Emergency Preparedness

Rare Events and the Persistence of Uncertainty*

Savitar Sundaresan†

Job Market Paper

updates at: http://www.columbia.edu/~svs2102/research.html

December 23, 2015

Abstract

Unexpected events can have lasting effects on financial uncertainty, which in turn affects the real economy. This paper uses a model in which the realizations of ex-ante unlikely 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. Uncertainty reduces asset demand and expected wealth, while increasing dispersion of beliefs. It also reduces investment and output, and results in higher credit spreads. Data on financial uncertainty, dispersion of beliefs, risk appetite, and credit spreads confirm the predictions of the model.

*Thanks to Ricardo Reis, Michael Woodford, Larry Glosten, and Jose Scheinkman for invaluable guidance, assistance and advice; to Geert Bekeart, Patrick Bolton, Robert Hodrick, Michael Johannes, Jennifer La'O, Emi Nakamura, Jaromir Nosal, Andrea Prat, Jon Steinsson, Tano Santos, Stephanie Schmitt-Grohe, Martin Uribe, Laura Veldkamp, and Neng Wang for discussions and comments that shaped the content of this paper; 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.

†Department of Economics, Columbia University, 420 West 118th Street, MC 3308, New York, NY 10027; Email: svs2102@columbia.edu; Website: http://www.columbia.edu/∼svs2102; Phone: (212) 662-1883

1 Introduction

Unexpected events of economic relevance can cause persistent jumps in uncertainty. Uncertainty has both real and financial implications: In the real economy, it can reduce investment and output through higher credit spreads or through firms' choosing to make production decisions later. 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 and tests a model that explains patterns in uncertainty and the effects of persistent uncertainty on real and financial variables.

Consider the examples of events marked by green lines in Figure 1 panel (a), 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 - thought to have originated with high frequency traders, the market plunged almost 10 percent and recovered within a few minutes. The fourth is the downgrade of U.S. treasury debt by the S&P ratings agency in 2011. In each case, the VIX[1] spikes upon the occurrence of the event and then slowly reverts to its pre-shock levels. In Figure 1 panel (b), we see that the uncertainty in panel (a) correlates to many other variables, including (i) increased credit spreads for U.S. Corporate Bonds, which measures borrowing costs; (ii) increased dispersion of analysts' forecasts of earnings per share, as constructed by Pinto (2010) and used by Li and Li (2014), which measures differences of opinion; and (iii) decreased margin debt as reported by the NYSE, which measures investors' appetite for risk.

I propose a model to explain three observations: (i) uncertainty spikes upon the realization of rare events; (ii) the spikes persist; and (iii) changes in uncertainty spill over into financial and real variables. In this model, agents can invest in information about 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 is assumed not to vary. In order to preserve informational asymmetries and prevent perfect revelation of information in prices, all trade takes place through a perfectly competitive market making sector. The value of the asset is also assumed to be indicative of productivity in the economy, so financial movements and uncertainty have implications on real decisions by firms,

1. The VIX is defined as the market's expectation of the annualized percentage standard deviation of the S&P index over a thirty day period.

such as borrowing, investment, and production.

This model produces three theoretical results. The first result is that agents choose to invest in more information for likely states of the world and less for unlikely states of the world. Consequently, 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). The inverse is also true. Thus, the occurrence of an ex-ante unlikely state results in increased short term and long term uncertainty due to poor signal quality.

Second, the model results in 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 and prolongs the initial spike in uncertainty.

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. Further, if the asset's value is correlated with aggregate productivity, signal quality can impact credit spreads, which in turn have real effects on production and investment.

The informational investment mechanism I use will deliver uncertainty persistence without long term 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. Therefore, taking the theory to the data is more direct in a financial setting.

I provide three sets of empirical results. First, I non-parametrically estimate the ex-ante probabilities of daily price movements in the S&P index. In a regression of the VIX on those probabilities, the coefficients are significantly negative, showing that rare events cause uncertainty. These results hold even when controlling for standard predictors of the VIX such as volume of trade, the size of price changes, and the level of the S&P index.

Second, I test the model's mechanism for persistence - that increased dispersion in an agent's ex-ante beliefs leads to a flatter distribution over future states. I test that increased dispersion in an agent's ex-ante beliefs lowers future probabilities on average, by showing that increases in the levels of the VIX are associated with lower 'highest' probabilities, 'average' probabilities, and dispersion of probabilities, and a smaller negative correlation between volatility and rarity. These results are all consistent with a with a theory of flatter future ex-ante distributions over price changes, as opposed to the alternative hypothesis of multimodality.

Third, a calibration and simulation tests the model's ability to match the VIX, credit spreads, dispersion of beliefs, and investors' risk appetite. Using one period's change in the S&P index as inputs, the model outputs the next period's opening value of the VIX at a monthly frequency. The model's predicted VIX is highly correlated with the data, both in levels and in first differences, and the patterns of credit spreads, dispersion, and risk match the data as well.

This model also provides a potential explanation of the 'pricing kernel puzzle.' This puzzle arises from a mismatch between an agent's theoretical and estimated stochastic discount factor. Theoretically, if an agent invests in an asset, then her stochastic discount factor (the amount that she values an additional future dollar of wealth) should monotonically decrease in the future value of that asset. In reality, an agent's estimated stochastic discount factor is slightly U-shaped. I show that this pattern can be explained by the increased uncertainty that arises from tail events. Uncertainty increases the marginal value of a dollar and can curl both ends of an otherwise monotonically decreasing stochastic discount factor, resulting in the U-shape seen in the data. I show this result theoretically and empirically, by non-parametrically estimating the stochastic discount factor and controlling for expected volatility.

1.1 Alternative Theories

Much of the existing literature on macroeconomic 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 (sometimes called black swans) - 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[2] (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 Chicago Board Options Exchange itself,[3] 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 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, 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 panel (a) 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 delivers rare events that lead to 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 affect 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

2. As calculated 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

3. Report (2010)

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 I interact 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) shows 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. 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. 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 at different prices. Bookstaber and Pomerantz (1989) relates 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. On the empirical side, Bleaney and Li (2014) and Yip et al. (2002) show that adverse selection is an important component of bid-ask spreads (with the latter finding that it is the single biggest component). This paper draws on all of these different strands, but seeks to answer a fundamentally different type of question: What effects do rare shocks have on a financial market *as communicated* through a particular market microstructure?

This paper also touches on the nature of information collection, as motivated by the notion of inattention, as 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 (2010) provides a host of results on price dynamics for inattentive sellers. In finance, Mondria (2010) uses inattention to discuss price comovement, Kacperczyk, Van 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. I add to this burgeoning literature by using elements of inattentive frameworks to address issues of uncertainty persistence.

The empirical section of this paper interacts with a large literature that wrestles with forecasting the VIX. Additionally, there is a large literature on using the VIX to predict stock returns, such as Bekaert and Hoerova (2014). The VIX forecasting literature focuses in part on parametric estimation of factors that could predict the value of the VIX, and formulates options trading strategies to profit off these predictions. Examples in this literature are Ahoniemi (2008), Harvey and Whaley (1992), Brooks and Oozeer (2002), and Konstantinidi, Skiadopoulos, and Tzagkaraki

(2008). This paper does not address issues of trading strategy in its notion of forecasting - rather it is interested in understanding movements in the VIX through a micro-founded understanding of the rarity of price movements.

This paper can converse with the literature in 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 iid underlying process.

Finally, examples in the 'pricing kernel puzzle' literature are Chabi-Yo, Garcia, and Renault (2008) who explain it with state dependence, and parametrically estimate a model. Christoffersen, Heston, and Jacobs (2009) use a two factor model, and Ziegler (2007) use heterogeneity of beliefs. More recently, Linn, Shive, and Shumway (2014) use a new nonparametric estimator based on an assumption of stationarity in the sdf.

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. The three stages of this simple model correspond to, in order: informational investment, informational payoff and trading, and accounting and conclusion.

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 η_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 π_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,' η_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, η_B indexes two random variables: η_H and η_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 η_B is determined, but it is not revealed until stage 2. Only one of η_H and η_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.

The possible values of η_H and η_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 η_H and η_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 η_B. The investment pays off by giving agents information about η_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 η_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 η_L will prove useless if $B = H$. Therefore, agents must choose whether to be informed about η_H, η_L, both or neither, knowing that only one informational investment will actually pay off. In stage 1, agents who invested in the correct η_B will receive a perfectly informative private signal about η_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 η_B in this model is exogenously given and constant across values of B. In stage 1, a public signal about whether η_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 η_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) is provided in Appendix C, which also allows private agents to be risk averse. The results are quite similar.

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 η_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.

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.

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, η_B is publicly revealed and agents receive their payoffs.

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:

$$\begin{aligned} P(\eta_B = 1|\text{buy order and MM Info}) + B &= ask \\ P(\eta_B = 1|\text{sell order and MM Info}) + B &= bid \end{aligned}$$

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:

$$
\begin{aligned}
\text{ask}_1 &= B + \frac{\left(\frac{T^*(s)}{2} + (1-T^*(s))\right)\beta}{\frac{T^*(s)}{2} + (1-T^*(s))\beta} && (1)\\
\text{bid}_1 &= B + \frac{\frac{T^*(s)}{2}\beta}{\frac{T^*(s)}{2} + (1-T^*(s))(1-\beta)} && (2)\\
\text{ask}_0 &= B + \frac{\left(\frac{T^*(s)}{2} + (1-T^*(s))\right)(1-\beta)}{\frac{T^*(s)}{2} + (1-T^*(s))(1-\beta)} && (3)\\
\text{bid}_0 &= B + \frac{\frac{T^*(s)}{2}(1-\beta)}{\frac{T^*(s)}{2} + (1-T^*(s))\beta} && (4)
\end{aligned}
$$

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 - ask_1 and bid_1 - 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.[4] 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 β, the trader decides whether purchase a signal for each η_B. The cost of acquiring a signal is fixed at c for all

4. If $\beta = 0.5$, then the public signal is completely uninformative and the two asks and the two bids would equal one another.

values of B. The expected benefit of acquiring a signal for B, given values T and π_B is:

$$\frac{\pi_B}{2}\left[\beta(1-\text{ask}_1(s))+(1-\beta)(1-\text{ask}_0(s))+(1-\beta)(\text{bid}_1(s))+\beta(\text{bid}_0(s))\right] \tag{5}$$

where an individual trader takes s as given in making her decision. Recall that π_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. β 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:

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

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 π_B increases, there is a higher likelihood that the agent will benefit from identifying η_B, and as a result, agents will buy signals for higher likelihood states. Although a simple case, 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 β 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 β, 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 π_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.*

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 η_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 theoretical section, I will simplify the signal acquisition process. Instead of heterogenous 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. As shown in Appendix B, 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 η_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 β:

$$V(\pi, 1-\pi) = \sum_{\pi} \max \left\{ \frac{\pi\beta(1-\beta)\frac{T}{2}}{\left(\frac{T}{2} + (1-T)(1-\beta)\right)\left(\frac{T}{2} + (1-T)\right)} - c + \delta\pi V(k, 1-k), \delta\pi V(0.5, 0.5) \right\} \tag{7}$$

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 η 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 β 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 π. 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 β 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 $\bar{c}$ such that if $\pi_t = 0.5$ and $c > \bar{c}$, $\pi_{t+1} = 0.5$.*

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 $\pi = 0.5$. 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 describe in this section. The x-axis corresponds to periods of time. The y-axis is the degree of uncertainty faced by agents over η_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.

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.

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 and Real Effects

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 by some non-calibrated simulations.

4.1 Structure

The structure of the economy will generalize to permit richer dynamics and predictions, and set up a link between the financial market and the real economy. To that latter end, I will now include two additional sectors: a perfectly competitive continuum of firms, and a perfectly competitive credit sector. 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, firms borrow, prices are set, and financial trade occurs. In stage 2, the final value of the asset is revealed, firms produce, 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 $V_t = B_t + \eta_{B,t}$. $B_t \sim \mathscr{N}(\mu_{B,t}, \sigma^2_{B,t})$ and $\eta_{B,t} \sim \mathscr{N}(0, \sigma^2_\eta)$ for each B. In stage 0, neither B_t nor $\eta_{B,t}$ is known, but $\mu_{B,t}$, $\sigma^2_{B,t}$, and σ^2_η are all known. In stage 1, B is publicly revealed, and private and public signals about the realization of $\eta_{B,t}$, with accuracy $\sigma^2_{\gamma,t}$ and $\sigma^2_{\beta,t}$, 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 $\epsilon_{t+1} \sim \mathscr{N}(0, \sigma^2_\epsilon)$. At the beginning of stage 0 of period $t+1$ a signal u_{t+1} is revealed that is distributed $\mathscr{N}(\epsilon_{t+1}, \sigma^2_{u,t+1})$ 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.

4.3 Stage 2 Problems

The model will again be solved backwards, and now the problems in stage 2 require some detail.

4.3.1 Traders

The representative trader in stage 2 now maximizes a utility function which is based on market returns *and* consumption:

$$U_t = \left(\int c_t(\omega)^\rho d\omega \right)^{\frac{1}{\rho}} + S_t \tag{8}$$

Where $c_t(\omega)$ is the consumption of good ω at time t, ρ is a measure of substitutability, and S is money holdings for the next period. The first order conditions yield that the the traders demand for good ω is:

$$p_t(\omega) = \frac{c_t(\omega)^{\rho-1}}{\rho} \tag{9}$$

The traders are always wealthy enough that their demand for goods does not suffer wealth effects (that is, wealth is high enough that $S_t > 0$ always - alternatively, one could assume that traders can carry over negative money holdings between periods). The only variation in S_t will come from the amount gained/lost in the stock market between stages 1 and 2 by financial activity. Given their stage-2 utility, traders are risk-neutral in their financial market participation. With this structure, the model can divorce firms' problems from the traders' financial issues. That is, the assumption makes the structure of the economy linear, as opposed to co-dependent.

4.3.2 Firms

Firms arrive in stage 2 with an amount of debt acquired in stage 1, q_t. Their stage 2 problem is:

$$\max_{q_{0,t}} A_t p_t q_{0,t}^\alpha - (1+r_t)q_t - f \qquad s.t. \qquad q_{0,t} \leq q_t \tag{10}$$

where $q_{0,t}$ is the amount of capital used in production (it must be weakly less than the amount borrowed), A_t is productivity, r_t is the pre-determined cost of borrowing, and f is a fixed cost of operating. Assume that A_t is directly correlated with V_t (conversely one could have assumed an input price was correlated with V). Given agents' demand, the operating profit for the firm is determined by:

$$\Pi_{F,t}(A_t, q_t, r_t) = \max \left\{ 0, \frac{A_t q_t^{\alpha+\rho-1}}{\rho} - (1+r_t)q_t, \frac{A_t q_t^{*\alpha+\rho-1}}{\rho} - (1+r_t)q_t + (q_t - q_{0,t}) \right\} - f \tag{11}$$

Where $q_t^* = \left(\frac{1+r}{\rho\alpha}\right)^{\frac{1}{\alpha+\rho-2}}$, the optimal level of capital desired ex-post. The three terms of the maximization are in turn: the option to default, the option to use all of the loan if the firm has borrowed weakly less than its optimal amount, and the option to use less than all of the loan, if the firm has borrowed more than its optimal amount.

4.3.3 Creditors

Creditors' returns depend on firms' outcomes:

$$\Pi_{C,t}(A_t, q_t, r_T) = \min\left\{(1+r_t)q_t, \left(\max\left\{\frac{A_t q_t^{\alpha+\rho-1}}{\rho}, \frac{A q_t^{*\alpha+\rho-1}}{\rho} + (q_t - q_t^*)\right\} - f\right)(1-k_c)\right\} \tag{12}$$

The creditor will receive the first argument of the minimum if the firm does not default. If the firm does default the creditor will receive the firm's profits, which are defined as the maximum in the firm's problem. k_c is the proportional loss in value upon default, and is used to proxy bankruptcy costs.

4.4 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 $\mu_{p,t} = \frac{\frac{1}{\sigma^2_{\beta,t}}x}{\frac{1}{\sigma^2_{\eta,t}} + \frac{1}{\sigma^2_{\beta,t}}}$, and $\sigma^2_{p,t} = \frac{1}{\frac{1}{\sigma^2_{\eta,t}} + \frac{1}{\sigma^2_{\beta,t}}}$.

4.4.1 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 the 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.

$$U(\text{public signal}, \text{private signal}) = \max\{(\text{bid}_t - x), (x - \text{ask}_t), 0\} \tag{13}$$

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

4.4.2 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 $\sigma^2_{\text{disp},t} = \frac{\frac{1}{\sigma^2_{\gamma,t}}}{\left(\frac{1}{\sigma^2_{\gamma,t}}+\frac{1}{\sigma^2_{p,t}}\right)^2}$. Conditional on a realization of B, a public mean μ_p, a public variance σ_p^2, and a dispersion of beliefs $\sigma^2_{\text{disp},t}$, the expected profits for a choice of a bid and an ask by the market making sector are:

$$\Pi_{\text{ask}} = \int \phi_{p,t}(x)\left((1-\Phi_{\text{disp},x}(ask))(1-T)+\frac{T}{2}\right)(ask-x)dx \tag{14}$$

$$\Pi_{\text{bid}} = \int \phi_{p,t}(x)\left(\Phi_{\text{disp},x}(bid)(1-T)+\frac{T}{2}\right)(x-bid)dx \tag{15}$$

where $\phi_{p,t} \sim \mathcal{N}(\mu_{p,t}, \sigma^2_{p,t})$ and $\phi_{\text{disp},x} \sim \mathcal{N}\left(\frac{\frac{\mu_{p,t}}{\sigma^2_{p,t}}+\frac{x}{\sigma^2_{\gamma,t}}}{\frac{1}{\sigma^2_{p,t}}+\frac{1}{\sigma^2_{\gamma,t}}}, \sigma^2_{\text{disp},t}\right)$. Perfect competition means that the bid and the ask must be selected to set the profits above to zero.

4.4.3 Creditors

Creditors are perfectly competitive, risk-neutral, and set a credit supply function that nets them zero profit in expectation. For a given level of borrowing, q, and risk-free rate r_f the credit supply function is:

$$\int \phi_p(x)\Pi_C(x, q_t, r_t)dx = (1+r_{f,t})q_t \tag{16}$$

Essentially, the expected return to the creditors must equal the return from investing in a risk-free asset. r_t is a function of μ_p, σ_p, and r_f.

4.4.4 Firms

Firms also must make zero profits in expectation, and so:

$$\int \phi_p(x)\Pi_F(x, q_t, r_t)dx = 0 \tag{17}$$

this is the credit-demand equation. q_t is a function of μ_p, σ_p, and r_f.

4.5 Stage 0 Problems

In stage 0, conditional private and public signals are selected.

4.5.1 Traders

The asset's value follows a random walk process both within and between periods. Within a period, the increment is η_t and between periods it is ϵ_t. The noise with which the public signals about η and ϵ 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:

$$\begin{aligned} V(\mu_{B,t}, \sigma^2_{B,t}, \sigma^2_{\beta,t}) &= \int \max_{\sigma^2_{\gamma,t}(B_t)} \phi(B_t) \int\int p(\text{public signal} = x)\left(p(\text{private signal} = y) U(x,y) dy\right) dx \\ &+\delta \left[\int\int \phi(B_t) V\left(\mu_{B,t+1}(u_{t+1}, \eta_t), \sigma^2_{B,t+1}, \sigma^2_{\beta,t+1}\right) du d\eta\right] - c\nu(\sigma^2_{\gamma,t}(B_t)) dB_t \end{aligned} \tag{18}$$

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

4.6 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* $\frac{\partial \sigma^*_\gamma(B)}{\partial \phi(B)} \leq 0$.

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 σ_γ and σ_β are unbounded, the variance of the posterior distributions are bounded.

4.6.1 Public Entity

It was shown in the simple model 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, σ_β for each potential value of B_t to maximize expected accuracy:

$$\int \min_{\sigma_\beta(B)} \phi_{B,p,t}(B) \left(\int p(\eta\text{-signal} = x)\sigma_p^2 dx \right) - c_p \nu_p(\sigma_\beta^2(B)) d\phi(B) \tag{19}$$

where ν_p is a convex decreasing function that satisfies the condition that $\frac{\partial^2 \nu_p}{\partial {\sigma_\beta^2}^2} > \frac{\partial^2 \sigma_p^2}{\partial {\sigma_\beta^2}^2}$ everywhere. The variance of the distribution over which the public entity solves its problem is time-invariant. One could say that the representative agent reveals the mean of her distribution to the public agent, but not the variance of signal. Initial simulations seem to show that the loosening of this restriction does not change much, but full simulation work is ongoing.

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. An alternative asymmetric objective is considered in the appendix, to show that the particular choice of objective is not driving the results. These changes in the structure of public information no longer allow us to describe patterns analytically, but simulations still permit insight into the dynamics.

4.6.2 Dynamic Equilibrium

Given values of $\{\mu_{B,t}, \sigma^2_{B,t}, c, c_p, \sigma^2_{\eta,t}, \sigma^2_{\epsilon,t}\}$, a dynamic equilibrium is defined by a choice of σ_β by the public entity that solves equation 19 a choice of a policy function $\sigma_\gamma(\mu_B, \sigma_B^2, \sigma_\beta^2)$ by the traders that satisfies equation 18, conditional on a choice q_t by Firms to solve equation 17, a choice of r_t by Creditors to solve equation 16, individual decisions to buy, sell, or abstain by traders to solve equation 13, prices p_t and r_t, and a choice of a bid and an ask by Market Makers to solve equations 14 and clear the goods and credit markets, given observed values of $\{\sigma_{\beta,t}, \sigma_{\gamma,t}, B_t, r_{f,t}, \sigma_{\eta,t}\}$, a public signal, and a set of private signals distributed $\mathcal{N}(\eta_t, \sigma^2_{\eta,t})$. .

4.7 Predictions and Spillovers

Figure 5 shows the results of the previous propositions. In this snapshot,[5] one can see how agents (both public and private) choose to invest in information. The blue dashed line is the ex-ante distribution of B. Then, conditional on these selections, one can see what the expectations for credit spreads, levels of borrowing, demand for financial assets, bid-ask spreads, uncertainty, volatility, and dispersion of beliefs are for each potential value of B.

As is expected by the propositions above, both σ_β and σ_γ are U-shaped, 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 (c). Most importantly here, uncertainty and volatility both spike for tail events. But it is important to note that bid-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 (d) 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.

The transmission of these effects to the real economy is also evident in panel (b). The 'counter cyclicality' of credit spreads is preserved here, as negative shocks result in higher credit spreads than positive ones, although spreads for rare positive shocks are still slightly higher than for 'normal' events.

4.8 Real and Financial 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.

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 - as uncertainty

5. The parameter values here are $\mu_{B,t} = 10$, $\sigma_\eta^2 = 4$, $\sigma_\epsilon^2 = 4$, $c_p = 0.0002$, $\delta = 0.95$, $c = 1$, $r_f = 0.03$, $\rho = 0.6$, $\alpha = 0.6$, $k_c = k_f = 0$, and $\nu_p(x) = \nu(x) = \frac{1}{x}$. α and ρ are selected with the calibration of section 5 in mind, while σ_η and σ_ϵ 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.

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 is the time it takes for agent's confidence in their forecasts to regain shape (for the variance of their beliefs to decline).

The spillover into the real economy can also be seen as credit spreads also spike and take time to decline, which causes output to stay below its initial value for several periods.

5 Taking the Model to Data

The model of this paper addresses the nature of informational investment. It has delivered three key predictions that I will now attempt to evaluate empirically.

First, propositions (2) and (4) both state low probability events lead to lower levels of information. The effects of these lower levels of information manifest in higher levels of uncertainty. Therefore, I first document that low probability events do lead to spikes in uncertainty.

Proposition (3) provides an explanation for the persistence of uncertainty - namely that more disperse beliefs lead to worse informational investment on average. I show that an initial low probability event flattens future probability distributions, which in conjunction with the first result, provides a channel for persistence.

Finally, and perhaps most illustratively, I use the model and actual S&P price changes to see how well the model can forecast values of the VIX, dispersion, risk taking and credit spreads. In Appendix D I provide some evidence that the channels through which the above effects take place is preparedness, as measured by a function of how well a representative consumer is hedged against different price movements.

5.1 Rare Events Trigger Uncertainty

First, I show that lower probability events cause higher levels of uncertainty. These are the predictions of Propositions (2) and (4).

5.1.1 Strategy

The first hurdle is deciding how to measure the probabilities of events. The seminal work of Breeden and Litzenberger (1978) allows recovery the ex-ante probabilities of price movements from options data. They show that the double derivative of call option prices on a range of strike prices gives the prices of Arrow-Debreu securities along the same range.[6] Taking the shortest horizon and the full set of strikes of call options, I construct a vector of implied volatilities, which I use to price one-day options. The double-difference of those one-day options yield a probability distribution over price movements in the following trading day.

A concern might arise from this methodology. Due to risk-aversion present in the market, the options data may be overestimating the probability of negative events and underestimating the probability of positive events. This would bias coefficients in any potential regression. To account for this, I also calculate the probabilities using the physical distribution based on fitting a nonparametric kernel density estimator to a physical distribution based on a moving window of historical largest intraday price movements using the novel method of Botev, Grotowski, Kroese, et al. (2010). Effectively, using a density of data acquired in a historical window (the results of this exercise are not very sensitive to the size of the window), the kernel density estimator is formed by assuming a linear-diffusion process (such as, in a simple case, the heat equation). There is then an estimator for each value of t, as the process gradually diffuses completely to its limiting distribution. The algorithm is a two-state process, which first models the dissipation of heat into a space of uniform diffusivity. This diffusion process depends on the nature of the data, and so the second stage again models dissipation but with nonuniform diffusivity, learned from the first stage. See Ramsey (2014) for an excellent summary.

A second hurdle is deciding how to measure financial uncertainty. The VIX is a measure constructed by the Chicago Board of Exchange to measure, based off of options prices, the market's expectation of volatility of the S&P index in the next 30 days, reported in annualized percentage terms. This metric reflects the market's uncertainty over future price paths, and so will serve as a useful proxy for *uncertainty*. Therefore, the regression that we're interested in will take the following form:

$$\Delta \log(VIX_t) = \alpha_0 + \alpha_1 \Delta \log(Prob_t) + \alpha_2 \Delta \log(X_t) + \epsilon_t \tag{20}$$

$Prob_t$ is the minimum of a vector with two values: the probabilities of the highest and lowest intraday values of the $S\&P$ during day t as calculated in the two methods discussed above. VIX_t

6. Of course, given the relative illiquidity of options markets, the data needs to be cleaned and prepared in order to apply this method. Starting with daily options prices for the S&P index from January 2nd, 1996 to March 15th, 2015, which totals about 5 million data points, I then eliminate, in order, the prices that allow the following arbitrage opportunities: (1) If the bid-ask spread on a particular security was negative. (2) If it was possible to buy and immediate exercise an option for profit. (3) If the value of options were not monotonic in the strikes. (4) If the double difference (price of Arrow-Debreu securities) were negative. After these adjustments there are about 2.5 million prices left. These are then arranged each into implied volatility surfaces, and I use a spline to interpolate between strikes and maturities.

is the closing value of the VIX index on day t. The hypothesis is that α_1 should be negative. Furthermore since the VIX and other control variables in this regression are serially correlated[7], the regression will be run in first differences of logs and will report Newey-West standard errors. X_t is a vector of control variables, such as the value of the S&P index, the volume of trade in the index, and the intraday volatility of the index, which should alleviate concerns of omitted variable bias. Additionally, I include the Skewness of the VIX in the control group, just to add a more rigorous test of the claim in the introduction, that the Skewness and level of the VIX are not positively correlated.

I use a standard set of regressors augmented with the probabilities. The value of controls such as the underlying and the volume of trade has been used to explain implied volatility by such papers as Simon (2003), Harvey and Whaley (1992), Franks and Schwartz (1991), and Fleming, Ostdiek, Whaley, et al. (1995). This regression is novel in that previous work have not tested the effect of the ex-ante probabilities on ex-post implied volatility.

5.1.2 Data

All options price data comes from WRDS. From 1996 to 2015 it includes daily closing bids, closing asks, implied volatilities, for all listed options on the $S\&P$ index by strike and maturity. I also have daily opening and closing values of the VIX from Yahoo!. All summary statistics are reported in table 1.

5.1.3 Results

Before we turn to the regression results, Figure 7 is a useful way of looking at the data. For each day, I have a probability of the intraday high point and a probability of the intraday low point. I take the minimum of those two probabilities, and arrange those along the x-axis as follows: the left half of the x-axis are all the days on which the intraday low was less likely than the intraday high. The low point probabilities are arranged from least likely to most likely. On the right half are all the days on which the intraday high was less likely than the intraday low. The high point probabilities are arranged from most likely to least likely. This is to mimic the x-axis seen in the graphs in section 4.6. Then corresponding to each probability I take the closing value of the VIX index for each day. As is visible from the fitted quadratic curve, the basic shape of the VIX seems to mimic that of σ_γ.

7. As decided by Dickey-Fuller tests on each variable.

As is evident from the table, and as was expected, the coefficients on the minimum probability and its lags are all negative in all columns. The results of the table are robust to many other control variables, and seem to confirm the finding of the model that ex-ante low probability events are followed by spikes in the VIX. Importantly, the spikes in uncertainty cannot solely be caused by the magnitude of the change in the S&P itself, nor in changes in volatility, as they are being controlled for. In terms of magnitude, a one percent increase in the size of intraday volatility increases the VIX by about 0.01 percent. On the other hand a one percent increase in the size of the probability of an event decreases the VIX by about 0.002 percent. This may seem small, but probabilities can fluctuate wildly, and can drop or increase many-fold in a day. This non-linearity is precisely why it is beneficial to look at this regression in log terms. Very rare events will be orders of magnitude smaller than the norm, and so will significantly increase the VIX. It is also important to note that the Skew coefficient is negative and significant, indicating that a fear of large negative events is not positively correlated with overall uncertainty.

Touching on the issue of causality here, an alternative hypothesis to the one posited here is that a rare event may in fact spike uncertainty, and that the spike in uncertainty is what causes large price fluctuations. The lagged terms should allay this concern somewhat, as an unlikely price movement not only impacts today's VIX but tomorrow's as well. However, the idea that an increase in uncertainty makes large price fluctuations more likely, is very much in step with the predictions of the model. The flattening of a distribution, referred to often here, means that not only does the center of a distribution get depressed, but the tails fatten somewhat as well, thus harking back to the 'self-exciting' process mentioned earlier. The model posits that both directions of causality are valid, but that timing is crucial - one rare event causes a flattening, *which in turn* may cause more events to be ill-identified.

5.2 Propagation

The next step tests whether the data is consistent with the model's explanation of persistence, namely, Proposition 3 and Corollary 3. Are spikes in uncertainty correlated with flattened distributions and lower future probabilities? This is a statistical test of a correlation as opposed to a causal argument. This test, in conjunction with the previous section's results, should display the desired patterns of uncertainty: namely a rare event spikes uncertainty, which in turn lowers the probabilities of price movements on average, thus raising future uncertainty in expectation.

5.2.1 Strategy

The previous section showed that the lower the probability of a price movement, the larger the increase in the VIX. Therefore, in order to justify the model's predictions, I must further show that increases in the level of the VIX flatten that period's probability distribution. I will test this by showing that given a pdf, $p(\cdot)$ over S&P price movements, an increase in the VIX will have four effects: first it will lower $\max(p)$; second it will lower $E_p[p]$, third it will lower $E_p[p^2]$; and fourth, that the effect of volatility on p is always negative. Intuitively, these four effects are that the likelihood of the most likely price movement declines, that the 'average' probability of price movements decline, that the dispersion of probabilities of price movements declines (thus showing that the effect also has the self-exciting aspect mentioned earlier) and finally that large price movements are not expected. All four of these results are consistent with a flattening of the probability distribution. The fourth additionally rules out potential alternative hypothesis such as bimodal or multimodal distributions, that could also be associated with increases in the VIX.

The last section already created a time series for p across strikes. I use the risk-neutral probabilities - the results are similar for the physical distribution, but since those distributions are already smoothed, by the kernel density estimator method, it is less surprising that those distributions would follow the patterns laid out. For each day, I normalize the state space of p by the opening value of the S&P index, in order to create a method of uniform comparison. Then I create three time series: one called max which is the maximum value of $p(\cdot)$, one called avg, which is defined as $E_p[p(\cdot)]$, and one called stdev, which is defined as $E_p[p^2]$. All three of these variable would take their maximum value when p is a point mass, and their minimum value when p is uniform across all strikes, and therefore provide measures of flatness. Finally, I regress volatility on the minimum probabilities of the intraday highs and lows:

$$prob_t = \phi_0 + \phi_1 volatility_t + \epsilon_t$$

This metric shows how large price changes are expected. One possible alternative theory to the pure flattening of the distribution, is that the VIX goes up in value, because investors have bimodal expectations: they expect a period of higher volatility. If ϕ_1 is negative and significant for each quantile, that would be evidence against multi-modal expectations, as large price changes would continue to be less expected than small ones.

I calculate the four variables above for 32 different quantiles of the VIX index, where each quantile contains 150 observations. Then I plot their values with confidence intervals against the quantiles, and also regress them as well to provide visual and statistical evidence that the relationships are as anticipated.

5.2.2 Results

As is evident from Figure 9, there is a strong negative relationship between the level of the VIX, and all of the measures of flattening described above. Further, Table 4 bears out this visual relationship statistically. The table indicates that a large movement in the VIX (of say, 10 quantiles) reduces the max probability by 1%, and the average probability by 0.6%, while the graphs show that the average probability declines by almost half over all quantiles.

Therefore, this empirical test, along with the tests of the effect of low probability events, shows that the data is consistent with the predictions of the model: lower probability events will increase the level of uncertainty, which in turn flattens out the distribution, making future price movements less likely, thus repeating the cycle and making uncertainty persistent.

5.3 Simulation and Forecasting

This section of the paper is devoted to testing how well the mechanisms described can actually match empirical patterns. To that end, all of the variables in the model have some real world parallel, that they can be calibrated to, with the exception of the costs of informational investment. Therefore, this section will calibrate all of the variables possible, and then estimate values of c and c_p that allow predicted uncertainty to come closest to the VIX.

5.3.1 Strategy

The list of variables that need to be calibrated are:

Variable	Value	Source/Justification
$\mu_{B,t}$	S&P Index	–
σ_η^2	23%	–
σ_u^2	23%	–
ρ	$\frac{4}{5}$	Hassan and Mertens (2014)
α	0.34	Hassan and Mertens (2014)
$r_{f,t}$	Federal Funds Rate	–
f	0.67	Fiedler (2012)
δ	0.95	–
T	0.38	Wang (2002)
k_c	0.3	Kozlowski, Veldkamp, and Venkateswaran (2015)

The VIX is directly tied to the S&P index. Therefore the asset in the model will track the value of the index. Both σ_η and σ_u will be measured in percentage terms, in order to match the units of the VIX. 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. Calibration for capital utilization, elasticity of substitution, the risk free rate, the cost of bankruptcy and the fixed cost of production are only relevant for the computation of credit spreads, and the values here are not overly important. The risk-free rate is taken to be the federal funds rate, which is the overnight rate as opposed to a monthly rate, but the results again are not sensitive to changes here. For an estimate of fixed cost, I use Fiedler (2012), which compares fixed to variable costs, and come up with an approximate ratio of the two, which allows me to estimate the fixed cost in this model at 0.67. 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.

Then the model takes the parameters above as inputs, and the previous period's maximum price change, and the previous period's VIX (either estimated or actual) and outputs uncertainty, calculated as $Var[\eta|\text{private and public information}]$. There are two simulations, in the first, the VIX each period is updated to its correct value in order to predict the next period's value (PbP); in the second, the actual value of the VIX is only input in the first period, and only price changes are input thereafter (PC).

Forecast	Mean	Var
VIX	6.12	5.19
VIX PbP	5.29	4.21
VIX PC	5.57	5.98

In order to construct a baseline for comparison, I also consider a naive forecaster, who uses the previous period's largest percentage deviation as the prediction for the next period's VIX. Additionally, given the inputs, I am also able to construct simulations for the path of credit spreads, the path of risk-taking, and the path of dispersion of beliefs. These can also be compared to actual time-series from the data.

5.3.2 Results

As is evident from Figure 10, the paths of the simulated VIXs and the actual VIX seem relatively close to one another, all spiking in the same places. Some summary statistics for the forecasts are below, the two simulations have a lower mean squared error and a higher R^2 than the naive forecast.

Forecast	MSE	R^2
Monthly PbP	2.88	0.476
Monthly PC	2.46	0.525
Monthly Naive	13.27	0.245

Further using a Diebold-Mariano test on the vectors of squared errors of each of the forecasts, we see that the best forecast is PC, followed by PbP, followed by the naive forecast. The intuition of the two sets of simulations from the last section should be evident here - the spikes after periods of large price movements, and the subsequent slow return to previous levels merely build on earlier intuition. Additionally, we can see in Figure 11 the forecasts (in blue and red) for credit spreads, dispersion of beliefs, and risk-taking, plotted against the same time series as in the first figure of the paper (in yellow)

Forecast	MSE	R^2
Credit Spread PbP	0.03	0.195
Credit Spread PC	0.04	0.069
Dispersion PbP	87.02	0.138
Dispersion PC	39.06	0.155
Debt Margin PbP	4.64	0.445
Debt Margin PC	4.77	0.394

Using only price changes in the S&P index, and an initial value of the VIX, the model is able to give predictions about these variables that, while not identical, still match the general shapes of each, and explain significant fractions of their variance.

6 Pricing Kernel Puzzle

If an agent purchases an asset, then her wealth increases in the value of an asset. One would expect then, that her stochastic discount factor would monotonically decrease in the same. However, empirically estimated stochastic discount factors are typically U-shaped in the value of the asset, violating this hypothesis. This contradiction is referred to as the 'pricing kernel puzzle'. This section will show how the framework of a model like that of Section 4 can deliver such a theoretical stochastic discount factor that matches the data.

6.1 Theory

Total production in the economy of this paper with and without uncertainty is shown in Figure 12. There are two lines - the first (blue) is total production in the economy when the public entity cannot perfectly invest in information. The second (red) is total production when there is no uncertainty (a perfectly informative public signal). As we can see, there is a sizable difference between the two lines, as poor public signals cause higher credit spreads and lower levels of borrowing and investment.

In the asset pricing literature, the law of one price implies that for a given asset that pays off a certain vector $\{x_1, ..., x_n\}$ in a subsequent period, the price of that asset today must equal the sum of the contingent claims:

$$P(\tilde{x}) = \sum_{1}^{n} p_c(i) x_i$$

We can convert this into probability terms, by multiplying and dividing each side by the probability of each state occurring (π_i):

$$P(\tilde{x}) = \sum_{1}^{n} \pi_i \frac{p_c(i)}{\pi_i} x_i = E[m\tilde{x}]$$

The stochastic discount factor, or pricing kernel, is defined as the ratio of the conditional price to the true probability. The price of a riskless asset should simply be: $1 + r_f = \frac{1}{E[m]}$. In order for the stochastic discount factor to not be degenerate, the conditional prices must not be exactly equal to the true probabilities, which would only occur under risk-neutrality.

The stochastic discount factor for an agent whose utility is a concave function of total production in the economy is shown in Figure 13. As is evident, uncertainty causes slightly U-shaped sdf, while the no-uncertainty case has a monotonically decreasing sdf. It therefore seems intuitive that conditional uncertainty would be able to explain the U-shaped pricing kernel found in nonparametric empirical studies.

6.2 Empirics

Here I want to test whether I can match the literature by non-parametrically estimating the pricing kernel smile from the data. Then, upon controlling for conditional expected uncertainty, I want to test whether the kernel is monotonically decreasing in the value of the underlying.

6.2.1 Strategy

I can non-parametrically estimate the stochastic discount factor implied by prices and historical returns of the S&P index and its options. Typically, in the literature, estimates of the stochastic discount factor are U-shaped and non monotonic. I already have all of the risk-neutral probabilities from the options data as described above. I can further construct daily physical probability distributions by using a moving window of historical returns and fitting the kernel density estimator of the previous section to it[8]. Then taking the ratio of the two pdfs will yield a vector of the sdf for each date.

I obtain the U-shaped pricing kernel as in other examples in the literature - a ratio of the physical and risk-neutral probabilities - and then further propose a novel resolution: the sdf is U-shaped in anticipation of conditional uncertainty. That is, when rare events occur, as we have seen, there is a spike in uncertainty. This spike is anticipated by market participants, leading to a higher value of the pricing kernel for those values.

I can test conditional uncertainty explanation by running the following cross-sectional regression:

$$sdf_{it} = \alpha_0 + \alpha_1 \text{changeund}_i + \alpha_2 \text{physprob}_{it} + \alpha_3 \text{physvol}_{it} + \alpha_4 SPVal_t + \epsilon_{it}$$

where sdf_{it} is the non parametrically estimated stochastic discount factor of percentage return i at time t, changeund is percentage return i at time t, physprob is the physical probability of the return i at time t, physvol is the expected volatility of return i at time t[9], and SPVal is the actual level of the S&P index at time t. The aim is to test whether $\alpha_1 < 0$, and $\alpha_3 > 0$.

6.2.2 Results

A graph of the mean and two-standard deviation bands of the pricing kernel is plotted below in Figure 15a. As expected, the shape of the pricing kernel is non-monotonic in the value of the underlying, with a slight U-shape. Next is a plot of the mean physical and risk-neutral distributions of price changes as well as the mean physical distribution of expected volatility. Now, controlling for the shape of the physical VIX, we can run the regression specified above to test whether the underlying is negatively correlated with the sdf. The results can be seen in Figure 15 and Table 5.

8. I use windows of 50, 100, 250, 500, 1000, and 1500 days - the results do not change significantly. In the reported tables I use a window of 1000 days.

9. Again here a rolling window is used to construct this series, with a size equal to the window used to calculate the physical probabilities of price movements. Instead of a distribution here, I use a spline to interpolate between the historical values of the VIX at different levels of returns.

The first regression shows that without controlling for the VIX, there is an insignificant and slight negative correlation between the sdf and the underlying. However, upon inclusion of the VIX, the size and significance of the effect increases. Controlling only for the effect of the VIX, the new mean of the sdf is as in Figure 15b which is clearly more monotonic than before. The results do not show a complete resolution of the puzzle, empirically, but do move towards a solution.

7 Conclusion

I have presented, solved, and simulated 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.

Further, I present a variety of empirical evidence that is consistent with the model, including that low probability events are a significant predictor of increases in the VIX, that such events have persistent effects on the VIX and lower the probabilities of subsequent price movements, and that forecasting simulations of the model are able to mimic the actual path of expected volatility and match patterns in credit spreads, dispersion of beliefs and debt margin.

References

Aghion, Philippe, George-Marios Angeletos, Abhijit Banerjee, and Kalina Manova. 2010. "Volatility and growth: Credit constraints and the composition of investment." *Journal of Monetary Economics* 57 (3): 246–265.

Ahoniemi, Katja. 2008. "Modeling and forecasting the VIX index." *Available at SSRN 1033812.*

Aït-Sahalia, Yacine, Julio Cacho-Diaz, and Roger JA Laeven. 2015. "Modeling financial contagion using mutually exciting jump processes." *Journal of Financial Economics.*

Aït-Sahalia, Yacine, Roger JA Laeven, and Loriana Pelizzon. 2014. "Mutual excitation in eurozone sovereign CDS." *Journal of Econometrics* 183 (2): 151–167.

Almeida, Heitor, and Thomas Philippon. 2007. "The risk-adjusted cost of financial distress." *The Journal of Finance* 62 (6): 2557–2586.

Arellano, Cristina, Yan Bai, and Patrick Kehoe. 2010. "Financial markets and fluctuations in uncertainty." *Federal Reserve Bank of Minneapolis Working Paper.*

Barro, Robert J. 2006. "Rare disasters and asset markets in the twentieth century." *The Quarterly Journal of Economics:*823–866.

Bekaert, Geert, and Marie Hoerova. 2014. "The VIX, the variance premium and stock market volatility." *Journal of Econometrics* 183 (2): 181–192.

Bekaert, Geert, and Alexander Popov. 2012. *On the Link between the Volatility and Skewness of Growth.* Technical report. National Bureau of Economic Research.

Bianchi, Francesco, and Leonardo Melosi. 2013. "Modeling the evolution of expectations and uncertainty in general equilibrium."

Bleaney, Michael, and Zhiyong Li. 2014. "A new spread estimator." *Review of Quantitative Finance and Accounting:*1–33.

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

Bond, Philip, Alex Edmans, and Itay Goldstein. 2011. *The real effects of financial markets.* Technical report. National Bureau of Economic Research.

Bookstaber, Richard M, and Steven Pomerantz. 1989. "An information-based model of market volatility." *Financial Analysts Journal* 45 (6): 37–46.

Botev, Zdravko I, Joseph F Grotowski, Dirk P Kroese, et al. 2010. "Kernel density estimation via diffusion." *The Annals of Statistics* 38 (5): 2916–2957.

Breeden, Douglas T, and Robert H Litzenberger. 1978. "Prices of state-contingent claims implicit in option prices." *Journal of business:*621–651.

Brooks, Chris, and M Currim Oozeer. 2002. "Modelling the implied volatility of options on long gilt futures." *Journal of Business Finance & Accounting* 29 (1-2): 111–137.

Chabi-Yo, Fousseni, Rene Garcia, and Eric Renault. 2008. "State dependence can explain the risk aversion puzzle." *Review of Financial Studies* 21 (2): 973–1011.

Christoffersen, Peter, Steven Heston, and Kris Jacobs. 2009. "The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well." *Management Science* 55 (12): 1914–1932.

Cogley, Timothy, Christian Matthes, and Argia M Sbordone. 2011. "Optimal disinflation under learning." *FRB of New York Staff Report,* no. 524.

Cúrdia, Vasco, Marco del Negro, and Daniel L Greenwald. 2014. "Rare shocks, great recessions." *Journal of Applied Econometrics* 29 (7): 1031–1052.

Evans, George W, and Seppo Honkapohja. 2001. *Learning and expectations in macroeconomics.* Princeton University Press.

Fajgelbaum, Pablo, Edouard Schaal, and Mathieu Taschereau-Dumouchel. 2014. *Uncertainty traps.* Technical report. National Bureau of Economic Research.

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

Fishman, Michael J, and Jonathan A Parker. 2015. "Valuation, adverse selection, and market collapses." *Review of Financial Studies:*hhv025.

Fleming, Jeff, Barbara Ostdiek, Robert E Whaley, et al. 1995. "Predicting stock market volatility: a new measure." *Journal of Futures Markets* 15 (3): 265–302.

Franks, Julian R, and Eduardo S Schwartz. 1991. "The stochastic behaviour of market variance implied in the prices of index options." *The economic journal:*1460–1475.

Gilchrist, Simon, Jae W Sim, and Egon Zakrajšek. 2014. *Uncertainty, financial frictions, and investment dynamics.* Technical report. National Bureau of Economic Research.

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.

Guerrieri, Veronica, and Robert Shimer. 2012. *Dynamic adverse selection: A theory of illiquidity, fire sales, and flight to quality.* Technical report. National Bureau of Economic Research.

Harvey, Campbell R, and Robert E Whaley. 1992. "Market volatility prediction and the efficiency of the S & P 100 index option market." *Journal of Financial Economics* 31 (1): 43–73.

Hassan, Tarek A, and Thomas M Mertens. 2014. "Information aggregation in a dsge model." In *NBER Macroeconomics Annual 2014, Volume 29.* University of Chicago Press.

Kacperczyk, Marcin T, Stijn Van Nieuwerburgh, and Laura Veldkamp. 2014. "A Rational Theory of Mutual Funds' Attention Allocation."

Konstantinidi, Eirini, George Skiadopoulos, and Emilia Tzagkaraki. 2008. "Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices." *Journal of Banking & Finance* 32 (11): 2401–2411.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2015. "The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation." *Manuscript.*

Kurlat, Pablo. 2010. "Lemons, market shutdowns and learning." *manuscript, Stanford University.*

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.*

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

———. 2010. "Business cycle dynamics under rational inattention."

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

Matějka, Filip. 2010. "Rationally inattentive seller: Sales and discrete pricing." *CERGE-EI Working Paper Series,* no. 408.

Mondria, Jordi. 2010. "Portfolio choice, attention allocation, and price comovement." *Journal of Economic Theory* 145 (5): 1837–1864.

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

Paciello, Luigi, and Mirko Wiederholt. 2013. "Exogenous Information, Endogenous Information and Optimal Monetary Policy*." *The Review of Economic Studies:*rdt024.

Parente, Stephen L, and Edward C Prescott. 1994. "Barriers to technology adoption and development." *Journal of political Economy:*298–321.

Pinto, Eugenio. 2010. *Forecasting investment with earnings expectations.* Unpublished manuscript. Federal Reserve Board of Governors.

Ramsey, Ford. 2014. "An Application of Kernel Density Estimation via Diffusion to Group Yield Insurance." In *2014 Annual Meeting, July 27-29, 2014, Minneapolis, Minnesota.* 170173. Agricultural and Applied Economics Association.

Ravn, Morten O, Stephanie Schmitt-Grohé, and Martin Uribe. 2007. *Explaining the effects of government spending shocks on consumption and the real exchange rate.* Technical report. National Bureau of Economic Research.

Report, CBOE. 2010. *The CBOE Skew Index - SKEW.* Technical report. Chicago Board Options Exchange.

Rietz, Thomas A. 1988. "The equity risk premium a solution." *Journal of monetary Economics* 22 (1): 117–131.

Routledge, Bryan R, and Stanley E Zin. 2009. "Model uncertainty and liquidity." *Review of Economic dynamics* 12 (4): 543–566.

Shannon, Claude Elwood. 2001. "A mathematical theory of communication." *ACM SIGMOBILE Mobile Computing and Communications Review* 5 (1): 3–55.

Simon, David P. 2003. "The Nasdaq volatility index during and after the bubble." *The Journal of Derivatives* 11 (2): 9–24.

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

Sundaresan, Savitar, and Sébastien Turban. 2014. *Inattentive valuation and belief polarization.* Technical report. Working paper.

Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2010. "Information acquisition and under-diversification." *The Review of Economic Studies* 77 (2): 779–805.

Wang, Changyun. 2002. "The effect of net positions by type of trader on volatility in foreign currency futures markets." *Journal of Futures Markets* 22 (5): 427–450.

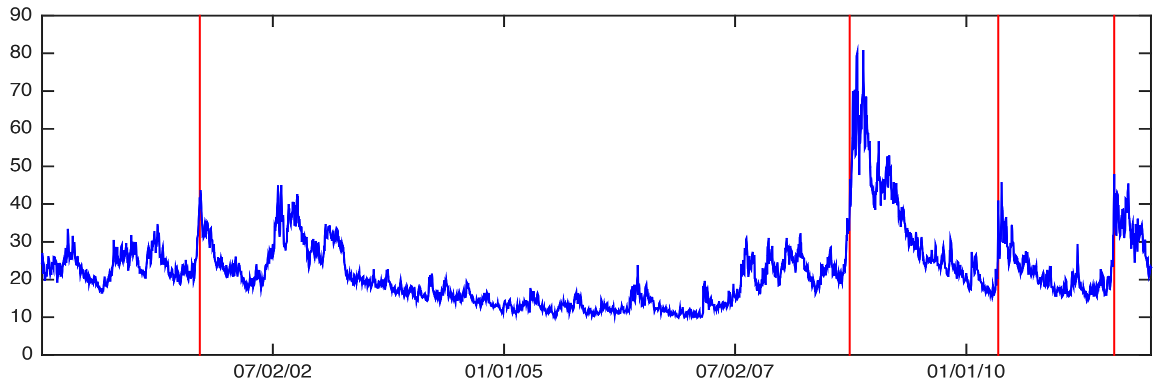
Woodford, Michael. 2012. "Inattentive valuation and reference-dependent choice." *Unpublished Manuscript, Columbia University.*

Yip, Henry YK, David Michayluk, Laurie Prather, and Li-Anne Woo. 2002. "Decomposing the bid-ask spread of a common stock: a cross-market approach."

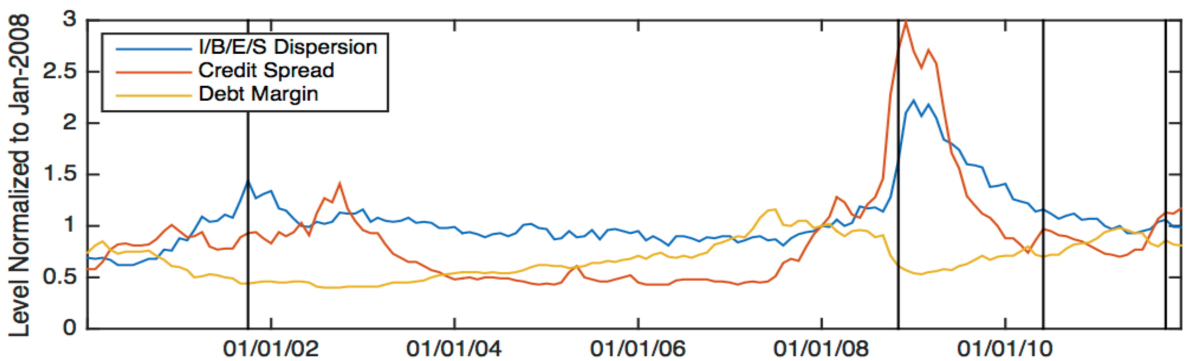
Zhang, Lu. 2005. "The value premium." *The Journal of Finance* 60 (1): 67–103.

Ziegler, Alexandre. 2007. "Why does implied risk aversion smile?" *Review of Financial Studies* 20 (3): 859–904.

Figures and Tables



(a) Daily values of VIX with markers for key rare events



(b) Monthly US Corporate Credit Spreads, NYSE Margin Debt, and I/B/E/S Dispersion of Forecasts

Figure 1: Motivating Figures

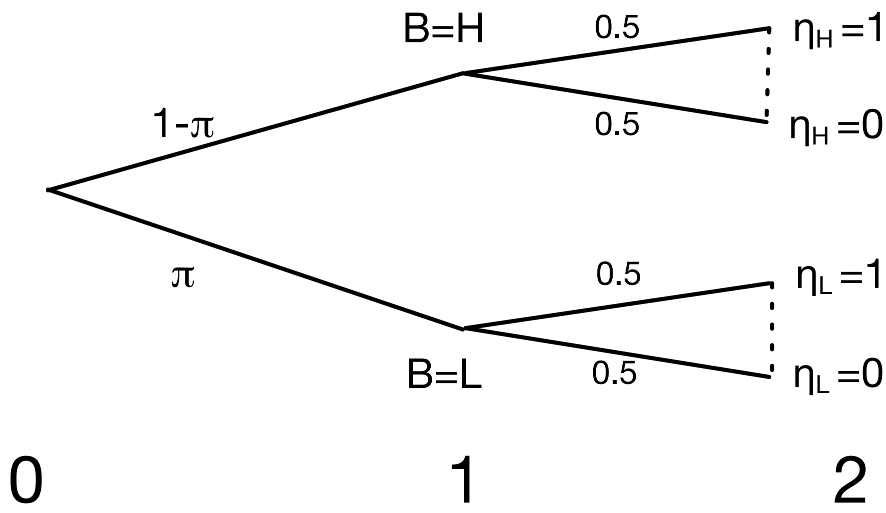


Figure 2: Structure of the Asset

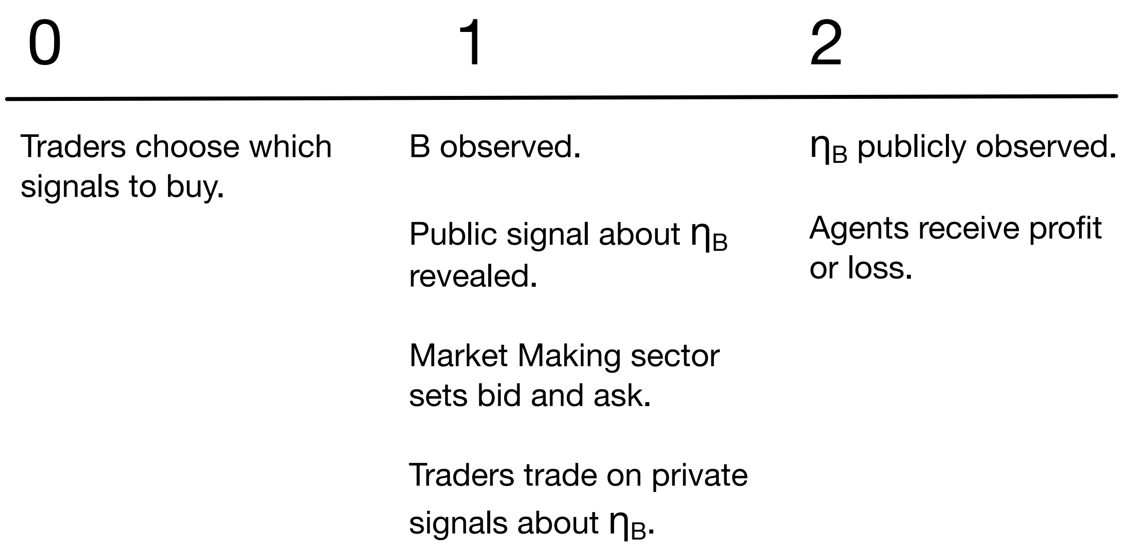
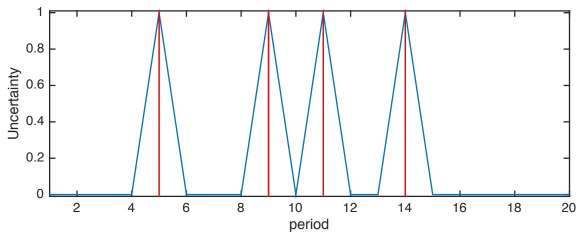
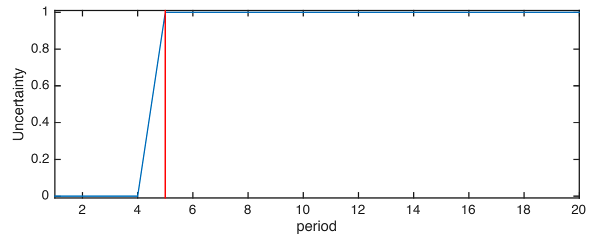


Figure 3: Timeline of within-period actions and decisions

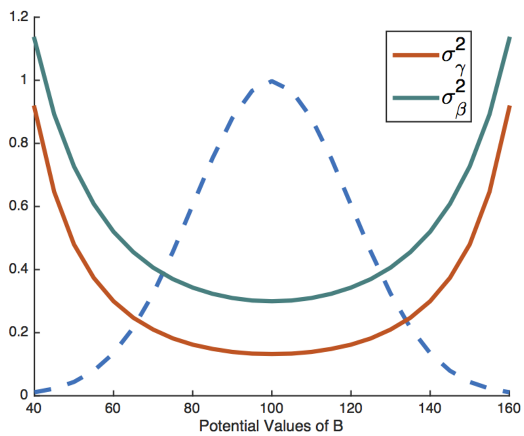


(a) Temporary Uncertainty for Low Information Costs

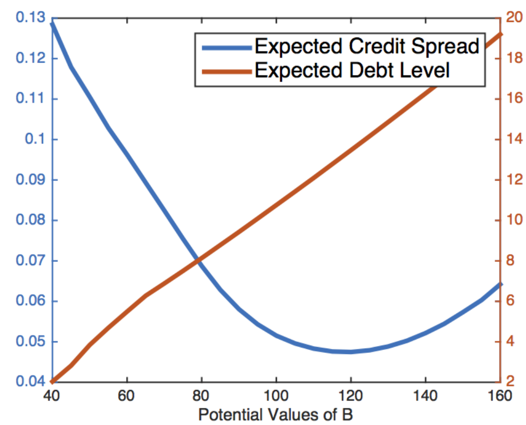


(b) Uncertainty Persists for High Information Costs

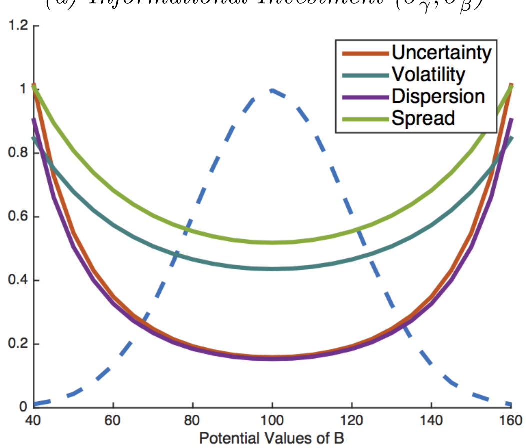
Figure 4: Uncertainty Reactions in Dynamic Model



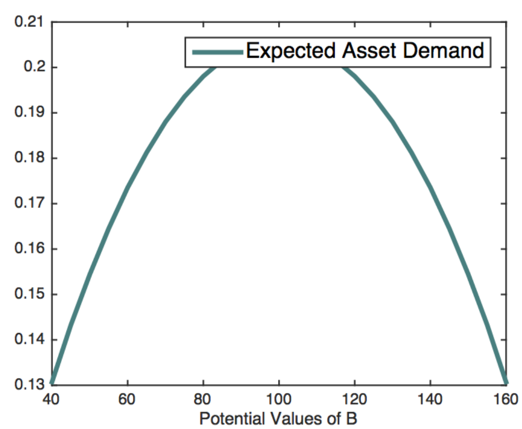
(a) Informational Investment ($\sigma^2_\gamma, \sigma^2_\beta$)



(b) Debt Level and Credit Spread



(c) Bid-Ask Spread, Uncertainty, Expected Volatility, and Dispersion



(d) Asset-Demand (Total number of shares bought/sold)

Figure 5: Simulation of Continuous State Model

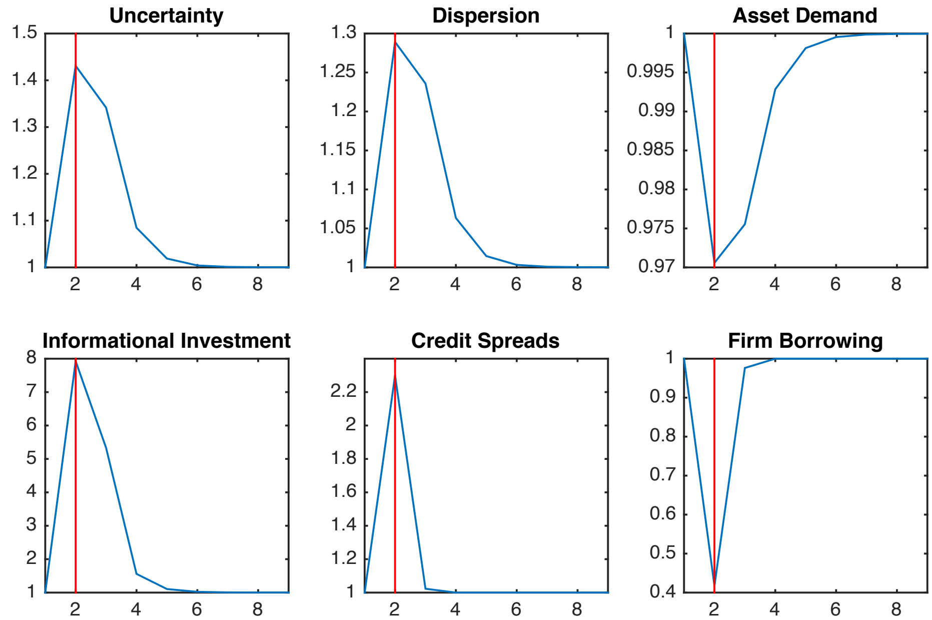


Figure 6: Impulse Responses of Continuous State Model to a three standard deviation drop and rebound in the asset's value.

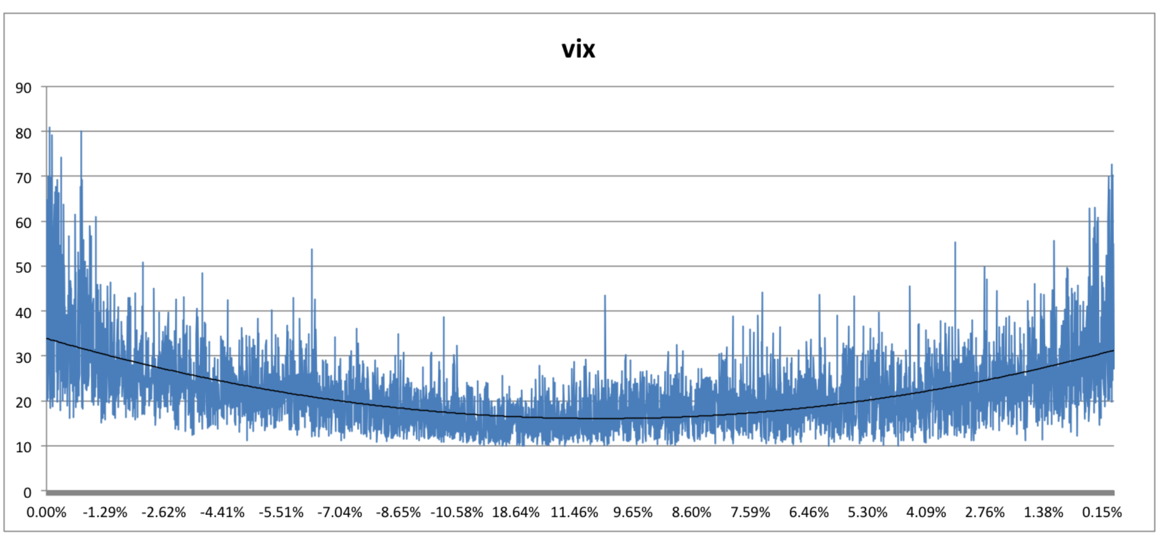


Figure 7: Closing levels of VIX as a function of the ex-ante probabilities of price changes. Probabilities are arranged such that most likely events are in the center, and least likely are in the tails.

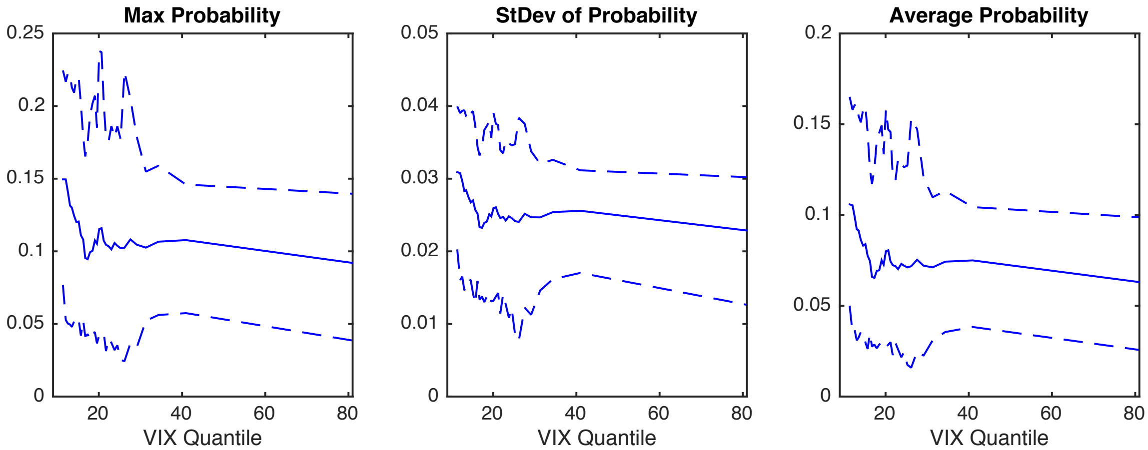


Figure 8: Probability statistics with 95% percentile bounds a function of VIX quantiles

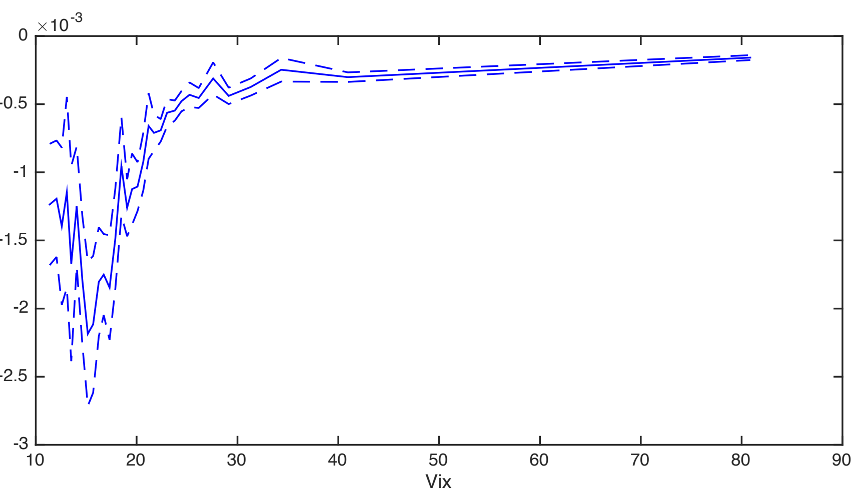


Figure 9: Coefficient of volatility on average intraday probability with 95% confidence interval as a function of VIX quintiles

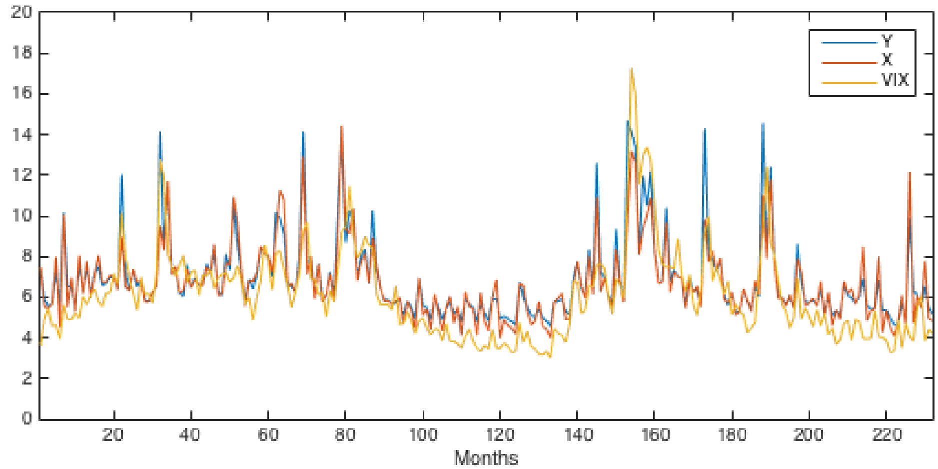
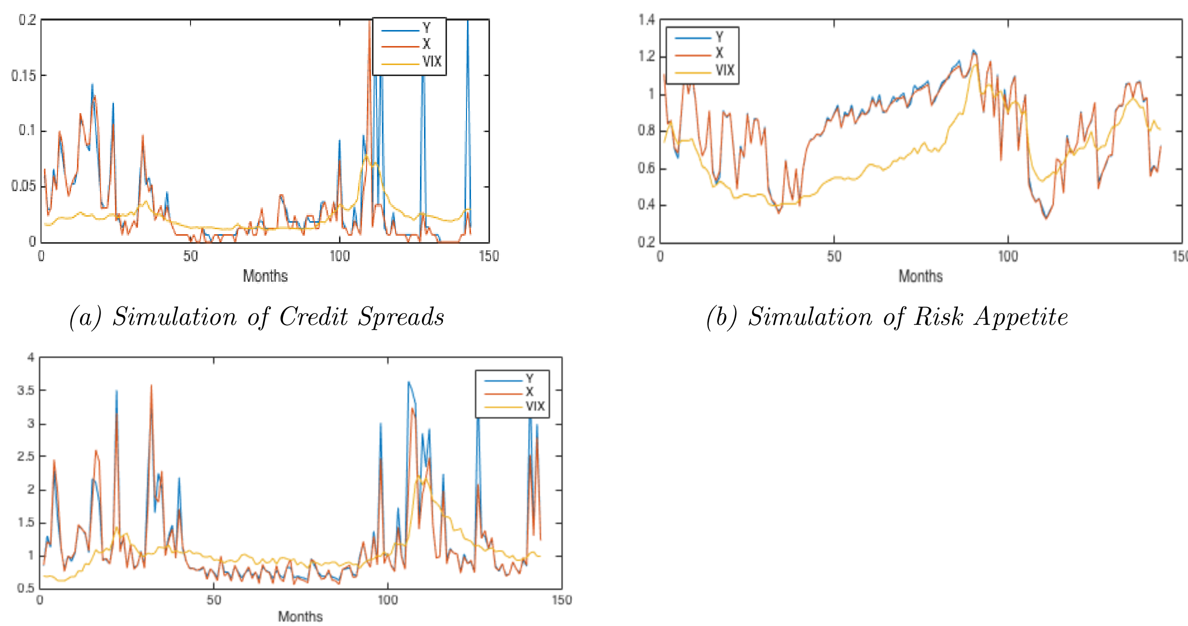


Figure 10: Monthly simulation of the VIX using probabilities as input



(a) Simulation of Credit Spreads

(b) Simulation of Risk Appetite

(c) Simulation of Dispersion of Beliefs

Figure 11: Additional Simulations

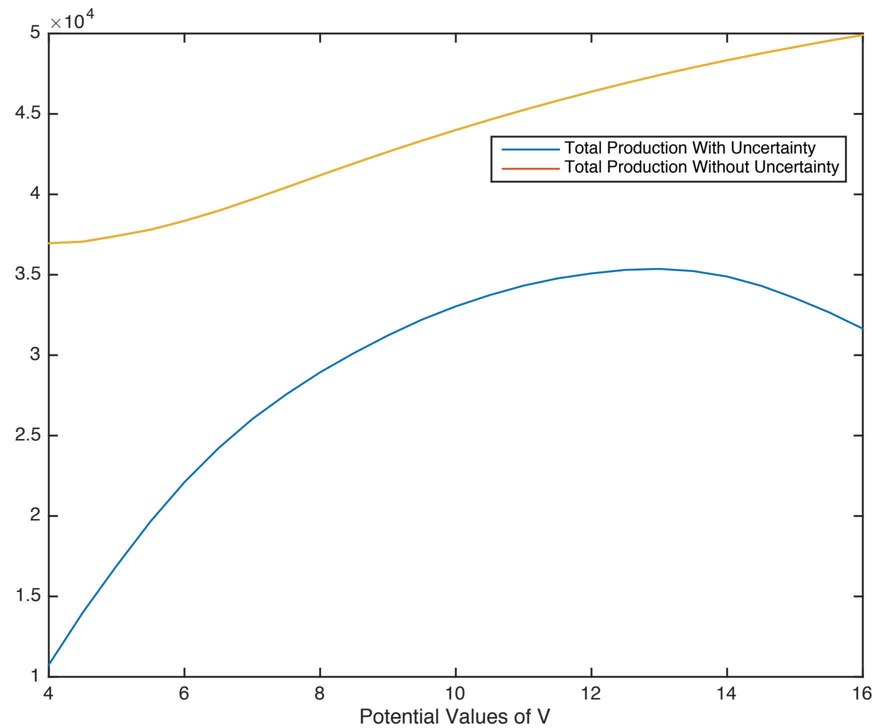
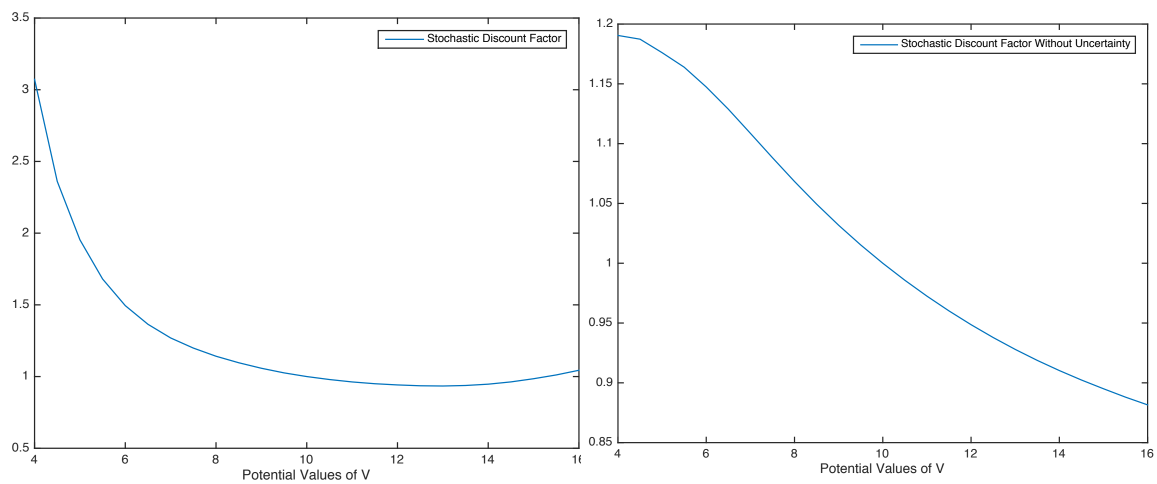


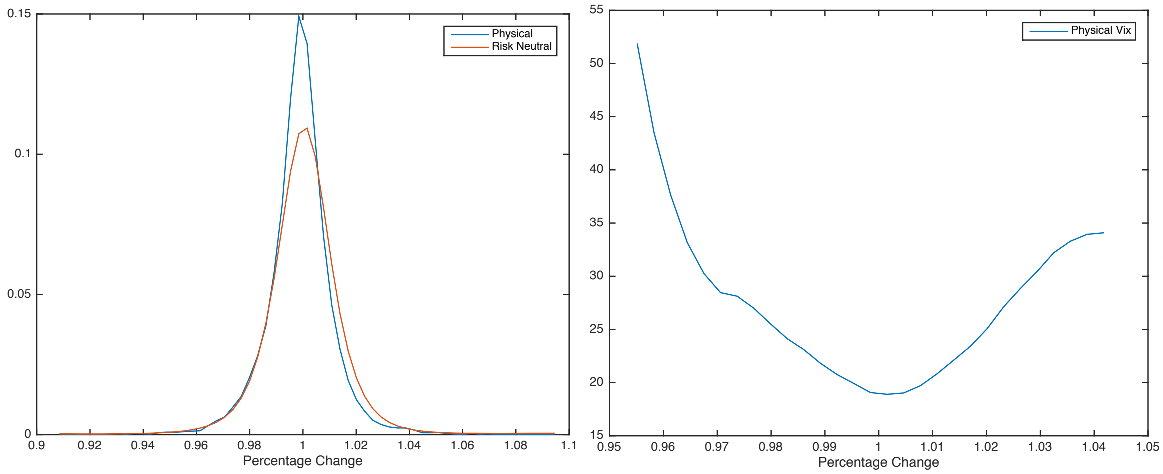
Figure 12: Total Production with and without Uncertainty



(a) SDF with Uncertainty

(b) SDF without Uncertainty

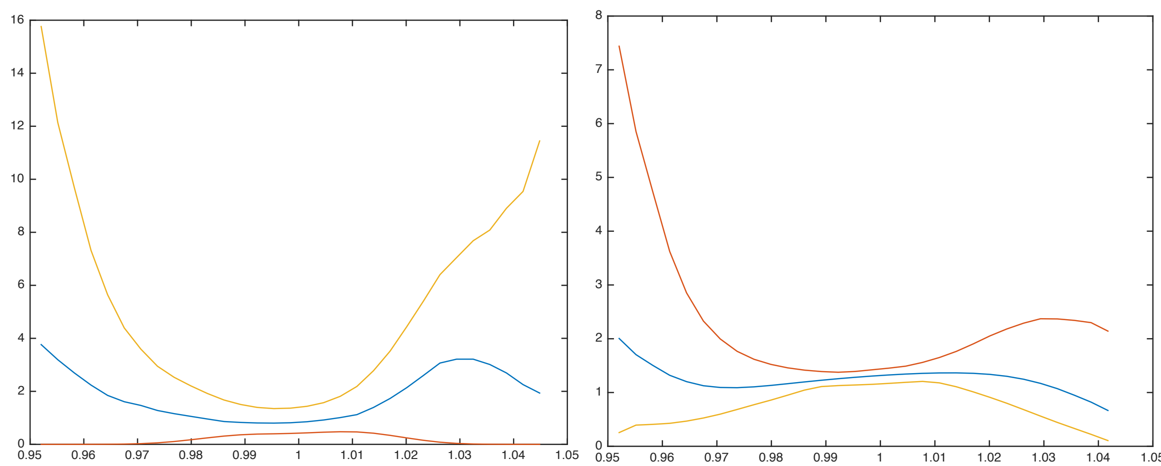
Figure 13: SDFs with and without Uncertainty



(a) Risk Neutral vs. Physical

(b) Physical Vix

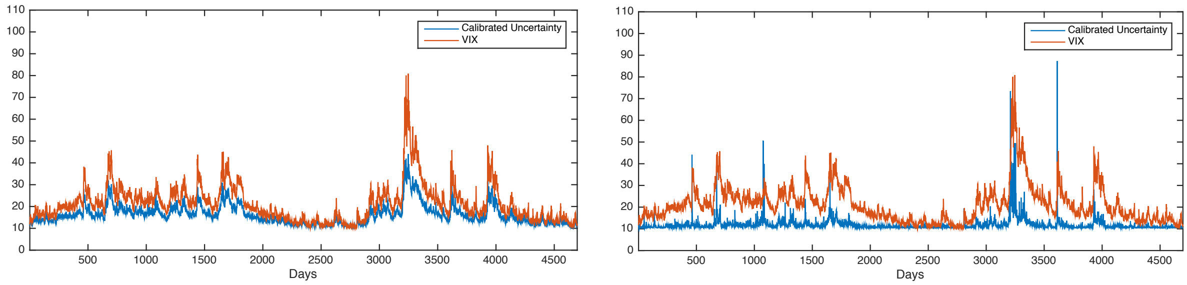
Figure 14: Physical and Risk-Neutral Distributions



(a) Non-Paramterically Estimated Stochastic Discount Factor

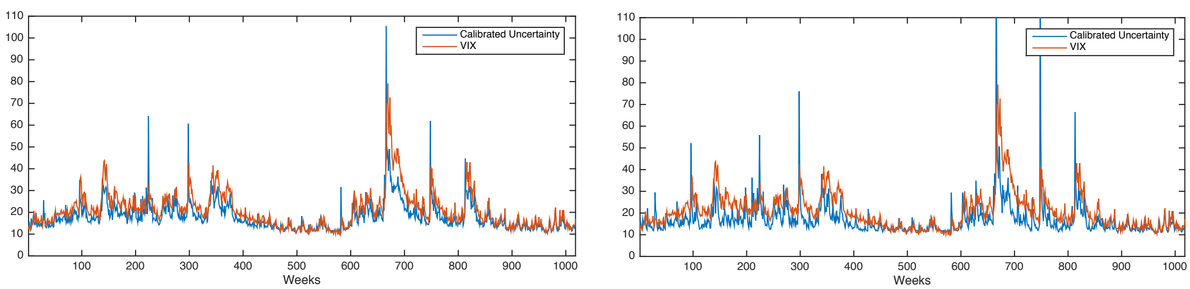
(b) SDF when controlling for conditional uncertainty

Figure 15: Nonparametric SDFs



(a) Daily using probabilities as input *(b) Daily using price changes as input*

Figure 16: Forecasting with Daily Data



(a) Weekly using probabilities as input *(b) Forecasting with Weekly Data*

Figure 17: Weekly using price changes as input

Variable	Mean	StDev	Min	Max
Vix	21.07	8.38	9.89	80.86
Volatility	16.67	10.65	1.79	110.04
Volume	244,924	171,056	1,499	1,145,623
Physical Prob of Intraday Low	10.26%	5.59%	3.20e-7%	23.17%
Physical Prob of Intraday High	9.27%	5.09%	3.02e-5%	24.44%
Risk-Neutral Prob of Intraday Low	8.56%	4.55%	6.14e-5%	31.12%
Risk-Neutral Prob of Intraday High	8.15%	4.54%	2.15e-4%	35.08%

Table 1: Summary Statistics

VARIABLES	(1) Vix	(2) Vix	(3) Vix	(4) Vix	(5) Vix	(6) Vix
Min	-0.00416***	-0.00556***	-0.00398***	-0.00361***	-0.00188***	-0.00188***
	(0.00108)	(0.00132)	(0.000648)	(0.000644)	(0.000669)	(0.000667)
L.Min		-0.00278***	-0.00242***	-0.00226***	-0.00240***	-0.00241***
		(0.00103)	(0.000636)	(0.000629)	(0.000635)	(0.000636)
Close			-3.670***	-3.668***	-3.640***	-3.614***
			(0.101)	(0.101)	(0.0993)	(0.0999)
Volume				0.0134***	0.00505	0.00534
				(0.00350)	(0.00331)	(0.00332)
Volatility					0.0115***	0.0114***
					(0.00154)	(0.00154)
Skew						-0.000495**
						(0.000242)
Constant	-0.000144	-0.000120	0.000755	0.000750	0.000753	0.000769
	(0.000870)	(0.000866)	(0.000578)	(0.000577)	(0.000572)	(0.000573)
Observations	4,535	4,527	4,527	4,527	4,527	4,509

Standard errors in parentheses

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

Notes: The regression is run in first differences, and Newey-West standard errors are shown in parentheses. Min is the lower of the ex-ante probability of the intraday high and the ex-ante probability of the intraday low. Volume is the volume of trade in the S&P index; Volatility is the size of the intraday price movement; Close is the closing value of the S&P index; Skew is the skewness of the VIX as calculated by the CBOE Data is at a daily frequency, and probabilities are risk-neutral, extracted from options data.

Table 2: Effect of price changes on the VIX

VARIABLES	(1) Vix	(2) Vix	(3) Vix	(4) Vix	(5) Vix	(6) Vix
Min	-0.00575***	-0.00670***	-0.00576***	-0.00529***	-0.00168	-0.00177
	(0.00118)	(0.00145)	(0.000885)	(0.000893)	(0.00117)	(0.00117)
L.Min		-0.00176	-0.00293***	-0.00277***	-0.00297***	-0.00300***
		(0.00119)	(0.000733)	(0.000730)	(0.000729)	(0.000729)
Close			-3.695***	-3.693***	-3.664***	-3.640***
			(0.100)	(0.100)	(0.0992)	(0.0996)
Volume				0.00956***	0.00356	0.00383
				(0.00329)	(0.00327)	(0.00328)
Volatility					0.0126***	0.0124***
					(0.00196)	(0.00196)
Skew						-0.000452**
						(0.000229)
Constant	0.000104	0.000133	0.000860	0.000853	0.000854	0.000872
	(0.000865)	(0.000863)	(0.000574)	(0.000573)	(0.000568)	(0.000570)
Observations	4,541	4,536	4,536	4,536	4,536	4,519

Standard errors in parentheses

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

Notes: The regression is run in first differences, and Newey-West standard errors are shown in parentheses. Min is the lower of the ex-ante probability of the intraday high and the ex-ante probability of the intraday low. Volume is the volume of trade in the S&P index; Volatility is the size of the intraday price movement; Close is the closing value of the S&P index; Skew is the skewness of the VIX as calculated by the CBOE; Data is at a daily frequency, and probabilities are physical, extracted from a kernel density estimation of historical distributions.

Table 3: Effect of price changes on the VIX

VARIABLES	(1) max	(2) avg	(3) stdev
quantile	-0.00115***	-0.000663***	-0.000155***
	(7.63e-05)	(2.22e-05)	(1.10e-05)
Constant	0.131***	0.0375***	0.0283***
	(0.00159)	(0.000540)	(0.000222)
Observations	4,534	4,656	4,534

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the results of three regressions. Quantile is a ordinal variable measuring the value of the VIX in 32 buckets. Max is the probability of the most likely price movement in a given day. Avg is the average weighted probability of a price movement. Stdev is the standard deviation of probabilities of price movements. All probabilities are risk-neutral extracted from options data. *** indicates statistical significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 4: Evidence on flattening distribution

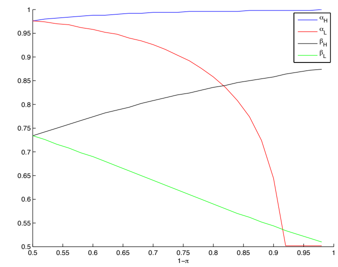
VARIABLES	(1) D.sdf1	(2) D.sdf1
Percentage Change in S&P	-0.000675	-0.00556
	(0.00171)	(0.00491)
Change in Physical VIX		0.0908***
		(0.0285)
Change in Probability		-80.47*
		(41.80)
Change in Starting S&P	0.00150	7.39e-07
	(0.00446)	(7.45e-07)
Constant	0.000535	0.00449
	(0.00167)	(0.00395)
Observations	125,519	125,519
Number of index	31	31

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

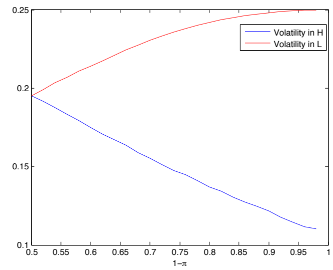
Notes: This table reports the results of two panel regressions. Each column corresponds to a different set of regressors. The dependent variable is the estimated stochastic discount factor, the time frequency is daily, and the cross sectional index is the size of a price movement in percentage terms (divided into 31 buckets). The regression is run in first differences, and Newey-West standard errors are shown in parentheses. *** indicates statistical significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 5: Effect of price changes on the VIX

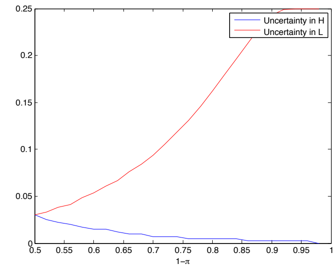
Appendices' Tables and Figures



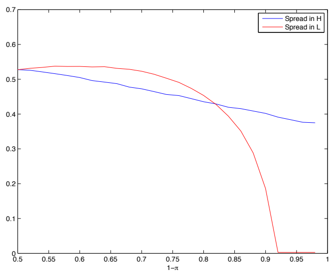
(a) Attention (α, β)



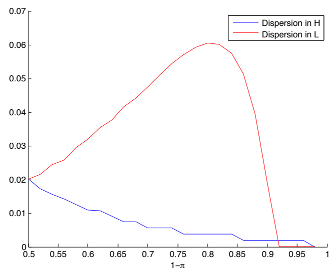
(b) Volatility $(\beta(1-\beta))$



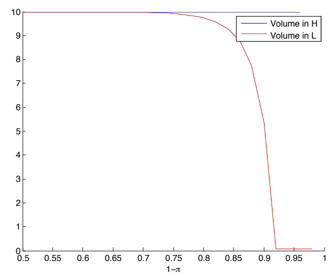
(c) Uncertainty $(\mu(1-\mu))$



(d) Spread

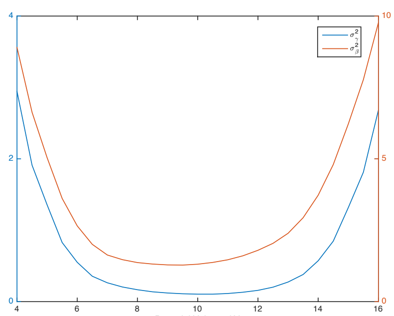


(e) Dispersion $(Var(\mu))$

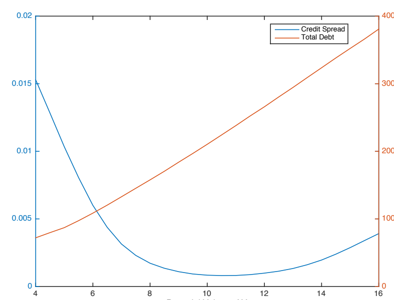


(f) Volume

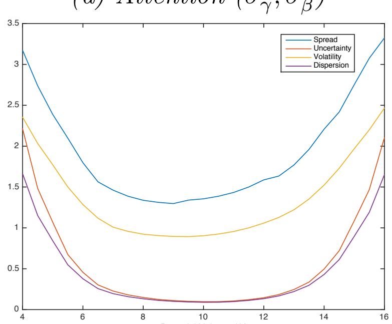
Figure 18: Results under Risk Aversion and Microfounded Cost of Attention



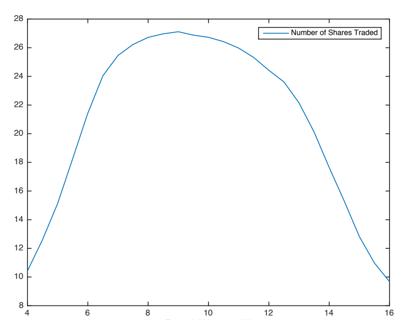
(a) Attention $(\sigma_\gamma^2, \sigma_\beta^2)$



(b) Debt Level and Credit Spread



(c) Bid-Ask Spread, Uncertainty, Expected Volatility, and Dispersion



(d) Asset-Demand (Total number of shares bought/sold)

Figure 19: Simulation of Continuous State Model

VARIABLES	(1) vix	(2) vix	(3) vix
volume	2.13e-05***	1.88e-05***	2.13e-05***
	(1.92e-06)	(2.10e-06)	(2.41e-06)
Open1	-5.16e-06**	-5.55e-06**	-2.68e-06
	(2.08e-06)	(2.26e-06)	(2.60e-06)
Open2	2.65e-06	5.07e-06	2.52e-06
	(2.87e-06)	(3.13e-06)	(3.52e-06)
Open3	-3.45e-05***	-3.31e-05***	-3.34e-05***
	(1.63e-06)	(1.87e-06)	(1.87e-06)
Open4	-9.26e-06***	-8.39e-06***	-7.90e-06***
	(9.07e-07)	(9.91e-07)	(1.16e-06)
Close1	-4.44e-05***	-4.10e-05***	-4.28e-05***
	(2.89e-06)	(3.20e-06)	(3.49e-06)
Close2	-2.88e-05***	-2.49e-05***	-2.70e-05***
	(3.28e-06)	(3.52e-06)	(4.07e-06)
Close3	-4.21e-05***	-4.06e-05***	-4.02e-05***
	(2.26e-06)	(2.59e-06)	(2.59e-06)
Close4	-2.69e-05***	-2.43e-05***	-2.62e-05***
	(1.47e-06)	(1.57e-06)	(1.97e-06)
SPdiff	-0.0193***	0.115***	-0.114***
	(0.00715)	(0.0151)	(0.0159)
volatility	0.207***	0.198***	0.178***
	(0.0201)	(0.0234)	(0.0255)
plow	-97.51***	-90.42***	-101.2***
	(9.049)	(10.37)	(12.58)
phigh	-37.64***	-21.14**	-53.01***
	(9.293)	(10.09)	(14.08)
Constant	17.89***	16.50***	17.71***
	(0.866)	(0.908)	(1.116)
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.

Table 6: Effect of price changes on the VIX

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 π_B such that agents will not purchase a signal. The condition for that π_B is:

$$\begin{aligned} \underline{\pi} \beta_B (1 - \beta_B) 2 - c &= 0 \\ \underline{\pi} \beta_B (1 - \beta_B) 2 &= c \\ \underline{\pi} &= \frac{c}{2\beta_B (1 - \beta_B)} \end{aligned}$$

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 π for which all agents will buy a signal. Such a result depends on the value of c. □

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 purchasing a signal is:

$$\begin{aligned} \frac{\overline{\pi} \beta_B (1 - \beta_B) \frac{T}{2}}{\left(\frac{T}{2} + (1-T)(1-\beta_B)\right)\left(\frac{T}{2} + (1-T)\right)} &> c \\ \frac{c\left(\frac{T}{2} + (1-T)(1-\beta_B)\right)\left(\frac{T}{2} + (1-T)\right)}{\beta_B (1 - \beta_B) \frac{T}{2}} &> \overline{\pi} \end{aligned}$$

Therefore if $\underline{\pi} < \pi < \overline{\pi}$, some but not all agents will purchase a signal, and the solution will be interior. Define the benefit of purchasing as signal as $P(s, \pi, \beta_B) \equiv \frac{\overline{\pi} \beta_B (1-\beta_B) \frac{T}{2}}{\left(\frac{T}{2} + (1-T)(1-\beta_B)\right)\left(\frac{T}{2} + (1-T)\right)}$. 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. $\square$

Proof. Proposition (3)

Define $P(\pi) \equiv \frac{\beta_B(1-\beta_B)\frac{T}{2}}{\left(\frac{T}{2}+(1-T)(1-\beta_B)\right)\left(\frac{T}{2}+(1-T)\right)}$ 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 $\frac{\pi_B\beta_B(1-\beta_B)\frac{T}{2}}{\left(\frac{T}{2}+(1-T)(1-\beta_B)\right)\left(\frac{T}{2}+(1-T)\right)} - c + \delta\pi_B V(k, 1-k) < 0 \leq \pi_B V(0.5, 0.5)$. $\square$

Proof. Proposition (4)

First, we can rewrite equations 14 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_γ^2. □

Proof. Corollary (2)
The necessary condition is that $\frac{dU}{d\phi(B_t)} => 0$. This is trivially true. □

Proof. Corollary (3)
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 $E_{B_1}[\phi_{B,t+1}(B)] > E_{B_2}[\phi_{B,t}(B)]$. □

B Dynamic Heterogeneous Signal Acquisition

In this section, I will consider the dynamic formulation of the simple model, while still allowing agents to behave heterogeneously with regards to purchasing signals. For simplicity I will set $\beta = 0.5$.

Assume that $\pi_{t+1} = 0.5 + (s/2)\rho$ where $\rho < 1$. Unlike the representative agent, heterogeneous agents do not internalize the long-term consequences of their individual decisions.Then an individual agent faces the following problem in terms of signal acquisition:

$$V(\pi, 1-\pi) = \sum_{\pi} \max\left\{ \frac{\pi_B \beta(1-\beta)\frac{T}{2}}{\left(\frac{T}{2} + (1-T)(1-\beta_B)\right)\left(\frac{T}{2} + (1-T)\right)} - c + \delta V(\pi_{t+1}, 1-\pi_{t+1}), \delta V(\pi_{t+1}, 1-\pi_{t+1}) \right\}$$

So the condition for buying a signal in the dynamic game is the same as in the static game:

$$\begin{aligned}
\frac{\pi_t \beta(1-\beta)\frac{T^*}{2}}{\left(\frac{T^*}{2} + (1-T^*)(1-\beta_B)\right)\left(\frac{T^*}{2} + (1-T^*)\right)} - c &> 0 \\
\frac{\pi_t \frac{T^*}{8}}{\left(\frac{T^*}{2} + \frac{(1-T^*)}{2}\right)\left(\frac{T^*}{2} + (1-T^*)\right)} &> c \\
\frac{\pi_t \frac{T^*}{4}}{\left(1 - \frac{T^*}{2}\right)} &> c \\
\frac{\pi_t T^*}{4 - 2T^*} &> c
\end{aligned}$$

Therefore the equilibrium level of s for interior solutions is defined as:

$$\begin{aligned}
\pi_t T^* &= c(4-2T^*) \\
T^*(\pi_t + 2c) &= 4c \\
\frac{T}{T+s(1-T)} &= \frac{4c}{\pi_t + 2c} \\
\frac{T(\pi_t+2c)}{4c(1-T)} - \frac{T}{1-T} &= s
\end{aligned}$$

This is linear in π_t. Therefore, the shape of π_{t+1} as a function of π_t is as follows. For very low levels of π_t, $\pi_{t+1} = 0.5$, for intermediate levels, $\pi_{t+1} = 0.5 + \frac{s\rho}{2}$, and depending on parameter values, for high values of π_t, $\pi_{t+1} = 0.5(1+\rho)$. Visually, it would look like a straight line at 0.5 kinked into a positively sloped line at $\underline{\pi} = 2c$, and rising until $\overline{\pi} = \frac{(4-2T)c}{T}$, at which point it remains at $\pi_{t+1} = 0.5(1+\rho)$ until $\pi_t = 1$.

Intersections of this 'policy function' with the 45-degree line determine steady states, of a sort. There are four types of intersections:

1. $\pi_t = \pi_{t+1} = 0.5$. For sufficiently high values of c, agents will be uncertain forever, caught in an uncertainty trap. This steady state is absorbing and inescapable.

2. $0.5 < \pi_t < \frac{1+\rho}{2}$, and $\frac{T\rho}{8c(1-T)} < 1$. This steady state is absorbing, but not inescapable.

3. $0.5 < \pi_t < \frac{1+\rho}{2}$, and $\frac{T\rho}{8c(1-T)} > 1$. This steady state not absorbing.

4. $\pi = \frac{1+\rho}{2}$. This steady state is absorbing but not inescapable.

The first and the last steady state are what remain when we switch from heterogeneous decision making to a representative agent.

C Inattention

Here the model will be restructured to allow for a more micro founded version of attention - namely the Inattentive Valuation structure of Woodford (2012). The basic versions - the timeline, environment and asset all are structured and behave in the same way - the change will come in the attentional choices and the risk aversion of informed traders.

C.1 Attentional Choices

In order to make micro founding attentional cost structure interesting, we need to expand the state-space of possible accuracies. Therefore, the definitions of signals for traders are:

$$\alpha_i = P(\text{signal} = 0|\eta_B = 0) = P(\text{signal} = 1|\eta_B = 1)$$

The public signal is again defined as:

$$\beta_i = P(\text{signal} = 0|\eta_B = 0) = P(\text{signal} = 1|\eta_B = 1)$$

We will assume, without loss of generality that α and β are both weakly greater than 0.5.

C.2 Informed Traders

We will solve the problem of the informed traders by working backwards. Informed traders, in period 1, have already made their attentional choices, and have received their individual perceptions of the value of η_i. At this point, they want to solve the following problem:

$$\max_z E[w^k] \quad s.t. \quad w = z(B_i + \eta_i - P) + N \quad \& \quad 0 < k < 1$$

Where P is the price at which agents can choose to trade (will be different depending on whether agents want to buy or sell), N is an initial endowment of cash, and k is a measure of risk-aversion. The solution for demand is:

$$z = \frac{N - N\gamma}{(B_i - P)\gamma - (B_i + 1 - P)} \quad \text{where} \quad \gamma = \left(\frac{(1-\mu)(P-B)}{(B+1-P)\mu}\right)^{\frac{1}{k-1}}$$

Where μ is an agent's idiosyncratic belief that $\eta_i = 1$ which is informed by his own attentional choice, his signal, and the price.

We will assume here for the sake of tractability, that noise traders will participate, on average, as much as informed traders. This is not a crucial assumption.

C.3 Equilibrium

Given the way we have set up the model so far, we know that, at the beginning of period 1, there are two groups of traders - one group comprised of agents who receive the signal '1', and the noise traders that mimic them, and one group comprised of agents who receive the signal '0', and the noise traders that mimic them.

First assume that market making sector receives a signal that '$\beta' = 1$ - and it knows the signal to have accuracy β_i:

$$\begin{aligned}
P(\text{buy}|\eta_i = 1) &= \frac{f(\cdot)T}{2} + (1-T)\alpha_i \\
P(\text{buy}|\eta_i = 0) &= \frac{f(\cdot)T}{2} + (1-T)(1-\alpha_i)
\end{aligned}$$

That is the number of uninformed buys $T/2$ is unrelated to the fundamental value of the asset, but the number of informed buys does vary. The market making sector weights this by its own belief:

$$P(\text{buy}) = \beta\left(\frac{f(\cdot)T}{2} + (1-T)\alpha_i\right) + (1-\beta)\left(\frac{f(\cdot)T}{2} + (1-T)(1-\alpha_i)\right)$$

Then we can calculate the following:

$$\begin{aligned}
P(\eta_i = 1|\text{buy}) &= \frac{P(\text{buy}|\eta_i = 1)P(\eta_i = 1)}{P(\text{buy})} \\
&= \frac{\frac{f(\cdot)T}{2} + (1-T)\alpha_i\beta_i}{\frac{f(\cdot)T}{2} + (1-T)(\beta_i\alpha_i + (1-\beta_i)(1-\alpha_i))}
\end{aligned}$$

This value, as proven by Glosten and Milgrom is exactly equal to the ask price charged by the market making sector under perfect competition. Therefore we can calculate the following four values:

$$\begin{aligned}
\text{ask}_1 &= B + \frac{\left(\frac{f(\cdot)T}{2} + (1-T)\alpha_i\right)\beta_i}{\frac{f(\cdot)T}{2} + (1-T)(\beta_i\alpha_i + (1-\beta_i)(1-\alpha_i))} \\
\text{bid}_1 &= B + \frac{\left(\frac{f(\cdot)T}{2} + (1-T)(1-\alpha_i)\right)\beta_i}{\frac{f(\cdot)T}{2} + (1-T)(\beta_i(1-\alpha_i) + (1-\beta_i)\alpha_i)} \\
\text{ask}_0 &= B + \frac{\left(\frac{f(\cdot)T}{2} + (1-T)\alpha_i\right)(1-\beta_i)}{\frac{f(\cdot)T}{2} + (1-T)((1-\beta_i)\alpha_i + \beta_i(1-\alpha_i))} \\
\text{bid}_0 &= B + \frac{\left(\frac{f(\cdot)T}{2} + (1-T)(1-\alpha_i)\right)(1-\beta_i)}{\frac{f(\cdot)T}{2} + (1-T)(\beta_i\alpha_i + (1-\beta_i)(1-\alpha_i))}
\end{aligned}$$

A first, quick result, is the movement of the spread (difference between the ask and the bid) with public and private information. As the market makers become better informed relative to the trader, spreads decline, and as the market makers become worse informed relative to the trader, spreads increase, as a protective measure.

Agents will buy if the ask is below their updated expectation of η, and will sell if the bid is above their expectation:

$$
\begin{aligned}
E[\eta = 1 | `\alpha' = 1, `\beta' = 1] &= \mu_{11} &= \frac{\alpha\beta}{\alpha\beta + (1-\alpha)(1-\beta)} \\
E[\eta = 1 | `\alpha' = 1, `\beta' = 0] &= \mu_{10} &= \frac{\alpha(1-\beta)}{\alpha(1-\beta) + \beta(1-\alpha)} \\
E[\eta = 1 | `\alpha' = 0, `\beta' = 1] &= \mu_{01} &= \frac{\beta(1-\alpha)}{\beta(1-\alpha) + \alpha(1-\beta)} \\
E[\eta = 1 | `\alpha' = 0, `\beta' = 0] &= \mu_{00} &= \frac{(1-\alpha)(1-\beta)}{\alpha\beta + (1-\alpha)(1-\beta)}
\end{aligned}
$$

The updated expectations satisfy the conditions of the equilibrium (see appendix for proof).

C.4 Attentional Choice for Informed Traders

Utility for the informed traders is defined as follows:

$$
\begin{aligned}
U_{11} &= \mu_{11}(\alpha_i, \beta_j)(z_{11}(\alpha_i, \beta_j)(B + 1 - \text{ask}_1(\alpha_i, \beta_j)) + N)^k \\
&\quad + (1 - \mu_{11}(\alpha_i, \beta_j))(z_{11}(\alpha_i, \beta_j)(B - \text{ask}_1(\alpha_i, \beta_j)) + N)^k
\end{aligned}
$$

We can show that $(U_{ij})_\alpha > 0$ for all i, j, which just means that overall utility:

$$U = (U_{11} + U_{00})(\alpha_i\beta_j + (1-\alpha_i)(1-\beta_j)) + (U_{10} + U_{01})(\alpha_i(1-\beta_j) + (1-\alpha_i)\beta_j)$$

depends positively on α. This means that agents want to get accurate signals. Informed traders know the value of π, and the value of θ, and therefore know what β^*_L and β^*_H are. Therefore their problem now becomes:

$$\max_{\alpha_L, \alpha_H} (1-\pi)U(\alpha_H, \beta_H) + \pi U(\alpha_L, \beta_L) - \theta_T I^*(\alpha_L, \alpha_H)$$

where θ_T is an exogenously given marginal cost of increasing their channel capacity. There is a marginal cost to increasing the channel capacity (as described by I^*) which would increase accuracy.

In this formulation of the market, the assumption to model attentional choice in this way allows us to have the attentional choice be decision theoretic as opposed to game theoretic - α will be a function of β but not vice-versa. The constraint function is defined as:

$$\begin{aligned} I((\alpha_H, \alpha_L); \pi) &= -E_{P(x,r)}[\ln(P(x)) - \ln(P(x|r))] \\ \pi^* &= argmax I((\alpha_H, \alpha_L); \pi) \\ I^*(\alpha_H, \alpha_L) &= I((\alpha_H, \alpha_L); \pi^*) \end{aligned}$$

where x is the state that realizes, and r is the representation perceived by the agent. In this particular case, the functional form will be:

$$I^*(\alpha_H, \alpha_L) = \ln(2) + \ln\left(e^{\alpha_L \ln(\alpha_L) + (1-\alpha_L)\ln(1-\alpha_L)} + e^{\alpha_H \ln(\alpha_H) + (1-\alpha_H)\ln(1-\alpha_H)}\right)$$

There is an important reason as to why we have chosen to constraint the channel capacity, as originally formulated by Shannon 2001 and Woodford 2012, as opposed to the more canonical constraint on mutual information as formulated by Sims 2003 and others. As was shown in Woodford 2012 and also in Sundaresan and Turban 2014, there is a large array of experimental and neurological evidence showing that low probability events are less well identified than high probability events. Using this methodology allows us to generate 'countercyclical' dispersion. For the technical properties, characteristics, and proofs of I^* such as convexity, monotonicity and the like, please refer to Woodford 2012 or Sundaresan and Turban 2014.

C.5 Results

Here we will simulate the model for the following parameters. $B_L = B_H = 0$ $N = 10$ $T = 0.35$ $\theta_T = 1$ $\theta_M = 1.3$ $k = 0.8$ The first set of variables that deserve attention are those of $\alpha_L, \alpha_H, \beta_L$ and β_H. These are plotted in Figure 7.

As we can see, both α and β are increasing in the likelihood of the state. So, as $1 - \pi$ increases, α_H and β_H increase and α_L and β_L decrease. As we can see, for very low values of π, $\alpha_L = 0.5$. Volatility obviously tracks β, but again, the general shape is that as the likelihood of a state increases, the expected volatility upon its realization decreases. As α and β are both increasing in $1 - \pi$, uncertainty will follow the opposite path - as we have seen, uncertainty is decreasing in both

α and β. The spread shows the tension between the two forces that drive the liquidity dry-ups. The spread decreases in the high state, as the market making sector receives better signals and requires less adverse-selection protection. in the low state, the spread initial increases for this same reason, but then starts to fall because both the market makers and the traders are receiving lower quality information, so traders demand less and less risk. Dispersion here is decreasing in the high state, as both α and β are increasing. It follows a hump-shape in the low state as agents at first start to receive more information from their own signals, as opposed to the public ones, but eventually declines, as all agents are equally uninformed.

D Options Volume

The key mechanism in the model of this paper is that of inattention, or informational investment. Unfortunately, there is no direct measure I am aware of for such a variable. However, in this appendix, I consider a proxy: hedging.

D.1 Interpretation

One could think of purchasing informational investment or paying attention, as a form of reducing conditional volatility or uncertainty. Therefore, purchasing securities that 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, a common pattern that emerged was that the way in which they prepared for the unexpected was to utilize options to develop a long position in volatility. Therefore, we will turn to options data, to answer these questions.

D.2 Dataset

The dataset used here is from MarketDataExpress, and provides information on the types and sizes of trades in the SPX options index at a daily frequency. Specifically, 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.

D.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 the ex-ante probabilities of events, 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.

D.4 Results

The three columns of table 6 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.

E Alternative Specification

Suppose that instead of trying to provide as accurate a signal as possible, the public entity had a more noble goal: minimizing credit spreads. Then the public entity's problem would be:

$$\int \min_{\sigma_\beta(B)} \phi_{B,t}(B) \left(\int p(\eta\text{-signal} = x)(r(\mu_p(x), \sigma_p(x), r_f) - r_f)dx \right) - c_p \nu_p(\sigma_\beta^2(B)) d\phi(B) \tag{21}$$

E.1 Predictions

Compare the predictions of this objective function to that of Section 4. We use the same parameters as before in Figure 19.[10] One can see how agents (both public and private) choose to invest in information. As before, conditional on these selections, one can see what the expectations for credit spreads, levels of borrowing, demand for financial assets, bid-ask spreads, uncertainty, volatility, and dispersion of beliefs are.

Overall, it is evident that the general patterns are very similar to the case considered in the main body of the paper. Let's examine each of these graphs in turn. First, it is evident that the the graphs in (a) are not symmetric. Rather σ_β is skewed towards the left, while σ_γ is skewed towards the right. The public entity is trying to minimize credit spreads which are naturally higher for lower values of V_t - so information is more valuable to the public entity at lower values of V_t. This doesn't imply that public information is good for unexpected negative shocks, rather that it is better than it is for equally unexpected positive shocks. Conversely, since public and private information are strategic substitutes, private agents will choose to be better informed, on average, for positive shocks than for negative ones. That is, for a normal distribution of shocks, informational investment will be highest near the mean, and will deteriorate quickly for the tails. The statistics are virtually identical to the other objective function, save for the skewness in the shapes.

10. The parameter values here are $B_t = 10$, $\sigma_\eta^2 = 4$, $\sigma_\epsilon^2 = 4$, $c_p = 0.0002$, $\delta = 0.95$, $c = 1$, $r_f = 0.03$, $\rho = 0.6$, $\alpha = 0.6$. α and ρ are selected with some calibration in mind, while σ_η and σ_ϵ 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.