Optimal Preparedness and Uncertainty Persistence*

Savitar Sundaresan†

November 20, 2018

Abstract

Unusual events trigger spikes in uncertainty that are persistent. Standard models cannot match such dynamic patterns of uncertainty. This paper presents a unified framework, motivated by the cognitive and economics literature on attention. Agents *prepare* for different possible states of the world by collecting information that would prove valuable should those states occur. Agents choose not to prepare for low probability events, so that the occurrence of a rare does not resolve uncertainty, but increases it. Uncertain agents have dispersed beliefs, and find it harder focus their preparation. For an endogenously inattentive agent, uncertainty begets uncertainty, causing endogenous uncertainty persistence, even when there is no persistence in the events themselves.

*Thanks to Ricardo Reis, Michael Woodford, Larry Glosten, and Jose Scheinkman for invaluable guidance, assistance and advice; to Nick Bloom, Patrick Bolton, Pedro Bordalo, Andrew Caplin, Chris Hansman, Rajkamal Iyer, Marcin Kacperczyk, Filip Matejka, Kristoffer Nimark, Elise Payzan-Lenestour, Joel Peress, Ansgar Walther, and Laura Veldkamp for discussions and comments that helped shape the content of this work; to seminar and conference participants at SLOAN-NOMIS, SITE, Columbia, Emory, UC Davis, Northwestern Kellogg, London Business School, Johns Hopkins Carey, the Federal Reserve Board, SFS Cavalcade and Alex Chinco (discussant) CEPR and David Schreindorfer (discussant). All errors are my own.

†Business School, Imperial College London; Email: s.sundaresan@imperial.ac.uk

1 Introduction

Rare or unexpected events repeatedly cause measures of political and financial uncertainty to spike up, *and stay persistently high* for weeks or even months.[1] For example, the VIX increased after the terrorist attacks of 9/11 and the collapse of Lehman Brothers, and policy uncertainty jumped after Brexit and Trump's election in 2016, as is shown in Figure 1. Uncertainty can dictate how economic agents make consumption, investment, pricing, and portfolio allocation decisions, so understanding the source and nature of uncertainty persistence is crucial in both macroeconomics and finance.

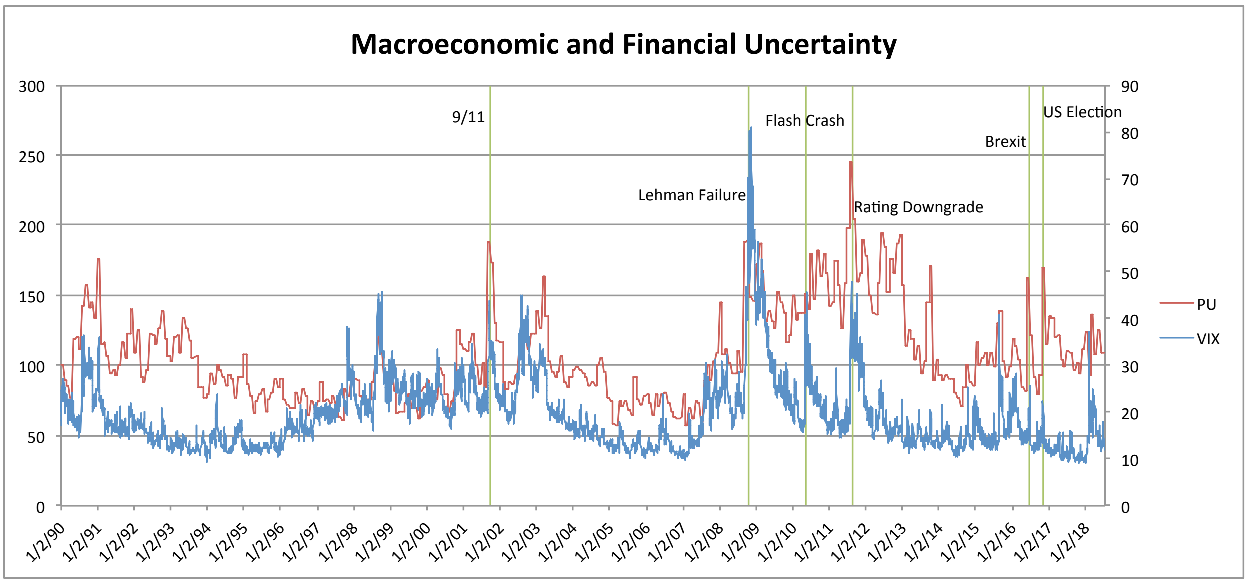


Figure 1: Plotted are monthly values of the Policy Uncertainty Index (PU) and daily values of the CBOE's Volatility Index (VIX) for the last 28 years. The VIX is a measure of expected volatility in the S&P 500 Index as calculated from options prices. The PU is a measure of macroeconomic uncertainty as calculated from news articles, policy expirations, and analyst disagreement.

Standard models cannot generate these patterns in uncertainty dynamics. The workhorse model of uncertainty dynamics is *parameter learning*,[2] where agents do not know the value of a parameter of a particular process. They can try to learn the value by observing the process, and the more data they collect, the better their estimate. If a very unusual draw is observed, they change their estimate sharply, and require many more observations for their estimates to stabilize, causing the change in their estimate to be persistent. However, the data show

1. As a rough measure, the VIX and PU exhibit autocorrelations of about 98% and 81%, respectively.
2. See Orlik and Veldkamp (2014), Kozlowski, Veldkamp, and Venkateswaran (2016), Nimark (2014) for persistence in a macroeconomic setting.

that uncertainty spikes up repeatedly, and that when it jumps, it jumps up and not down. Parameter-learning rational-expectations models need a lot of assumptions on the data to generate these and other patterns. This paper, as part of a recent literature on dynamic inattention,[3] builds on the general framework of rational inattention to offer a theory that can explain these patterns parsimoniously. In applying the model to a financial market, the theory is able to match patterns in bid-ask spreads, expected volatility, dispersion of beliefs, and volume of trade.[4]

The Mechanism

In this model, there is an underlying stationary process. Agents can choose to exercise attention to collect information about some potential states of the world one period in advance. Given that information collection is costly, they cannot attend to all future states. Which states will they choose to learn about? Agents will devote their resources towards states of the world that are, risk-weighted, more likely to occur, and will ignore states of the world that are, risk-weighted, less likely to occur. To fix ideas, consider the following illustration: prior to the 2016 US presidential election, investors could have prepared for how markets would react to either a Clinton presidency or a Trump presidency. Most forecasts predicted a Clinton win, so most investors would have devoted their resources to preparing for that outcome. When Trump won, Clinton-based preparation would have been useless, and unprepared investors would have been uncertain of where to invest. This uncertainty would make it difficult for investors to figure out what policies would be pushed forward next, how international relations would change, etc. Due to the dispersed set of possible future states, investors *wouldn't know what to prepare for next.* Therefore, *no matter what event followed*, our inability to sufficiently prepare would lead to high levels of uncertainty again - uncertainty would persist endogenously.

The mechanism of preparation is motivated by the experimental literature. Cognitive experiments have shown that allocating attention in a preparatory manner is a part of

3. See, for example: Steiner, Stewart, and Matějka (2017), Maćkowiak, Matějka, and Wiederholt (2018), Nimark and Sundaresan (2018), Ilut and Valchev (2017)

4. Other papers that use dynamic information structure to think about financial variables include Bolton and Faure-Grimaud (2009),Banerjee and Breon-Drish (2017), and Banerjee and Kremer (2010)

our attentional makeup. In a seminal experiment by Shaw and Shaw (1977), agents were presented with visual stimuli. The subjects were told how likely the stimuli were to occur in different parts of their visual field. When stimuli were then shown in parts they were told were likely, subjects were able to identify them accurately. When stimuli were shown in parts they were told were unlikely, subjects could not identify them accurately. The authors concluded that agents are able to arrange their peripheral attention to prepare for different events, and that they prepare more for likely events than for unlikely events. To deliver this finding theoretically, I make one key assumption about the preparation process: *the marginal cost for preparation is the same for every state of the world.*[5] Such an assumption is not really necessary, and the results of the paper are robust to may intuitive alternatives. The cost structure that would weaken or even negate the findings of the paper is one where it is *harder* to prepare for common events than it is to prepare for rare events. Such a structure is not intuitive.[6]

The Results

Using preparation as the key mechanism to explain uncertainty dynamics, this paper's model delivers two results linked by one crucial mechanism. The first result matches that delivered by the experiment described above: *agents prepare more for events they deem likely and less for events they deem unlikely.*[7] This result is also well established in the experimental literature,[8]. When an anticipated event occurs, uncertainty is resolved and the precision of agents' beliefs increases; whereas when an unexpected event occurs the precision of agents' beliefs decreases[9].

The key of this paper is the mechanism that links the previous result to the following one, showing important dynamic implications. Namely, *agents' incentives to prepare change*

5. This assumption is a reduced form simplification of the channel capacity concept introduced to economics by Woodford (2012), which was motivated in part by the experimental evidence of Shaw and Shaw (1977). The results of this paper are robust to many intuitive alternatives.

6. It is however, as shown by Woodford (2012) a relatively undesirable property of a constraint on mutual information.

7. This result was shown first by Maćkowiak and Wiederholt (2018) in a static model.

8. See Shaw and Shaw (1977), Posner, Snyder, and Davidson (1980), Eriksen and James (1986)

9. See Tenenbaum et al. (2011), Gershman, Horvitz, and Tenenbaum (2015), Posner, Snyder, and Solso (2004), and Peters, McEwen, and Friston (2017) for the neuroscientific on preparation and uncertainty.

when the agents are uncertain. By the first result, agents choose not to prepare for low probability events, so the occurrence of low probability events causes agents to be uncertain. Being uncertain means that agents' beliefs over future states of the world are more dispersed, and *when beliefs are more dispersed, agents preparation is more dispersed*. To understand this intuitively, consider two different cases. In the first case, agents have strong, certain beliefs, with a very small variance. In such a case, agents would be able to prepare for the few events that are very likely according to their beliefs, and essentially ignore the rest. In the second case, agents have weak, uncertain beliefs, with a very large variance. In such a case, agents won't be able to focus her preparation on a few events, because no events are sufficiently likely in their beliefs to warrant such a focus. Agents will instead try to prepare a little (if at all) about a host of different possible states.

The second result of this paper is to show that *if agents are uncertain today, they will continue to be uncertain tomorrow, because they cannot prepare well enough for any event*. Putting the first result together with the mechanism, one unexpected or rare event will trigger a high level of uncertainty because agents will not have prepared for it. High uncertainty makes it difficult for agents to prepare because preparation is costly, and their beliefs are too diffuse to focus their preparations. Conversely, an expected or likely event results in a low level of uncertainty because agents will have prepared for it. Low uncertainty makes it easy for agents to prepare because their beliefs are tight, so agents can focus their preparations on the few, likely states in their beliefs. Therefore, uncertainty (and certainty) persist endogenously. This result differs from standard rational inattention models, which place constraints or costs on mutual information. In those models, costs are proportional to the probabilities of events, so ex-post uncertainty is identical in every state - there could, in fact, be no uncertainty dynamics.

This mechanism is supported by the fact that *economic and financial agents do, in fact, prepare for various events*. Central banks have instructed financial institutions to undergo stress-testing in anticipation of another financial crisis. Martin and Pindyck (2015) considers the mitigating properties of preparation for rare events. Such preparation is, at least in part, useful to mitigate the uncertainty that would stem from a financial crisis, by having a plan in place to deal with some of its effects. Additionally, this model supports the intuition that *agents invest more in preparation when they are uncertain.* Uncertain agents are sometimes modeled as wanting information more, or scrambling for information. As I show in one of the final results of the paper, despite the fact that high uncertainty leads to worse preparation *on average*, it also leads to more spending on preparation *in total* due to the wealth of possible states that are likely enough to warrant preparation.

Other Applications

The mechanism in this paper is quite general, and can be applied to other settings. Dispersion, and persistence in dispersion have been noted in many economic variables. Consider the following examples:

The mechanism can be used to think about the persistence and countercyclicality of *price dispersion* or *wage dispersion*. If firms prepare more for good times than for bad, then according to this paper's mechanism, firms will receive noisy signals on how to set prices and wages in bad times. Therefore, price dispersion and wage dispersion will be higher in bad times. By the mechanism of the paper, the dispersion will also be persistent.

Persistence in performance could also be explained with this mechanism. Firms or individual who perform well in certain states of the world in one period would be better positioned to prepare for events in the next period. Therefore, they would be more likely to face low levels on uncertainty subsequently, which would improve their performance as well. Firms that perform poorly would be unsure of how to proceed, and would continue to perform poorly.

Understanding uncertainty dynamics is crucial for financial and economic analysis.[10] Uncertainty Because risk-aversion is a fundamental assumption in economics, reducing uncertainty provides first-order welfare benefits. Therefore, finding the mechanism that allows uncertainty to persist and propagate facilitates policies that reduce uncertainty and temper its effects. This paper addresses this inquiry by showing that tools from the cognitive and economics literature on attention yield a mechanism that causes uncertainty to spike and remain persistently high after rare events.

10. See the literature spawned by Bloom (2009).

Literature Review

Works such as Sims (2003) and Woodford (2012) argue that even if information is plentiful, attention is a scarce resource. This notion is borne out by the experimental literature mentioned above, and has been explored in numerous fruitful settings, where a cost is typically placed on the reduction in entropy agents obtain from receiving noisy information. See, for example Matějka (2015a), Maćkowiak and Wiederholt (2009), Matějka (2015b), Matejka and McKay (2015).

Bolton and Faure-Grimaud (2009) uses a dynamic information model to show that agents might not fully work through contingencies up front, instead choosing to prioritize their costly information collection, and postponing deliberations on low-probability or low-risk events. The focus of that paper lies less with dynamic spillovers and persistence of beliefs, and more with the tradeoff between reacting quickly to an event and thinking through future implications carefully.

There is also a large literature that uses informational frictions to analyze the macroeconomic impacts of changes in beliefs. Such papers typically use a parameter-learning setting. Orlik and Veldkamp (2014) and Kozlowski, Veldkamp, and Venkateswaran (2016) study the impact of changes in tail beliefs. Nimark (2014) looks at the impact of extraordinary news events, and generates persistence in volatility by showing that agents are more sensitive to information after such events.

Because this paper emphasizes the importance of attention across many possible states, there is a natural parallel to the concept of salience. Salience in economics and finance was introduced by a sequence of papers: Bordalo, Gennaioli, and Shleifer (2012), Bordalo, Gennaioli, and Shleifer (2013a), Bordalo, Gennaioli, and Shleifer (2013b), and shows that agents will choose to focus their attention on a subset of possible outcomes based on what they deem salient. While this paper does not consider alternative reasons for why some states might be more attractive to agents than others, the notion that different states carry differing levels of importance, and therefore will command differing levels of attention is crucial for this paper's results.

This paper is part of a larger agenda in the inattention literature, that is interested in exploring the *dynamics* of inattention. More specifically, this paper is placed in a sequence of papers that analyze how attending to the world *today* impacts one's ability and relative willingness to attend to the world later. Steiner, Stewart, and Matějka (2017) study sluggish responses to a slow-moving state, Maćkowiak, Matějka, and Wiederholt (2018) propose analytical methods to study dynamic attention problems. Nimark and Sundaresan (2018) show that ex-ante identical agents can diverge to opposite ends of a belief spectrum due to rationally confirmatory and complacent behavior. Ilut and Valchev (2017) propose a dynamic framework where agents can pay attention to an entire policy *function*. In the finance literature, papers such as Banerjee and Breon-Drish (2017) and Banerjee and Kremer (2010) consider dynamic implications of asymmetric information, and disagreement on financial variables.

The rest of the paper is as follows: Sections 2 and 3 present the model in binary and continuous settings. Section 4 applies the continuous form to an economy, deriving implications for uncertainty, volatility, dispersion of beliefs, and trade. Section 5 concludes.

2 Simple Model

This section presents the simplest possible model in which the key mechanism of the paper can be described. This model has one agent, two possible states of nature, and an infinite number of periods. Using costly, state contingent preparation via information acquisition, the model delivers persistent spikes in uncertainty in response to unlikely events.

2.1 Model Structure

State Structure: There are two possible states s of the world each period, A and B. To fix intuition, suppose that one state of the world, A, is 'the market increases in value' while the other, B, is 'the market decreases in value'. The probability of the state taking the value A each period, unconditional on any information, is $\pi = P(s_{t+1} = A)$. For simplicity, I assume that $\pi = 1 - \pi = \frac{1}{2}$. Relaxing this assumption would produce quantitative but not qualitative changes to the results.

Agent: There is a single agent. The agent enters period t with an information set I_t that informs her beliefs. Her beliefs in period t about period $t+1$'s state are given by $p_t = \max(P(s_{t+1} = A|I_t), P(s_{t+1} = B|I_t) \geq \frac{1}{2}$. Note that p_t is *not* the probability of the market going up in value, but is a value indicating the likelihood of the (weakly) more likely state occurring. The likely state could either be the market increasing or decreasing. But one state has to be at least as likely as the other, so p_t refers to the probability of the more likely event. p_t can therefore be viewed as measuring the *strength* of the agent's beliefs.

I will index the more likely state for period $t+1$ with L in period t. I will index the less likely (or rare) state for period $t+1$ with R in period t. Put differently, $p_t = P(s_{t+1} = L) = 1 - P(s_{t+1} = R)$.

Preparation Conditional on her beliefs p_t, the agent can choose in period t to *prepare* for each of the two possible states that could occur in period $t+1$. That is, the agent can *prepare* for the market either increasing or decreasing in value in the next period, by collecting information about what will happen conditional on each of those two possibilities. Preparation takes the form of picking information sets $I_{t+1,L}$ and $I_{t+1,R}$. When state L or state R occurs in period $t+1$, the agents beliefs will be:

$$\begin{aligned} p_L &= \max(P(s_{t+2} = A|s_{t+1} = L, I_{t+1,L}), P(s_{t+2} = B|s_{t+1} = L, I_{t+1,L})) \\ p_R &= \max(P(s_{t+2} = A|s_{t+1} = R, I_{t+1,R}), P(s_{t+2} = B|s_{t+1} = R, I_{t+1,R})) \end{aligned}$$

When one of the possibilities occurs, the agent's preparation for that state comes to fruition. Information forms her beliefs, conditional on the state. For simplicity, I model this decision as the agent directly selecting two possible values for p_{t+1} in period t: $p_L \geq \frac{1}{2}$ is the belief distribution she will have if the *likely* state occurs in $t+1$, and $p_R \geq \frac{1}{2}$ is the belief distribution she will have if the *rare* state occurs in $t+1$. In a binary distribution, the probability parameter p comoves with the precision of the distribution $p(1-p)$, so long as $p \geq \frac{1}{2}$, as has been assumed. Therefore, for the simplicity of the proofs, I will have the agent directly select p_L and p_R, which is isomorphic to the agent picking $p_L(1-p_L)$ and $p_R(1-p_R)$ or $I_{t+1,L}$ and $I_{t+1,R}$.

The intuition behind this structure is as follows. An investor might have beliefs about whether the market will go up or down tomorrow. And she might have already placed trades to take advantage of those beliefs. But she can also prepare for what comes next. If the market goes up tomorrow, will it go up again the next day? If the market goes down tomorrow, will it rebound afterwards? Collecting information about those possibilities will allow the agent to react to each quickly and profitably. Therefore, she would want to spend resources to be better prepared. In the appendix I consider an extension where agents can also collect in formation in period t about the events of period $t+1$ and show that all of the results continue to hold.

Agent's Costs: Improving the quality of preparation is costly, and the agent can choose not to prepare at all. There is no fixed cost of preparing, but there is a cost function $c(p_{t+1})$ associated with increased accuracy of conditional beliefs. That cost function is the *same* for the likely state and the rare state.

Agent's Objective: The agent has a period-by-period utility function, which has one input: the strength of the agent's beliefs. As discussed above, Under a binary distribution, as in this setup, the value p_t comoves with the precision of the distribution $p_t(1-p_t)$, so long as $p_t \geq \frac{1}{2}$, as has been assumed. Therefore, for simplicity, I assume that the agent's *utility function* is given by $U(p)$. The agent benefits from having precise beliefs. Such a benefit is a common feature of any utility function of a risk-averse agent.

Note that the utility function assumes the agent is indifferent as to whether one state or another occurs, and cares only about the accuracy of her predictions going forward. The state itself is not an input into the utility function. Intuitionally, the agent does not care if the market rises or falls, and can trade beneficially long or short as long as her beliefs are precise. Relaxing this assumption would have quantitative but not qualitative effects on the results. One way to interpret the paper to allow for state-dependent payoffs is to think of all the probabilities describe here as *risk-neutral* or *risk-weighted* probabilities.

Graphical Interpretation: Figure 2 shows the setup of the model. An agent in period t

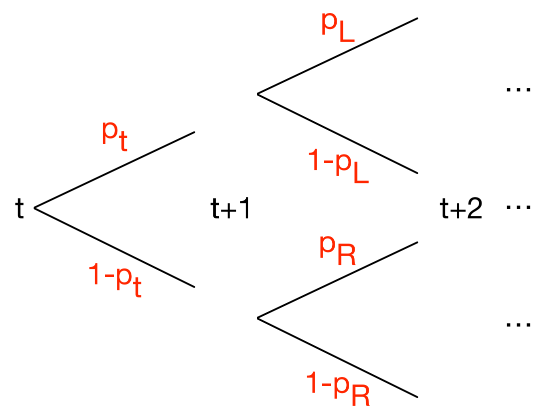


Figure 2: Plotted is a sample of three time periods in the binary state version of the model. Agents take their beliefs in time t, $p_t, 1-p_t$, about the relative likelihood of the two possible states in period $t+1$ as given, and can choose p_L and p_R, subject to a cost. Those choices will form their beliefs in period $t+1$, conditional on one of the two possible states occurring, about the relative likelihood of the two possible states in period $t+2$.

takes her beliefs about the states in period $t+1$, p_t, as given. Then, using costly information acquisition, the agent chooses her conditional beliefs for period $t+1$ about period $t+2$, p_L and p_R. Now that the structure of the economy has been described, I can present the agent's problem.

2.2 Statement of Problem

Formally, the agent's problem in period t is given by the following Bellman equation:

$$V(p_t) = \max_{p_L, p_R} U(p_t) - c(p_L) - c(p_R) + \beta p_t V(p_L) + \beta(1-p_t)V(p_R) + \mu_R \left(p_R - \frac{1}{2}\right) + \mu_L \left(p_L - \frac{1}{2}\right) \tag{1}$$

The agent begins period t with a certain set of beliefs p_t. The value associated with a particular set of beliefs can be broken up into several components. First, there is the *utility function*, U, a real-valued function on $\left[\frac{1}{2}, 1\right]$, that is increasing in the strength of the agent's beliefs. As discussed above, because $p_t \geq \frac{1}{2}$ comoves negatively with the variance of the distribution, having U be an increasing function of p_t is a common feature of almost any standard utility function. Then there is the *cost function*, c. The agent pays a cost to improve the strength of her future beliefs. That cost is additively separable across states, and is the same regardless of the likelihood for the state being prepared for. These assumptions are strong for tractability, but not essential, as will be discussed further below. The agent discounts the future at a rate $\beta < 1$. With probability p_t, the *likely state* will occur, and the agent's conditional beliefs will be p_L. With probability $1 - p_t$, the *rare state* will occur, and the agent's conditional beliefs will be p_R. The agent's choices are subject to inequality constraints, without loss of generality, that her subjective beliefs be weakly stronger than $\frac{1}{2}$. The Kuhn-Tucker conditions associated with those constraints are represented by the last two terms.[11] Intuitively, an agent values her beliefs based on two things: (i) how uncertain those beliefs make her, today, and (ii) the relative difficulty those beliefs cause her in preparing for the next period's events.

Deviation from the standard: This Bellman formulation differs in one crucial way from a standard problem. The state variable today has *two* purposes. The first, standard purpose is that it enters the utility function U. The second purpose, which is the innovation of this paper (as well as of Nimark and Sundaresan (2018), albeit under different circumstances) is that the state variable affects the *distribution of the continuation value*. My information set today impacts how I want to collect information. Therefore, my motives in collecting information today are twofold: first to maximize my utility next period, and second to change *how I want to collect information in the next period*. This second channel emphasizes the dynamic spillover of attention.

2.3 Solution

Assumptions: To guarantee a solution, I need to impose some structure on the functions mentioned above. Namely I assume that:

1. $U'(\cdot) > 0$. This assumption states that the agent prefers stronger beliefs to weaker ones.

2. $c'(\cdot) > 0$. This assumption states that stronger beliefs are more costly to obtain than weaker ones.

11. A proof that the Bellman equation 1 is a contraction is in the appendix. The fact that the Bellman is a contraction mapping implies that there is a unique solution to the optimal choice that can be found by iteration.

3. $c''(\cdot) > U''(\cdot)$. This assumption states that the marginal cost of stronger beliefs increases faster than the marginal benefit - a technical assumption to guarantee an interior solution.

4. $c'(1) > U'(1)$. This assumption states that certainty is too costly to ever be completely attained.

These assumptions are relatively standard, in that they are satisfied by most existing utility and costly information forms, and are sufficient to guarantee that interior solutions $p_L(p_t)$ and $p_R(p_t)$ exist.

2.4 Results

The first result is that the agent will choose to collect *more* information about the *likely* state, and *less* information about the *rare* state.

Proposition 1. *$p'_R(p_t) \leq 0$ and $p'_L(p_t) \geq 0$.*

Proof: See Appendix

The value of preparing for an event is proportional to its likelihood, but the cost is independent. Therefore, an agent will focus her preparation on the state of the world she believes likely to occur, as it has a good chance of proving useful. The converse of this result is that the agent will tend to prepare less for states of the world she believes less likely to occur, as such information would have a small chance of proving useful. Put more simply, agents will tend to pay less attention to rarer events. A version of this result for a static decision is shown in Maćkowiak and Wiederholt (2018).

The second result is that if the agent's beliefs are sufficiently strong, she will not collect *any* information about the *rare* state.

Lemma 1. *If $\frac{c'\left(\frac{1}{2}\right)}{U'\left(\frac{1}{2}\right)\beta} = 1 - p^* > 1 - p_t$, then $p_R = \frac{1}{2}$.*

Proof. It is straightforward to see that $p_R\left(\frac{1}{2}\right) = p_L\left(\frac{1}{2}\right)$. Then the first-order conditions with respect to the choice variables can be written as:

$$\frac{c'\left(\frac{1}{2}\right)}{\beta(1-p)} = U'\left(\frac{1}{2}\right) + \mu_R$$

The value p^* for which $\mu_R = 0$, is the point past which the agent stops collecting information. For all values of p larger than p^*, p_R is constrained at $\frac{1}{2}$. Therefore,

$$\frac{c'\left(\frac{1}{2}\right)}{\beta(1-p)} \geq U'\left(\frac{1}{2}\right)$$

□

If the marginal benefit of collecting information about the rare state is lower than the marginal cost of collecting information, *when the agent hasn't collected any information yet*, the agent won't bother collecting any information about the rare state. Because $U'(\cdot)$ and $c'(\cdot)$ are positive everywhere, there is always a set of beliefs about the likely state that are strong enough such that the rare state will be ignored. Put slightly differently, an agent doesn't prepare for sufficiently unlikely scenarios. This is obviously true in reality - for example, most people don't prepare for a meteor hitting their home. Such an event is sufficiently unlikely that preparing for it would be a waste or resources.

Agents completely ignore *sufficiently* rare events. An dynamic implication of this result comes from noticing that the term *sufficiently* is relative: The third result shows that under certain conditions, the agent will not collect any information about either state when her beliefs are the unconditional distribution.

Corollary 1. *(High uncertainty can be permanent). If* $\frac{c'\left(\frac{1}{2}\right)}{U'\left(\frac{1}{2}\right)\beta} \geq \frac{1}{2}$ *then* $p_s = \pi_s \ \forall s > t$.

Proof. Follows immediately from Lemma 1. □

If the agent doesn't collect information, her beliefs will be the unconditional probability distribution. But if the unconditional probabilities are *sufficiently* unlikely according to the above definition, that problem becomes permanent. If they don't collect information under the unconditional probability distribution, they will then again face the unconditional probability distribution in the subsequent period. Therefore, if the agent, for any reason, chooses not to acquire information for a particular state, and then that state occurs, the agent will opt *never to collect information again.* Such a condition could be satisfied if the marginal benefit of information is quite low at the unconditional distribution, or if the marginal cost is quite high.

But the corollary is stronger than that: it also implies that if the conditions are met, a uncertainty trap is *inevitable*. If the agent doesn't collect information for the unconditional probability distribution, the agent won't collect information for any state less likely than the unconditional 50-50 distribution, by Proposition 1. But $1 - p_t$, the probability of the rare state is weakly less than $\frac{1}{2}$. Therefore, if the above conditions hold, the agent will not collect information for the rare state, *ever*. The rare state must occur *at some point*, and when it does, the agent will *permanently* cease information collection.

The final result of this section shows that under certain conditions, there is a 'steady state' level of uncertainty to which the agent will converge.

Proposition 2. *(Uncertainty is persistent) If* $\frac{c'\left(\frac{1}{2}\right)}{U'\left(\frac{1}{2}\right)\beta} < \frac{1}{2}$, *then there is a 'convergent' fixed point* $\underline{p}$ *where* $\underline{p} = p_L = p_t$.

Proof. Proposition (2) Suppose $p^* > \frac{1}{2}$. We know that $p_R'(p_t) < 0$. Therefore $p_R\left(\frac{1}{2}\right) = p_L\left(\frac{1}{2}\right) > \frac{1}{2}$. We also know that $p_L'(p_t) > 0$ and $p_L(1) < 1$ due to the limit conditions on c and U. Therefore, there must exist a point $\tilde{p}$ such that $p_L(\tilde{p}) = \tilde{p}$. □

There is a level of uncertainty at which the agent's beliefs will remain *conditional on the likely state continuing to occur*. If the rare state occurs, uncertainty will spike up, and will only decline as the agent collects information again.

The best way to see the implications of these result is graphically. Consider Figure 3. On the x-axis of this figure is the agent's beliefs today. The right-hand side of the axis is p_t, while the left-hand side of the axis is $1 - p_t$. The kinked, curved line is the *policy function* - the agent's information collection as a function of her beliefs today. The left-hand side of the curve (the part that corresponds to the left-hand side of the x-axis) is p_R, and the right-hand side of the curve (the part that corresponds to the right-hand side of the x-axis) is p_L. The straight, solid line, is a *45-degree line* - the set of points where beliefs in the next period have the same distribution as beliefs today. There is an intersection between the policy function and the 45-degree line at the green dot. Call the point of intersection $\tilde{p}$. If $p_t = \tilde{p}$, then $p_L(\tilde{p}) = \tilde{p}$, and as long as the likely state occurs, the agent's beliefs will always be $\tilde{p}$. With beliefs of $\tilde{p}$, the agent's choices of information collection will be at the green dot

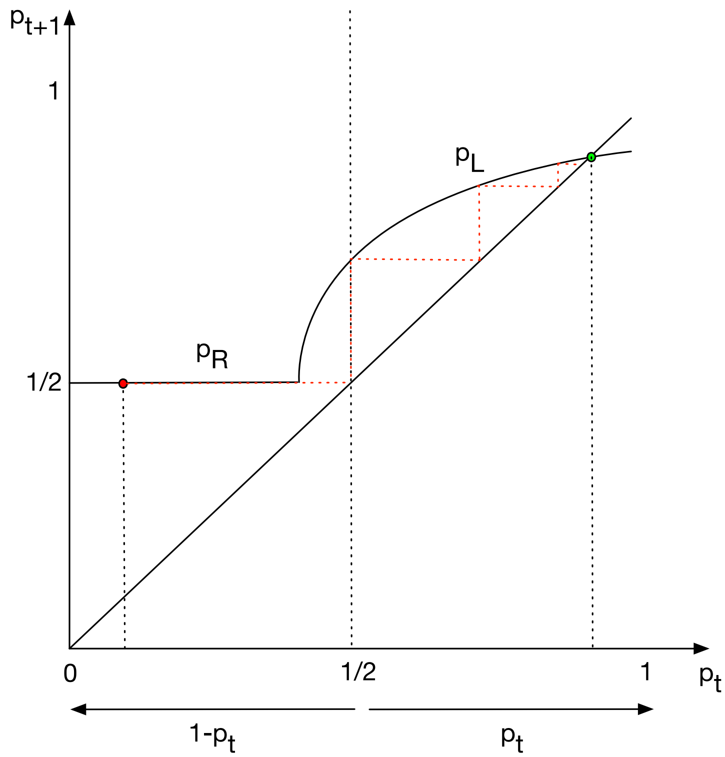


Figure 3: Plotted is the policy function of the strength of beliefs in period $t+1$ as a function of the probability of a particular event occurring in period t. The kinked-curved line is the policy function, while the straight line is a 45-degree line. The red dashed line shows a possible evolution of beliefs between the red-rare event level of precision and the green-steady state level of precision.

and the red dot. The reason there is a kink in the curved line is because of Lemma 1 to the left of the kink, the rare state is sufficiently unlikely, and the agent will not prepare for it - hence $p_R = \frac{1}{2}$. To the right of the kink, the rare state is sufficiently likely, and the agent will prepare for it - hence $p_R > \frac{1}{2}$.

If the rare state occurs, as is indicated by the red dot, the agent will not have collected any information, and will therefore be very uncertain. The policy function, as illuminated by the red-dashed line, shows that when the agent is uncertain, she will collect information about both states (as they are equiprobable). The red dashed line, shows the evolution of the agent's beliefs in subsequent periods as the likely state continues to occur. If the rare state occurred at some point before the agent's beliefs had converged back to $\tilde{p}$, the process would be set back, but would then start to converge again. One can think of this exercise as being a two-state version of an impulse response function. Given one 'shock' or rare event, followed by no other deviations, what happens to beliefs? An illustration of the evolution of

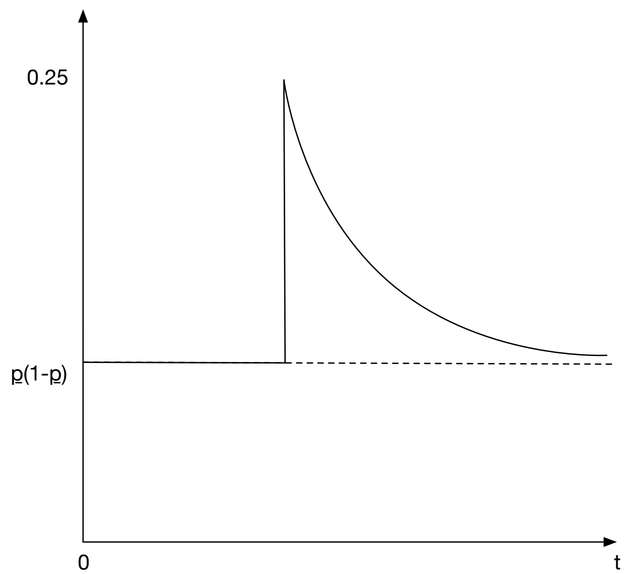


Figure 4: Plotted is an impulse response exercise of uncertainty as a function of time. After several periods of steady-state certainty, a rare event occurs. Upon such an occurrence, uncertainty spikes up, and beliefs will trace out the path of the red-dashed line in the previous figure, assuming that only likely states occur subsequently.

the agent's beliefs can be seen in Figure 4.

This graph shows what happens to the agent's uncertainty if the likely state happens for several periods, the rare state happens once, and then the likely state happens thereafter. Uncertainty remains low while the likely state occurs. There is a spike when the rare state occurs, because the agent did not collect any information about it beforehand. Due to the increase in uncertainty, the agent will start to collect information again, and as the likely state continues to occur, the agent's uncertainty will start to decline. Notice the similarity of the dynamics of this exercise with the plots in Figure 1 - there is persistence in uncertainty!

Uniqueness The solution does not guarantee *uniqueness*, yet uniqueness is not required for its results. The agent's beliefs can only move in two ways. Uncertainty can suddenly jump up if the rare state occurs. Uncertainty can decline slowly, according to the policy function, as the likely state occurs. As the likely state occurs, beliefs will converge to the first intersection between the policy function and the 45-degree line seen in Figure 3. If there are other, additional intersections for higher values of p_t, they can never be reached, as beliefs converge to the first intersection, and cannot move above it.

The key link between periods - in fact, the *only* link between periods, is that collecting information in one period *affects how information can be collected* in the next. It is the sole mechanism that delivers persistence in information collection, and therefore, persistence in uncertainty. This is the sole dynamic link for all versions of the model presented in this paper.

2.5 Discussion of Key Assumptions

There are two key assumptions in this model. The first is the structure of information collection and revelation. The second is the cost structure of information. Both are important, but the model's results can withstand loosenings of both.

Information Structure: The big assumption in the structure of information collection is that the agent takes her beliefs in period t about the likelihood of states in period $t+1$ as *given*. This precludes the possibility that the agent can collect information about $t+1$ in period t - she can only collect information about period $t+2$ in period t, even though that information will only be revealed in $t+1$. As I show in the appendix, this assumption is for tractability purposes - the results will hold even if it is relaxed.

From a *theoretical* standpoint, this paper is interested in *contingent* information acquisition, so this is the simplest setting to *isolate* that channel. Relaxing this restriction, and allowing the agent to collect information in time t about the state in $t+1$ as well as contingent information about $t+2$, would weaken the propositions quantitatively, but not qualitatively. Ultimately, contingent information about $t+2$ can only be collected once beliefs about $t+1$ are established. If the agent can collect information about $t+1$ at time t, she will merely start from a higher baseline for her contingent information collection.

From an *intuitional* standpoint, it is not unreasonable to think that agents will try to form contingent plans (be they investment plans, spending plans, etc), taking their beliefs about the future as given. In planning for the results of elections, referendums, etc, investors make plans for any outcome. They may try to reduce their uncertainty about the event, but they will allocate their resources in conditional information collection based on their beliefs about the outcome itself.

Cost Structure: The main assumption in the cost of information collection is that it is equally costly to collect information about either future state of the world. From a *theoretical* standpoint, this assumption improves tractability. This assumption is also justified by the more detailed framework of Woodford (2012). There is, as was mentioned earlier, also substantial support in the cognitive experimental literature for the static results that obtain from such an assumption. From an *intuitional* standpoint, one might think that it is actualy *costlier* to collect information about low-probability states, and *cheaper* to collect information about high-probability states. Such an alteration of the cost structure, would actually strengthen the results by making it even less likely that the agent would collect information about rare states. The cost structure that would weaken, or even negate the results of the paper is one where it is *costlier* to collect information about high-probability states. Such a cost structure does not seem intuitive, so I do not consider it.

This section presented the simplest setting for the model's mechanism: that an agent being uncertain today makes it likely that the agent will be uncertain in the next period. However, in order to make such a mechanism portable to more general models, I will need to demonstrate that it works under more general, continuous distributions.

3 Continuous-State Model

The previous section laid out an illustration of the paper's key mechanism. In this section, I extend the model both to generalize the results, and to show how the mechanism could be used in larger, economic models. The portability achieved with this section ensures that the results of this paper are not so specialized as to be unhelpful to standard economic models.

3.1 Model Structure

State Structure: The most important difference between this section's model and the previous section's model is the state space. There are an infinite number of possible states, indexed by the real line. To fix intuition in this setting, suppose that each state of the world corresponds to a different possible price change - that is the state x corresponds to 'the market changes by x next period'. The probability density function over the states, is given by a normal distribution with mean 0 and variance σ^2. Therefore $P(s_{t+1} = x) = \phi(x)$, where ϕ is the probability density function of a normal with variance σ^2. I will generally refer to the distribution in terms of its *precision*, $\tau = \frac{1}{\sigma^2}$.

Agent: As before, there is one agent. The agent enters period t with an information set I_t that informs her beliefs. Given the change in the state structure, her beliefs in period t about period $t+1$'s state are given by a normal distribution with variance $\widehat{\sigma}_t^2 \leq \sigma^2$. The mean of the agent's beliefs depends on the information set I_t: $E[s_{t+1}|I_t] = \mu_t$. Therefore $P(s_{t+1} = x|I_t) = \varphi_t(x)$ where φ is the probability density function of a normal with variance $\widehat{\sigma}_t^2$. I will refer to the agent's beliefs in terms of their *precisions* $\widehat{\tau}_t = \frac{1}{\widehat{\sigma}_t^2}$.

Preparation: Preparation is very similar to the previous setup. Conditional on her beliefs $\varphi_t(x)$, the agent can choose in period t to *prepare* for each of the possible states that could occur in period $t+1$. That is, the agent can *prepare* for the market changing in value in the next period, by collecting information about what will happen conditional on each of those possibilities. Preparation takes the form of picking information sets $I_{t+1,x}$. When state x occurs in period $t+1$, the agents beliefs will be:

$$\varphi_{t+1}(x) = V[s_{t+2}|I_{t+1,x}, s_{t+1} = x] \quad \forall x$$

When one of the possible states occurs, the agent's preparation for that state comes to fruition. Information forms her beliefs, conditional on the state. For simplicity, I model this decision as the agent directly selecting the precision of their posterior beliefs, $\widehat{\tau}_{t+1}(x)$, which is isomorphic to them choosing $I_{t+1,x}$.

The intuition behind this structure is that an agent has some beliefs about what will happen to the market in the next period. Having placed trades that make use of those beliefs, she can also prepare for what will come next. If the market crashes tomorrow, will it recover? If the market experiences a modest gain, what will happen next? Collecting information about those possibilities will allow the agent to react to each quickly and profitably.

Therefore, she would want to spend resources to be better prepared.

Agent's Costs: Costs are modeled in the same way as the previous setup. Improving the quality of preparation is costly. There is no fixed cost of preparing, but there is a cost function $c(\widehat{\tau}_{t+1}(x))$ associated with increased accuracy of future beliefs. That cost function is the *same* for all states x.

Agent's Objective: The agent's objective is modeled very similarly to the previous setup. The agent has a period-by-period utility function, which has one input: the precision of the agent's beliefs. As before, the utility function is independent of the conditional mean of beliefs, which, as before, is not a crucial assumption.

Now that the setup of the model has been laid out, I can state the agent's problem. Formally, the agent's problem in period t is given by the following Bellman equation:

$$V(\widehat{\tau}_t) = \max_{\widehat{\tau}_{t+1}(x)} U(\widehat{\tau}_t) - \int c(\widehat{\tau}_{t+1}(x))dx + \beta \int \varphi(x)V(\widehat{\tau}_{t+1}(x))dx - \int \mu_x \left(\tau_{t+1}(x) - \tau\right) dx$$

The agent begins period t with a certain precision of beliefs $\widehat{\tau}_t$. As before, the value is broken into several components. First, there is the *utility function*, U, that depends on the strength of the agent's beliefs. Then there is the *cost function*, c, that depends on the amount the agent prepares for state x. Finally, there is the *continuation value*, which is how the agent's beliefs in the next period will affect her preparation in the next period. The agent's choices are subject to inequality constraints, that her subjective beliefs be weakly stronger than the unconditional distribution. The Kuhn-Tucker conditions associated with those constraints are represented by the last integral. Notice again, that the key difference from a more typical Bellman formulation is that the state variable determines the distribution over which the continuation value is calculated.

3.2 Solution

Assumptions: To guarantee a solution, I need to impose some structure on the functions mentioned above. Namely I assume that:

1. $U'(\cdot) > 0$. This assumption states that the agent prefers stronger beliefs to weaker ones.

2. $c'(\cdot) > 0$. This assumption states that stronger beliefs are more costly to obtain than weaker ones.

3. $c''(\cdot) > U''(\cdot)$. This assumption states that the marginal cost of stronger beliefs increases faster than the marginal benefit - a technical assumption to guarantee an interior solution.

These assumptions are relatively standard, in that they are satisfied by most existing utility and costly information forms, and are sufficient to guarantee that interior solutions exist.

3.3 Results

All the major results of the previous section continue to hold here. The agent will collect *more* information about a state, the *more likely* it is.

Proposition 3. *If $\varphi_t(x) > \varphi_t(y)$, then $\widehat{\tau}_{t+1}(x) > \widehat{\tau}_{t+1}(y)$*

Proof: See Appendix

An agent focuses her preparation on states of the world she believes likely to occur, as those preparations are likely to prove useful. Agents prepare for events relative to how likely they are. The more likely, the more they prepare; the less likely, the less they prepare.

The second result is that if the agent believes a certain state is sufficiently unlikely, she will not collect *any* information about it.

Lemma 2. *If $\varphi(x) < \frac{c'(\tau)}{\beta V'(\tau)}$, then $\widehat{\tau}_{t+1}(x) = \tau$.*

Proof. If the above condition holds, then the first order condition

$$\beta\varphi(x)V'(\widehat{\tau}_{t+1}(x)) = c'(\widehat{\tau}_{t+1}(x)) + \mu_x$$

will only hold if $\mu_x > 0$, which means that the inequality constraint is binding. □

If the marginal benefit of collecting information about the a state is lower than the marginal cost of collecting information, *when the agent hasn't collected any information yet*, then the agent won't collect any information about that state. The agent won't prepare for sufficiently unlikely scenarios.

Agents completely ignore *sufficiently* rare events. An dynamic implication of this result comes from noticing that the term *sufficiently* is relative: The third result shows that under certain conditions, the agent will not collect any information about either state when her beliefs are the unconditional distribution.

Corollary 2. *(High uncertainty can be permanent). If* $\phi(0) < \frac{c'(\tau)}{\beta V'(\tau)}$*, then* $\widehat{\tau}_{t+s}(x) = \tau$*,* $\forall s > 0, \forall x$*.*

Proof. Follows immediately from Lemma 2. □

If the agent doesn't collect information, her beliefs will be the unconditional probability distribution. But if the most likely event under the unconditional distribution does not warrant any preparation, then the problem becomes permanent. If the agent doesn't collect information for the most likely event under the unconditional distribution, she won't collect information about any event under the unconditional distribution, which means that no matter what state occurs, she will face the unconditional distribution again in the next period, and every period thereafter.

The final result of this section mirrors that of the last. Under certain conditions, there is a 'steady state' level of uncertainty to which the agent will converge.

Proposition 4. *(Uncertainty is persistent) There is a* $\tilde{\tau}$ *such that if* $\widehat{\tau}_{t+1}(x) = \tilde{\tau}$*, then* $\widehat{\tau}_{t+2}(\varphi(\mu_t)) = \tilde{\tau}$*.*

Proof: See Appendix

The best way to see the implications of these result is graphically. Consider Figure 5. On the x-axis of this figure is the probability of an event according to the agent's beliefs. Points close to the origin are less likely according to her beliefs, and points farther away are more likely according to her beliefs. The kinked, curved line is the *policy function* - the agent's information collection as a function of a state's likelihood according to her beliefs today. It

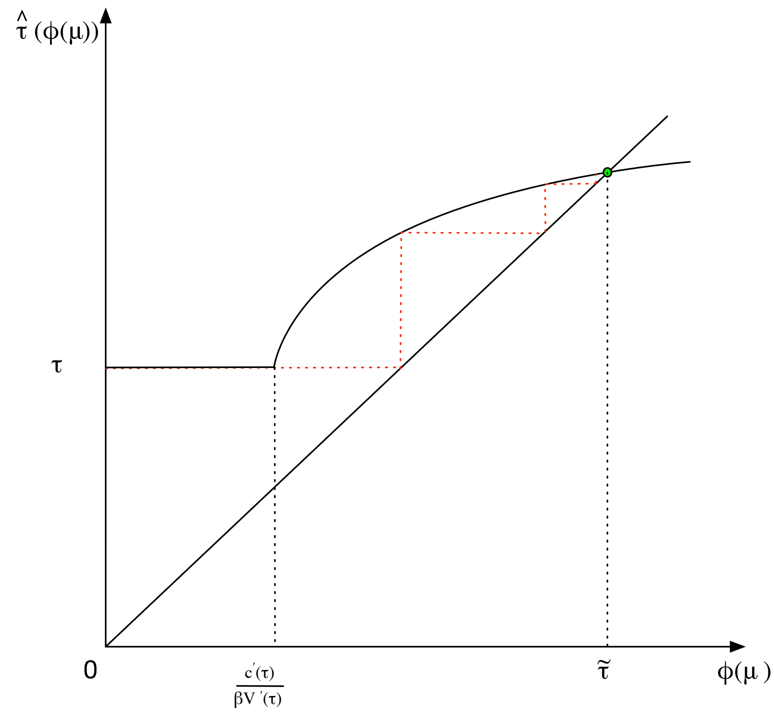


Figure 5: Plotted is the policy function of the precision of beliefs in period $t+1$ as a function of the ex-ante probability of an event in period t. The kinked-curved line is the policy function, the straight line is a 45-degree line. The red dashed line shows a possible evolution of beliefs when, conditional on one rare event occurring, only mean events occur subsequently (a standard impulse-response exercise).

plots the probability of the most likely event in period $t+2$ as a function of the agent's beliefs in period $t+1$. In a standard 'impulse response' exercise, one would examine the impact of one shock, followed by no further shocks. In this case, no shock is equivalent to the average, or expected event occurring each period. Therefore, an impulse response would involve some unlikely event occurring, and then the policy function guiding the agent's beliefs back to their steady state levels. The function is qualitatively very similar to the two-state case of the previous section.

This section presented a more portable setting for the model's mechanism: that an agent being uncertain today makes it likely that the agent will be uncertain in the next period. By using normal distributions on a continuous state space, the mechanism is ready to be brought into a larger model. In the next section, I present one example of such a model.

4 Application

Having described a general form of the mechanism, we now import it into a larger, applied-theory setting. In this section we embed the the continuous-state model of uncertainty persistence from the previous section in a financial market. In doing so, we can see the effects of persistence in uncertainty on financial and economics variables such as bid-ask spreads, volatility, dispersion of beliefs, and volume of trade. This application shows the ability of the mechanism to deliver patterns observed in data. This application is illustrative, and is by no means the only one to which the mechanism is suited.

4.1 Model Structure

State Structure: The information structure and cost structure of the previous two sections still hold in this section. The correlate of the agent in the previous two sections is an *informed trader* in this section. She takes her beliefs in time t over the possible states of the world in $t+1$ as given. At time t, she can prepare by collecting information about period $t+2$ for each possible future state of the world in $t+1$.

Financial Market: The key identifying structure to this model is that the states of the world are defined as possible fundamental values of an asset. The unconditional distribution of fundamental values at time $t+1$ is given by $F_{t+1} \sim N(F_t, \sigma^2)$. The unconditional volatility of fundamental values is constant across time, and the unconditional process of the fundamental value is a random walk.

Agents: Unlike the previous sections, there are three types of agents in this model. There are a continuum of perfectly competitive *market makers*, who observe public information, and set bids and asks. The remaining two types of agents are traders, noise and informed. Combined they form a unit continuum. One fraction, comprising a measure T, are *informed traders*. These traders are the correlate of the agent in the previous section[12]. They are able to collect state-contingent information, and can trade with the benefit of that information. The remaining fraction, comprising a measure $1-T$, are *noise traders*. These traders are

12. Here we treat the continuum of informed traders as a representative trader, who trades in a block. Alternatively, one could think of this as enforcing a symmetric equilibrium. The two alternatives are, ex-ante, identical.

present to provide liquidity in the model - they participate without heeding public or private information.

Trading: Each type of trader can buy one unit of the asset, sell one unit of the asset, or abstain from trade each period. Half the noise traders will *always* buy one unit of the asset, and half the noise traders will *always* sell one unit of the asset. The informed traders will base their trading decision on their information and on prices. If their beliefs lie above the market makers' asks, they will buy one unit of the asset. If their beliefs lie below the market makers' bids, they will sell one unit of the asset. If their beliefs lie within the market makers' bid-ask spread, they will not trade.

Signals: There are two types of signals, public and private. The public signal, y_t is distributed $y_t \sim N\left(F_{t+1}, \sigma^2_{y,t}\right)$. It is revealed to all agents (market makers, informed traders, and noise traders). All agents will have updated beliefs over F_{t+1} after seeing y_t that can be characterized as: $N\left(\mu_{pub,t}, \sigma_{pub,t}\right)$ where $\mu_{pub,t} \equiv \frac{\frac{1}{\sigma^2_{y,t}} y_t}{\frac{1}{\sigma^2} + \frac{1}{\sigma^2_{y,t}}}$, and $\sigma^2_{pub,t} \equiv \frac{1}{\frac{1}{\sigma^2} + \frac{1}{\sigma^2_{y,t}}}$. We label the probability distribution function associated with the updated beliefs $\phi_{pub,t}$. There is a private signal z_t, which is distributed $x_t \sim N\left(F_{t+1}, \sigma^2_{x,t}\right)$. It is revealed only to informed traders.

4.2 Statements of problems

There are two sets of decisions in the model. There are *pricing and trading decisions* that need to be made after all information has been revealed to the relevant agents. There are *information decisions* that need to be made beforehand. I will proceed backwards through these decisions.

4.3 Pricing and Trading Decisions

There are two stage 1 decisions - a pricing decision by market makers, and a trading decision by traders.

Informed Traders Trading Decisions: The informed traders will buy one unit of the asset if their beliefs lie above the market makers' asks, and sell one unit if their beliefs lie above the market makers bids. Therefore, their trading problem is:

$$\max\left\{(\text{bid}_t - s), (s - \text{ask}_t), 0\right\}$$

Where $s = E[F_{t+1}|\text{public signal}, \text{private signal}]$. Given the public and private signals, we can express s as

$$s = \frac{\frac{\mu_{pub,t}}{\sigma^2_{pub,t}} + \frac{x_t}{\sigma^2_{x,t}}}{\frac{1}{\sigma^2_{pub,t}} + \frac{1}{\sigma^2_{x,t}}}$$

Notice that agent's beliefs about an asset's fundamentals in a subsequent period are distributed normally, which allows for a quick relation, empirically to the VIX. The VIX measure's agents beliefs about volatility within a 30-day period.

Market Makers' Pricing Decisions: The market makers will observe public information and set prices - a bid and an ask. They are modeled after Glosten and Milgrom (1985). The bid is the price at which market makers buy the asset. The ask is the price at which market makers sell the asset. In order to understand their decision, we must understand the dispersion of possible beliefs. The dispersion allows us to understand the probabilities that the market maker will receive a buy order and a sell order. The dispersion is $\sigma^2_{\text{disp},t} = \frac{\frac{1}{\sigma^2_{x,t}}}{\left(\frac{1}{\sigma^2_{x,t}} + \frac{1}{\sigma^2_{pub,t}}\right)^2}$. The market making sector is perfectly competitive, so each market maker sets prices to satisfy a zero-profit condition. These zero-profit conditions are written as follows:

$$\Pi_{\text{ask}} = \int \phi_{pub,t}(x)\left((1 - \Phi_{\text{disp},x}(ask))(1 - T) + \frac{T}{2}\right)(ask - x)dx = 0 \quad (2)$$

$$\Pi_{\text{bid}} = \int \phi_{pub,t}(x)\left(\Phi_{\text{disp},x}(bid)(1 - T) + \frac{T}{2}\right)(x - bid)dx = 0 \quad (3)$$

where $\phi_{\text{disp},x} \sim \mathcal{N}\left(\frac{\frac{\mu_{pub,t}}{\sigma^2_{pub,t}} + \frac{x}{\sigma^2_{x,t}}}{\frac{1}{\sigma^2_{pub,t}} + \frac{1}{\sigma^2_{x,t}}}, \sigma^2_{\text{disp},t}\right)$. The market makers make zero profit in expectation.

4.4 Information Decisions

There's only one stage 0 decision - a state contingent information decision made by informed traders. Their decision in stage 0 of period t, given the quality of public signals, is to select a quality of a private signal on each possible realization of F_{t+1}. Formally, given $\sigma^2_{y,t}$, the informed trader's decision is:

$$
\begin{aligned}
V(F_t, \sigma^2_{y,t}) \quad = \quad & \int \max_{\sigma^2_{x,t}(F_t)} \phi(F_{t+1}) \int \int p(\text{public signal} = x) \left(p(\text{private signal} = y) U(x,y) dy \right) dx \\
& + \beta \left[\phi(F_{t+1}) V \left(F_{t+1}, \sigma^2_{y,t+1} \right) \right] - c\nu(\sigma^2_{x,t}(F_{t+1})) dF_{t+1} \qquad (4)
\end{aligned}
$$

where ν is a convex decreasing function such that $\frac{\partial^2 \nu}{\partial {\sigma^2_\delta}^2} > \frac{\partial^2 U}{\partial {\sigma^2_\delta}^2}$ everywhere. Note the parallel between this setup and that of the previous two sections. The relative convexity assumption on ν is meant to ensure interior solutions. Given the parallel of the setups, we can derive some of the same results.

4.5 Propositions

The first result shows that the utility function satisfies the assumption made in the previous sections - namely, that agents value information. All of the theoretical results of this section assume a known, exogenous process for public signal precision.

Proposition 5. *Utility decreases in the noise of private signals:* $\frac{\partial U}{\partial \sigma_x} < 0$.

The better the precision of a private signal, the higher the agent's utility. This result means that we can use the mechanism of the previous two sections - one of the fundamental assumptions in each, was that utility was increasing in signal precision. Now, we can also show that the accuracy of a signal is inversely proportional with the likelihood of the state occurring.

Corollary 3. *For any given c, δ, signal noise decreases in the liklihood of the state:* $\frac{\partial \sigma^2_x(F)}{\partial \phi(F)} \leq 0$.

This result confirms the intuition from the previous sections: information is a good, albeit a costly one. Next we show that the average probabilities of events are lower after rare events than they are under common events. Once that's shown, given the corollary above, we can infer that a rare event will trigger poor subsequent information collection on average, while a common event will trigger good subsequent information collection on average.

Corollary 4. *For any given c, δ, and for any two points F_1 and F_2, such that $\phi_{F,t}(F_1) > \phi_{F,t}(F_2)$. Then at time $t+1$, $E_{F_1}[\phi_{F,t+1}(F)] > E_{F_2}[\phi_{F,t}(F)]$.*

Thus the main results of the previous sections hold: first, statically, that agents collect more information about states the more likely they are, and second, dynamically, that poor information collection in a state reduces the condition average state-contingent information in subsequent periods. Now, for the purposes of simulation, we introduce one other agent and decision.

4.5.1 Public Entity

For the purposes of simulation, I now allow the precision of the public signal to be endogenous. The question here is what happens to financial variables, when public information suffers from the same patterns as private information. Public preparation is likely to be poor for rare events. But information types are strategically substitutable, so perhaps poor public information would encourage traders to prepare more for rare events, counteracting the previous results. To analyze this, I postulate a public entity who chooses the quality of state-contingent public information. The public entity tries to maximize state-contingent information subject to a cost - similar to the problems faced by the privately informed traders. The public entity works to select conditional signal quality, $\sigma^2_{y,t}$ for each potential value of F_t to maximize expected accuracy:

$$\int \min_{\sigma_{y,t}(F)} \phi_{F,pub,t}(F)\sigma^2_{pub,t}(\sigma^2_{y,t}(B)) - c_{pub}\nu_{pub}(\sigma^2_{pub,y,t}(F))d\phi_{F,pub,t} \tag{5}$$

where ν_{pub} is a convex function that satisfies the condition that $\frac{\partial^2 \nu_{pub}}{\partial \sigma_y^{2^2}} > \frac{\partial^2 \sigma^2_{pub}}{\partial \sigma_y^{2^2}}$ everywhere. The public entity seeks to be accurate, but the particular objective function is not overly important. One could think of the public information as being reports or actions from public institutions like the Federal Reserve or the government, or research reports published by financial institutions. These changes in the structure of public information no longer allow us to describe patterns analytically, but simulations still permit insight into the dynamics.

4.5.2 Dynamic Equilibrium

Given values of $\{F, \sigma^2, \sigma^2_{F,t}, c, c_{pub}\}$, a dynamic equilibrium is defined by a choice of σ_y by the public entity that solves equation 5 a choice of a policy function $\sigma_x(\mu_F, \sigma^2_F, \sigma^2_x)$ by the traders that satisfies equation 4, individual decisions to buy, sell, or abstain by traders to solve equation 4.3, and a choice of a bid and an ask by Market Makers to solve equations 2, given observed values of $\{\sigma_{y,t}, \sigma_{x,t}\}$, and a public signal.

4.6 Predictions and Spillovers

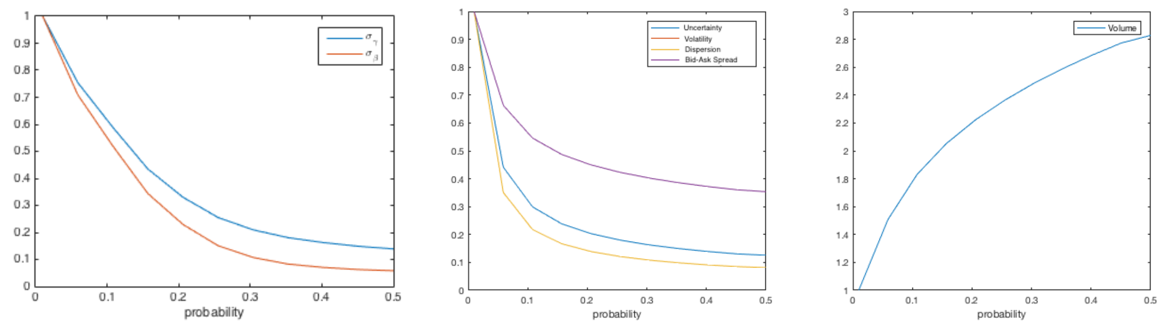
I present the simulation of the model statically and dynamically.[13] First, statically, Figure 6 shows the values of several financially relevant variables as functions of the ex-ante probability of events. The x-axis on all three panels is $\phi(F_t)$. Panel (a) shows the variance of public and private signals, which are both, unsurprisingly, declining in the probability of the event: the more likely the event, the higher the quality of public and private information. This initial finding spills over to other variables. Panel (b) shows that Volatility,[14] Uncertainty,[15] Dispersion, and Bid-Ask spreads are all lower after relatively likely events, than they are after rare events. Finally, panel (c) shows that agents trade more aggressively after high probability events, because they are better prepared for what happens next, and can trade strongly on their relatively accurate beliefs.

These results all correlate to the static version of the model. Unlikely events have different properties than likely events *within a period*. The next presentation has to do with the dynamic implications of those properties. For this I present an impulse-response exercise. I suppose that F_t makes a three standard deviation move between periods one and two. Then I assume that F_{t+1} immediately moves back to it's previous level, where it subsequently remains. Put differently, $F_1 = F_i$, $\forall i \neq 2$.

13. The parameter values here are $\sigma^2 = 10$, $c_{pub} = 0.4$, $T = 0.35$, $\beta = 0.95$, $c = 25$, and $\nu_{pub}(x) = \nu(x) = \frac{1}{x}$. This is not a calibration, but is meant only to be illustrative.

14. The expected variation of the fundamental conditional only on public information

15. The expected variation of the fundamental conditional on public *and* private information



(a) *Informational Investment* (b) *Financial Variables* (c) *Asset-Demand*

Figure 6: Plotted is a static snapshot of the applied model. The x-axis of all three plots is the ex-ante probability of an event. Figure (a) captures the variance of private and public signals. Figure (b) captures Uncertainty, Volatility, Dispersion, and Bid-Ask Spreads. Figure (c) captures the volume of trade. The first six variables all decline the more 'expected' an event was. The last increases.

The purpose of this exercise is to show how one rare event in an otherwise unconditionally random walk process can trigger lasting effects in uncertainty and volatility. The results are shown in Figure 7. Unsurprisingly, given the previous graph all the variables spike up in period 2, which when the first big change in the fundamental occurs. Uncertainty remains high when the second big change occurs, but drops slightly. The reason for this is important: when uncertainty is high in period 2, the variance of agents' conditional beliefs increase, which makes subsequent large movements *relatively more likely*. Therefore an identically sized move would be relatively more expected, lowering uncertainty slightly in period 3. This slight reduction in uncertainty is a distinguishing feature of this model from a standard parameter-learning model, as was discussed in the introduction. Between periods 3 and 9, the fundamental stays at exactly the same value, which means that any subsequent dynamics in these graphs is due to the natural processes of the mechanism of the model. As we can see the high levels of uncertainty in periods 2 and 3 persist, as agents find themselves unable to prepare well for even the most likely (no change in F) events. However, with each passing period, they find that the event they prepared the most for (again, no change in F) occurs, thus placing them in a slightly better position to prepare subsequently. This continues as the agents converge to their steady state levels of uncertainty and risk. All of these first five panels, whose variables were introduced in the previous Figure, exhibit the same dynamics.

A point of particular interest comes from the last panel. A standard hypothesis about

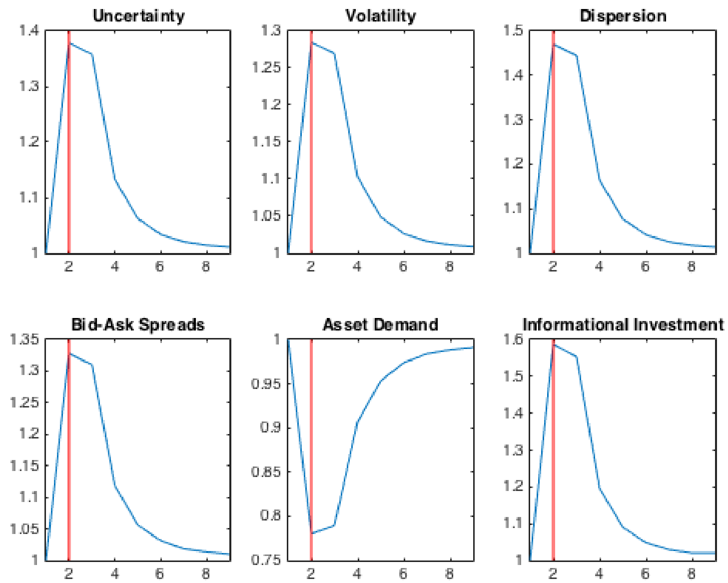


Figure 7: Plotted are six variables responding to a three-standard deviation move in the price of the asset and subsequent recovery. In period 2 the price moves by three standard deviations. In period 3, the price reverts to its period 1 level, and remains there subsequently. All variables respond very strongly to the first movement, and slightly less strongly to the second movement. However, starting in period 4, although the price is no longer moving, the mechanism of the model nonetheless generates persistence in the variables.

human behavior is that when we are uncertain, we try harder to learn. That intuition is borne out by this paper's mechanism. The last panel plots the total amount spent on information collection period by period. The point of emphasis thus far has been that when agents are uncertain, they don't know which states of the world to prepare for, so *on average* they are less prepared than when they are certain. However, when they are uncertain, they actually spend more *in sum* on information than when they are certain. This finding comes because agent's can't ignore anything when they are uncertain - anything *could* happen! Therefore, they must prepare at least a little for many many more states than they would if they were certain.

One of the benefits of this type of application is that it contains a measure of uncertainty and expected volatility that follow an autoregressive pattern, which is borne out empirically by the VIX, while the underlying process itself follows a random walk, which is a standard assumption. This paper's mechanism is able to wed those two seemingly conflicting concepts relatively tightly.

5 Discussion

Comparison to Parameter-Learning

As was mentioned in the introduction, parameter learning can also yield spikes and persistence in uncertainty. If an agent who does not know how volatile a process (for example, stock prices) is. She can try to learn what the volatility is by observing the process for a while. The more data she collects, the better her estimate of the volatility. If a very unusual draw is observed, the agent will change her volatility estimate sharply, which will look like a big change in uncertainty. The agent will require many more observations for her estimate to stabilize, so the change will persist for some time. Parameter learning is an elegant and intuitive way to understand how agents update beliefs, and it is undoubtedly a real part of our decision-making processes. However, the new theory laid out in this paper differs from parameter learning in several ways:

First, this model *requires fewer assumptions about the data-generating process.* One feature of parameter-learning agents is that they learn very quickly. Therefore, in order for a parameter-learning model to generate repeated, persistent spikes in uncertainty, the parameter the agents learn about must change over time. Otherwise, the agents would learn its value quickly and never be uncertain again. On the other hand, almost no assumptions on the data are needed in this paper's model to generate such repeated spikes. In fact, for the model in this paper, the *data generating process can even be unconditionally i.i.d. and still generate repeated, persistent uncertainty spikes.*

Second, this paper's formulation of *uncertainty can only jump up, and never down, which is a feature of the world.* When agents are uncertain in this model, they can only gradually regain certainty through preparation. Even if volatility vanishes when agents are uncertain, the mean outcome in a dispersed distribution is still not *very* likely, so agents will not be *much* better prepared for it than others. In a parameter-learning context, a sudden drop in volatility should cause a sudden drop in the estimation of the volatility parameter. We do not typically observe sudden reductions in uncertainty, so this difference is a feature of this paper's model.

Third, this model matches patterns in uncertainty and risk during *high volatility periods.*

According to a parameter-learning framework, sustained periods of high volatility should see sharp and continued increases in estimated volatility. The more high volatility a parameter-learning agent sees, the more confident they will be that their estimate of volatility should be high. However, in this paper's model sustained periods of volatility will cause spikes in such measures that subside as well. The basic intuition for why, is that uncertain agents in this paper will view future large shocks as being relatively more likely than when they were certain, and will thus prepare for them slightly more for them when they are uncertain. Such a pattern is observable in Figure 1 even during periods of sustained volatility like the financial crisis.

Definition of Uncertainty

Finally, this paper is testing a *fundamentally different type of uncertainty* from that generated by parameter-learning models. Parameter-learning uncertainty refers to a lack of understanding of how the world in which they live works. Seeing an unlikely event will make parameter learning agents reevaluate how likely that event is in the first place. Therefore, uncertainty in such a model refers to variance in agents' beliefs about a parameter's value that changes each period as their information sets change. Put more formally, uncertainty in a parameter learning model refers to $V_t[X|I_t]$, the variance in agents' beliefs at time t, conditional on any and all information available to them in time t about a deep parameter X that will never be directly observed.

By contrast, agents in this paper's model know exactly how the world in which they live works. Seeing an unexpected event, will not cause them to revise their beliefs about how likely the event was. However, due to their lack of preparation for unexpected events, such events will make them very uncertain about *what happens next*. Uncertainty in this paper refers to the variance in agents' beliefs about a variable that will be revealed to them imminently. More formally, uncertainty in this paper refers to $V_t[X_{t+1}|I_t]$, the variance in agents' beliefs at time t, conditional on any and all information available to them in time t about a variable X that will be revealed to them in the subsequent period. This latter definition of uncertainty conforms to how uncertainty is described in the uncertainty shocks literature, started by Bloom (2009).

6 Conclusion

I have presented and solved a simple, tractable model that uses a cognitive feature of our attentional makeup to deliver observed patterns in uncertainty dynamics. Specifically, I employ the concept of preparation - that agents distribute their attention over possible states of the world to be better prepared for their occurrence. When unlikely or rare states occur, agents have not prepared for them, and face high levels of uncertainty. High uncertainty makes it harder for agents to prepare subsequently, as they do not know where to focus their preparations. Therefore, uncertainty begets itself, persisting endogenously, even when the underlying states of the world exhibits no persistence. I showed this mechanism for persistence in a two-state case for intuition, and then generalized to a continuous state case. Finally, I embedded it in a larger model to show its portability and tractability. I then discussed alternative applications, distinguishing features of this new model from more standard parameter-learning models, and definitions of uncertainty. I believe that this mechanism provides a fundmanetally new perspective from which to analyze changes in uncertainty.

References

Banerjee, Snehal, and Bradyn M Breon-Drish. 2017. "Dynamic information acquisition and strategic trading."

Banerjee, Snehal, and Ilan Kremer. 2010. "Disagreement and learning: Dynamic patterns of trade." *The Journal of Finance* 65 (4): 1269–1302.

Bloom, Nicholas. 2009. "The impact of uncertainty shocks." *econometrica* 77 (3): 623–685.

Bolton, Patrick, and Antoine Faure-Grimaud. 2009. "Thinking ahead: the decision problem." *The Review of Economic Studies* 76 (4): 1205–1238.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012. "Salience theory of choice under risk." *The Quarterly journal of economics* 127 (3): 1243–1285.

———. 2013a. "Salience and asset prices." *American Economic Review* 103 (3): 623–28.

———. 2013b. "Salience and consumer choice." *Journal of Political Economy* 121 (5): 803–843.

Eriksen, Charles W, and James D St James. 1986. "Visual attention within and around the field of focal attention: A zoom lens model." *Perception & psychophysics* 40 (4): 225–240.

Gershman, Samuel J, Eric J Horvitz, and Joshua B Tenenbaum. 2015. "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines." *Science* 349 (6245): 273–278.

Glosten, Lawrence R, and Paul R Milgrom. 1985. "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders." *Journal of financial economics* 14 (1): 71–100.

Ilut, Cosmin, and Rosen Valchev. 2017. "Economic agents as imperfect problem solvers."

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2016. "The tail that wags the economy: Belief-driven business cycles and persistent stagnation."

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. 2018. "Dynamic rational inattention: Analytical results." *Journal of Economic Theory* 176:650–692.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal sticky prices under rational inattention."

Maćkowiak, Bartosz, and Mirko Wiederholt. 2010. "Business cycle dynamics under rational inattention."

———. 2011. "Inattention to rare events."

———. 2018. "Lack of Preparation for Rare Events." *forthcoming, Journal of Monetary Economics.*

Martin, Ian WR, and Robert S Pindyck. 2015. "Averting catastrophes: the strange economics of Scylla and Charybdis." *American Economic Review* 105 (10): 2947–85.

Matějka, Filip. 2015a. "Rationally inattentive seller: Sales and discrete pricing." *The Review of Economic Studies:* rdv049.

———. 2015b. "Rigid pricing and rationally inattentive consumer." *Journal of Economic Theory* 158:656–678.

Matejka, Filip, and Alisdair McKay. 2015. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." *American Economic Review* 105 (1): 272–98.

Nimark, Kristoffer. 2014. "Man-bites-dog business cycles." *The American Economic Review* 104 (8): 2320–2367.

Nimark, Kristoffer, and Savitar Sundaresan. 2018. *Inattention and Belief Polarization.* Technical report. Cornell University.

Orlik, Anna, and Laura Veldkamp. 2014. *Understanding uncertainty shocks and the role of black swans.* Technical report. National Bureau of Economic Research.

Peters, Achim, Bruce S McEwen, and Karl Friston. 2017. "Uncertainty and stress: Why it causes diseases and how it is mastered by the brain." *Progress in neurobiology* 156:164–188.

Posner, Michael I, Charles R Snyder, and Brian J Davidson. 1980. "Attention and the detection of signals." *Journal of experimental psychology: General* 109 (2): 160.

Posner, Michael I, Charles R Snyder, and R Solso. 2004. "Attention and cognitive control." *Cognitive psychology: Key readings* 205.

Shaw, Marilyn L, and Peter Shaw. 1977. "Optimal allocation of cognitive resources to spatial locations." *Journal of Experimental Psychology: Human Perception and Performance* 3 (2): 201.

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

Steiner, Jakub, Colin Stewart, and Filip Matějka. 2017. "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." *Econometrica* 85 (2): 521–553.

Tenenbaum, Joshua B, Charles Kemp, Thomas L Griffiths, and Noah D Goodman. 2011. "How to grow a mind: Statistics, structure, and abstraction." *science* 331 (6022): 1279–1285.

Woodford, Michael. 2012. "Inattentive valuation and reference-dependent choice." *Unpublished Manuscript, Columbia University.*

A Proofs

Proof. Proof that the Bellman is a contraction. The Bellman equation is:

$$V(p_t) = \max_{p_L, p_R \in \left[\frac{1}{2}, 1\right]} U(p_t) - c(p_L) - c(p_R) + \beta p_t V(p_L) + \beta(1 - p_t)V(p_R)$$

I show here that the equation satisfies Blackwell's sufficient conditions and thus describes a contraction mapping. The value function V is bounded, as $\frac{1}{2} \leq p_t \leq 1$, and U is a real-valued function. The Value function is a mapping from $\left[\frac{1}{2}, 1\right]$ to $\left[\frac{U\left(\frac{1}{2}\right)}{1-\beta}, \frac{U(1)}{1-\beta}\right]$. Therefore the Bellman equation describes a self-map on the space of bounded functions $B(X)$. To show that the mapping T from space of functions onto itself is a contraction, we must show that:

1. Monotonicity: if $f, g \in B(X)$, and $f(x) < g(x)$ for all $x \in X$, then $Tf(x) \leq Tg(x)$ for all $x \in X$.
2. Discounting: For $\gamma \in \mathbb{R}$, there exists a β such that for all $f \in B(X)$, and all $x \in X$, $T(f + \gamma)(x) \leq Tf(x) + \beta\gamma$.

First, monotonicity:

$$\begin{aligned}
Tf(p_t) &= \max_{p_L, p_R \inf \left[\frac{1}{2}, 1\right]} U(p_t) - c(p_{L,f}) - c(p_{R,f}) + \beta p_t f(p_{L,f}) + \beta(1 - p_t) f(p_{R,f}) \\
&\leq \max_{p_L, p_R \inf \left[\frac{1}{2}, 1\right]} U(p_t) - c(p_{L,f}) - c(p_{R,f}) + \beta p_t g(p_{L,f}) + \beta(1 - p_t) g(p_{R,f}) \\
&\leq \max_{p_L, p_R \inf \left[\frac{1}{2}, 1\right]} U(p_t) - c(p_{L,g}) - c(p_{R,g}) + \beta p_t g(p_{L,g}) + \beta(1 - p_t) g(p_{R,g}) \\
&= Tg(p_t)
\end{aligned}$$

Where the first equality is a definition, the second follows from the fact that $f < g$ everywhere, and the third follows from the the agent's optimization. The last equality is again a definition.

Second, discounting:

$$\begin{aligned} T(f+\gamma)(p_t) &= \max_{p_L,p_R \inf\left[\frac{1}{2},1\right]} U(p_t) - c(p_{L,f}) - c(p_{R,f}) + \beta p_t f(p_{L,f}) + \beta(1-p_t) f(p_{R,f}) + \beta\gamma \\ &= Tf(p_t) + \beta\gamma \end{aligned}$$

□

Proof. Proposition(1)

The first order conditions for the agent's problem are:

$$\begin{aligned} V'(p_t) &= U'(p_t) + \beta(V(p_L) - V(p_R)) \\ c'(p_L) &= \beta p_t V'(p_L) + \mu_L \\ c'(p_R) &= \beta(1-p_t)V'(p_R) + \mu_R \\ \mu_L\left(p_L - \frac{1}{2}\right) &= 0 \\ \mu_R\left(p_R - \frac{1}{2}\right) &= 0 \end{aligned}$$

Substituting, we get that:

$$\frac{c'(p_L) - \mu_L}{\beta p_t} - U'(p_L) - \beta(V(p_L^*(p_L)) - V(p_R^*(p_L))) = \frac{c'(p_R) - \mu_R}{\beta(1-p_t)} - U'(p_R) - \beta(V(p_L^*(p_R)) - V(p_R^*(p_R))) = 0$$

If μ_L and μ_R are both non-zero, then the solution is trivial. Let us then consider the case where $\mu_L = \mu_R = 0$.

$$\frac{c'(p_L)}{\beta p_t} - U'(p_L) - \beta(V(p_L^*(p_L)) - V(p_R^*(p_L))) = \frac{c'(p_R)}{\beta(1-p_t)} - U'(p_R) - \beta(V(p_L^*(p_R)) - V(p_R^*(p_R))) = 0$$

Given that $c'' > U''$, the left hand side of the above equation increases for a marginal increase in p_L (by the envelope theorem, since the Vs are already optimized, a marginal increase in p_L does not change their value), and the right hand side is similarly marginally increasing in p_R. Therefore if either p_L and p_R are interior solutions, $p_L'(p_t) \geq 0$, and $p_R'(p_t) \leq 0$.

Now consider the case where $\mu_R = 0$ and $p_L = \frac{1}{2}$.

$$\frac{c'\left(\frac{1}{2}\right) - \mu_L}{\beta p_t} - U'\left(\frac{1}{2}\right) = \frac{c'(p_R)}{\beta(1-p_t)} - U'(p_R) - \beta(V(p_L^*(p_R)) - V(p_R^*(p_R)))$$

Because $p_t \geq \frac{1}{2}$, in order for the above equality to be satisfied, it must be the case that $\mu_L < 0$, which is a contradiction. Therefore it can never be the case that $\mu_R = 0$ while $\mu_L > 0$. Next consider the case where $\mu_L = 0$, and $p_R = \frac{1}{2}$.

$$\frac{c'(p_L)}{\beta p_t} - U'(p_L) = \frac{c'\left(\frac{1}{2}\right) - \mu_R}{\beta(1-p_t)} - U'\left(\frac{1}{2}\right)$$

In the absence of the binding constraint that $p_R \geq \frac{1}{2}$, increasing p_t would lead to a decrease in p_R, all else constant. Therefore, the higher p_t, the lower the 'shadow value' of p_R, and the higher the shadow cost. If $\mu_R'(p_t) > 0$, then $p_L'(p_t) > 0$.

Under all conditions, $p_L'(p_t) \geq 0$, and $p_R'(p_t) \leq 0$, with equality only when the constraint is binding. □

Proof. Proposition (3) The first order conditions to this problem are:

$$\begin{aligned} V'(\widehat{\tau}_t) &= U'(\widehat{\tau}_t) + \int \beta \varphi_\tau(x) V(\widehat{\tau}_{t+1}(x)) dx \\ \beta \varphi(x) V'(\widehat{\tau}_{t+1}(x)) &= c'(\widehat{\tau}_{t+1}(x)) + \mu_x \end{aligned}$$

Substituting as before we get:

$$U'(\widehat{\tau}_t) + \int \beta \varphi_\tau(x) V(\widehat{\tau}_{t+1}^*(x)) dx = \frac{c'(\widehat{\tau}_t(x)) - \mu_x}{\beta \varphi(x)}$$

By the same envelope theorem argument as the binary state case, along with the relative convexities of c and U, we get that $\widehat{\tau}_t'(\varphi(x)) \geq 0$. □

Proof. Proposition (4) If $\lim_{\widehat{\tau} \to \infty} \frac{c''(\tau)}{\varphi_\tau(x)} > U''(\tau)$, then $\lim_{\varphi \to \infty} \widehat{\tau}'(f) < 1$. Consider the variable $g(\widehat{\tau}) = \varphi(\mu, \tau)$. Then, $g'(\widehat{\tau}) < 1$ in the limit. We also know that $g'(0) = \tau > 0$, which means that there is a point τ^* such that $g(\tau^*)$ intersects the 45-degree line at least once, yielding a fixed-point. □

Proof. Proposition (5)

First, we can rewrite equations 2 as zero-profit conditions as follows:

$$\begin{aligned}\frac{T}{2}(\text{ask}-\mu_p) &= \int \phi_p(x)(\text{ask}-x)(1-\Phi_{\text{disp},x})(\text{ask})dx(T-1)\\ \frac{T}{2}(\mu_p-\text{bid}) &= \int \phi_p(x)(x-\text{bid})\Phi_{\text{disp},x}(\text{bid})dx(T-1)\end{aligned}$$

The left-hand side of these expressions is the profit earned from noise traders - $\frac{T}{2}$ purchases or sales, which in expectation are equal to μ_p. On the right hand side is the expected loss from adverse selection to informed traders. Market Maker profit is increasing in the ask and decreasing in the bid, both by extracting more from noise traders, and giving away less to informed traders.

Focusing just on the bid equation (the ask follows similarly):

$$\begin{aligned}& \frac{\partial}{\partial\sigma_\gamma^2}\int \phi_p(x)(x-\text{bid})\Phi_{\text{disp},x}(\text{bid})dx(1-T)\\ \propto\ & \frac{\partial}{\partial\sigma_\gamma^2}\int \phi_p(x)(x-\text{bid})\Phi_{\text{disp},x}(\text{bid})dx\\ =\ & \frac{\partial}{\partial\sigma_\gamma^2}\int_{-\infty}^{\text{bid}} \phi_p(x)(x-\text{bid})\Phi_{\text{disp},x}(\text{bid})dx + \frac{\partial}{\partial\sigma_\gamma^2}\int_{\text{bid}}^{\infty} \phi_p(x)(x-\text{bid})\Phi_{\text{disp},x}(\text{bid})dx\\ =\ & \int_{-\infty}^{\text{bid}} \underbrace{\phi_p(x)(x-\text{bid})}_{<0}\underbrace{\int_{-\infty}^{\text{bid}}\frac{\partial}{\partial\sigma_\gamma^2}\phi_{\text{disp},x}(y)dy}_{>0}\,dx\\ & +\int_{\text{bid}}^{\infty} \underbrace{\phi_p(x)(x-\text{bid})}_{>0}\underbrace{\int_{\text{bid}}^{\infty}\frac{\partial}{\partial\sigma_\gamma^2}\phi_{\text{disp},x}(y)dy}_{<0}\,dx\\ <\ & 0\end{aligned}$$

Therefore, by similar logic we have that an decrease in σ_γ^2, holding all other variables fixed, results in lower profits for the market maker.

In order to increase profits again, the market maker will need to lower the bid and raise the ask, to make up the difference from noise traders. Thus, an decrease in σ_γ^2 increases the transfer from noise traders to informed traders in expectation. So U depends negatively on σ^2_γ. □

Proof. Corollary (3)

The necessary condition is that $\frac{dU}{d\phi(B_t)}$ => 0. This is trivially true. □

Proof. Corollary (4)

If $\phi_{B,t}(B_1) > \phi_{B,t}(B_2)$, then $\sigma_{\gamma,t}(B_1) < \sigma_{\gamma,t}(B_2)$. Then $V_{B_1}[\phi_{B,t+1}(B)] < V_{B_2}[\phi_{B,t}(B)]$. So $E_{B_1}[\phi_{B,t+1}(B)] > E_{B_2}[\phi_{B,t}(B)]$. □

Extension to contemporaneous preparation

In this extension, I consider the possibility that an agent in period t can change her beliefs about the events in period $t+1$ by collecting information. In the original continuous-state formulation, the agent's problem was:

$$V(\widehat{\tau}_t) = \max_{\widehat{\tau}_{t+1}(x)} U(\widehat{\tau}_t) - \int c(\widehat{\tau}_{t+1}(x))dx + \beta \int \varphi(x,\widehat{\tau})V(\widehat{\tau}_{t+1}(x))dx - \int \mu_x\left(\tau_{t+1}(x) - \tau\right)dx$$

If the agent could change her beliefs about period $t+1$, that means she could change $\widehat{\tau}_t$. That variable shows up in two places - the utility function, and the continuation value. I will continue to assume that changing her beliefs about $t+1$ today will not affect her utility. That is because the assumption of the benefit to the utility function is that the agent can use her information quickly. If she must collect information today, she cannot also take actions that benefit her today. However, by collecting information today, it *can* aid in her preparation for the next period. Therefore, her new problem will look like this:

$$V(\widehat{\tau}_t) = \max_{s,\widehat{\tau}_{t+1}(x)} U(\widehat{\tau}_t) - \int c(\widehat{\tau}_{t+1}(x))dx + \beta \int \varphi(x,\widehat{\tau}+s)V(\widehat{\tau}_{t+1}(x))dx - \int \mu_x\left(\tau_{t+1}(x) - \tau\right)dx - c_s(s)$$

The first order conditions here are:

$$U'(\widehat{\tau}_t) + c_s'(s) \quad = \quad \frac{c'(\widehat{\tau}_t(x)) + \mu_x}{\beta\varphi(x,\widehat{\tau},s)}$$

The envelope argument of the previous proofs still hold, and it is still the case that $\widehat{\tau}_{t+1}$ is an increasing function of φ for any values of $\widehat{\tau}$ and s. As long as $c_s', c_s'' > 0$, it is also true that agents will not eliminate future uncertainty, which means that the persistence result will hold as well. How strongly it holds will be a function of how easy it is to scramble for information.