the agent's total information and thus her material payoffs.

First, we consider the effect of a policy easing information access (the "supply" side). Second, we consider how changing increasing the stakes of the decision can affect the demand for information (the "demand" side). Note that the comparative statics we conduct need not be about information, but should be understood more broadly as changes in the return (change in material value) of attention. We show that with emotional inattention these policies do not always lead to improvements in material outcomes (because of the complementarity of the consequence of attention and the ostrich effect, respectively). This implies policymakers need to be careful about changing the environment when anticipatory emotions drive inattention.

[9]A welfare-based justification can be made if consumption utility from the tasks is discounted due to impatience and anticipatory utility consumed contemporaneously. In this case, a benevolent policymaker should be more patient (Caplin and Leahy, 2004), i.e., put more weight on the material value.

2.3.1 Providing Information

A policymaker may want to provide information, or otherwise change the environment of the decision, to improve material and total payoff. The success of such policies is mixed; in this section, we study how the endogenous response of our agent interacts with the provision of information.

Consider the following policy of providing information. The agent is provided with some free information, β, that she is forced to attend to. This information is equivalent to what the agent would acquire by devoting β attention to the decision. The information increases the weight on the material value of decision in the agent's anticipatory utility by $\beta' < \beta$, that is less attention is required to attain a certain level of material value.

Formally, attention α given provided information β leads to material value of $V^D(\alpha + \beta)$. Furthermore, the weight of the anticipatory utility of the decision given attention α is given by $\beta' + \alpha$ (and $\alpha \in [0, 1 - \beta']$.

There are two mechanical reasons why such information policy information may lead to higher material value: 1) the agent's maximum material value potentially increases (see Section 2.4 for a more general discussion of extreme attention allocations), and 2) the agent may be forced to attend to more information than she initially does. However, in the absence of these two reasons, total information, $\alpha + \beta$, and thus material value decreases.

Proposition 2. *Denote the optimal level of attention without the information policy by α^* and with the information policy by α^*_β and suppose both are unique.*

If $\beta < \alpha^*$ and $\alpha^*_\beta + \beta \leq 1$, then

$$\alpha^*_\beta + \beta \leq \alpha^*.$$

The key driver behind the aforementioned result is a complementarity between increasing the material value and the anticipatory utility. Here, suppose the agent responds to the free information by reducing attention to exactly offset β. In this case, the cost of devoting attention is the same as the material value of the decision is unchanged. However, this occurs at a lower level of attention (reduced by exactly $\beta - \beta'$), and thus a lower weight on the material value of the decision. Hence, the agent may in fact further decrease attention.

We thus provide a cautionary tale. Information provision is not an innocuous intervention that at worst will not have any effect on material payoffs. Instead, as We show quite generally, informational intervention can decrease material payoffs.

2.3.2 Increasing the value of information

Does increasing the return of the value of information increase attention? A policymaker may do so by increasing the reward of making good decisions or the penalty of making poor ones. On a standard agent, these two interventions would have similar effects: The return of information (attention) increases so that she devotes more attention to the decision. To our agent, however, these two "feel" rather different: Increasing rewards raise the overall material value of the decision, and the emotional response from devoting attention increases; increasing penalties has the opposite effect. From our discussions of the ostrich effect in Section 2.2, we expect the agent's attention response to differ.

What do we mean by increasing the reward and the penalty? We can write the material value of the decision equivalently as

$$V^D(\alpha) = V^D(0) + \underbrace{[V^D(\alpha) - V^D(0)]}_{\text{reward}}, \quad \text{or} \quad V^D(\alpha) = V^D(1) - \underbrace{[V^D(1) - V^D(\alpha)]}_{\text{penalty}}.$$

More generally, fix H^D (the *height*) and $\tilde{V}^D(\alpha)$, and consider $V^D(\alpha) = H^D + \gamma \tilde{V}^D(\alpha)$ for $\gamma > 0$. If $\tilde{V}^D(\alpha) \geq 0$, increasing γ corresponds to increasing the reward, and if $\tilde{V}^D(\alpha) \leq 0$, increasing γ corresponds to increasing the penalty.

Increasing the reward will always lead to an increase in attention. Note that in both cases, the return of the value of information increases with γ so that a standard agent would devote more attention. However, the changes in material value differ: for a fixed level of attention, the material value increases when increasing the reward, and decreases when increasing the penalty making the effects on the emotional response drastically different.

Proposition 3. *If $\tilde{V}^D \geq 0$ (reward), then attention devoted to the decision is increasing in γ.*

If $\tilde{V}^D \leq 0$ (penalty), then attention devoted to the decision is decreasing (increasing) in γ if

$$\tilde{V}^D(\alpha)[(1-\tau) + \tau\alpha] \tag{3}$$

is decreasing (increasing) in α.

The emotional response dominates, i.e., (3) *is decreasing in α if $\tilde{V}^D$ is semi-differentiable and at least one of the two following conditions is satisfied*

1. *$\tilde{V}^D(1)$ is low enough and $\tau > 0$;*

2. τ *is large enough* ($\tau < 1$) *and* $\tilde{V}^D(1) < 0$.

The increased value of information dominates, i.e., (3) *is increasing in* α *if*

1. there exists $C > 0$ *so that for all* α', α *with* $\alpha' > \alpha$,

$$\frac{\tilde{V}^D(\alpha') - \tilde{V}^D(\alpha)}{\alpha' - \alpha} \geq C.$$

and τ *small enough.*

Some evidence, such as Böheim et al. (2019) and Ariely et al. (2009) indicates that individuals "choke", i.e., perform worse, when the stakes go up. This is clearly at odds with standard rational models. However, it can be consistent with our model, where increased stakes increase the downside of getting a decision wrong, and so reduce attention (ostrich effect), and overall performance. Of course, key to our mechanism is that the increased stakes reduce the worst payoffs, not just increase the best payoffs. If only the latter were true, then we would expect increased performance. Thus, our model may also be able to rationalize why we observe choking in some environments but not others.

Our model also suggests that the effects of penalties and rewards depend on whether the individual can reallocate her attention. A laboratory experiment consisting of a decision problem where subjects are treated with either increased penalties or rewards may provide dramatically different results depending on whether attention is held constant or not.

2.4 Attentional Volatility

Attention can change dramatically with small changes in the environment. Olafsson and Pagel (2017) find a discrete jump in the propensity to log in on to one's financial account when the balance turns from negative to positive and the investors' logins in Vanguard data of Karlsson et al. (2009) as grouped by the change in S&P 500 of the previous 4 days seems to resemble a bi-modal distribution (see their Figure 4). While optimizers changing discretely in parameters is common, there is a particular sense in which attention is volatile in our model; as we explain now.

The two roles of attention (assisting better decisions and generating anticipatory utility) are complementary. Higher levels of attention generate higher material value, which increases the value of anticipatory utility. In particular, so long as the payoff function from the decision, as a function of attention, isn't too concave, then the utility function is a convex function of attention (a precise statement follows in the ensuing proposition)[10]

Moreover, the total payoff may not be increasing in attention devoted to the decision, e.g., the agent may maximize her total payoff by not devoting attention to the decision as in the ostrich effect discussed in Section 2.2.

Convexity the total payoff function implies that the agent's optimal strategy is a boundary point — they should either devote no or full attention to the problem. In these circumstances the decision-maker may experience attentional volatility — small changes in the decision problem can lead to large observed changes in behavior, and decision-utility (although the change in utility will still be smooth).

The next proposition provides a particular example of how volatility may occur: when individuals have bounds on their attention. Previously, there was a unit of attention to be allocated across the two tasks. Now, we presume that the agent has some exogenous bounds on their attention: $\alpha \in [\underline{\alpha}, \bar{\alpha}]$. These bounds potentially represent the fact that the agent must pay at least some attention to either the decision or the outside option, perhaps due to salience, reminders outside their control, or outside distractors.[11]

It is relatively clear that increasing $\underline{\alpha}$ must weakly increase the chosen attention level while decreasing $\bar{\alpha}$ will weakly decrease the chosen attention level. However, when the utility function is globally convex and non-monotone in attention, we may observe the agent engaging in attentional volatility. In other words, a small change in the required amount of attention the agent must pay to a particular task leads to dramatic overall allocations in attention and dramatic shifts in observable payoffs.

Proposition 4. *The optimal level of attention to the decision is weakly increasing in $\underline{\alpha}$ and $\bar{\alpha}$. Moreover, V^D is twice-differentiable and not too concave, that is*

$$\frac{2\tau}{(1-\tau)+\alpha\tau} > -\frac{\frac{\partial^2}{\partial\alpha^2}V^D(\alpha)}{\frac{\partial}{\partial\alpha}V^D(\alpha)} \text{ for all } \alpha \in [0,1],$$

then the total payoff is strictly convex and for any $\underline{\alpha}, \bar{\alpha}$, the optimal level of attention is $\alpha^* \subseteq \{\underline{\alpha}, \bar{\alpha}\}$.

In those situations, i.e., when the total payoff is strictly convex, the optimal level of attention

[10]Throughout the rest of this sub-section we will assume V^D is twice differentiable. Similar results hold without the assumption, but conditions are more complicated to state.

[11]It could also represent a reduced form way of capturing third tasks that also require attention. Let α_3 denote the attention devoted to said third task. Then for any fixed level of attention α_3 devoted to the third task, the agent allocates $1-\alpha_3 \equiv \bar{\alpha}$ between the decision and the trivial task. Changes in $\bar{\alpha}$ can then be understood as changing the parameters of the third task so that the agent changes the amount of attention devoted to it.

may jump from $\underline{\alpha}$ to $\bar{\alpha}$ as $\underline{\alpha}$ and $\bar{\alpha}$ increase.

We provide two examples to illustrate these effects. First, consider the left panel in Figure 1, which shows the agent's total payoff (1) as a function α. Even though V^D is concave in α ((2) defines a concave function), overall utility is convex in α due to the complementarities between information acquisition and emotional focus. As a consequence, the agent's solution is at either extreme. Small changes in the upper bound on α can thus cause attention devoted to the decision to drop to 0. In the figure, we can see that for bounds above the blue line, the agent would choose the maximum amount of attention to devote to the decision, while for bounds less than below it, the agent would choose to devote no attention to the decision.

(a) V^D is given by (2) with $\beta = 10, \gamma = 2000, H^D = 87.5, V^O = 10, \tau = 0.8, \psi = 13$.

(b) V^D is given by (2) with $\beta = 10, \gamma = 2000, H^D = 87.5, V^O = 20, \tau = 0.8, \psi = 13$.

Figure 1

Similarly, shifts in the minimum of attention the agent needs to devote to the decision can cause her to switch from only devoting the minimum level of attention to devoting all her attention to it. The right panel of Figure 1 provides an example; we can see that for bounds below the blue line, the agent would choose the minimal amount of attention to devote to the decision, while for bounds above it, the agent would choose to devote full attention to the decision.

Comparative statics studied in previous sections can also generate attentional volatility. For example, recall that we showed increasing the value of the outside task can generate an ostrich effect — attention to the decision is reduced. This reduction, when the convexity condition holds, can then be dramatic — we can get a reduction of attention paid to the decision all the way to 0. Similarly, the changes in attention in response to raising the penalties in an underlying decision problem can also be dramatic: A slight increase in the stakes of the decision can cause dramatic decreases in attention and the material value of the decision.

Attentional volatility can explain why individuals may not devote any attention to a task, even when doing so would improve their material outcomes — as in poverty traps. Although much of the poverty trap has focused on money as being the relevant constraint for poor individuals, many papers have noted that cognitive scarcity, which is often linked to monetary scarcity, can also be very relevant. For example, in a recent review Ghatak (2015) discusses "scarcity driven" poverty traps, where the individual has a scarce resource, potentially attention. Many explanations for poverty traps rely on some kind of non-convexity in the production technology, which can imply increasing returns to scale. Our model can be thought of as generating increasing returns to attention, which are non-monotonic, even when the material payoff does not feature such increasing returns.[12]

Attentional volatility can also help explain why individuals who are otherwise very similar in terms of resources can end up very differently, including whether individuals remain poor. In a recent paper, Balboni et al. (2021) conduct a field experiment to test whether individuals remain poor because of differences in fundamental abilities or opportunities in which case giving them resources should help escape the poverty trap. They find that if the program pushes individuals' resources above some threshold, then they escape poverty, but if it does not, then individuals end up staying poor. Although their focus is on transferring assets, one might imagine that additional assets relax an attentional constraint. In this case, finding threshold effects is consistent with the predictions of our model.

Mullainathan and Shafir (2013) and Dean et al. (2019) discuss how cognitive functions can play an important role in maintaining poverty, and that poverty may itself impede cognitive function. Lybbert and Wydick (2019) suggests that poverty may shift preferences. Our model represents a synthesis of the cognitive and preference-based approaches while suggesting a new pathway: if being poor implies that other tasks take up lots of attention, the optimal policy for an individual might change from devoting lots of attention to an unpleasant, but ultimately life-improving task, to devoting no attention to it.

3 Dynamic Model

We extend our static model to a dynamic one. This allows us to study preferences over the timing of information (early vs late). This extension also increases the set of environments our model can be applied to.

The timing is as follows. There are two periods, $t = 1, 2$. In period 1, the agent decides how much

[12] Other mechanisms can also create increasing returns to attention. In Banerjee and Mullainathan (2008), the individual allocates attention between "solving problems" at home and at work. The objective is convex because of a feedback loop: attention at work leads to income allowing the individual to purchase distraction-saving goods making a lack of attention at home less costly.

attention α_1 to devote to the decision. Some uncertainty is subsequently resolved. This uncertainty may pertain to the realized signals when attention leads to information as described in Section B.1, which choices become available as described in Section B.2 or simply some information about a future payoff. Among other things, it may thus affect the (expected) material value associated with the decision. We denote the realization of such randomness as ϵ_1. In period 2, the agent's problem is reminiscent to that in the static environment; now parameterized by ϵ_1: the agent decides again how much attention, now denoted by α_2, to devote to the decision where α_2 may depend on the resolved uncertainty, ϵ_1.

In period 2, the material value of the decision is denoted by $V_2^D(\alpha_2(\epsilon_1)|\epsilon_1)$, and satisfies the same properties as V^D in the static environment.

In period 1, the material value of the decision is given by

$$V_1^D(\alpha_1, \alpha_2(\cdot)) = E_{\epsilon_1 \sim F(\alpha_1)}[V_2^D(\alpha_2(\epsilon_1)|\epsilon_1)],$$

where F is the distribution of the uncertainty that is resolved at the end of period 1 which may depend on α_1.

The agent weights the anticipatory utility in period t by τ_t and the consumption utility by $1 - \tau_1 - \tau_2$ (restricted to be positive).[13] Given attention level α_1 in period 1, and a strategy for the attention level in period 2, $\alpha_2(\cdot)$, the agent's payoff in period 1 is given by

$$V(\alpha_1, \alpha_2(\cdot)) = V_1^D(\alpha_1, \alpha_2(\cdot))[\alpha_1\tau_1] + E_{\epsilon_1 \sim F(\alpha_1)}[V_2^D(\alpha_2(\epsilon_1)|\epsilon_1)\alpha_2(\epsilon_1)\tau_2] + V_1^D(\alpha_1, \alpha_2(\cdot))[1 - \tau_1 - \tau_2]$$
$$+ V^O[(1 - \alpha_1)\tau_1 + (1 - \alpha_2)\tau_2 + (1 - \tau_1 - \tau_2)]. \quad (4)$$

In period 2, without commitment the agent maximizes (4) with $\tau_1 = 0$.

Lemma 2. *For every fixed α_1, there exists $\alpha_2(\cdot)$ that maximizes V.*

As previously mentioned, the agent may learn about the material value of the decision at the end of period 1. This need not be the case and to isolate different forces in the model, we define a class of environments that lack this feature.

[13] This formulation encapsulates (conventional) discounting of future payoffs, e.g., fix τ and discount factor δ and let $\tau_1(1+\tau) = \tau, \tau_2(1+\tau) = \delta\tau$ and $(1+\tau)(1-\tau_1-\tau_2) = \delta(1-\tau)$, but also discounting of anticipatory utility if consumption is far into the future (Loewenstein, 1987), e.g., fix τ and "discount factor" δ and let $\tau_1(1 + \delta\tau) = \delta\tau, \tau_2(1 + \delta\tau) = \tau$ and $(1 + \delta\tau)(1 - \tau_1 - \tau_2) = 1 - \tau$.

Definition 1. *The environment is realization-independent if for all* α_1, α_2,

$$V_2^D(\alpha_2|\epsilon_1)$$

is constant across $\epsilon_1 \in \text{support}(F(\alpha_1))$. *(With slight abuse of notation) we write* $V^D(\alpha_1, \alpha_2)$ *as the material value of the decision, taking* α_1 *as given in period 2.*

Examples of realization-independent environments are those where $F(\alpha_1)$ is degenerate, i.e., there is no uncertainty, an example of which is given Section B.2 (appropriately extended to multiple periods) where attention (deterministically) increases the available choices (e.g., allow the agent to deviate from a default). However, realization-independent environments can include uncertainty: When the agent tries to match a normally distributed state and receives a normally distributed signal as in Section B.1, the variance of her posterior is independent of the signal realization and so is her material value.

3.1 Attention traps and time inconsistency

Because, as the agent moves forward through time, they are no longer concerned about the anticipatory utility of past selves, she may be time-inconsistent. In particular, without commitment, in period 2, she will choose some $\alpha^*(\epsilon_1)$ that maximizes $V_2^D(\alpha_2|\epsilon_2)$; with commitment, she maximizes (4), without this constraint. We thus refer to period-1 and period-2 selves.

In general, period-1 self prefers her period-2 self to devote more attention to the decision that her period-2 self does: An increase in period-2 attention leads to an increase in anticipatory utility in period 1, an effect that period-2 self does not internalize.[14]

Lemma 3. *Fix any* α_1 *and* ϵ_1. *The optimal* α_2 *chosen by period-2 self is less than that period-1 would choose (i.e., with commitment).*

The time inconsistency, in particular period-2 self devoting low levels of attention, has knock-on effects. Period-1 self anticipates period-2 self not to devote a lot of attention, relative to what period-1 self would do, to the decision. As a result, 1) the material return to devoting additional attention (changes in V^D's) changes, and 2) the emotional response from devoting attention worsens; period-1 self may, in turn, readjust her attention.

[14]The agent may also want to adjust the timing of consumption to increase or decrease the time of anticipation. This can also lead to time inconsistency (or "reverse time inconsistency") but is studied elsewhere (Loewenstein, 1987).

We study how period-1 self may readjust her attention to overcome the commitment problem, when doing so is effective, and in what situations the commitment problem has severe consequences. The different forces at play are most easily presented in a stylized environment that we introduce next.

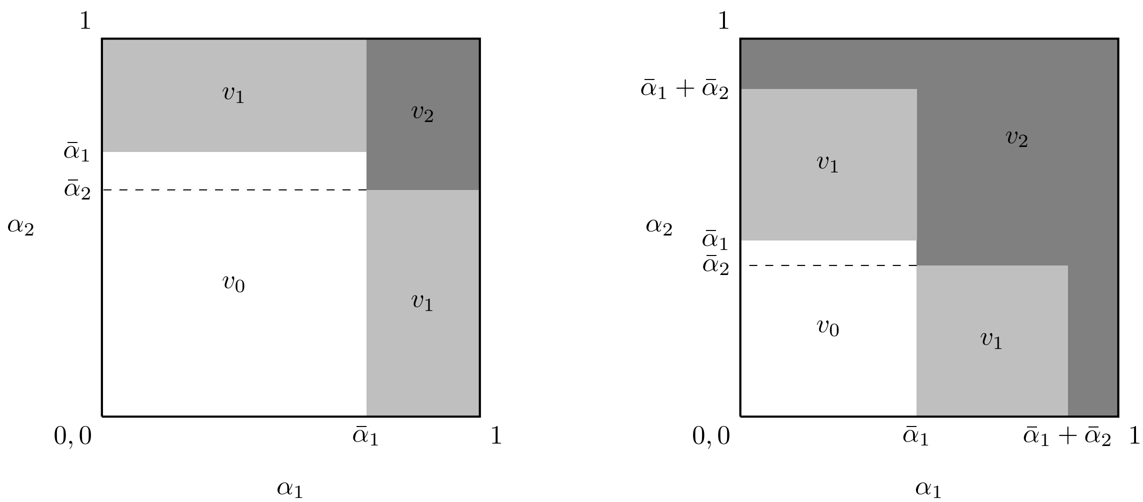
The agent faces a "two-step problem." There is a minimum amount of attention, $\bar{\alpha}_1 > 0$, required to complete the first step which increases the material value of the decision by some constant. The second step is completed analogously, requires an additional level of attention denoted by $\bar{\alpha}_2 > 0$, and increases the material value by some other constant. Depending on how much attention is required at each step, the agent can complete both steps in one period or only one step at a time. A step needs to be completed fully in a given period, i.e., it is not possible to complete a step by devoting some attention in both periods that only together sum up to the required level. Lastly, the agent has to complete the first step before the second step.

Formally, we say that V^D is a two-step problem if (the environment is realization-independent) and V^D can be written as

$$V^D(\alpha_1, \alpha_2) = \begin{cases} v_0 & \text{if } \alpha_1 < \bar{\alpha}_1, \alpha_2 < \bar{\alpha}_2 \\ v_1 & \text{if } (\underbrace{\bar{\alpha}_1 \leq \alpha_1 < \bar{\alpha}_1 + \bar{\alpha}_2 \text{ and } \alpha_2 < \bar{\alpha}_2}_{\text{step 1 in period 1 and nothing in period 2}}) \text{ or } (\underbrace{\alpha_1 < \bar{\alpha}_1 \text{ and } \bar{\alpha}_1 \leq \alpha_2 < \bar{\alpha}_1 + \bar{\alpha}_2}_{\text{nothing in period 1 and step 1 in period 2}}) \\ v_2 & \text{otherwise,} \end{cases}$$

for some $\bar{\alpha}_1, \bar{\alpha}_2 \in (0, 1]$ and $v_0 < v_1 < v_2$.

We are visualizing a two-step problem in Figure 2. Each panel depicts a unit square, i.e., the possible pairs of levels of attention (α_1, α_2). In the left panel, $\bar{\alpha}_1 + \bar{\alpha}_2 > 1$ and the agent needs to devote a minimum of $\bar{\alpha}_1$ attention in period 1 and $\bar{\alpha}_2$ in period 2 to achieve the highest material payoff of v_2, i.e., attention levels in the dark gray region. She can also only complete the two-step problem partially, i.e., only the first step, which is achieved for pairs of levels of attention in the light gray regions. The right panel differs from the left in that it depicts the case $\bar{\alpha}_1 + \bar{\alpha}_2 \leq 1$ which allows the agent to complete both steps of the two-step problem in a single period.



(a) Case $\bar{\alpha}_1 + \bar{\alpha}_2 > 1$. (b) Case $\bar{\alpha}_1 + \bar{\alpha}_2 \leq 1$.

Figure 2: Two-step problem V^D.

Let us compare now the outcome with and without commitment given such a two-step problem. Suppose that period-1 self strictly prefers to attend to the problem fully but does not devote extra attention, i.e., $\alpha_1 = \bar{\alpha}_1$ and $\alpha_2 = \bar{\alpha}_2$, or $\alpha_t = \bar{\alpha}_1 + \bar{\alpha}_2$ for some t. Observe that if each of the two steps does not require a lot of attention relative to the agent's attention budget, then the agent can implement her preferred solution without commitment. Specifically, she can "front-load," that is choose $\alpha_1 = \bar{\alpha}_1 + \bar{\alpha}_2$, in which case period-1 self cannot reduce her attention as the optimal level of attention under commitment is already minimal. The agent can also "back-load," that is she does not devote attention in period 1 and chooses $\alpha_2 = \bar{\alpha}_1 + \bar{\alpha}_2$, in which case period-2 self's choice of attention has no effect on period-1 anticipatory payoff so that period-2 and period-1 selves' objective functions coincide and the commitment solution is again implementable. Thus, when the agent can complete both steps in a single period, the time inconsistency problem can be muted.

Let us now consider the case when the agent has to devote attention in both periods to complete the two-step problem, e.g., if $\bar{\alpha}_1 + \bar{\alpha}_2 > 1$. This corresponds to "large problems." Alternatively, we can hold the required levels of attention to complete the two-step problem fixed and vary the degree of cognitive scarcity, i.e., whether the agent has a unit of attention per period to allocate or less. It may be the case that some of her attention is captured by a third task and she cannot devote a unit of attention. In either case, the time-inconsistency issue may have bite and can even make every self worse off compared to the commitment outcome, i.e., the agent may find herself in an attention trap. There are essentially two conditions on the two-step problem that need to be satisfied for an attention trap to occur. First, v_1, the payoff when the first part only is solved, cannot be too high, otherwise, the solution under commitment may be to just devote partial attention and no trap can occur. Second, v_1 cannot be too low, otherwise period-2 self would in fact devote attention $\alpha_2 = \bar{\alpha}_2$ when period-1 self chooses $\alpha_1 = \bar{\alpha}_1$ in order to capture the large gains in material payoff.

Proposition 5 formalizes these conditions.

Proposition 5. *Fix any two-step problem V^D with $\bar{\alpha}_1, \bar{\alpha}_2$ and v_0, v_1 and v_2, and $\tau_1 = \tau_2 = \tau > 0$. If*

$$\bar{\alpha}_1 + \bar{\alpha}_2 \leq 1,$$

then the agent can implement the commitment solution without commitment. If

$$\bar{\alpha}_1 + \bar{\alpha}_2 > 1,$$

then there exist $\underline{v}_1, \bar{v}_1(\tau)$ with $\underline{v}_1 < \bar{v}_1(\tau)$ and $\bar{v}_1(\tau)$ strictly increasing in τ, and material value of the outside option V^O so that if $v_1 \in (\underline{v}_1(\tau), \bar{v}_1(\tau))$, then

1. *Period-1 self strictly prefers to commit to $\alpha_1 = \bar{\alpha}_1$ and $\alpha_2 = \bar{\alpha}_2$ to any other outcome;*

2. *The unique outcome is no attention devoted in either period;*

3. *Period-2 self strictly prefers the commitment outcome to the outcome without commitment.*

Attention traps can occur in less extreme ways, in situations in which the lack of commitment does not lead to a drop in attention all the way to no attention. Consider a realization-independent environment (but not a two-step problem) with $\tau_1 = \tau_2 = 0.1$, V^O normalized to zero and

$$V^D(\alpha_1, \alpha_2) = H^D + \sqrt{\alpha_1} + \sqrt{\alpha_2},$$

with $H^D = -7$. Period-1 self's objective, $V(\alpha_1, \alpha_2)$, is maximized by equalizing attention across periods; and plotted as a function of attention in Figure 3(a).

Consider period-1 self's total payoff under commitment. Evidently, it is concave and maximized at $\alpha_1 = \alpha_2 = 1$ ($= \alpha^c$). In the absence of commitment, Lemma 3 suggests that period-2 self may want to deviate from this prescribed level of attention devoted to the decision; indeed, Figure 3(b) shows that period-2 self's total payoff, $V_2(\alpha_2|\alpha^c)$, is maximized at a lower level of attention, $\alpha_2(\alpha^c) < \alpha^c$. Figure 3(c) plots period-2 self's optimal level of attention devoted to the decision as a

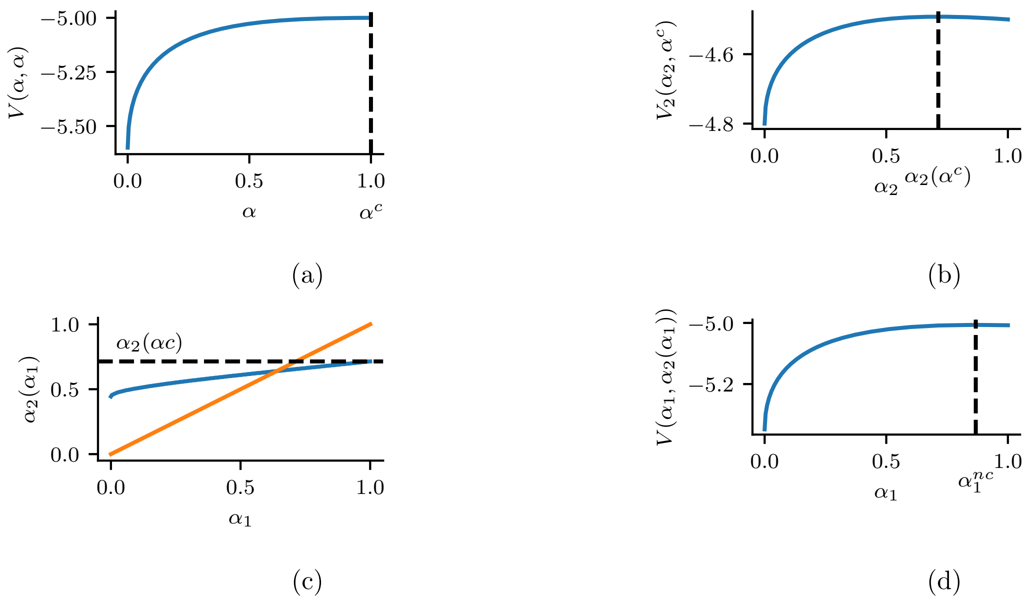


Figure 3

function of period-1 self's attention (blue line). Again, we can see that $\alpha_2(\alpha^c) < \alpha^c$ as $(\alpha^c, \alpha_2(\alpha^c))$ lies below the 45-degree line (orange line). Period-1 self takes period-2 self's optimal response into account and maximizes her total payoff as a function of α_1 as depicted in Figure 3(d). Without commitment, period-1 self then chooses a lower level of attention, α_1^{nc}, than with commitment. However, she does not stop devoting any attention.

Period-1 self is obviously worse off without commitment. Here, period-2 self is also worse off (as it is in Proposition 5): the loss in total payoff due to the decrease in attention by period-1 self outweighs the benefit from flexibly choosing own attention. In particular, period-2 self's payoff with commitment is -4.5 and ≈ -4.55 without. In other words, the lack of commitment makes both selves worse off.

An emotional inattentive agent thus demands commitment against falling into attention traps, or more generally being time-inconsistent; e.g., she could set automatic reminders that make the decision salient to her at a future date. In what follows, to keep the analysis tractable, we assume that the agent has access to such commitment devices.

Our attentional traps represent a dynamic extension of the attentional volatility we found in Section 2.4. As should be apparent from the examples, we may observe threshold effects: where if attention in different periods falls just below a particular level, then attention drops discontinuously in all periods (as can utility); thus these results are also consistent with the threshold evidence of

Balboni et al. (2021) for scarcity in physical assets. But, of course, the results we obtain are stronger, because as we have shown, if agents could commit, they could often achieve payoffs that are a Pareto improvement for all selves.

Thus, our dynamic extension can help shed even more light on issues surrounding poverty traps as discussed in Section 2.4. In particular, even tasks that can ultimately generate large returns in helping individuals escape poverty may not be undertaken, because they require attention spread out over many periods. Due to dynamic inconsistency agents may simply not take on these tasks at all. Thus, building on Mullainathan and Shafir (2013), it is not just scarce cognitive resources that act as a constraint, but the inability to coordinate the use of those resources over time. Thus, either simplifying tasks so they can be accomplished in one period (thus eliminating the need for temporal coordination as in the first part of Proposition 5 which required that the total attention need was less than what could be devoted in a single period), assisting agents with commitment devices, or providing them with enough resources to accomplish a task quickly, can all potentially serve to help individuals escape poverty.

3.2 Early vs late attention

It is well known that when individuals have non-linear anticipatory utility they can have "non-standard" preferences towards information. For example, a classic question is to what extent individuals have preferences over information even when they cannot take any actions (e.g., Kreps and Porteus, 1978; Epstein and Zin, 1989). Experimental evidence has been strongly consistent with the hypothesis that individuals have a preference for earlier resolution of information (see Masatlioglu et al., 2017 and Nielsen, 2020 for recent papers exploring this).

In our setting, we can ask an equivalent question. Because information is endogenously generated via attention, we can ask when it is the case that agents want to acquire information earlier, rather than later, even when the information does not influence any action they are taking?[15]

We again focus on a simple two-period setting. In order to focus on intrinsic reasons for informational preferences (in contrast to the class instrumental reasons), we focus on a particular set of environments: we rule out that the agent can condition an observed action on the realization of the information.

[15] A distinct question asks about the timing of consumption — does the agent prefer consumption to occur in period 1 or period 2. If anticipatory utility drives the emotional response, then the agent may prefer to delay consumption of pleasurable goods while consuming unpleasant goods immediately as to vary the duration of anticipatory feelings (see Loewenstein (1987)).

Definition 2. *There is no instrumental value of attention if $V_1^D(\alpha_1, \alpha_2(\cdot))$ is constant.*

An example of an environment in which there is no instrumental value of attention is one where the agent's material value is determined by a random state but she does not take any action or make any choice. By devoting attention to the decision, the agent may still learn about the material value.

Our environment not only rules out information being instrumentally valuable after period 1 but also instrumentally valuable after period 2. This eliminates a classic motivation for the timing of learning in dynamic models: with costly experimentation, there is a benefit to sequential learning (i.e., learning some in period 1, and then conditioning learning in period 2 on the outcome of information in period 1) — e.g., in the setting of Moscarini and Smith (2001). In these settings, the agent would benefit from devoting some attention early (in period 1), with the option of devoting further attention late (in period 2).

Instead, our setting allows us to focus on a second, distinct form of informational preferences: The agent can not only condition her attention devoted to the decision in period 2 on how much she learned in period 1, but also on how it shifted her belief about the eventual payoff up or down: If the payoff is likely to be high, the agent continues to devote attention to the decision; otherwise, she reduces attention. In particular, upon observing different values for the material value in period 1, the agent may allocate her attention in period 2 depending on such values because of the emotional focus: she may devote more attention to the decision if the material value turns out to be high.

Proposition 6.

- *For every V^O and H^D, there exists V_1^D, with $V_1^D(0,0) = H^D$, and no instrumental value of attention, so that the agent devotes all her attention to the decision in period 1, $\alpha_1 = 1$.*

- *For every V_1^D and V_2^D such that $V_2^D(0|\epsilon_1)$ is not a constant with probability $\delta > 0$ when $\epsilon_1 \sim F(\alpha_1')$ for some α_1', and no instrumental value of attention, there exists $V^O > V_1^D(0,0)$ such that $\alpha_1 > 0$.*

Allowing for instrumental value of attention does not change the substance of the proposition. For example, suppose the agent can acquire information, say a single signal, in either period by devoting at least $\underline{\alpha}$ of attention to the decision of (she can always devote more attention and attention in the period without information). Then the agent prefers the signal to be available in period 1.

Classic models of anticipatory utility generate a preference for early resolution of information via an assumption that utility is convex in anticipatory utility (e.g., Kreps and Porteus, 1978; Caplin and Leahy, 2001). In contrast, anticipatory utility enters total utility in a linear fashion in our model. Despite this, we have a preference for early attention. This is because agents have hidden actions (e.g., Ergin and Sarver, 2015). Because of this, information generates option value for the agent, and so the payoff function is convex information, even without observable actions.

4 Applications

4.1 A consumption-savings model

We study the consumption and savings behavior of an agent in the presence of emotional inattention. There are two consumption periods, $t = 1, 2$. At the beginning of period 1, the agent needs to decide how much to consume in period 1 which is denoted by c_d (for 'default'). *After* making this decision, she receives income y which is distributed according to some continuous distribution F.[16] She can then either devote attention to the income, revise her consumption, and consume $c_1^*(y)$ in period 1 and $y - c_1^*(y)$ in period 2, or she can devote attention to an outside option that gives a payoff of V^O in which case she consumes c_d in period 1 and $y - c_d$ in period 2. The agent is risk-averse and equipped with a CARA within-period utility function denote by u. Let y_{inf} denote the infimum of the support of y. c_d is restricted to be at most y_{inf} so that the agent need not devote attention to consume in period 2. Given her risk-aversion, $c_1^*(y) = y/2$.

Her total payoff if she devotes attention is given by

$$2u(\frac{y}{2}) + V^O[1 - \tau],$$

and if she does not devote attention

$$(u(c_d) + u(y - c_d))[1 - \tau] + V^O.$$

For a given y, she pays attention to her consumption if and only if the former exceeds the latter.

Let y_{sup} denote the supremum of the support of y and suppose that $2u(\frac{y_{\text{sup}}}{2}) > V^O$ so that the agent pays attention for some income shocks.

[16] Continuity is not a substantive assumption but simplifies the statement of the results.

Proposition 7. *Given any c_d, there exist $\underline{y}, \bar{y}$, so that the agent devotes attention if $y < \underline{y}$ or $y > \bar{y}$, does not devote attention if $y \in (\underline{y}, \bar{y})$ and is indifferent otherwise.*

When V^O is large enough, then for any c_d, $\underline{y} = -\infty$, i.e., the agent never pays attention for low-income realization, but may otherwise be finite. In contrast, $\bar{y}$ is always finite.

When $\underline{y} = \infty$, the agent responds *asymmetrically* to income shocks: low-income shocks are ignored and she does not adjust her planned income, whereas she pays attention to positive income shocks and increases from her planned consumption.

The agent anticipates this and has a precautionary savings motive. Let c_{na} (for 'no adjustment') denote her optimal level of consumption in period 1 if the agent could not readjust her consumption after observing y. c_{na} solves:

$$u'(c_{\text{na}}) = E[u'(y - c_{\text{na}})],$$

where u' is the first derivative of u.

When $\underline{y} = -\infty$ and the agent can adjust consumption by devoting attention to it, the agent chooses a (tentative) consumption level c_d that is at most c_{na}. c_d solves:

$$\int_{y=-\infty}^{\bar{y}} u'(c_d) dF = \int_{y=-\infty}^{\bar{y}} u'(y - c_d) dF.$$

Thus, an emotionally inattentive agent may exhibit several potential biases simultaneously. First, ex-ante pessimism, because she wants to ensure that she does not lose too much utility in states she fails to pay attention to. Second, she exhibits a default effect — for some states of the world the agent fails to adjust her plan. Third, she exhibits ex-post optimism for the worst states — because she fails to pay attention for the worst realizations and overconsumes when times are bad. (She may also overconsume for income realizations just below $\bar{y}$.)

Within this setting, we can consider the effects of changing τ (while remaining in the regime with $\underline{y} = -\infty$). We focus on situations where $\tau > 0$ because when $\tau = 0$ the solution for c_1 is set-valued.

Lemma 4. *Suppose $\underline{y} = -\infty$ and consider an increase in $\tau \in (0, 1)$. Then both $\bar{y}$ and c_1 increase.*

In other words, an agent who puts more weight on anticipatory utility will avoid paying attention for more income realizations. Consequently, she will choose a consumption level that is more suitable for those higher-income realizations.

4.2 Portfolio choice

We next study the effects of emotional inattention on the agent's portfolio choice in an asset allocation problem. Among others, our results highlight the dynamic implications of the previously discussed ostrich effect. Our goal will be to compare an agent who has no anticipatory utility $\tau = 0$, to an agent with anticipatory utility $\tau > 0$. We refer to the former as the standard agent and to the latter as an emotionally inattentive agent.

There are three ways in which emotional inattention affects the agent's portfolio choice. First, the agent likes risk as it allows her to exercise option value about what to pay attention to: If the asset is performing well, the agent devotes attention to it, if it performs poorly, she chooses to ignore it. Hence, she values mean-preserving spreads, just as in Proposition 6.

Formally, consider a two-period setting. In period 1, the agent chooses to allocate her wealth w into a safe asset, that returns 1 for sure, and a risky asset with variable return $1 + r$. At the end of period 1, she learns the return of the risky asset but cannot readjust her portfolio. In period 2, she then decides to devote attention to her portfolio or an outside option. At the end of period 2, she consumes the outside option and her wealth, evaluated with some felicity function u. She weights consumption utility by $1 - \tau$ and anticipatory utility in period 2 by τ. If she invests a fraction f in the risky asset, her return is r and she devotes attention in period 2, her payoff is

$$u(w + wfr) + V^O[1 - \tau],$$

and if she does not devote attention, her payoff is

$$u(w + wfr)[1 - \tau] + V^O.$$

Anticipating her optimal attention choice in period 2, the agent chooses f optimally. In this situation, an emotionally inattentive agent invests more in the risky asset than a standard agent.

Proposition 8. *The share invested in the risky asset is increasing in τ.*

Second, the agent dislikes assets that require attention. In general, the agent is always free to devote attention even though it is not needed to increase material value (as is the case for the outside option). Assets that do not require (but permit) attention are thus better, ceteris paribus. For example, consider two assets with the same monetary payoff profile. However, one asset requires some attention of the agent, say because of periodic paperwork, while the other can be safely ignored. An emotionally inattentive agent would then prefer the latter: when the asset is performing well, she can still (voluntarily) devote attention, and when it is performing poorly, she can ignore it as no paperwork is due; a standard agent is instead indifferent. Emotionally inattentive investors may thus prefer funds that are managed by someone else and that do not require constant attention by the investor.

Lastly, an emotionally inattentive agent may choose allocations that perform well even if the market is down. She anticipates that she will not want to re-optimize her portfolio in response to evolving market conditions if those conditions are poor. Hence, she will choose a portfolio that precisely performs well when she will not attend to it. If risky assets perform poorly when market conditions are bad, more so than safe assets, then emotionally inattentive agents show more risk-aversion than their felicity functions suggest. We make this point in a simple stylized example, in particular one that 1) turns off the channel for risk-seeking behavior (Proposition 8), and 2) correlates bad market conditions with risky assets performing poorly.

Consider a setting similar to that from the beginning of this section. The agent chooses a fraction f of her wealth to invest in the risky asset. The risky asset generates a payoff at two points in time and the payoffs are independent. The first payoff of the risky asset is realized and immediately consumed by the agent so that the agent's wealth and asset allocation stays fixed. The agent additionally learns whether the market is "at risk of being down" (or simply down) which occurs with probability p_{risk} in which case the future payoff of the risky asset will be wiped out with probability $\rho > 0$. When she learns that the market is down, she can devote attention to her portfolio allocation problem and reallocate her assets. However, this, as usual, creates an emotional response by increasing the weight on the anticipatory utility from consuming her expected wealth. For simplicity, there is no period in which she devotes attention conditional on the second realized return.

For simplicity, we compare the emotionally inattentive agent with $\tau > 0$ with the standard agent and assume that the parameters are such that the emotionally inattentive agent optimally chooses to devote attention if and only if the market is not down. We have the following proposition.

Proposition 9. *The emotionally inattentive agent initially invests less in the risky asset than the standard agent.*

When the market is up, both the standard agent and the emotionally inattentive agent invest the same fraction of their wealth in the risky asset.

When the market is down, the inattentive agent does not re-optimize and invests more in the *risky asset than the standard agent.*

Here also, emotionally inattentive investors would value automation, a technology that adjusts her portfolio in response to market conditions without requiring attention.

Researchers and policymakers thus need to be careful when inferring risk-aversion from observed data. As we have shown, it is confounded by emotional inattention in perhaps three distinct ways. Moreover, our results highlight that individuals may appear to be too risk-loving, or too risk-averse, depending on the context; but that neither of these deviations is actually due to changes in risk preferences, but rather due to attentional concerns. The fact that attention is costly, and more costly when states are bad, leads to the fact that agents in our model have a preference for assets that do not require attention when performing poorly.

5 Discussion

In this section, we will discuss how our current approach relates both at a conceptual, as well as substantive, level to existing models. We compare our model of emotional inattention to four types of models. First, to the literature that assumes information acquisition is exogenously costly — the rational inattention literature (e.g., Sims 2003). This literature supposes that individuals must pay a cost to acquire information, but that this cost is not related to anticipatory feelings. Second, we compare our model to the literature on anticipatory emotions, where there is no explicit cost of information acquisition, although individuals can potentially choose their information source. Third, we consider the literature on news utility and reference-dependent preferences, where individuals gain flow utility not by how likely they think outcomes are, but how those likelihoods have changed recently. Last, we discuss how our model relates to a small subset of papers that have tried to understand attention and anticipation simultaneously.

One distinction between our approach and the first three literatures are "income effects." In particular, in our model three things are true simultaneously: (i) if we increase the value of the decision problem, the cost of information acquisition (due to the emotional response) shrinks, (ii) if we change the value of the decision problem and the trivial task keeping overall expected utility constant, we will change the cost of information acquisition for the decision, (iii) if we increase the value of the trivial task, the cost of information acquisition increases.[17] In models of rational

[17](i) and (iii) generate our asymmetric ostrich effect: individuals exhibit the ostrich effect for bad situations, but not good ones.

inattention, anticipatory utility, or news utility these three phenomena imply that there must be income effects (because utility levels affect the cost of information acquisition). But to match the fact that informational attitudes vary with changes to both the payoff of the decision and the payoff of the trivial task individuals must broadly bracket. However, to accommodate the fact that an increase in utility from just the decision problem is perceived differently than the same increase in utility to the trivial task, individuals must also narrowly bracket.[18]

Rational Inattention

Our model captures costly attention, and so is very similar in many ways to the approach of rational inattention (e.g., Sims (2003)). However, as discussed, it differs because in these models the cost of attention is an exogenous function that depends only on the informativeness of the signal and prior over the states. In contrast, in our model of emotional inattention the cost of attention is endogenous — it depends on future payoffs, which themselves depends4 on the amount of attention that is paid. As we have shown, this endogeneity drives many of our important results and serves to distinguish us from rational inattention. For example, when the value of the decision is concave in the amount of information acquired, and the exogenous cost of information acquisition is convex, then we should expect that the optimum amount of information acquired (and so material payoffs) smoothly varies with the parameters of the model. In contrast, in our model, because of feedback between the two roles of attention, even with a concave V^D, we can have a payoff function that is convex in attention, where small changes in parameters lead to large changes in attention and material utility. Moreover, most existing models of rational inattention assume dynamic consistency — leading to no value of commitment, and so no attentional traps, contrary to our model. Last, existing models of rational inattention cannot generate the intrinsic preferences for early resolution that emotional inattention generates — we should observe timing indifference when actions do not affect payoffs.

Anticipatory Emotions

We can also compare our model models of anticipatory utility (e.g., Caplin and Leahy, 2001). The key conceptual distinction is that in anticipatory models agents must gain (or lose) anticipatory utility in accordance with their beliefs, regardless of how much they focus on a decision. In contrast,

[18] In fact, most models of rational inattention, with the exception of Liu et al. (2019), make the cost of attention additively separable from the benefits. This rules out "income effects" in the cost of information acquisition.

in emotional inattention agents can control the extent to which they gain or lose anticipatory utility by controlling where they direct their attention. After receiving utility from beliefs in a given period, our agent has a choice as to whether to continue receiving those anticipatory benefits (even without beliefs changing) solely by adjusting their attention.

In order for standard models of anticipatory utility to generate the asymmetric ostrich effect our model delivers, it needs to be the case that individuals' utility over beliefs about future outcomes is concave for low expected payoffs (which generates information aversion), and convex for high expected payoffs (which generates information loving) — i.e., it is inverse S-shaped (see Kreps and Porteus, 1978 and Dillenberger and Raymond, 2020 for a discussion of how the shape of the utility function over beliefs affects preferences towards learning). In anticipatory models, this immediately implies that policies that raise the expected payoff of the agent should induce her to become more information-loving. In contrast, we have shown that an emotionally inattentive agent may become more information averse under policies that raise expected payoffs — e.g., giving free information.

Similarly, the anticipatory model with inverse S-shaped belief-based utility should, for high initial levels of utility, always prefer earlier resolution, and for low levels of utility, prefer later resolution of uncertainty. In a realization-independent world, where everything is symmetric, the predictions of anticipatory models are even starker — individuals shouldn't exhibit any informational preferences (because anticipatory utility is generated via utility differences across states which do not exist). In contrast, our results setting show that regardless of the payoff level emotionally inattentive agents always weakly prefer earlier resolution when actions do not affect payoffs.

Reference Dependence

Distinct from models of anticipatory utility, where flow utility depends on the beliefs about future outcomes, there are also models of changing beliefs, news utility, and reference dependence, where utility depends on changes in beliefs due to recent information (e.g., Kőszegi and Rabin, 2009). In these models (unlike our model) flow utility from emotions is 0 if and only if nothing has been learned. In our model, flow utility is equal to the normalized value of the outside option if and only if either all attention is directed at the trivial task, or the decision has the same expected utility as the outside option. Models like this cannot generate the feedback loops that generate our attentional volatility or attention traps.

However, these models can generate various forms of information avoidance. (Olafsson and Pagel, 2017) explore how reference-dependence utility could generate ostrich-type behavior. They show that individuals will want to avoid paying attention to their bank accounts and that this attention aversion will decrease when wealth is large. Despite this, using reference-dependent preferences to capture the ostrich effect has potential pitfalls. First, as Olafsson and Pagel (2017) explain "it cannot rationalize an increase in attention at a fully expected income payment or a jump in the probability of logging in when balances turn from negative to positive." In contrast, attentional volatility can lead to large swings in attention despite small changes in parameters and can rationalize paying attention to even fully anticipated changes. Second, the asymmetric ostrich effects generated by reference-dependent models are "second-order" — they emerge because of decreases in the concavity over money at higher levels of wealth. In emotional inattention such shifts are first-order — they occur because of the changing the height of payoffs in the decision problem. Third, like anticipatory models, in reference-dependent models, because utility is based on changes in expected future payoffs, in a realization-independent setting with equal priors across states, reference-dependence models predict indifference about the timing of information.

Models Combining Attention and Belief-Based Utility

Last, we compare our model of emotional inattention to existing models that capture both anticipation and attention simultaneously. Perhaps the earliest paper to discuss the potential role of attention and anticipation is Loewenstein (1987). Although he focuses his formal analysis on anticipation alone, he briefly discusses an extension to his model which allows for the vividness of an experience (which can be related to the amount of attention an experience attracts). Building on this insight, Karlsson et al. (2009) propose that attention amplifies the emotional impact of information. They model attention as a discrete choice and suppose that individuals experience not anticipatory utility but rather reference-dependent utility, which is based on current expectations relative to prior expectations. Paying attention increases the relative impact of gain-loss utility, and also speeds up the adjustment of the reference point. They find that in their setting that investors should pay more attention to their finances after good news than after bad news.

Our model extends their approach — for example, relaxing the assumption of discrete attention (which imposes substantive restrictions on the shape of V^D, and rules out important behaviors). Our model is also conceptually distinct because the endogenous cost for us is not due to changes in beliefs, but rather levels of beliefs. This is important because in reference-dependent models, generating ostrich effects can be sensitive to the specification — e.g., simple modifications to the model of Karlsson et al. (2009) show that although these models robustly predict a symmetric ostrich effect, but much less robustly an asymmetric effect.[19]

More broadly, because flow emotional utility is generated via changes in beliefs, Karlsson et al. (2009) cannot capture the complementary roles of attention as (i) improving decisions and (ii) generating anticipatory utility, which rules out many of our important results, such as attentional volatility and attentional traps.

Golman and Loewenstein (2018) and Golman et al. (2021) provide models that incorporate attention, as well as anticipatory utility, but model attention provision as automatic, rather than as a choice variable. Because attention is exogenous, their analysis excludes most of the features and results that we emphasize.

Tasoff and Madarasz (2009)'s model is closest to our own. Agents face a decision problem with multiple dimensions (analogous to our different decisions), and receive anticipatory utility as a function of utility in that dimensions and the attention they pay to that dimension. Like us, they model attention as a fixed resource. Unlike us, they assume that attention is a discrete choice. They suppose that if information is gained about payoffs in a particular dimension, then attention is directed there. There are many similarities between our approach and theirs. Formally, their Proposition 2 (and Corollary 1) draws on the same intuition as our results on the ostrich effect: fixing the instrumental value of information, bad situations attract little attention, while good situations do. Their Proposition 3 implies that there will be "default" effects, that individuals will not always adjust their decisions, and that this happens more often after bad signals than good, similar to what we find in our applications on consumption-savings problems and asset allocations. Like us, they highlight the important fact that in realization independent environments (which they say capture situations in which information is only about how to choose the right action) that intrinsic informational preferences can occur, unlike pre-existing anticipatory approaches.

However, there are also important differences. Two of these differences are conceptual. First, they model the acquisition of information as causing attention. In other words, while in our model the choice of attention can determine the information structure, in theirs the choice of information structure determines (in expectation) the attention. Thus, for example, they do not focus on the situation where attention can be directed even when no information has come to light, a key

[19] The implications of the Karlsson et al. (2009) model are generated in part because the first-period gains or losses (relative to the reference point) are not affected by attention, they are just scaled up or down by attention. Only the second-period reference point is affected. If instead both the realized gains/losses in the first period, and the reference point in the second period are equally affected by attention, then their major finding, of the asymmetric ostrich effect, disappears, and instead, a symmetric ostrich effect emerges. One could also try to rely on the concavity of utility over wealth to generate asymmetries, as in Olafsson and Pagel (2017), but as noted before, this only generates second-order asymmetries.

driving force for our results regarding intrinsic preferences for the timing of learning. Second, they separately consider different ways that attention could affect material payoffs — e.g., thinking about attentions and changing decisions, or attention related to information as distinct objects. We consider only the "combined" effect of this through our general V^D function.

Substantively our paper also considers important behaviors they do not address. For example, they do not discuss potential policy interventions, nor attentional volatility.[20] Similarly, they do not address dynamic issues, such as intrinsic informational preferences or dynamic inconsistency. Our applications are also distinct from theirs, as their focus is on a monopolist manipulating information provided to (and thus the attention of) consumers.

6 Conclusion

This paper introduces a model of emotional inattention in which attention serves a dual role leading to instrumental value but also prompting an emotional response. The emotional response from attention then immediately predicts the well-documented ostrich effect of individuals not attending to negative situations (to avoid the negative emotional response). A key driver of our results is the complementarity between the instrumental and emotional consequences of attention. We study preferences over the timing of uncertainty and highlight potentially perverse consequences of policies designed to increase attention. We also show how dynamic inconsistency and volatile attention levels can lead to attention traps. We relate some of our predictions to existing empirical evidence and leave others to be tested in the future.

We show the implications of emotional inattention for consumption-savings decisions and portfolio choice, again providing novel explanations and predictions, but also highlighting the flexibility of our modeling approach showing its ripeness for applied work.

[20] Proposition 4 discusses a form of complementary between consumption and attention. Although reminiscent of our complementarities in the dual roles of attention, it is distinct, and they do not explore the effects, such as volatility.

References

Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar, "Large stakes and big mistakes," *The Review of Economic Studies*, 2009, *76* (2), 451–469.

Balboni, Clare A, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil, "Why do people stay poor?," Technical Report, National Bureau of Economic Research 2021.

Banerjee, Abhijit V. and Sendhil Mullainathan, "Limited Attention and Income Distribution," *American Economic Review*, May 2008, *98* (2), 489–93.

Becker, Marshall H and Lois A Mainman, "Sociobehavioral determinants of compliance with health and medical care recommendations," *Medical Care*, 1975, *1* (13), 10–24.

Benedek, Mathias, Emanuel Jauk, Roger E. Beaty, Andreas Fink, Karl Koschutnig, and Aljoscha C. Neubauer, "Brain mechanisms associated with internally directed attention and self-generated thought," *Scientific Reports*, Mar 2016, *6* (1), 22959.

Böheim, René, Dominik Grübl, and Mario Lackner, "Choking under pressure–Evidence of the causal effect of audience size on performance," *Journal of Economic Behavior & Organization*, 2019, *168*, 76–93.

Brosch, Tobias, Klaus Scherer, Didier Grandjean, and David Sander, "The impact of emotion on perception, attention, memory, and decision-making," *Swiss medical weekly*, May 2013, *143.*

Caplin, Andrew and John Leahy, "Psychological Expected Utility Theory and Anticipatory Feelings*," *The Quarterly Journal of Economics*, 02 2001, *116* (1), 55–79.

_ **and** _ , "The Social Discount Rate," *Journal of Political Economy*, 2004, *112* (6), 1257–1268.

Chen, Yuming, Dandan Zhang, and Donghong Jiang, "Effects of Directed Attention on Subsequent Processing of Emotions: Increased Attention to Unpleasant Pictures Occurs in the Late Positive Potential," *Frontiers in Psychology*, 2018, *9*, 1127.

Currie, Janet, "The Take-Up of Social Benefits," in Alan J. Auerbach, David Card, and John M. Quigley, eds., *Public Policy and the Income Distribution*, Russell Sage Foundation, 2006, pp. 80–148.

Dean, Emma Boswell, Frank Schilbach, and Heather Schofield, "2. Poverty and Cognitive Function," in "The economics of poverty traps," University of Chicago Press, 2019, pp. 57–118.

Dillenberger, David and Collin Raymond, "Additive-Belief-Based Preferences," Working Paper 20-020, SSRN July 2020.

DiMattero, Robin M, Kelly B Haskard, and Summer L Williams, "Health beliefs, disease severity, and patient adherence: a meta-analysis," *Medical Care*, 2007, *6* (45), 521–8.

Epstein, Larry G and Stanley E Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption," *Econometrica*, 1989, *57* (4), 937–969.

Ergin, Haluk and Todd Sarver, "Hidden actions and preferences for timing of resolution of uncertainty," *Theoretical Economics*, 2015, *10* (2), 489–541.

Falk, Armin and Florian Zimmermann, "Beliefs and Utility: Experimental Evidence on Preferences for Information," Working Paper 10172, Institute for the Study of Labor August 2016.

Ganguly, Ananda and Joshua Tasoff, "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future," *Management Science*, 2017, *63* (12), 4037–4060.

Ghatak, Maitreesh, "Theories of poverty traps and anti-poverty policies," *The World Bank Economic Review*, 2015, *29* (suppl_1), S77–S105.

Golman, Russell and George Loewenstein, "Information gaps: A theory of preferences regarding the presence and absence of information," *Decision*, 2018, *5* (3), 143–164.

_ , Nikolos Gurney, and George Loewenstein, "Information gaps for risk and ambiguity," *Psychological Review*, 2021, *128* (1), 86–103.

Karlsson, Niklas, George Loewenstein, and Duane Seppi, "The ostrich effect: Selective attention to information," *Journal of Risk and Uncertainty*, 2009, *38* (2), 95–115.

Katsuki, Fumi and Christos Constantinidis, "Bottom-Up and Top-Down Attention: Different Processes and Overlapping Neural Systems," *The Neuroscientist*, 2014, *20* (5), 509–521. PMID: 24362813.

Kreps, David M. and Evan L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 1978, *46* (1), 185–200.

Kőszegi, Botond and Matthew Rabin, "Reference-Dependent Consumption Plans," *American Economic Review*, June 2009, *99* (3), 909–36.

Lindberg, Nangel M. and David Wellisch, "Anxiety and compliance among women at risk for breast cancer," *Annals of Behavioral Medicine*, 2001, *4* (23), 298–303.

Liu, Ce, Christopher Chambers, John Rehbeck et al., "Costly Information Acquisition," Technical Report 2019.

Loewenstein, George, "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 1987, *97* (387), 666–684.

Lybbert, Travis J and Bruce Wydick, *4. Hope as Aspirations, Agency, and Pathways. Poverty Dynamics and Microfinance in Oaxaca, Mexico*, University of Chicago Press, 2019.

Masatlioglu, Yusufcan, A Yesim Orhun, and Collin Raymond, "Intrinsic information preferences and skewness," *Ross School of Business Paper*, 2017.

Moscarini, Giuseppe and Lones Smith, "The optimal level of experimentation," *Econometrica*, 2001, *69* (6), 1629–1644.

Mrkva, Kellen, Jairo Ramos, and Leaf Van Boven, "Attention influences emotion, judgment, and decision making to explain mental simulation," *Psychology of Consciousness: Theory, Research, and Practice*, 2020, *7* (4), 404–422.

Mullainathan, Sendhil and Eldar Shafir, *Scarcity: Why having too little means so much*, Macmillan, 2013.

Nielsen, Kirby, "Preferences for the resolution of uncertainty and the timing of information," *Journal of Economic Theory*, 2020, *189*, 105090.

Olafsson, Arna and Michaela Pagel, "The Ostrich in Us: Selective Attention to Financial Accounts, Income, Spending, and Liquidity," Working Paper 23945, National Bureau of Economic Research October 2017.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, April 2013, *103* (2), 804–30.

Schelling, T. C., "THE MIND AS A CONSUMING ORGAN," 1988.

Sherbourne, Cathy Donald, Ron D. Hays, Lynn Ordway, M. Robin DiMatteo, and Richard L. Kravitz, "Antecedents of adherence to medical recommendations: Results from the medical outcomes study," *Journal of Behavioral Medicine*, 1992, (15), 447–468.

Shouldson, Ira and Anne Young, "Milestones in huntington disease," *Movement Disorders*, 2011, *6* (26), 1127–33.

Sicherman, Nachum, George Loewenstein, Duane J. Seppi, and Stephen P. Utkus, "Financial Attention," *The Review of Financial Studies*, 11 2015, *29* (4), 863–897.

Sims, Christopher A, "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Tasoff, Joshua and Kristof Madarasz, "A Model of Attention and Anticipation," Working Paper, SSRN November 2009.

Wiens, Stefan and Elmeri Syrjänen, "Directed attention reduces processing of emotional distracters irrespective of valence and arousal level," *Biological Psychology*, 2013, *94* (1), 44 – 54.

Yamaguchi, Shuhei and Keiichi Onoda, "Interaction between Emotion and Attention Systems," *Frontiers in neuroscience*, Sep 2012, *6*, 139–139. 23055954[pmid].

A Proofs Omitted from the Text

Proof of Lemma 1. V can only be discontinuous when V^D is; V^D increasing then ensures that V is increasing at such points of discontinuity. Let $c < 1$ be such a point of discontinuity. Right continuity and boundedness of V^D ensure that V is right-continuous and bounded and so $\lim_{x \to c^+} V(x) = V(c)$. $\square$

Proof of Proposition 1. The numbered claims follow immediately from supermodularity of V in (α, H^D) and $(\alpha, -V^O)$, and Topkis's Monotonicity Theorem.

For the second part, we need to show that $V(0) < V(\alpha)$ for all $\alpha > 0$, i.e., find H^D such that

$$[H^D + \tilde{V}^D(0)][(1-\tau)] + V^O > \sup_{\alpha>0}\{[H^D + \tilde{V}^D(0)][(1-\tau) + \tau\alpha] + V^O[(1-\tau) + \tau(1-\alpha)]\}.$$

Rearranging gives

$$H^D < \sup_{\alpha>0}\left\{-\frac{1-\tau}{\tau}\frac{\tilde{V}^D(\alpha) - \tilde{V}^D(0)}{\alpha} - \tilde{V}^D(\alpha) + V^O\right\}.$$

To ensure that the right-hand side is finite, we need to check that the following limit exists and is finite:

$$\lim_{x\to 0}\left\{-\frac{1-\tau}{\tau}\frac{\tilde{V}^D(\alpha) - \tilde{V}^D(0)}{\alpha} - \tilde{V}^D(\alpha) + V^O\right\}$$

As $\tilde{V}^D$ is differentiable, and thus continuous, at 0, the limit indeed exists and is finite.

For the final claim, we consider two cases: First, suppose that $V^D(1) > V^O$. The unique optimal level of attention devoted to the decision is $\alpha = 1$. Second, suppose that $V^D(1) \leq V^O$; we show that the derivative of V with respect to τ is decreasing in α establishing supermodularity of V in $(\alpha, -\tau)$. Consider $\alpha' > \alpha$. We have

$$\frac{\partial}{\partial\tau}V(\alpha) = V^D(\alpha)[-1+\alpha] + V^O[-1-\alpha].$$

Consider $\alpha^H > \alpha^L$. We have

$$\begin{aligned}
&\frac{\partial}{\partial\tau}V(\alpha)\mid_{\alpha=\alpha^H} - \frac{\partial}{\partial\tau}V(\alpha)\mid_{\alpha=\alpha^H} \\
=&V^D(\alpha^H)[-1+\alpha^H] + V^O[-1-\alpha^H] - \left(-V^D(\alpha^L)[-1+\alpha^L] + V^O[-1-\alpha^L]\right) \\
\leq&V^D(\alpha^L)[-1+\alpha^H] + V^O[-1-\alpha^H] - \left(-V^D(\alpha^L)[-1+\alpha^L] + V^O[-1-\alpha^L]\right) \\
=&V^D(\alpha_L)[\alpha_H - \alpha_L] - V^O[\alpha_H - \alpha_L] \\
\leq&0.
\end{aligned}$$

□

Proof of Proposition 2. Let $I = \alpha + \beta$. The total payoff as a function of I is given by

$$V^D(I)[(1-\tau) + (I-\beta)\tau] + V^O[(1-\tau) + (1-(I-\beta))\tau].$$

It is easy to check that the above has increasing differences in $(I, -\beta)$. As $\beta < \alpha^*$ and $\alpha^*_\beta + \beta \leq 1$, it is without loss to maximize the above over $I \in [\beta, 1+\beta]$ and the result follows by Topkis' Monotonicity Theorem. □

Proof of Proposition 3. (3) is the derivative of agent's objective with respect to γ. Thus, in general, if it is increasing (decreasing) in α, then the attention devoted to the decision is increasing (decreasing) in γ by Topkis' Monotonicity Theorem. Clearly, if $\tilde{V}^D \geq 0$, then (3) is increasing in α.

Suppose $\tilde{V}^D$ is semi-differentiable. Semi-differentiability implies differentiability almost everywhere. Whenever it exists, the derivative of (3) with respect to α is given by

$$\tilde{V}'^D(\alpha)[(1-\tau) + \tau\alpha] + \tilde{V}^D(\alpha)\tau.$$

We have

$$\begin{aligned}
&\tilde{V}'^D(\alpha)[(1-\tau) + \tau\alpha] + \tilde{V}^D(\alpha)\tau \\
\leq& \sup_{\alpha\in\{\alpha:\tilde{V}'^D(\alpha)\text{ exists}\}} \tilde{V}'^D(\alpha)[(1-\tau) + \tau\times 1] + \tilde{V}^D(1)\tau.
\end{aligned}$$

The sup is finite by semi-differentiability. Thus, for either low enough $\tilde{V}^D(1)$ and $\tau > 0$ or large enough $\tau < 1$ and $\tilde{V}^D(1) < 0$, the derivative of (3) exists almost everywhere and is negative. As $\tilde{V}$ is semi-differentiable, it is continuous; thus, its derivative existing almost everywhere, and being negative implies that it is decreasing in α.

Suppose there exists a constant $C > 0$ such that for all α, α' with $\alpha' > \alpha$, we have

$$\frac{\tilde{V}^D(\alpha') - \tilde{V}^D(\alpha)}{\alpha' - \alpha} \geq C.$$

Take any α, α' with $\alpha' > \alpha$. We have

$$\begin{aligned}
&\tilde{V}^D(\alpha')[(1-\tau) + \tau\alpha'] - [\tilde{V}^D(\alpha)[(1-\tau) + \tau\alpha]] \\
=&[\tilde{V}^D(\alpha') - \tilde{V}^D(\alpha)](1-\tau) + \tilde{V}^D(\alpha)\tau(\alpha' - \alpha) + \alpha'[\tilde{V}^D(\alpha') - \tilde{V}^D(\alpha)]\tau\alpha' \\
>&[\tilde{V}^D(\alpha') - \tilde{V}^D(\alpha)](1-\tau) + \tilde{V}^D(\alpha)\tau(\alpha' - \alpha) \\
\geq&C(\alpha' - \alpha)(1-\tau) + \tilde{V}^D(0)\tau(\alpha' - \alpha).
\end{aligned}$$

Whether the last expression is nonnegative is independent of α, α' and as $\tilde{V}(0) < 0$ ensured when

$$\tau \leq \frac{C}{-\tilde{V}^D(0) + C}.$$

□

Proof of Proposition 4. It is obvious that the optimal level of attention to the decision is weakly increasing in $\underline{\alpha}$ and $\bar{\alpha}$. As

$$\frac{\partial^2}{\partial\alpha^2}V(\alpha) = \frac{\partial^2}{\partial\alpha^2}V^D(\alpha)[(1-\tau) + \alpha\tau] + 2\frac{\partial}{\partial\alpha}V^D(\alpha)\tau,$$

the condition in the proposition implies that the second derivative is strictly positive, i.e., the total payoff is strictly convex. Of course, a strictly convex function is maximized at extreme points. □

Proof of Lemma 2. For fixed α_1, an optimal $\alpha_2(\cdot)$ is found by maximizing the agent's objective pointwise as in Lemma 1. □

Proof of Lemma 3. Fix any α_1 and ϵ_1. Dropping terms independent of α_2, period-1 self maximizes

$$V_2^D(\alpha_2|\epsilon_1)[\alpha_1\tau_1 + \alpha_2\tau_2 + (1 - \tau_1 - \tau_2)] + V^O[(1-\alpha_2)\tau_2].$$

Period-2 self maximizes the same expression dropping the term involving $\alpha_1\tau_1$. As V_2^D is increasing in α_2, the result follows. $\square$

Proof of Proposition 5. Clearly, the agent can implement the commitment solution when it can be achieved in one period. We thus focus on the case in which $\bar{\alpha}_1 + \bar{\alpha}_2 > 1$.

We proceed as follows: we find V^O so that period-1 self is indifferent between $\alpha_1 = \bar{\alpha}_1$ and $\alpha_2 = \bar{\alpha}_2$ and no attention devoted in either period. (A small decrease in V^O will then lead period-1 self to strictly prefer the former to the latter.) We then derive conditions on v_1 so that i) period-1 self strictly disprefers devoting attention in one period only, and ii) that period-2 self does not devote attention if period-1 self does so that the unique outcome is no attention in either period. Those conditions will correspond to $v_1 \in (\underline{v}_1, \bar{v}_1(\tau))$. The fact that period-2 self strictly prefers the commitment outcome to the outcome without commitment follows from period-1 self's indifference and a comparison between period-1 and period-2 selves' objectives.

Note that if V^D is a two-step problem, then the environment is realization-independent. Algebra reveals that period-1 self (under commitment) strictly prefers $\alpha_1 = \bar{\alpha}_1$ and $\alpha_2 = \bar{\alpha}_2$ to no attention devoted in either period, i.e., $V(\bar{\alpha}_1, \bar{\alpha}_2) > V(0,0)$, if and only if

$$V^O < \underline{V}^O \equiv v_2 + \frac{1-2\tau}{\bar{\alpha}_1\tau + \bar{\alpha}_2\tau}(v_2 - v_0).$$

We next find conditions so that period-1 self prefers $\alpha_1 = \bar{\alpha}_1$ and $\alpha_2 = \bar{\alpha}_2$ to any other levels of attention. Note that for $V^O < \underline{V}^O$ large enough, the agent would never devote "excess attention," i.e., she would only choose $\alpha_1, \alpha_2 \in \{\bar{\alpha}_1, \bar{\alpha}_2\}^2$. As we will consider V^O close to $\underline{V}^O$, it thus suffice to find conditions so that $V(\bar{\alpha}_1, \bar{\alpha}_2) > V(\bar{\alpha}_1, 0) = V(0, \bar{\alpha}_1)$ (the inequality follows as $\tau_1 = \tau_2$). For $V^O = \underline{V}_O$, the above condition is equivalent to

$$v_1 < \bar{v}_1(\tau) \equiv v_0 \frac{\omega_1}{\omega_1 + \omega_2} + v_2 \frac{\omega_2}{\omega_1 + \omega_2}$$

with $\omega_1 = \bar{\alpha}_2 - 2\bar{\alpha}_2\tau$ and $\omega_2 = \bar{\alpha}_1^2\tau + \bar{\alpha}_1\bar{\alpha}_2\tau$. Algebra shows that $\frac{\partial}{\partial\tau}\frac{\omega_1}{\omega_1+\omega_2} > 0$. Furthermore, whenever c_1 satisfies the above condition, then we still have $V(\bar{\alpha}_1, \bar{\alpha}_2) > V(\bar{\alpha}_1, 0) = V(0, \bar{\alpha}_1)$ for $V^O < \underline{V}^O$ large enough.

We next show that period-2 self does not devote any attention. Suppose first the period-1 self devotes attention (i.e., if $\alpha_1 = \bar{\alpha}_1$). Then we need to find conditions so that $V_2(0|\bar{\alpha}_1) > V_2(\bar{\alpha}_2|\bar{\alpha}_1)$.

Again, algebra shows that this is satisfied for $V^0 = \underline{V}^O$ if

$$v_1 > v_0 \frac{\bar{\alpha}_2}{\bar{\alpha}_1 + \bar{\alpha}_2} + v_2 \frac{\bar{\alpha}_1}{\bar{\alpha}_1 + \bar{\alpha}_2}.$$

Furthermore, whenever v_1 satisfies the above condition, then we still have $V_2(0|\bar{\alpha}_1) > V_2(\bar{\alpha}_2|\bar{\alpha}_1)$ for any $V^O < \underline{V}^O$.

Now suppose that period-1 self does not devote any attention. By construction, for large enough $V^O < \underline{V}^O$, period-1 self strictly prefers no attention to $\alpha_1 = 0$ and $\alpha_2 = \bar{\alpha}_2$. Note that this also implies $V_2(0|0) > V_2(\bar{\alpha}_1|0)$.

Finally note that $V(\bar{\alpha}_1, \bar{\alpha}_2) > V(0,0)$ implies that period-2 self strictly prefers the outcome under full commitment to not devoting any attention. □

Proof of Proposition 6. For the first claim, let $F(1)$ be uniform on $\{\epsilon_l, \epsilon_h\}$ and define $v_l^D \equiv V_2^D(0|\epsilon_l)$ and $v_h^D \equiv V_2^D(0|\epsilon_h)$. By no instrumental value of attention, we have $\frac{v_l^D + v_h^D}{2} = V^O$. Consider α_2 given by $\alpha_2(\epsilon_l) = 0$ and $\alpha_2(\epsilon_h) = 1$. For high enough v_h^D (and low enough v_l^D), devoting full attention to the decision in period 1 and according to α_2 described above is uniquely optimal.

For the second claim, suppose the agent devotes attention α_1' in period 1 and α_2 is chosen optimally. For $V^0 > V_1(0,0)$ and V^O small enough, the variation in $V_2^D(0|\epsilon_l)$ leads to α_2 to be positive with a probability bounded away from 0. For such a case, the agent benefits from devoting attention in period 2. □

Proof of Proposition 7. Let u' and u'' denote the first and second derivatives of u, respectively. Consider the ratio of the derivatives of two payoffs (when devoting attention and when not) with respect to y:

$$\frac{u'(\frac{y}{2})}{u'(y - c_d)[1 - \tau]}.$$

We show that the above is monotonically increasing in y implying the result. We have

$$\frac{\partial}{\partial y} \left(\frac{u'(\frac{y}{2})}{u'(y - c_d)[1 - \tau]} \right) = \frac{\frac{1}{2}u''(\frac{y}{2})u'(y - c_d) - u'(\frac{y}{2})u''(y - c_d)}{u'(y - c_d)^2} \frac{1}{1 - \tau}$$

which is positive if and only if

$$-\frac{\frac{1}{2}u''(\frac{y}{2})}{u'(\frac{y}{2})} < -\frac{u''(y - c_d)}{u'(y - c_d)},$$

which holds as u has CARA.

When $2u(\frac{y}{2}) > V^O$, the agent devotes full attention so $\bar{y} < \infty$. Uniqueness follows from the continuity of the two expressions preceding the proposition.

We now show that $\underline{y}$ may be negative infinity. Recall that $y_{\inf}$ denotes the infimum of the support of y and $c_d \leq y_{\inf}$. A necessary and sufficient condition for $\underline{y} = -\infty$ given c_d is that

$$2u(\frac{y_{\inf}}{2}) + V^O[1-\tau] < (u(c_d) + u(y_{\inf} - c_d))[1-\tau] + V^O.$$

The right-hand side is minimized for $c_d \in \{y_{\inf}, 0\}$, so the agent never devotes attention to low income realizations if

$$2u(\frac{y_{\inf}}{2}) + V^O[1-\tau] < (u(0) + u(y_{\inf}))[1-\tau] + V^O,$$

i.e.,

$$2u(\frac{y_{\inf}}{2}) - (u(0) + u(y_{\inf})[1-\tau] < \tau V^O.$$

□

Proof of Lemma 4. The agent's objective is given by

$$\max_{c_1, \bar{y}} \int_{y=-\infty}^{\bar{y}} (1-\tau)(u(c_1)+u(y-c_1)))dF(y) + F(\bar{y})(1+\tau)V^O + \int_{y=\bar{y}}^{\infty} (1+\tau)2u(\frac{y}{2})dF(y) + (1-F(\bar{y}))(1-\tau)V^O.$$

Divide the objective by $(1-\tau)$.

Notice that c_1 is independent of τ and given any $\bar{y}$, c_1 is increasing in $\bar{y}$. To see the latter claim, note that given $\bar{y}$, c_1 uniquely solves

$$\int_{y=-\infty}^{\bar{y}} u'(c_1) - u'(y-c_1)dF(y) = 0,$$

and that increasing $\bar{y}$ while holding c_1 constant makes the above positive, hence requiring c_1 to increase.

To finish the proof, it thus suffices to show increasing differences in the outer maximization, i.e., between $\bar{y}$ and τ. Note that it is never optimal to choose $\bar{y}$ greater than y' solving $V^O = \frac{1}{2}u(\frac{\bar{y}}{2})$, as the agent prefers to devote attention for such values of y. It is thus without loss to restrict the domain of $\bar{y}$. The cross-partial derivative of τ and $\bar{y}$ is given by

$$f(\bar{y}\frac{2}{(1-\tau)^2})V^O - f(\bar{y})\frac{2}{(1-\tau)^2}2u(\frac{\bar{y}}{2})$$

and positive for all $\bar{y}$ and τ completing the proof. $\square$

Proof of Proposition 8. The agent's objective can be written as

$$E[u(w+wfr)] + \tau \max\{E[u(w+wfr)], V^O\}.$$

Evidently, the above is supermodular in (f, τ) and so the proposition follows by Topkis's Monotonicity Theorem. $\square$

Proof of Proposition 9. When the market is up, both the standard agent and the emotionally inattentive agent choose f as to maximize $E[u(w+wfr)]$, i.e., independent of τ. The standard agent also chooses f to maximize this objective in the first period. The emotionally inattentive agent, however, maximizes

$$E[u(w+wfr)] + p_{\text{risk}}[(1-\rho)E[u(w+wfr) + \rho u(w(1-f))],$$

which implies that her initial investment in the risky asset is less than that of the standard agent. This is also her investment in the risky asset when the market is down. Thus, as the standard agent chooses f to maximize

$$(1-\rho)E[u(w+wfr) + \rho u(w(1-f),$$

the final claim follows. $\square$

B Examples of decision problems

B.1 Information acquisition

The agent takes an action $a \in A$, whose payoff depends on a random state of the world ω of which she has a (possibly diffuse) prior $\mu_0 \in \Delta(\Omega)$. The agent's payoff is

$$u(a|\omega).$$

Given any belief μ about the state of the world ω, the agent chooses an action to maximize her expected payoff from the decision. Her payoff is then

$$U(\mu) = \max_{a \in A} E_{\omega \sim \mu}[u(a, \omega)]$$

The agent can acquire information through devoting attention. Let $\Delta(\Delta(\Omega))$ denote all probability distributions over beliefs about the state of the world and define $B_{\mu_0} = \{\tau : \tau \in \Delta(\Delta(\Omega)), E_{\mu \sim \tau}[\mu] = \mu_0\}$ denote all probability distributions over beliefs about the state of the world whose mean equals the agent's prior. For $\mu, \mu' \in B_{\mu_0}$, we say $\mu \geq \mu'$ if μ is a mean-preserving spread of μ'.

The agent acquires information according to $\pi : [0, 1] \to B_{\mu_0}$, where π is monotone, that is if $\alpha \geq \alpha'$, then

$$\pi(\alpha) \geq \pi(\alpha'),$$

and right-continuous.

When the agent devotes attention α to the decision, her material value from the decision is then given by

$$V^D(\alpha) = E_{\mu \sim \pi(\alpha)}[U(\mu)].$$

Notice that $V^D(\alpha)$ is indeed increasing in α, as U is convex, and right-continuous, as π is.

We provide an example.

Normally distributed state. The state is a real number and normally distributed, i.e., $\omega \in \Omega$ and $\omega \sim \mathcal{N}(0, 1/\beta)$. β is the precision of the prior; when $\beta = 0$, we take the agent's prior to be diffuse. The agent chooses an action on the real line, to match the state, $a \in A = \mathbb{R}$. In particular, we have

$$u(a|\omega) = -\gamma(a - \omega)^2 + H^D$$

with $\gamma \geq 0$ and H^D as the "height" of the decision. When the agent devotes attention α to the decision, she acquires a signal that is normally distributed around the state of the world and has precision $\psi\alpha$. Her posterior will then also be normal; let s denote the signal, her posterior is given by $\mu = \mathcal{N}(\frac{\psi\alpha}{\beta+\psi\alpha}s, \frac{1}{\beta+\kappa\alpha}$ giving payoff

$$U(\mu) = -\gamma \frac{1}{\beta + \psi\alpha} + H^D.$$

(For completeness, $\pi(\alpha)$ is a distribution over normal distributions $\mathcal{N}(x, \frac{1}{\beta+\psi\alpha})$ where $x \sim \mathcal{N}(0, \frac{1}{\beta+\psi\alpha}.)$

We thus have

$$V^D(\alpha) = -\gamma \frac{1}{\beta + \psi\alpha} + H^D.$$

B.2 Increasing available choices

The agent takes an action $a \in A$ leading to payoff

$$u(a).$$

The available actions, A, are random. Let $\mathcal{A}$ be a set. Let P_α denote the probability measure over subsets, denoted by $2^{\mathcal{A}}$, of $\mathcal{A}$. Subsets are ordered by the strong set order giving a partial order. Let $F(\cdot|\alpha)$ be the cumulative distribution function over $\mathcal{P}(\mathcal{A})$ associated with this partial order, i.e.,

$$F(A|\alpha) = P_\alpha(\{A' \in \mathcal{P}(\mathcal{A}) : A' \leq A\}).$$

Through devoting attention to the decision, the agent increases her available actions: for any α, α' with $\alpha' > \alpha$, $F(\cdot|\alpha')$ first-order stochastically dominates $F(\cdot|\alpha)$. The agent can always make a choice so for all α, $F(\emptyset|\alpha) = 0$.

Her material value is given by

$$V^D(\alpha) = E_{A \sim F(\cdot|\alpha)}[\max_{a \in A} u(a)].$$

We mention two examples.

Attending to the decision. The agent needs to devote a minimum amount of attention $\underline{\alpha} \in (0, 1]$ to the decision in order to change a default choice. When she does, she chooses from the whole choice set. Formally, let $a_0 \in \mathcal{A}$ denote the default option, i.e., for $\alpha < 1\underline{\alpha}$, $P_\alpha(\{a_0\}) = 1$. For $\alpha \geq \underline{\alpha}$, $P_\alpha(\mathcal{A}) = 1$.

The material value of the decision is then a step function:

$$V^D(\alpha) = \begin{cases} u(a_0) & \text{for } \alpha < \underline{\alpha} \\ \max_{a \in \mathcal{A}} u(a) & \text{for } \alpha \geq \underline{\alpha}. \end{cases}$$

Choosing from memory. The agent's choice set consists of some actions that are always available, $A_0 \neq \emptyset$, and whatever actions the agent can remember, A_1. In particular, let $\mathcal{A}_1 = \mathcal{A} \setminus A_0$ be the agent's memories. When devoting attention α, the agent makes $d = \lfloor \alpha\kappa \rfloor$ draws with replacement from $\mathcal{A}_1$, the memories that come to mind. She can then choose from the actions that are always available or from the ones associated with the memories she draws (recalls). Let $\mathcal{A}_1$ be finite with cardinality M; let A_1 have cardinality m. If $d < m$, $P_\alpha(A_1) = 0$. Otherwise,

$$P_\alpha(A_1) = \frac{d \times \cdots \times (d-(m-1)) \times (d-m)^m m!}{M^d}.$$

Note that as $\lfloor \alpha\kappa \rfloor$ is right-continuous, so is P_α and hence V^D.