Rare Events and the Persistence of Uncertainty*

Savitar Sundaresan†

April 19, 2017

Abstract

Unexpected events can have lasting effects on uncertainty. This paper presents a model in which occurrences of rare events endogenously result in lower levels of private information. Lower levels of information propagate within the model, as uncertainty makes it harder for agents to acquire information about future periods, resulting in uncertainty persistence. This model of uncertainty is applied to an economy with a financial market, yielding implications for the dynamics of financial uncertainty, asset demand, expected wealth, dispersion of beliefs, bid-ask spreads, and volatility.

*Thanks to Ricardo Reis, Michael Woodford, Larry Glosten, Jose Scheinkman, and Jennifer La’O for invaluable guidance, assistance and advice; to Geert Bekeart, Patrick Bolton, Robert Hodrick, Michael Johannes, Marcin Kacperczyk, Emi Nakamura, Jaromir Nosal, Andrea Prat, Jon Steinsson, Tano Santos, Stephanie Schmitt-Grohe, Suresh Sundaresan, Martin Uribe, Laura Veldkamp, and Neng Wang for discussions and comments that shaped the content of this work; to seminar and conference participants at CEPR, Columbia, Columbia GSB, Emory, UC Davis, Northwestern Kellogg, London Business School, Imperial, Johns Hopkins Carey, and the Federal Reserve Board; and to Cynthia Mei Balloch, Keshav Dogra, Stephane Dupraz, Sergey Kolbin, Shaowen Luo, Antonio Miscio, Paul Piveteau, Evan Plous, Samer Shousha, Xing Xia, and Daniel Villar for all their help. All errors are my own.

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

Unexpected events of economic relevance can cause persistent jumps in uncertainty. Uncertainty has financial implications: In financial markets, spikes in uncertainty are often accompanied by increased dispersion of beliefs, increased bid-ask spreads, increased volatility, and decreased leverage. By focusing on the aftermath of rare events, this paper presents a model that explains patterns in uncertainty and the effects of persistent uncertainty on financial variables.

Consider the examples of events marked by red lines in Figure 1, each of which could be considered ex-ante unlikely. The first is the terrorist attacks in September 2001. The second is Lehman Brothers’ bankruptcy in September 2008. The third is the flash crash in May 2010. The fourth is the downgrade of U.S. treasury debt by the S&P ratings agency in 2011. The fifth is a sharp downturn in China’s manufacturing activity in August 2015. In each case, the VIX\(^1\) spikes upon the occurrence of the event and then slowly reverts to its pre-shock levels.

![Figure 1: Daily values of VIX with markers for key rare events](image)

I propose a model to explain three observations upon the realization of rare events: (i) that uncertainty spikes; (ii) that the spikes persist; and (iii) that changes in uncertainty spill over into other financial variables. In this model, agents can invest in information about

---

1. The VIX is the market’s expectation of the annualized percentage standard deviation of the S&P index over a thirty day period. It is calculated by the CBOE, using options prices, and is often used to proxy financial uncertainty.
future states of the world. Upon the occurrence of any state, agents receive signals about the future path of an asset’s price. Therefore, the quality of signals received is determined by their informational investment decision. The benefit of investing in information varies positively with the ex-ante likelihood of the state, but the cost of investing is assumed not to vary. To preserve informational asymmetries and prevent perfect revelation of information in prices, all trade takes place through a perfectly competitive market making sector.

This model produces three theoretical results. The first result is that the occurrence of an ex-ante unlikely state results in increased short term and long term uncertainty due to poor signal quality. The agents’ desire for information is weighted by the likelihood of that information paying off, so agents choose to invest in more information about how to act in likely states of the world and less in unlikely states of the world. When an agent invests in information about one state of the world, and that state occurs, signals about future states are less noisy (or, of a higher quality).

Second, upon the occurrence of a rare event, the model delivers increased long term uncertainty, which persists endogenously. Poor signals result in a more dispersed distribution over possible states in the subsequent period. The average quality of signals purchased under this dispersed distribution will be worse than under a tight distribution because signal quality is positively correlated with the probability of a state’s occurrence. Therefore, an initial unlikely event generates uncertainty, which persists and propagates by preventing agents from concentrating the allocation of their informational resources.

The third set of theoretical results stem from the interaction of agents’ signal quality with the market makers’ information set. Because the market making sector prevents perfect price revelation, this interaction causes spillovers into financial variables such as volatility, bid-ask spreads, dispersion of beliefs, and asset demand.

The informational investment mechanism I use will deliver uncertainty persistence without long term parameter-learning under any circumstances where information is a good. Therefore, there are several environments in which to explore the effects of this mechanism. In this paper, I consider a framework where information gives agents an advantage in trade. Choosing this environment to explore the spillover is helpful because it allows the model to deliver predictions about financial uncertainty, while maintaining portability into
a macroeconomic setting. Fundamentally the patterns of uncertainty are universal (that is, not specific to finance), but financial variables have plentiful and frequently observed data.

Persistence in financial uncertainty has been well documented in the literature, and is often modeled using ARCH or GARCH processes. One of the benefits of the framework I develop in this paper is that it takes an underlying i.i.d. process and endogenously generates an autoregressive-style shape to uncertainty.

I provide some suggestive empirical evidence for the mechanism in the model. First, I provide a calibration and simulate the model’s ability to match the VIX, dispersion of beliefs, and investors’ risk appetite. The model’s predicted VIX overlays closely with the data. I then use options volume data to examine whether higher levels of ‘preparation’ mitigate the effects of large price movements on the VIX. Finally, I also provide some empirical implications and theoretical questions that can serve as a basis for future work.

1.1 Alternative Theories

Much of the existing literature on macroeconomic and financial uncertainty employs one of two theoretical underpinnings - rare disasters theory and Bayesian learning. This paper relates to both and contributes to each. Rare disasters theory shows that the realization of disasters$^2$ - or usually the increased fear of their occurrence - makes agents more uncertain. This theory (as described by, for example, Orlik and Veldkamp (2014)) achieves a good description of macroeconomic (or, real) uncertainty. The importance of the skew is also documented in Bekaert and Popov (2012). However, if the black swan story were true for financial uncertainty, then the SKEW$^3$ (or fatness of the left tail) of the VIX should be positively correlated with the level of the VIX. Fatness of the left tail of the distribution would be a measure of fear of disasters and should cause generalized uncertainty. In fact, as shown by the CBOE itself,$^4$ the two are not correlated and at a daily frequency, they are actually negatively correlated - at about -0.13. This indicates that fear of rare events does

---

2. Sometimes called black swans
3. As calculated from options prices by the CBOE, it is a non-parametric measure of the skewness of the distribution of the market’s expectation of annualized percentage price changes of the S&P index over the next 30 days
not constitute a full explanation of changes in financial uncertainty. The model in this paper assumes no skewness in either the underlying process or in beliefs and, as a result, does not have to explain this negative correlation.

The second standard explanation of uncertainty patterns is Bayesian learning, such as with Drechsler (2013) under which a Bayesian agent learns and updates beliefs about the true parameters of the economy through shocks. Bayesian agents learn very quickly, so when using Bayesian learning, there is a tension between making the shocks big enough to match the significant changes in economic dynamics and having the learning process be slow enough that uncertainty is not resolved too quickly. It is evident from Figure 1 that because the level of the VIX repeatedly spikes, a purely Bayesian story might be incomplete. Other papers have proposed using anticipated utility equilibria, limited memory, or regime changes (such as is found in Bianchi and Melosi (2013)) to get repeated spikes. The agents in this paper use Bayes rule to update expectations; however, this paper shows that rare events can lead to persistent uncertainty even with an infinite amount of previous data.

This paper’s explanation of persistent uncertainty emerges along with a growing literature on the subject. I will discuss the methods of two recent and notable contributions. The first is Kozlowski, Veldkamp, and Venkateswaran (2015), in which the authors posit that as agents use standard econometric tools to estimate the distribution of aggregate shocks, the occurrence of an extremely unlikely event changes their beliefs suddenly and increases their uncertainty. The uncertainty persists as the shock lives permanently in the agents’ datasets. Using this setup, they are able to match the downward shift in trend output of the great recession. The second is Fajgelbaum, Schaal, and Taschereau-Dumouchel (2014). In that paper’s framework, uncertainty over underlying parameters of the economy affects firms’ decisions about how much to invest. The results of those decisions inform other firms, which can learn from the distribution of all investment returns. An initial spike in uncertainty causes fewer firms to invest, resulting in lower levels of learning, and therefore uncertainty traps. This paper joins this literature in trying to explain the persistence of uncertainty from a complementary angle: Instead of trying to understand agents’ learning of economic parameters, this model focuses on endogenous information acquisition.
1.2 Literature Review

This paper borrows modeling techniques from two sources. First, it uses a reduced form of the inattention structure built in Woodford (2012), which shows how rare events cause poor conditional identification. Second, in order to preserve uncertainty when trading and prevent price revelation, it uses a simple batch-order version of the market makers designed by Glosten and Milgrom (1985). The initiation and preservation of uncertainty is key, as it allows for the informational and adverse selection dynamics necessary for all of the results of the paper.

There are several strands of the theoretical literature that the paper interacts with. The first is that of uncertainty shocks. As proposed in Bloom (2009), and subsequently expanded upon by many others (including Aghion et al. (2010); Arellano, Bai, and Kehoe (2010); and Gilchrist, Sim, and Zakrajšek (2014)), downturns in output or investment can be thought of as originating from spikes in macroeconomic uncertainty. The causality can either run through the increased option value of waiting or (as in this work) the increased cost of borrowing caused by higher credit spreads. I endogenize the connection between the actual event and the increase in uncertainty, while maintaining flexibility and portability into macro settings.

The second strand is that of uncertainty being tied to rare disasters. Orlik and Veldkamp (2014) show that tail events (or even fear of tail events) can cause spikes in real uncertainty. Cúrdia, Negro, and Greenwald (2014) show that fattening the tails on the distribution of shocks has better results in matching data on volatility. There is also a sizable literature in finance on the equity premium puzzle started by Rietz (1988), and reinvigorated by Barro (2006), showing that a fear of tail events can generate excess returns in equities. Backus, Chernov, and Martin (2011) rigorously estimate probabilities of disasters from financial data. Learning models, such as Evans and Honkapohja (2001) and Cogley, Matthes, and Sbordone (2011) take anticipated utility approaches. This paper differs in three main respects from the rare events literature: First, there is no skewness in the model - conditional uncertainty can arise from positive shocks as well as negative ones. This symmetry is imposed deliberately to push the central mechanism of the information collection as far as possible, without
worrying about risk aversion (Segal, Shaliastovich, and Yaron 2015) discuss asymmetry. Second, there is no long-term Bayesian learning about the parameters of the economy. Third, it treats the level of informedness as endogenous.

The third strand is that of market microstructure, informational asymmetries, and imperfect price revelation. Hassan and Mertens (2014) use noise traders to create a DSGE that results in imperfect price aggregation. Fishman and Parker (2015) study a market where adverse selection can occur and lead to multiple equilibria. Bond, Edmans, and Goldstein (2011) show transmission of price information from the financial sector to the real economy. Kurlat (2010) and Guerrieri and Shimer (2012) show that adverse selection can lead to illiquidity even when sellers offer different prices. Bookstaber and Pomerantz (1989) relate discrete information packet collection to financial volatility. Routledge and Zin (2009) show that adverse selection can lead to illiquidity when the market makers have uncertainty aversion.

This paper also touches on the nature of information collection, as motivated by the notion of inattention, first brought to light by Sims (2003). Woodford (2012) provides neuro-scientific evidence that traditional formulations of rational inattention are incompatible with actual attentional behavior and proposes an alternative formulation. This paper takes that alternative as a starting point; although the model does not microfound agents’ attentional decisions, the motivation is to stay in step with Woodford (2012). An application of these methods to a macro setting is seen in Maćkowiak and Wiederholt (2009) to explore price stickiness, Paciello and Wiederholt (2013) for optimal monetary policy, and Maćkowiak and Wiederholt (2010) to explore sluggish responses. As in this paper, the only source of adjustment in Maćkowiak and Wiederholt (2010) is attention. Matějka (2015) provides a host of results on price dynamics for inattentive sellers. In finance, Mondria (2010) uses inattention to discuss price co-movement, Kacperczyk, Nieuwerburgh, and Veldkamp (2014) looks at how fund managers add value to clients, and Van Nieuwerburgh and Veldkamp (2010) match portfolio selection patterns. Rational inattention is used in Maćkowiak and Wiederholt (2011) along with limited liability and strategic complementarity in actions to get inattention to rare events. Kacperczyk, Nosal, and Stevens (2014) show how information collection can endogenously lead to different levels of participation in markets, causing cap-
ital income inequality. I add to this burgeoning literature by using elements of inattentive frameworks to address issues of uncertainty persistence.

The mechanism of the model in this paper generates autoregressive-style beliefs off of i.i.d. underlying data. Thus, the paper could be thought of as an endogenous modeling of ARCH or GARCH processes. Surveys of the literature in these areas are provided by Bauwens, Laurent, and Rombouts (2006) and Bollerslev, Engle, and Nelson (1994).

This paper can also converse with the literature on self-exciting processes: papers such as Aït-Sahalia, Cacho-Diaz, and Laeven (2015) and Aït-Sahalia, Laeven, and Pelizzon (2014) construct models in which large jumps in one market can make jumps more likely in the same market as well as in other markets. This paper can deliver similar results in beliefs, though with an i.i.d. underlying process.

2 Rare Events and Uncertainty

I present a simple model in this section, which formally introduces the underlying mechanism through which rare events increase uncertainty.

2.1 Setup

The simple model has three types of agents, one period with three stages (0, 1, and 2), and one asset. The agents are noise traders, informed traders, and market makers. Noise traders function as hedgers - investors who are not sensitive to prices. Informed traders function as speculators - investors who have an opinion over the future path of an asset’s price and seek to profit from it. Market makers are intermediaries through which noise traders and informed traders trade. In this model, the market maker provides a friction that prevent buyers and sellers from interacting directly with each other over the course of the period. The model focuses on how informed traders collect information in an imperfectly informative market and how they use that information to exploit market prices.

The simple model includes only one period, which allows clear identification of the channel through which an unexpected event can lead to increases in uncertainty. Once the mechanism for uncertainty is established, I will introduce dynamics in later sections to explore
uncertainty persistence.

2.1.1 Asset

The asset has a stage 2 value that is comprised of two random variables: \( B \), the value of which is revealed in stage 1, and \( \eta_B \), the value of which is revealed in stage 2. \( B \) can be thought of as the ‘event’ - a price change for which investors can prepare. \( B \) is realized and publicly revealed in stage 1, when it can take one of two values: \( H \) or \( L \). In stage 0, before \( B \) is revealed, all agents have a common prior distribution on the values of \( B \): \( P(B = H) = \pi_H \) and \( P(B = L) = 1 - \pi_H = \pi_L \). If it is ex-ante unlikely for \( B \) to take the value \( L \), then \( \pi_L \) will be small, and if \( B \) were revealed to equal \( L \) in stage 1, that could be considered a rare event.

If \( B \) can be thought of as the ‘event,’ \( \eta_B \) can be thought of as the ‘reaction’ - a price change that occurs upon the value of \( B \) realizing. Just as \( B \) indexes \( H \) and \( L \), \( \eta_B \) indexes two random variables: \( \eta_H \) and \( \eta_L \). These two random variables are independent of one another, and all agents know that each is distributed \( P(\eta_B = 1) = P(\eta_B = 0) = 0.5 \). When \( B \) realizes in stage 1, the relevant variable \( \eta_B \) is determined, but is not revealed until stage 2. Only one of \( \eta_H \) and \( \eta_L \) will realize, as \( B \) will take the value \( H \) or \( L \) in stage 1. The structure of the asset is best seen in Figure 2.

\[ \begin{align*}
0 & \xleftarrow{\pi} \quad 1 - \pi \\
B = H & \xrightarrow{0.5} \quad \eta_H = 1 \\
B = L & \xrightarrow{0.5} \quad \eta_L = 1 \\
& \quad \eta_H = 0 \\
& \quad \eta_L = 0
\end{align*} \]

Figure 2: Structure of the Asset
The possible values of $\eta_H$ and $\eta_L$ are independent of one another. Although this assumption is not meant to be a pure reflection of reality in a financial market, it is used here to illustrate that different macroeconomic events may affect asset prices through different and not necessarily dependent channels. It is not important what the relative value of $H$ and $L$ are; rather it is important that $\eta_H$ and $\eta_L$ are independent so that collecting information about one does not help identify the value of the other.

2.1.2 Information

There are two types of information in this model: private and public. In stage 0, informed traders can pay to invest in information about each $\eta_B$. The investment pays off by giving agents information about $\eta_B$ before it is revealed in stage 2, allowing them to trade advantageously with respect to the asset’s terminal value: $B + \eta_B$. Investment in a particular value of $\eta_B$ is advantageous only if that value of $B$ realizes, and provides no benefit if a different value of $B$ realizes. For example, investing in information about $\eta_L$ will prove useless if $B = H$. Therefore, agents must choose whether to be informed about $\eta_H$, $\eta_L$, both or neither, knowing that only one informational investment will actually pay off. In stage 1, agents who invested in the correct $\eta_B$ will receive a perfectly informative private signal about $\eta_B$’s value before it is publicly revealed in stage 2. This information allows them to trade advantageously against the market makers, who only have access to public information.

Public information about $\eta_B$ in this model is exogenously given and constant across values of $B$. In stage 1, a public signal about whether $\eta_B$ is 1 or 0 is revealed to all agents and has accuracy $\beta \geq 0.5$. This means that $P(\text{public signal} = 1|\eta_B = 1) = P(\text{public signal} = 0|\eta_B = 0) = \beta$. In later sections this assumption will be relaxed, and public information will be allowed to vary across states.

Intuitively, allowing agents to purchase signals about $\eta_B$ mirrors traders doing research on conditional trading decisions for potential events. Theoretically, this assumption is motivated by the inattention literature. For tractability, the full model of inattention is not treated here. A more rigorous treatment of inattention with the full entropy formulation of Woodford (2012) and risk averse agents has similar results.
2.1.3 Trading Mechanics

Traders inhabit a unit continuum, a fraction $T > 0$ of which are noise traders. The noise traders are defined by their behavior: Regardless of prices, half of the noise traders always buy one unit of the asset and half always sell one unit of the asset. Their sole purpose is to provide liquidity, and without them, the results of the simple model would be degenerate. The model’s results are not very sensitive to the behavior of the noise traders, and the choice described above is the simplest.

The remaining $1 - T$ traders are risk-neutral informed traders. Informed traders want information because information allows them to update their private valuations of the asset and to trade advantageously with the market makers. In order to acquire information, each trader can purchase a signal at cost $c$ in stage 0. A signal gives perfect information about $\eta_B$ in stage 1 to the trader who bought it. If no signal is purchased, no information is gained. After a trader has received her signal, she can choose to buy one unit of the asset, sell one unit of the asset, or abstain from trade in stage 1. I limit asset purchase decisions to the set \{-1, 0, 1\} and employ risk-neutrality for simplicity. If the traders were risk-averse and had free volume choices it would not significantly change the results of the model.

It is important to note that because of this risk-neutrality, there is no asymmetry between ‘good’ events and ‘bad’ events. The values of $B$ are irrelevant for traders’ decisions. All that matters is the ex-ante likelihood of those values. This paper does not attempt to include this asymmetry, because the model is pushing the informational mechanism alone to see how much it can deliver. For reasons of intuition or applicability, it might be helpful to think of the distribution of $B$ as being risk-neutral, or risk-weighted probabilities.

The most important assumption, however, is that of the cost structure. I assume that it is equally difficult to collect information about conditional price movements (the $\eta$s) no matter how likely the state of the world. This is a strong but simple assumption, and it improves tractability. Realistically, it may seem more intuitive to say that it is harder to collect information about low-probability states than it is to collect information about normal, or high-probability states. Such an alteration, would only strengthen the results of the model, as it would further reduce the incentive to learn about unlikely events. The
cost structure that would weaken my results would make it cheaper, or easier to learn about low-probability events, but such a structure does not seem intuitive.

The **market makers** broker trades among noise and informed traders. They are perfectly competitive. Market makers observe public information and set a bid and an ask. A bid is the price at which the market maker is willing to buy the asset, and an ask is the price at which the market maker is willing to sell the asset. Conversely, the bid is the price at which traders can sell, and the ask at which traders can buy. All trade must go through market makers - agents cannot trade directly with one another. In reality, most financial trade is carried out through intermediaries, so this assumption is not unrealistic. The particular mathematical formulation of the market making sector I use in this model is based off Glosten and Milgrom (1985). Their structure allows this paper to analyze the effects of changes in public and private information. This type of sector will preserve informational asymmetries and prevent price revelation. Other types of models of trade, such as those that include search frictions could also be employed with this mechanism.

Figure 3 shows the order of events of the simple model. In stage 0, agents decide whether or not to purchase signals. In stage 1, $B$ realizes and is publicly revealed. Agents receive their private signals, and the public signal is revealed. The market making sector sets a bid and an ask, and agents trade. In stage 2, $\eta_B$ is publicly revealed and agents receive their payoffs.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Traders choose which signals to buy.</td>
</tr>
</tbody>
</table>
| 1     | $B$ observed.  
|       | Public signal about $\eta_B$ revealed.  
|       | Market Making sector sets bid and ask.  
|       | Traders trade on private signals about $\eta_B$. |
| 2     | $\eta_B$ publicly observed.  
|       | Agents receive profit or loss. |

*Figure 3: Timeline of within-period actions and decisions*
2.2 Solving the Model

The model is solved backwards. Stage 2 is trivial, as it is merely an accounting exercise. The problems of interest are in stage 1 and stage 0.

2.2.1 Pricing Problem

Consider the market makers’ stage 1 problem once the public signal has been revealed and private signals have been seen by informed traders. Market makers see that \( B \) has been revealed, so the terminal value of the asset will be \( B \) or \( B + 1 \). The zero-profit conditions, as proven by Glosten and Milgrom (1985) are:

\[
P(\eta_B = 1|\text{buy order and MM Info}) + B = \text{ask}
\]

\[
P(\eta_B = 1|\text{sell order and MM Info}) + B = \text{bid}
\]

Since the public signal could be either 1 or 0, there are four potential prices that could be set (two bids and two asks). These prices depend on the fraction \( s \) of informed agents that chose to buy the relevant signal in stage 0:

\[
\text{ask}_1 = B + \frac{\left( \frac{T^*(s)}{2} + (1 - T^*(s)) \right) \beta}{\frac{T^*(s)}{2} + (1 - T^*(s)) \beta} \tag{1}
\]

\[
\text{bid}_1 = B + \frac{T^*(s) \beta}{\frac{T^*(s)}{2} + (1 - T^*(s)) (1 - \beta)} \tag{2}
\]

\[
\text{ask}_0 = B + \frac{\left( \frac{T^*(s)}{2} + (1 - T^*(s)) \right) (1 - \beta)}{\frac{T^*(s)}{2} + (1 - T^*(s)) (1 - \beta)} \tag{3}
\]

\[
\text{bid}_0 = B + \frac{T^*(s) (1 - \beta)}{\frac{T^*(s)}{2} + (1 - T^*(s)) \beta} \tag{4}
\]

where \( T^*(s) = \frac{T}{T+s(1-T)} \). \( T^*(s) \) can be thought of as the effective fraction of noise traders because \( 1 - s \) informed traders will not buy a signal and will abstain from trade.

The first two prices - \( \text{ask}_1 \) and \( \text{bid}_0 \) - correspond to those the market making sector would set if the public signal were a 1 and the second two correspond to those the market making
sector would set if the public signal were a 0.\footnote{If $\beta = 0.5$, then the public signal is completely uninformative and the two asks and the two bids would equal one another.} All four prices will always lie between $B$ and $B + 1$ by definition. Further the ask will always be weakly larger than the bid (otherwise there would be an arbitrage opportunity). The difference between the bid and the ask is called the bid-ask spread, and is an indication of how much the market making sector fears adverse selection.

An increase in $T^*(s)$ due to fewer informed traders buying signals would narrow the bid-ask spread because a higher percentage of noise traders reduces the adverse selection problem faced by market makers. If no traders purchase signals, there would be no adverse selection for market makers to protect against, and spreads would go to zero. Conversely, because $T > 0$, all bids and asks will lie strictly between $B$ and $B + 1$, which means that perfectly informed traders have an incentive to buy or sell.

Knowing how prices will be calculated in stage 1, the signal acquisition problem of stage zero may be discussed.

### 2.2.2 Traders’ Signal Problem

Each trader faces an identical problem in stage 0. Conditional on the strength of $\beta$, the trader decides whether purchase a signal for each $\eta_B$. The cost of acquiring a signal is fixed at $c$ for all values of $B$. The expected benefit of acquiring a signal for $B$, given values $T$ and $\pi_B$ is:

$$
\pi_B \left\{ \beta (1 - \text{ask}_1(s)) + (1 - \beta)(1 - \text{ask}_0(s)) + (1 - \beta)(\text{bid}_1(s)) + \beta(\text{bid}_0(s)) \right\} \quad (5)
$$

where an individual trader takes $s$ as given in making her decision. Recall that $\pi_B$ reflects the probability that the state for which the signal is bought will occur and thus the probability that the signal will be useful. $\beta$ represents the probability that the public signal will be correct. If traders see through their private signal that $\eta_B = 1$, then they will buy at the ask, which is lower than $B + 1$; if they see that $\eta_B = 0$, they will sell at the bid, which is
higher than $B$. Rearranging, and substituting terms, we can see that the trader will choose:

$$
\max \left\{ 0, \pi_B \beta_B (1 - \beta_B) \frac{T^*(s)}{2} \left( \frac{T^*(s)}{2} + (1 - T^*(s))(1 - \beta_B) \right) \right\} - c \right\} \quad (6)
$$

The above expression shows that the agent is more likely to acquire a signal if, all else being equal, $\pi_B \uparrow$, $c \downarrow$, $T^*(s) \uparrow$ or $\beta \downarrow$. As $\pi_B$ increases, there is a higher likelihood that the agent will benefit from identifying $\eta_B$, and as a result, agents will buy signals for higher likelihood states. Although simple, this model shows that less likely states are less likely to attract informed agents, as the cost of acquiring a signal in a state is not correlated with the probability of that state occurring. Similarly as the cost of purchasing a signal decreases, more agents will invest in information generally. As $T^*(s)$ increases, bid-ask spreads will decline due to a reduction in the adverse selection problem, but the informativeness of the price will not change. Therefore, the expected profit of being informed will increase, as the prices at which agents can trade are more attractive. As $\beta$ increases, market makers can still set tighter spreads, as the level of adverse selection has gone down, but now prices are more accurate. Therefore information is less valuable to traders and so a signal is less likely to be purchased. This means that public and private information are strategic substitutes.

An equilibrium, given a value of $\beta$, is defined as a set of prices $\{\text{ask}, \text{bid}\}$, and a purchase decision $Pur \in \{1, 0, -1\}$ that satisfy equations (1)-(4), and a set of purchase decisions $\{\text{buy}, \text{don’t buy}\}$ by the measure $(1 - T)$ informed traders that satisfy equation (6) in stage 0.

2.3 Key Predictions

Given the equilibrium definitions, there are two key propositions:

**Proposition 1.** $s^*$, the equilibrium fraction of informed traders who purchase a signal, is non-decreasing in $\pi_B$.

**Proposition 2.** For any given $c$, there is a sufficiently low $\pi_B \in (0, 1)$ such that all traders will not buy a signal for state $B$. 

15
All proofs of propositions and corollaries are in Appendix A. These results follow directly from equation (6), and illustrate the key mechanism of this section: as the probability of a state goes down, agents are less informed on average upon its occurrence. Further, there is always a sufficiently unlikely state such that all agents will be uninformed when it occurs.

This model was set up to uncover the mechanism under which low probability events can trigger spikes in uncertainty. As is evident, private signals are more likely to be purchased when the probability of the state increases. These propositions show that when rare events occur, uncertainty increases through an informational investment mechanism. The structure that drives this result is that the value of investment varies with the ex-ante probability of the state, but the cost of investment does not. As a result, rare events result in poor signal quality, lower levels of information, and uncertainty.

3 Rare Events and Persistence

I will now extend the simple model to multiple periods to show how uncertainty can persist endogenously. Informational investment is the inter-temporal choice variable, and will be the only inter-temporal aspect of the model. I will not consider consumption smoothing, investment adjustment costs, or other such methods. As seen in the previous section, investment in information can affect trading decisions today. As we will see in this section, information investment today can also affect informational investment decisions tomorrow. This multi-period model will test how much an informational investment mechanism can cause uncertainty to persist.

3.1 Dynamic Model

Suppose that the model described in the previous section runs for many periods, each of which has three stages to match the initial setup. I will assume that informational investment in time $t$ impacts not only agents’ beliefs over $\eta_B$ in time $t$, but also agents’ beliefs over $B$ in time $t + 1$. In order to find the distribution of $B_{t+1}$ given $B_t$, $\eta_{B,t}$ and choices by private traders, assume that $B_{t+1} = B_t + \eta_B + u_{t+1}$. That is, the value of $B$ in period $t + 1$ is simply the terminal value of the asset in period $t$, plus a noisy variable $u$. This variable
exists so that, at the outset of each period, agents do not know which value $B$ will take, and the structure of informational investment can be used again. A trader’s information investment decision in one period is affected by the informational investment decision of the period before. As before, the underlying process of the asset is not dependent on decisions by agents, but the distributions agents have over the variables that define the asset are dependent on the agents’ decisions.

For the rest of the paper, I will simplify the signal acquisition process. Instead of heterogeneous agents making individual decisions to acquire information, I will use a representative agent who either acquires a signal or does not. This agent could be thought of as a single trader purchasing information and then commanding power in the market equal to a share $1 - T$ of trade, or as a way to enforce a symmetric decision-making among a continuum of traders. This simplification has no effect on the utility of traders in the static or dynamic version of the model, and it eliminates additional unstable equilibria in the dynamic version. Since characterizing the nature of these equilibria is not a priority of this paper, introducing a representative agent allows focus on the result of changes in signal quality across states rather than the result of changes across agents within a state.

As with $\eta_B$, the agent can receive a signal about $u$ with an accuracy determined by the signal acquisition decision. $\pi_{t+1} = 0.5$ if the agent does not buy a signal and $\pi_{t+1} = k > 0.5$ if the agent does buy a signal. To simplify notation, I will define $\pi_t \equiv \max\{\pi_H, \pi_L\}$. The representative trader now needs to solve the following problem given a value of $\beta$:

$$V(\pi, 1 - \pi) = \sum_\pi \max\left\{ \frac{\pi \beta (1 - \beta) T}{\left( \frac{T}{2} + (1 - T)(1 - \beta) \right) \left( \frac{T}{2} + (1 - T) \right)} - c + \delta \pi V(k, 1 - k), \delta \pi V(0.5, 0.5) \right\}$$

(7)

In the multi-period case, there are two effects: Selecting a signal impacts trade of the asset in $t$, but it also impacts information collection in $t + 1$. Now, a signal is more valuable, as it has long term informativeness as well as short. I have deliberately separated the long term and short term effects of signal acquisition, although they could have been combined without much trouble. By separating them, we can see that reduction of noise in $\eta$ is useful to agents because it allows them to trade advantageously against market makers, and reduction of noise in $u$ is useful because it allows agents to collect information more efficiently going
forward.

The definition of a **dynamic equilibrium** for a given $\beta$ and $k$ in stage 0 is the representative trader’s choice to buy a signal or not to satisfy equation (7) and a choice of prices by market makers to satisfy equation (1)-(4).

### 3.1.1 Information Over Time

The multi-period model produces uncertainty persistence because a trader’s lack of information in one period carries forward into future periods. This is best shown through the step-wise incentive structure that informed traders face for different costs of information. At the extremes, if information is cheap and $c$ is close to 0, traders will always purchase a signal, as it costlessly provides information. Conversely, if information is exorbitantly expensive and $c$ goes to infinity, traders will never purchase a signal, and $V(\pi, 1 - \pi) = 0$ for all $\pi$. In between, there are two intermediate thresholds:

1. Purchase signals for state $B$ if $\pi_B = k$ and if $\pi_B = 0.5$ but not if $\pi_B = 1 - k$. In this case, the agent will be perfectly informed until the unlikely state occurs, at which point they will be uninformed. Recall that when no signal is purchased, $\pi_{t+1} = 0.5$. Once they reach this state, they will purchase signals for both (now equiprobable states in $t + 1$) and uncertainty will last for only one period.

2. Purchase signals for state $B$ if $\pi_B = k$ but not if $\pi_B = 0.5$ or $\pi_B = 1 - k$. In this case the agent will be perfectly informed until the unlikely state occurs, at which point they will be *permanently* uninformed. This type of persistence is a stark version of what will be developed in the next section.

**Proposition 3.** For any $\beta$ and any $\pi_t \in (0, 1)$, there is a sufficiently high value of $c$, such that the private trader will not purchase a signal and $\pi_{t+1} = 0.5$.

**Corollary 1.** There is a threshold $\overline{c}$ such that if $\pi_t = 0.5$ and $c > \overline{c}$, $\pi_{t+1} = 0.5$.

**Proposition 3** shows that for sufficiently high costs, or equivalently, for sufficiently rare events, signal quality is poor, and the ex-ante distribution of $B_{t+1}$ will be disperse. The corollary shows that there is a cost such that the agent will not choose to buy a signal for
Thus, for certain threshold cost, if an event occurs with a very low probability \((1 - k)\), the agent will be permanently uninformed thereafter.

Figure 4 illustrates the function described in this section. The \(x\)-axis corresponds to periods of time. The \(y\)-axis is the degree of uncertainty faced by agents over \(\eta_B\). The red vertical lines correspond to events that occur with a low ex-ante probability. In the first panel, the agent purchases signals when \(\pi_t = 0.5\) and when \(\pi_t = k\), but not when \(\pi_t = 1 - k\). When rare events occur, the agent is uncertain, faces a uniform distribution in the next period, and is able to purchase signals to guarantee certainty in the following period. In the right panel, the agent only purchases signals when \(\pi_t = 1 - k\). An initial rare event makes the agent uncertain, and when she faces a uniform distribution, she can no longer buy signals, and is trapped in uncertainty forever.

(a) Temporary Uncertainty for Low Information Costs  
(b) Uncertainty Persists for High Information Costs

\hspace{2cm} Figure 4: Uncertainty Reactions in Dynamic Model

Intuitively, the nature of the model’s trap can be related to the real world as follows: one could imagine an investor trying to make informational investment decisions about sporting events in a tournament in order to place bets. Sporting events have binary outcomes, so it is a natural point of comparison to the simple model. Suppose that early in the tournament, a low quality team plays a high quality team. The high quality team is expected to win, so an investor spends more time researching and simulating the high quality team’s chances in potential future games. If the low quality team wins, the investor will not have as much information about how the low quality team will perform in future games and may not have the time or energy to do additional research. As a result, the investor will not have strong enough opinions to place bets for the subsequent round and as a result, will not do research about either team to predict the outcomes of the rounds that follow. The investor might
choose, after an initial upset, not to place any additional bets for the rest of the tournament - the effort it would take become informed about future outcomes outweighs the potential benefit of making informed bets.

There are two purposes to information. Collecting information about $\eta$, allows traders to make money in the financial market. Collecting information about $B$ allows traders to collect information more accurately in future periods. This problem could be shortened. It could be that $\eta_t \equiv B_{t+1}$, which means that agents would simultaneously trade in the financial market and allocate attention for the next period. I have separated this into two variables to emphasize the dual nature of information in this model. One could also allow the agent to collect information about $\eta$ and $B$ separately (instead of together, as it currently is formulated). Doing so changes none of the qualitative results, as agents will always rationally favor collecting information in likely states as opposed to unlikely states.

The starkness of the result of this simple model - that uncertainty can be a trap that lasts forever - can be attributed to the binary nature of the simple model. In an all-or-nothing investment framework, it is possible to get permanent non-investment. The simple model shows that agents will choose to be uninformed if their distributions are sufficiently uniform - if there are too many events that could potentially occur, it is not worth paying attention to any single one. As we will see in the following section, once informational investment decisions are made freely, uncertainty will persist, but it will not become permanent.

4 General Setup

The simple model generates spikes and persistence in uncertainty, but it has stark predictions, due to the all-or-nothing form of informational investment. This section loosens the assumptions of the model to allow for more nuanced predictions. The analytical results of Section 2 continue to hold. Additionally, the assumptions of Section 3 are loosened, and instead of uncertainty traps that last forever, the general model delivers uncertainty persistence. Relaxing the structure of informational investment, the state space, and the nature of public information allows for a richer set of predictions about dispersion of beliefs and asset-demand, as well as implications for volatility and bid-ask spreads, which are illustrated
with some simulations.

4.1 Structure

The structure of the economy is the same as the previous section. As before, there are many periods. Within each period there are again three stages. In stage 0 the representative agent decides how accurate future signals should be. In stage 1, the public signal of the asset’s value is revealed, prices are set, and financial trade occurs. In stage 2, the final value of the asset is revealed and agents purchase and consume.

4.2 Asset

There are, as before, two components of the asset, $B_t$ and $\eta_{B,t}$, and the terminal value of the asset in stage 2 is $V_t = B_t + \eta_{B,t}$. $B_t \sim \mathcal{N}(\mu_{B,t}, \sigma_{B,t}^2)$ and $\eta_{B,t} \sim \mathcal{N}(0, \sigma_{\eta}^2)$ for each $B$. In stage 0, neither $B_t$ nor $\eta_{B,t}$ is known, but $\mu_{B,t}, \sigma_{B,t}^2$, and $\sigma_{\eta}^2$ are all known. In stage 1, $B_t$ is publicly revealed, and private and public signals about the realization of $\eta_{B,t}$, with accuracy $\sigma_{\gamma,t}^2$ and $\sigma_{\beta,t}^2$, respectively, are gathered as well. In stage 2, $\eta_{B,t}$ is revealed and the gains and losses are realized. The data-generating process is a random walk: $\mu_{B,t+1} = V_t + \epsilon_{t+1}$, where $\epsilon_{t+1} \sim \mathcal{N}(0, \sigma_{\epsilon}^2)$. At the beginning of stage 0 of period $t + 1$ a signal $u_{t+1}$ is revealed that is distributed $\mathcal{N}(\epsilon_{t+1}, \sigma_{u,t+1}^2)$ where $\sigma_{u,t+1} \propto \sigma_{\gamma,t}$. One advantage of generalizing the asset’s structure is that the model can view forecasts of the asset’s value as continuous distributions, which allows a quick relation, empirically, to the VIX.

The model will again be solved backwards, and stage 2 is again trivial.

4.3 Stage 1 Problems

$\sigma_{\beta,t}(B_t)$ is the accuracy of the public signal about $\eta_{B,t}$ conditional on a realization of $B_t$, and $\sigma_{\gamma,t}(B_t)$ is the accuracy of private signals about $\eta_{B,t}$ for every $B_t$. Then the ex-post public mean and variance over $\eta_{B,t}$ after seeing a public signal $x$ are $\mu_{p,t} = \frac{1}{\sigma_{\gamma,t}^2 + \sigma_{\beta,t}^2} x$, and $\sigma_{p,t}^2 = \frac{1}{\sigma_{\gamma,t}^2 + \sigma_{\beta,t}^2}$.
4.3.1 Traders

Traders have access to private and public information and, as they are risk neutral, will choose to buy one unit of the asset if their private mean lies above the ask and will sell one unit of the asset if their private mean lies below the bid. One can think of this problem in two different ways: first, that the representative trader receives one signal and is able to make a choice between buying \(1 - T\) units, selling \(1 - T\) units, or abstaining; or second, that a continuum of \(1 - T\) traders each make individual signals based off of idiosyncratic signals purchased on their behalf by the representative trader in stage 0. The two cases are mathematically equivalent ex-ante to risk-neutral agents.

\[
U(\text{public signal}, \text{private signal}) = \max \{ (\text{bid}_t - x), (x - \text{ask}_t), 0 \} \tag{8}
\]

Where \(x = E[V|\text{public signal}, \text{private signal}] = \frac{\frac{\mu_{p,t}}{\sigma_{p,t}^2} + \text{private signal}}{\frac{\sigma_{p,t}^2}{\sigma_{p,t}^2} + \frac{\sigma_{\gamma,t}^2}{\sigma_{\gamma,t}^2}}.\) Traders’ decisions in stage 1 are independent of those in stage 2, and are not inter-temporal.

4.3.2 Market Makers

Market makers have access to all public information about the asset and also know the function \(\sigma_{\gamma}(B)\). The dispersion of private beliefs (assuming the signal is purchased on behalf of a continuum of investors) can be calculated as \(\sigma_{\text{disp},t}^2 = \frac{\frac{1}{\sigma_{\gamma,t}^2} + \frac{1}{\sigma_{p,t}^2}}{\sigma_{p,t}^2 + \sigma_{\gamma,t}^2}.\) Conditional on a realization of \(B\), a public mean \(\mu_p\), a public variance \(\sigma_{p,t}^2\), and a dispersion of beliefs \(\sigma_{\text{disp},t}^2\), the expected profits for a choice of a bid and an ask by the market making sector are:

\[
\Pi_{\text{ask}} = \int \phi_{p,t}(x) \left( (1 - \Phi_{\text{disp},x}(\text{ask}))\left(1 - T\right) + \frac{T}{2} \right) (\text{ask} - x) dx \tag{9}
\]

\[
\Pi_{\text{bid}} = \int \phi_{p,t}(x) \left( \Phi_{\text{disp},x}(\text{bid})\left(1 - T\right) + \frac{T}{2} \right) (x - \text{bid}) dx \tag{10}
\]

where \(\phi_{p,t} \sim \mathcal{N}(\mu_{p,t}, \sigma_{p,t}^2)\) and \(\phi_{\text{disp},x} \sim \mathcal{N}\left(\frac{\mu_{p,t}}{\sigma_{p,t}^2} + \frac{x}{\sigma_{p,t}^2 + \sigma_{\gamma,t}^2}, \frac{\sigma_{\gamma,t}^2}{\sigma_{p,t}^2 + \sigma_{\gamma,t}^2}, \sigma_{\text{disp},t}^2\right).\) Perfect competition means that the bid and the ask must be selected to set the profits above to zero.
4.4 Stage 0 Problems

In stage 0, conditional private signals are selected.

4.4.1 Traders

The asset’s value follows a random walk process both within and between periods. Within a period, the increment is \( \eta_t \) and between periods it is \( \epsilon_t \). The noise with which the public signals about \( \eta \) and \( \epsilon \) are observed are proportional to one another. Traders take \( \sigma_\beta(B) \) as given, and pick \( \sigma_\gamma(B) \) to maximize expected future profits subject to a proportional cost \( c \):

\[
V(\mu_{B,t}, \sigma_{B,t}^2, \sigma_{\beta,t}^2) = \int \max_{\sigma_{\gamma,t}^2(B_t)} \phi(B_t) \int \int p(\text{public signal} = x) (p(\text{private signal} = y) U(x, y) dy) dx \\
+ \delta \left[ \int \int \phi(B_t) V(\mu_{B,t+1}(u_{t+1}, \eta_t), \sigma_{B,t+1}^2, \sigma_{\beta,t+1}^2) dud\eta \right] - c \nu(\sigma_{\gamma,t}^2(B_t)) dB_t \tag{11}
\]

where \( \phi \sim \mathcal{N}(\mu_{B,t}, \sigma_{B,t}^2) \), \( u_t \) is the signal about \( \epsilon \), \( \mu_{B,t+1}(u_{t+1}, \eta_t) = \frac{\sigma_{\epsilon,t+1}^2 u_{t+1} + \sigma_{\epsilon,u,t+1}^2 (B_t + \eta_t)}{\sigma_{\epsilon,t+1}^2 + \sigma_{\epsilon,u,t+1}^2} \), \( \sigma_{B,t+1}^2 = \frac{\sigma_{\epsilon,t+1}^2 \sigma_{\epsilon,u,t+1}^2}{\sigma_{\epsilon,t+1}^2 + \sigma_{\epsilon,u,t+1}^2} \), and \( \nu \) is a convex decreasing function such that \( \frac{\partial^2 \nu}{\partial \sigma_{\beta}^2} > \frac{\partial^2 U}{\partial \sigma_{\beta}^2} \) everywhere.

4.5 Propositions

**Proposition 4.** \( \frac{\partial U}{\partial \sigma_{\gamma}} < 0 \).

**Corollary 2.** For any given \( c, \forall \delta > 0 \), if \( \sigma_\beta(B) = 0 \), then \( \frac{\partial \sigma_\gamma(B)}{\partial \sigma_\beta(B)} \leq 0 \).

This result confirms the intuition from Proposition 2 holds: attention is a good, albeit a costly one. As such, more attention is paid to higher probability states than to low probability states. Therefore, lower probability events reduce identification and increase uncertainty.

**Corollary 3.** For any given \( c, \forall \delta > 0 \), if \( \sigma_\beta(B) = 0 \), then for any two points \( B_1 \) and \( B_2 \), such that wlog, \( \phi_{B,t}(B_1) > \phi_{B,t}(B_2) \). Then at time \( t + 1 \), \( E_{B_1}[\phi_{B,t+1}(B)] > E_{B_2}[\phi_{B,t}(B)] \).

The rarer the event that occurs, the longer the subsequent uncertainty is likely to last. There is an upper bound on how uncertain the agents can be, as \( \sigma_\eta < \infty \) and \( \sigma_\epsilon < \infty \), so even though \( \sigma_\gamma \) and \( \sigma_\beta \) are unbounded, the variance of the posterior distributions are bounded.
4.5.1 Public Entity

The simple model showed that public and private information are substitutes. Therefore, it would be of interest to find out whether deteriorating public information changes private traders’ incentives. To analyze this, I introduce a public entity who chooses the quality of public information to maximize a given objective function (the particular function is not very important, so long public information is treated as a good).

The public entity works to select conditional signal quality, \(\sigma_{\beta}\) for each potential value of \(B_t\) to maximize expected accuracy:

\[
\int \min_{\sigma_{\beta}(B)} \phi_{B,p,t}(B) \left( \int p(\eta-\text{signal} = x)\sigma_{p}^2 dx \right) - c_p\nu_p(\sigma_{\beta}(B))d\phi(B) \tag{12}
\]

where \(\nu_p\) is a convex decreasing function that satisfies the condition that \(\frac{\partial^2 \nu_p}{\partial \sigma_{\beta}^2} > \frac{\partial^2 \sigma_{\beta}^2}{\partial \sigma_{\beta}^2}\) 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 \(\{\mu_{B,t}, \sigma_{B,t}^2, c, c_p, \sigma_{\eta,t}^2, \sigma_{\epsilon,t}^2\}\), a dynamic equilibrium is defined by a choice of \(\sigma_{\beta}\) by the public entity that solves equation 12 a choice of a policy function \(\sigma_{\gamma}(\mu_B, \sigma_B^2, \sigma_{\beta})\) by the traders that satisfies equation 11, individual decisions to buy, sell, or abstain by traders to solve equation 8, and a choice of a bid and an ask by Market Makers to solve equations 9, given observed values of \(\{\sigma_{\beta,t}, \sigma_{\gamma,t}, \sigma_{\eta,t}\}\), a public signal, and a set of private signals distributed \(\mathcal{N}(\eta_t, \sigma_{\eta,t}^2)\).
4.6 Predictions and Spillovers

Figure 5 shows the results of the previous propositions. In this snapshot,\(^6\) one can see how agents (both public and private) choose to invest in information. The \(x\)-axis is the ex-ante probability of different occurrences \(B\). Then, conditional on these selections, one can see what the expectations for demand for financial assets, bid-ask spreads, uncertainty, volatility, and dispersion of beliefs are for each potential value of \(B\). All variables are normalized to start at 1.

![Figure 5: Simulation of Continuous State Model](image)

As is expected by the propositions above, both \(\sigma_\beta\) and \(\sigma_\gamma\) are decreasing, with better signal quality for high probability states and worse signal quality for low probability states.

The results of these attentional choices are shown in panel (b). Most importantly here, uncertainty and volatility both spike for tail events. But it is important to note that bid-

---

\(^6\) The parameter values here are \(\mu_{B,t} = 0, \sigma_\eta = 10, \sigma_\epsilon = 10, c_\rho = 0.4, T = 0.35, \delta = 0.95, c = 25,\) and \(\nu_p(x) = \nu(x) = \frac{1}{2}\). \(\alpha\) and \(\rho\) are selected with the calibration of section 5 in mind, while \(\sigma_\eta\) and \(\sigma_\epsilon\) are set equal to indicate that the uncertainty from period to period is not more or less than stage to stage. The rest of the parameters will deliver similar results.
ask spreads, and dispersion of beliefs also are higher for tail events, and are comparatively lower for higher probability events. This bears out the first hypothesis in the introduction - that attentional choices, or investment in information ex-ante can lead to spikes in second moments that are reflected in a financial setting. Similarly panel (c) shows that volume of trade, or balance sheet size, or leverage, plummets for tail events as investors are less well informed in those states and therefore less likely to want to take on risk.

4.7 Persistence

The above was an illustration of the comparative statics. Consider now the effect of the persistence mechanism. In Figure 6 the asset’s value takes a three standard deviation drop in period 2 and then immediately rebounds in period 3 to its prior level, where it then remains for several periods.

![Figure 6: Impulse Responses of Continuous State Model to a three standard deviation drop and rebound in the asset’s value.](image)
There are several aspects of this figure of note: first, the large drop in the asset’s value corresponds to a low-probability event ex-ante, and therefore large increases in all the financial variables (drop in asset demand); second the rebound is just as large in size as the drop, but due to the flattening of the distribution in period 2, it actually lowers uncertainty somewhat - uncertainty increases agents’ ex-ante beliefs about the likelihood of large price movements; third, as the asset’s value does not move at all after period 3, the only source of persistence of the spike is the time it takes for agent’s confidence in their forecasts to regain shape (for the variance of their beliefs to decline). The last frame of the figure shows the inverse of the integral over alpha - which is a measure of the total amount of attention paid to the Bs. This spikes along with uncertainty, showing that although agents collect less information on average, they collect more information in total when they are uncertain. That is, they try harder to collect information but are, on average, worse at it.

5 Empirical Implications

The model has delivered three sets of theoretical results. In this section, I will now present a calibration and a reduced form regression as evidence in support of those results as well as empirical and theoretical tests that could serve as the basis for future work. First, a simple calibration shows how the model could be used to match movements in the VIX. Second, a reduced form regression analysis using options volume data shows how preparedness could be assessed empirically. Third, I discuss several theoretical and empirical ideas that could be developed from this paper.

5.1 A Simple Calibration

All of the variables in the model, with the exception of the costs of informational investment, have some real world parallel that they can be calibrated to. Therefore, this section will calibrate all of the variables possible and then estimate values of $c$ and $c_p$, the private and public costs of attention, respectively, that allow predicted uncertainty to come closest to the VIX.
5.1.1 Strategy

The list of variables that need to be calibrated are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{B,t}$</td>
<td>S&amp;P Index</td>
</tr>
<tr>
<td>$\sigma^2_\eta$</td>
<td>23%</td>
</tr>
<tr>
<td>$\sigma^2_u$</td>
<td>23%</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.95</td>
</tr>
<tr>
<td>$T$</td>
<td>0.38</td>
</tr>
</tbody>
</table>

In the model, movements in uncertainty are tied to the price movements in an underlying asset. The VIX reflects perceived volatility in the S&P index, so the index will serve as that underlying asset. Therefore, I set $\mu_{B,t}$ to the opening value of the S&P index each period.

Both $\sigma^2_\eta$ and $\sigma^2_u$ will be measured in percentage terms, as is the convention with the VIX.\(^7\)

Calibrating in percentage terms means that the results will not be driven by the S&P index. Since they provide an upper bound on the amount of uncertainty faced by agents, I calibrate them to be equal to the maximum level of the VIX in the period observed. This value likely overstates the true underlying volatility process, due to the lack of risk aversion and volatility premia. For the purposes of this simple exercise though, we are more concerned with matching the shape of the VIX as opposed to the exact level. The discount factor is set at 0.95, and the ratio of noise traders is taken from Wang (2002) who measures the volume and number of noise traders in foreign exchange markets. Data on this variable is difficult to come by, but thankfully its value not crucial in estimation.

Each period, the model requires a price change for the asset, and a start-of-period uncertainty for the agent. The model will then output a level an end-of-period uncertainty, which I compare to the VIX. I simulate the model in two ways. In the first, the start-of-period uncertainty will be the actual VIX, so the model runs, period-by-period. In the second, the start-of-period uncertainty is the model’s output for the previous iteration’s end-of-period uncertainty, so the model runs dynamically. I call these versions $X$ and $Y$, respectively. $X$, therefore, takes a vector of price changes, and a vector of values of the VIX as inputs, while $Y$ takes a vector of price changes, and only a starter value of the VIX for the first period as

---

\(^7\) However, unlike the VIX, the uncertainty will not be ‘annualized’ but will be period-by-period - and therefore the reported VIX will be divided by the square root of 12.
inputs. $c$ is estimated to be 25, while $c_p$ is estimated to be 0.4 for simulation $X$ and 0.8 for $Y$. The final values for the costs are not as important as the shape of the uncertainty they determine, which is qualitatively similar for many values.

5.1.2 Results

As shown in Figure 7, the paths of the simulated VIXs and the actual VIX are close to one another, and spike in the same places. The simulations spike more often than the actual VIX, but there are no points at which the VIX spikes when the simulations do not.

![Figure 7: Monthly simulation of the VIX using probabilities as input](image)

The reason that this simple calibration should be compelling is two-fold: first, that the only input into the model is price-deviations of the S&P index. Typically one would assume that many different factors would be required to simulate the VIX, but the mechanism of this paper gets pretty far with just one. The second that the only inter-temporal aspect of this model is that collecting information today, makes it easier to collect information tomorrow. There is no consumption-smoothing, habit formation, or investment adjustment
in the model. Again, instead of needing to specify an ARCH or GARCH process, the model is able to deliver persistence through a purely inter-temporal informational channel.

5.2 Options Data

The key mechanism in the model of this paper is that of inattention, or informational investment. Unfortunately, I am not aware of any direct measure for such a variable. There are proxies, however, and I consider a simple exercise with options hedging here.

5.2.1 Interpretation

One could think of informational investment or paying attention as a form of reducing conditional volatility or uncertainty. Buying options, which reduce exposure to volatility, could be seen to have a direct correlation with the mechanism utilized in the model. After speaking with several investors, including some portfolio managers and a hedge fund manager, I found a common pattern in the way in which they prepared for the unexpected was by using options to develop a long position in volatility. Therefore, by looking at the options trade data, perhaps we can answer whether the market’s aggregate position in options, or preparedness, mutes spikes in the VIX.

5.2.2 Dataset

I use data from MarketDataExpress, which provides information on the types and sizes of trades in the SPX options index at a daily frequency. For each day, the dataset provides volumes for all non market-making entities four different types of trade - open buy, open sell, close buy, and close sell - for puts and calls at every strike. An open trade is one where the entity is taking on risk, while a close trade is one where the entity is closing out an existing position.

Control variables are: volatility, which is defined as the difference between the intraday high and intraday low; SPDiff, which is the daily return on the S&P; and Volume, which is the number of shares traded in the index on a given day.
5.2.3 Strategy

Given the data above, a useful way to proxy how prepared agents are for different events would be to construct net long positions in options (either puts or calls). Given four different buckets (strikes less than 90% of the value of the underlying, strikes between 90% and 100% of the underlying, strikes between 100% and 110% of the underlying, and strikes over 110% of the underlying) one could create two variables per bucket - subtracting the number of open sells from open buys in each bucket to create open interest and subtracting the number of close sells from close buys in each bucket to assess reduction of risk. Then for each date, the sum of those buckets for all options traded before that date that expire after that date would be a measure of preparedness. I drop the first 500 days to allow for the buildup of a portfolio. Then the following regression could be estimated:

\[
vix_t = \alpha_0 + \alpha_i \{volume_{it}\} + controls + \epsilon_t
\]

The desire is to understand, controlling for changes in volatility and the underlying, etc. whether an extra unit of preparedness - an extra unit of volatility protection negatively correlated with the level of the vix. So the null hypothesis is that \( \alpha_i = 0 \) against the alternative that it is significantly negative.

5.2.4 Results

The three columns of table 1 correspond to days on which the S&P moved down, the days on which it moved up, and all days. With the exception of open long positions in options that have strikes between 100% and 110% of the underlying, every single other type of long position in options mitigates - is negatively correlated with - the level of the VIX. Controlling for this, we still see that the underlying is negatively correlated with the VIX, and that the probabilities of the intraday low and intraday high are negatively correlated as well - with the negative event having a larger impact than the positive event - maintaining the central result that ex-ante unlikely events spike the VIX. Therefore we can reject the null hypothesis, and conclude that conditional uncertainty is negatively correlated with the degree to which agents prepare for events ahead of time.
5.3 Extensions and Other Applications

The two exercises above are indicative of some empirical support for the results developed in the model. There are also other theoretical and empirical implications of the paper that I will now discuss.

5.3.1 Uncertainty vs. Risk

The first is also the most challenging, and relates the model back to the original discussion of long term learning of an economy’s parameters, and period-by-period changes in perceptions of the second-moment of a distribution. The former is a Knightian formulation of uncertainty, while the latter is closer to risk. This paper assumes that there is no Knightian uncertainty, but there is a natural question that arises from its results: How much of the fluctuations we see in indices like the VIX result from uncertainty, and how much from changes in risk-perception?

Financial data on some event-driven markets may help to answer this question - to capture and quantify the relative effects of uncertainty and risk. There is nothing in the model of this paper that precludes long-term Bayesian learning. So a natural extension would be to take a model such as the one described in this paper, and introduce long-term learning about the underlying parameters of the model, such as $\sigma_\eta$ or $\sigma_B$. Empirically distinguishing orthodox definitions uncertainty and risk would bring a lot of clarity to a fundamental question, as the two terms are often used fungibly.

5.3.2 Cross-sectional Implications

Starting with section 3 this paper considered only a representative agent when discussing uncertainty dynamics. But the implications of the paper could generalize to a heterogenous agent setup. In particular, it is worth considering the cross-sectional implications of the model. If some agents are prepared for ‘rare’ events, and others aren’t, then some agents will not only be able to profit off of those events, but also be better prepared for subsequent events, while agents that are unable to profit, will also struggle to recover.

For example, if two firms had different opinions over which way the Brexit vote would turn
out, one might have been much better prepared for the eventual vote to leave. The better prepared firm might then be able to devote time and resources to analyzing future political and economic variables, such as whether David Cameron would step down, or whether the UK would have access to the Common Market. Other, less certain firms would be less able to analyze those same variables. Financial data on mutual funds would help to distinguish the degree of persistence in performance. Funds that do well after unlikely events should continue to do well, while funds that perform poorly will continue to underperform.

5.3.3 Normative Implications

On a purely theoretical front, this paper only discussed positive implications of the model. All the agents act rationally, and there is no reason any agent should change her behavior. However, another relevant question would be how agents should pay attention, when this dynamic aspect is present. Several papers have considered this question in a static setting with a single variable of interest, where agents face an informational externality of their attentional choices. Given the larger state space of this paper and the dynamics interplay, there are two additional dimensions along which to attain normative implications.

6 Conclusion

I have presented and solved a model that describes the process of informational investment by private investors. The solution to the model leads to contingent information sets that are most accurate near the mean of a distribution and most inaccurate in the tails. Therefore, upon the realization of a tail event, uncertainty spikes. The interaction of the information sets lead to concurrent spikes in bid-ask spreads and dispersion of beliefs, as well as drops in asset demand. However, an initial increase in uncertainty results in a flatter distribution, which in turn, leads to lower levels of information investment on average. As a result, initial spikes in uncertainty can can persist for several periods before resolving.
References


Fiedler, Angela. 2012. “Fixed Costs and Inter-Industry Wage Differentials.”


Li, Dan, and Geng Li. 2014. “Are Household Investors Noise Traders: Evidence from Belief Dispersion and Stock Trading Volume.”
Linn, Matthew, Sophie Shive, and Tyler Shumway. 2014. “Pricing kernel monotonicity and conditional information.” *Available at SSRN 2383527.*


——. 2011. “Inattention to rare events.”


Table 1: Effect of price changes on the VIX

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume</td>
<td>2.11e-05***</td>
<td>1.86e-05***</td>
<td>2.10e-05***</td>
</tr>
<tr>
<td></td>
<td>(1.94e-06)</td>
<td>(2.12e-06)</td>
<td>(2.46e-06)</td>
</tr>
<tr>
<td>Open1</td>
<td>-6.49e-06***</td>
<td>-6.73e-06***</td>
<td>-4.05e-06</td>
</tr>
<tr>
<td></td>
<td>(2.12e-06)</td>
<td>(2.30e-06)</td>
<td>(2.05e-06)</td>
</tr>
<tr>
<td>Open2</td>
<td>2.63e-06</td>
<td>5.14e-06</td>
<td>2.64e-06</td>
</tr>
<tr>
<td></td>
<td>(2.88e-06)</td>
<td>(3.14e-06)</td>
<td>(3.58e-06)</td>
</tr>
<tr>
<td>Open3</td>
<td>-3.58e-05***</td>
<td>-3.40e-05***</td>
<td>-3.49e-05***</td>
</tr>
<tr>
<td></td>
<td>(1.63e-06)</td>
<td>(1.87e-06)</td>
<td>(1.89e-06)</td>
</tr>
<tr>
<td>Open4</td>
<td>-1.09e-05***</td>
<td>-9.61e-06***</td>
<td>-9.76e-06***</td>
</tr>
<tr>
<td></td>
<td>(8.94e-07)</td>
<td>(9.97e-07)</td>
<td>(1.15e-06)</td>
</tr>
<tr>
<td>Close1</td>
<td>-4.86e-05***</td>
<td>-4.40e-05***</td>
<td>-4.80e-05***</td>
</tr>
<tr>
<td></td>
<td>(2.86e-06)</td>
<td>(3.17e-06)</td>
<td>(3.53e-06)</td>
</tr>
<tr>
<td>Close2</td>
<td>-3.07e-05***</td>
<td>-2.64e-05***</td>
<td>-2.89e-05***</td>
</tr>
<tr>
<td></td>
<td>(3.21e-06)</td>
<td>(3.49e-06)</td>
<td>(4.04e-06)</td>
</tr>
<tr>
<td>Close3</td>
<td>-4.33e-05***</td>
<td>-4.15e-05***</td>
<td>-4.14e-05***</td>
</tr>
<tr>
<td></td>
<td>(2.24e-06)</td>
<td>(2.59e-06)</td>
<td>(2.59e-06)</td>
</tr>
<tr>
<td>Close4</td>
<td>-2.72e-05***</td>
<td>-2.45e-05***</td>
<td>-2.63e-05***</td>
</tr>
<tr>
<td></td>
<td>(1.47e-06)</td>
<td>(1.59e-06)</td>
<td>(2.00e-06)</td>
</tr>
<tr>
<td>SPdiff</td>
<td>-0.0185**</td>
<td>0.120***</td>
<td>-0.118***</td>
</tr>
<tr>
<td></td>
<td>(0.00724)</td>
<td>(0.0151)</td>
<td>(0.0164)</td>
</tr>
<tr>
<td>volatility</td>
<td>0.278***</td>
<td>0.202***</td>
<td>0.249***</td>
</tr>
<tr>
<td></td>
<td>(0.0155)</td>
<td>(0.0193)</td>
<td>(0.0195)</td>
</tr>
<tr>
<td>Constant</td>
<td>13.78***</td>
<td>12.95***</td>
<td>13.21***</td>
</tr>
<tr>
<td></td>
<td>(0.554)</td>
<td>(0.586)</td>
<td>(0.708)</td>
</tr>
</tbody>
</table>

Observations 4,166 2,225 1,940

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results of three regressions. Each column corresponds to a different subset of the data - these first is for the whole sample, the second is for days on which the S&P index closed up, the second is for days on which it closed down. The dependent variable is the estimated stochastic discount factor, the time frequency is daily. The dependent variable is the change in the level of the VIX - the regression is run in first differences and Newey West Standard Errors are reported. Volume is the volume of trade in the S&P index, SPdiff is the first difference of the S&P index’s value, volatility is the intraday price change, plow and phigh are the ex-ante probabilities of the intraday low and high respectively, calculated with the risk neutral distribution. The Open variables are the net positions of the market in different ranges of strikes for positions that have taken on risk. The Close variables are the net levels of positions that were closed out in different ranges of strikes.

*** indicates statistical significance at the 1% level; ** at the 5% level; and * at the 10% level.
A Proofs

Proof. Proposition (2)
It will actually be easier to start with the intuition for this proposition and then move to
Proposition 1. It is quick to see that since the benefit of acquiring a signal in equation 6 is
increasing in \( T^*(s) \), that agents’ informational choices are strategic substitutes. Therefore,
an agent receives the most benefit if she is the only one purchasing a signal (i.e. \( s = 0 \)).
Therefore, consider the maximal benefit possible: \( T^*(s) = 1 \):

\[
\pi_B \beta_B (1 - \beta_B) 2 - c
\]

We want to show that for any \( c \), there is a \( \pi_B \) such that agents will not purchase a signal.
The condition for that \( \pi_B \) is:

\[
\pi_B (1 - \beta_B) 2 - c = 0
\]
\[
\pi_B (1 - \beta_B) 2 = c
\]
\[
\pi = \frac{c}{2 \beta_B (1 - \beta_B)}
\]

This condition is necessary and sufficient for no agent to purchase a signal. Note that it is
not necessarily the case that there is a sufficiently large value of \( \pi \) for which all agents will
buy a signal. Such a result depends on the value of \( c \).

\[ \square \]

Proof. Proposition (1)
Let us first find the conditions under which all agents would purchase a signal. If all agents
purchase a signal then \( T^* = T \). The necessary and sufficient condition for all agents pur-
chasing a signal is:

\[
\frac{\pi_B (1 - \beta_B)^T}{\left( \frac{T}{2} + (1 - T)(1 - \beta_B) \right) \left( \frac{T}{2} + (1 - T) \right)} > \frac{c}{\beta_B (1 - \beta_B)^{\frac{T}{2}}}
\]

\[ \pi \]
Therefore if $\pi < \pi < \bar{\pi}$, some but not all agents will purchase a signal, and the solution will be interior. Define the benefit of purchasing as signal as $P(s, \pi, \beta_B) \equiv \frac{\pi\beta_B(1-\beta_B)T}{(\frac{T}{2}+(1-T)(1-\beta_B))(\frac{T}{2}+(1-T))}$. At interior solutions, it must be the case that the marginal investor is indifferent between purchasing a signal and not, so $P = c$. Since $c$ is fixed, and since $P_s < 0$ and $P_\pi > 0$, we have that $s^*_\pi > 0$, and so we are done. \[
abla
\]

Proof. Proposition (3)
Define $P(\pi) \equiv \frac{\beta_B(1-\beta_B)\pi}{(\frac{T}{2}+(1-T)(1-\beta_B))(\frac{T}{2}+(1-T))}$ to be the benefit of purchasing a signal if the state for which information was purchased realizes. It must be the case that the lower bound on $V(\pi, 1-\pi)$ is 0, as a negative payoff could be avoided by abstention from purchase. Further, the upper bound on $V(\pi, 1-\pi)$ is achieved when $c = 0$, and equals $Y \equiv \frac{P}{1-\delta}$. Since $V$ is bounded between 0 and $Y$, there must exist a $c$ such that $\frac{\pi\beta_B(1-\beta_B)T}{(\frac{T}{2}+(1-T)(1-\beta_B))(\frac{T}{2}+(1-T))} - c + \delta\pi_B V(k, 1-k) < 0 \leq \pi_B V(0.5, 0.5)$.

Proof. Proposition (4)
First, we can rewrite equations 9 as zero-profit conditions as follows:

\[
\frac{T}{2}(\text{ask} - \mu_p) = \int \phi_p(x)(\text{ask} - x)(1 - \Phi_{\text{disp},x})(\text{ask})dx(T-1)
\]

\[
\frac{T}{2}(\mu_p - \text{bid}) = \int \phi_p(x)(x - \text{bid})\Phi_{\text{disp},x}(\text{bid})dx(T-1)
\]

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 $\mu_p$. 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):

\[
\frac{\partial}{\partial \sigma^2} \int \phi_p(x)(x - \text{bid}) \Phi_{\text{disp},x}(\text{bid}) dx (1 - T)
\]

\[
\propto \frac{\partial}{\partial \sigma^2} \int \phi_p(x)(x - \text{bid}) \Phi_{\text{disp},x}(\text{bid}) dx
\]

\[
= \frac{\partial}{\partial \sigma^2} \int_{-\infty}^{\text{bid}} \phi_p(x)(x - \text{bid}) \Phi_{\text{disp},x}(\text{bid}) dx + \frac{\partial}{\partial \sigma^2} \int_{\text{bid}}^{\infty} \phi_p(x)(x - \text{bid}) \Phi_{\text{disp},x}(\text{bid}) dx
\]

\[
= \int_{-\infty}^{\text{bid}} \phi_p(x)(x - \text{bid}) \left[ \int_{-\infty}^{\text{bid}} \frac{\partial}{\partial \sigma^2} \Phi_{\text{disp},x}(y) dy dx \right] < 0
\]

\[
+ \int_{\text{bid}}^{\infty} \phi_p(x)(x - \text{bid}) \left[ \int_{\text{bid}}^{\infty} \frac{\partial}{\partial \sigma^2} \Phi_{\text{disp},x}(y) dy dx \right] > 0
\]

\[
< 0
\]

Therefore, by similar logic we have that an decrease in \( \sigma^2 \), 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 \( \sigma^2 \) increases the transfer from noise traders to informed traders in expectation. So \( U \) depends negatively on \( \sigma^2 \).

\[ \square \]

**Proof. Corollary (2)**

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

\[ \square \]

**Proof. Corollary (3)**

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

\[ \square \]