

# Heterogeneity in Beliefs and Volatility Tail Behavior\*

Gurdip Bakshi,<sup>†</sup> Dilip Madan,<sup>‡</sup> and George Panayotov<sup>§</sup>

June 3, 2014

---

---

\*The efforts of Hank Bessembinder (the editor) and the referee have greatly improved the paper. The authors acknowledge helpful feedback from Charles Cao, Peter Carr, Fousseni Chabi-Yo, Xiaohui Gao, Steve Heston, Nikunj Kapadia, Tao Li, Huston McCulloch, Georgios Skoulakis, René Stulz, Liuren Wu, Ingrid Werner, and Jin Zhang. An earlier version of the paper was presented at the Chinese University of Hong Kong, Hong Kong University, HKUST, City University of Hong Kong, University of Maryland, and Ohio State University. We welcome comments, including references to related papers we have inadvertently overlooked. Any remaining errors are our responsibility alone.

<sup>†</sup>Smith School of Business, University of Maryland, College Park, MD 20742, USA; tel: 301-405-2261; email: gbakshi@rhsmith.umd.edu .

<sup>‡</sup>Smith School of Business, University of Maryland, College Park, MD 20742, USA; tel: 301-405-2127; email: dlm@rhsmith.umd.edu .

<sup>§</sup>School of Business, Hong Kong University of Science and Technology, Clearwater Bay, Hong Kong; tel: +852-2358-5049; email: panayotov@ust.hk .

## **Abstract**

We propose a model of volatility tail behavior, in which investors display aversion to both low volatility and high volatility states, and, hence, the derived pricing kernel exhibits an increasing and decreasing region in the volatility dimension. The model features investors who have heterogeneity in beliefs about volatility outcomes, and maximize their utility by choosing volatility-contingent cash flows. Our empirical examination suggests that the model is better suited to reproduce data features in the left tail of the volatility distribution, both qualitatively and quantitatively.

**KEY WORDS:** traded volatility, VIX option returns, tails of pricing and physical distributions, heterogeneity about volatility outcomes, belief distributions.

# I. Introduction

The purpose of this paper is to propose a new framework for characterizing volatility tail behavior. Our central insight is that by allowing for heterogeneity in beliefs about volatility outcomes, one can obtain pricing kernels that are better suited to reconcile features, germane to traded market volatility, and to mimic the data counterparts, both qualitatively and quantitatively.

To motivate our modeling approach, let  $v \in (0, \infty)$  be the market return volatility, where, for example,  $v^2$  could be an estimate of the quadratic variation in the log of the market index. Suppose that the physical density of  $v$ , denoted by  $\Phi[v]$ , and the pricing (or risk-neutral) density, denoted by  $q[v]$ , are both lognormal. Specifically,  $\Phi[v] = \frac{1}{v\sigma_p\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_p^2}(\ln(v) - \mu_p)^2\right)$ ,  $q[v] = \frac{1}{v\sigma_q\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_q^2}(\ln(v) - \mu_q)^2\right)$ , and hence, the change of measure (or pricing kernel) is  $d\mathbb{Q}/d\mathbb{P} = \psi_0 \exp(\psi_1 (\ln(v))^2 + \psi_2 \ln(v))$ , where  $\psi_0$  is a constant,  $\psi_1 = 1/(2\sigma_p^2) - 1/(2\sigma_q^2)$ , and  $\psi_2 = \mu_q/\sigma_q^2 - \mu_p/\sigma_p^2$ . It is seen now that if  $\sigma_q = \sigma_p$ , as in many standard models, then  $d\mathbb{Q}/d\mathbb{P}$  could be monotonically increasing in the volatility dimension, which would imply geometrically that  $q[v]$  crosses  $\Phi[v]$  only once, and would entail an asset-pricing model where deeper out-of-the-money volatility puts have higher expected returns. On the other hand, if, for example,  $\sigma_q > \sigma_p$ , then  $d\mathbb{Q}/d\mathbb{P}$  exhibits a decreasing region for low volatility levels, and increasing region for high volatility levels, implying that  $q[v]$  crosses  $\Phi[v]$  twice. Out-of-the-money volatility puts and calls in this setting manifest low expected returns. Besides, expected returns could be decreasing for deeper out-of-the-money volatility puts.

Our interest in a model where investors display aversion to both low and high volatility states, giving rise to a non-monotonic change of measure for volatility, is guided, intuitively, by certain

specifics of the volatility market. In particular, (i) the underlying volatility is not traded directly, but only via volatility contingent claims, and (ii) unlike equities, these claims are in zero net supply, and offer exposure to volatility movements both on the upside and the downside, for different investors (see also CBOE (2009)). The model prediction that both puts and calls on volatility will have low expected returns can be tested using newly available data from the VIX option market.

In our model of the volatility market, investors display heterogeneity in beliefs about volatility outcomes. Posited within the model are two types of investors who anticipate volatility to increase, or to decline, respectively. In our setup, investors maximize their utility by choosing volatility-contingent cash flows, given their beliefs about the distribution of volatility, and we derive the pricing kernel in closed form.

We also contrast our model to a conventional approach, where the pricing kernel is monotonically declining in market returns and increasing in market return volatility, which reflects aversion to volatility. Under certain conditions, we show that a model of this type, where high volatility states are disliked, predicts that the expected returns of puts (calls) on volatility are decreasing in strike and are positive (negative) for sufficiently low (high) strikes. Comparing these predictions to those from our model can be important, as they can be traced to ideas that are often at the core of finance theory, including mean-variance theory.

Elaborating, we explore whether there is empirical support for our model with heterogeneity in beliefs versus the conventional approach to modeling volatility tail behavior. We show that the two models share the same qualitative prediction about the pattern of expected returns of volatility calls, and, hence, focus on the hypotheses that the expected returns of volatility puts either (i) decrease in

strike, or (ii) increase in strike. The results support, more generally, the model with heterogeneity in beliefs in our sample, which appears to be better suited for capturing the left-tail behavior of traded volatility.

Our paper brings insights similar to Song and Xiu (2013), who estimate a U-shaped volatility pricing kernel, confirming empirically that investors display aversion to both low and high volatility states. The study of Amengual and Xiu (2013) investigates the origin and differences in the risk pricing specifically of volatility jumps in the two tails, while Aït-Sahalia, Karaman, and Mancini (2013) study the risk premium from the term structure of variance swaps. Furthermore, Jackwerth and Vilkov (2013) examine the joint behavior of the risk-neutral distributions of the market index and its expected volatility, and show how to extract non-parametrically the expected risk-neutral correlation between them. We differ from many extant studies by demonstrating that a possible driver of the risk pricing for low and high volatility states is heterogeneity in beliefs about volatility outcomes, which could generate low average returns of out-of-the-money volatility calls and puts, that are also decreasing in moneyness.<sup>1</sup>Xiong (2013) provides an overview of the literature on the importance of modeling heterogeneity in beliefs in many empirically relevant contexts.

The modeling and empirical issues that are broached in our paper are likely to gain traction,

---

<sup>1</sup>In the spirit of the index options literature, there is interest in understanding how alternative models of volatility dynamics fare in fitting the prices of VIX options and other volatility derivatives, and in producing accurate forecasts of quadratic variation. Among others, this promising line of research includes Andersen, Bollerslev, and Meddahi (2005), Buehler (2006), Zhu and Zhang (2007), Carr and Lee (2008), Gatheral (2008), Sepp (2008), Aït-Sahalia and Mancini (2008), Cont and Kokholm (2009), Lin (2009), and Egloff, Leippold, and Wu (2010).

given the advent of active trading in volatility. Besides, our discussions could complement the insights from studies that rely on realized volatility, index-option inferred volatility, and variance swaps (e.g., Rubinstein (1994), Andersen, Bollerslev, Diebold, and Labys (2003), Carr and Wu (2009), Todorov (2010), Bollerslev and Todorov (2011), and Christoffersen, Heston, and Jacobs (2013)). Our contribution aims to fill some gaps from the perspective of the left tail of the volatility distribution.

## II. A Model with Traded Volatility and Heterogeneity in Beliefs

This section first states our model assumptions, then describes the choice problem of the investors. Next, we present the solution for the pricing kernel and study its properties.

### A. A Model of the Volatility Market

**Assumption 1** *The model is two-period, with three dates (denoted  $0, T, T'$ ).*

In analogy with the VIX, and to provide a link to our data on returns of volatility options, volatility  $v_T \in (0, \infty)$  reflects uncertainty over  $[T, T']$ , and, without loss of generality, we set the current level of volatility to  $v_0 = 1$ . There is a continuum of realizations of volatility.

It is worth noting that while our model is cast in a two-period setting, we have also considered an analogue that treats volatility exposure in continuous time. Under the assumption that volatility follows a diffusion process, our analysis in Appendix B shows that the derived pricing kernel is monotonic in volatility. On the other hand, we recognize that under a richer volatility dynamics in

continuous time, that allows for two-sided jumps, a non-monotonic kernel could be obtained (e.g., Amengual and Xiu (2013) and Song and Xiu (2013)).

To examine the implications of non-monotonic pricing kernels in the setting of a volatility market with heterogeneity in beliefs about volatility outcomes, we pursue a discrete-time model that is suited to our objectives and that could offer the flexibility of using general change of measure densities (see, for instance, Neftci (2007)).

**Assumption 2** *Options on volatility of all strikes are traded in the volatility market.*

Any function of volatility  $v_T$  can be synthesized with a portfolio of options on volatility. The market for volatility is complete.

**Assumption 3** *The contingent claims on volatility are in zero net supply in the aggregate.*

This assumption ensures that the contingent cash flows, purchased by some investors, are sold by others.

**Assumption 4** *Investors can be divided into two types, denoted  $F$  and  $G$ , which are differentiated by their beliefs about volatility. They have personalized densities about volatility outcomes, denoted  $\Phi_f[v_T]$  and  $\Phi_g[v_T]$ , respectively. In the model, the two investor types are present with mass  $0 < \phi < 1$  and  $1 - \phi$ , respectively.*

Differences in beliefs, and possibly risk aversion, can give rise to trading in contingent claims on volatility. We suppose the existence of a physical density  $\Phi[v_T]$ , and a pricing density  $q[v_T]$ .

For simplicity, the interest rate is set equal to zero, and hence, we omit discounting. The pricing density  $q[v_T]$  is determined endogenously, so as to clear the volatility market. Our approach generalizes to multiple types of investors.

Importantly, one should *not* construe our framework to imply that the future level of volatility is treated by investors as a variable independent from the past. On the contrary, we describe the investors' beliefs about future volatility via time 0 conditional densities, which implicitly incorporate all past information. Therefore, we adhere to the standard approach for representing the role of the past.

When the probability measures corresponding to  $\Phi_f[v_T]$  and  $\Phi_g[v_T]$  are absolutely continuous with respect to  $\Phi[v_T]$ , the change of measure densities for the two types of investors can be defined as follows:

$$(1) \quad \mathcal{A}[v_T] = \frac{\Phi_f[v_T]}{\Phi[v_T]}, \quad \mathcal{B}[v_T] = \frac{\Phi_g[v_T]}{\Phi[v_T]}.$$

It is assumed that the densities  $\Phi_f[v_T]$ ,  $\Phi_g[v_T]$ , and  $\Phi[v_T]$  belong to a finite parameter class, are twice continuously differentiable, and have finite moments up to order four.

**Assumption 5** *Investors F and G hold volatility contingent cash flows, denoted  $f[v_T]$  and  $g[v_T]$ , respectively, that are self-financed, i.e.,*

$$(2) \quad 0 = \int f[v_T]q[v_T]dv_T \quad \text{and} \quad 0 = \int g[v_T]q[v_T]dv_T,$$

with  $f \in C$  and  $g \in C$ , where  $C = \{c|c : \mathfrak{R}^+ \rightarrow \mathfrak{R}\}$ .



Equation (2) reflects the cost of synthesizing the volatility contingent cash flows  $f[v_T]$  and  $g[v_T]$ .

**Assumption 6** *Both types of investors have a utility function, defined over their respective volatility contingent cash flow, and in the constant absolute risk aversion class (with  $\gamma_f > 0$  and  $\gamma_g > 0$ ):*

$$(3) \quad U(f[v_T]) = \left( \frac{-1}{\gamma_f} \right) e^{-\gamma_f f[v_T]}, \quad U(g[v_T]) = \left( \frac{-1}{\gamma_g} \right) e^{-\gamma_g g[v_T]}.$$

In the specification (3), the exponential function guarantees real valued utility, even when the volatility contingent cash flow is negative. The self-financing condition in equation (2) implies that both  $f[v_T]$  and  $g[v_T]$  exhibit a range of negative values, hence, for example, adopting an isoelastic utility does not result in a real-valued utility.

## B. Choice Problem of Investors

It may be possible to characterize  $f[v_T]$  and  $g[v_T]$  endogenously, and in analytical closed form. The choice problem of investors F and G can be stated as,

**Problem 1** *Choose  $f \in C$  and  $g \in C$  to maximize expected utility under the personalized densities, subject to the self-financing constraint (under zero interest rate):*

$$(4) \quad f^* = \arg \max_{f \in C} \int \left( \frac{-1}{\gamma_f} \right) e^{-\gamma_f f[v_T]} \mathcal{A}[v_T] \Phi[v_T] dv_T \quad \text{subject to} \quad \int f[v_T] q[v_T] dv_T = 0,$$

$$(5) \quad g^* = \arg \max_{g \in C} \int \left( \frac{-1}{\gamma_g} \right) e^{-\gamma_g g[v_T]} \mathcal{B}[v_T] \Phi[v_T] dv_T \quad \text{subject to} \quad \int g[v_T] q[v_T] dv_T = 0,$$

and the market clearing condition for volatility contingent claims,

$$(6) \quad \phi f^*[v_T] + (1 - \phi) g^*[v_T] = 0.$$

When the change of measure densities are parameterized in a way such that  $\mathcal{A}[v_T]$  ( $\mathcal{B}[v_T]$ ) is increasing (decreasing), then investors F (G) anticipate volatility to rise (fall).

In equation (4), observe that maximizing  $\int \left(\frac{-1}{\gamma_f}\right) e^{-\gamma_f f[v_T]} \mathcal{A}[v_T] \Phi[v_T] dv_T$  is equivalent to minimizing  $\int \frac{1}{\gamma_f} e^{-\gamma_f f[v_T]} \mathcal{A}[v_T] \Phi[v_T] dv_T$ , which is (i) a convex function of  $f[v_T]$ , and (ii) bounded below by zero. At the same time, a strategy of zero cash flow satisfies the self-financing constraint, with an expected utility of  $-1/\gamma_f$ . It follows that the optimal expected utility lies between  $-1/\gamma_f$  and zero.

Using the optimal volatility contingent cash flows  $f[v_T]$  and  $g[v_T]$  that solve the first-order conditions in equation (A-1) and satisfy equation (6) (and suppressing the \* superscript on  $f^*[v_T]$  and  $g^*[v_T]$  from now on), our goal is to derive the pricing kernel

$$(7) \quad m[v_T] \equiv \frac{q[v_T]}{\Phi[v_T]},$$

under suitable assumptions about the change of measure densities  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$ , and then to investigate the model properties, especially the tail behavior of the volatility pricing distribution.

### C. Non-Monotonic Pricing Kernel in the Volatility Market

The following theorem presents the solution for the pricing kernel from Problem 1, together with conditions under which the pricing kernel is non-monotonic in volatility. Subsection D builds on our

results and develops examples that explicitly parameterize the physical and personalized densities of the investors.

**Theorem 1** *The following statements hold:*

I. *The pricing kernel is related to the change of measure densities as,*

$$(8) \quad m[v_T] = m_0 (\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta}, \quad \eta \equiv \frac{\phi \gamma_g}{(1-\phi)\gamma_f + \phi \gamma_g},$$

where  $m_0$  is a constant of integration, chosen to ensure that  $\int m[v_T] p[v_T] dv_T = 1$ . The optimal volatility contingent cash flows are,

$$(9) \quad f[v_T] = f_0 + \left( \frac{(1-\phi)}{(1-\phi)\gamma_f + \phi \gamma_g} \right) \ln \left( \frac{\mathcal{A}[v_T]}{\mathcal{B}[v_T]} \right) \quad \text{and} \quad g[v_T] = - \left( \frac{\phi}{1-\phi} \right) f[v_T],$$

where  $f_0$  in equation (9) is chosen to satisfy the self-financing constraint  $\int f[v_T] q[v_T] dv_T = 0$ .

II. *Suppose that over some interval  $v_T \in [\underline{v}, \bar{v}]$ ,*

$$(10) \quad (a) \quad \mathcal{A}'[v_T] \equiv \frac{d\mathcal{A}[v_T]}{dv_T} > 0 \quad \text{and} \quad \mathcal{B}'[v_T] \equiv \frac{d\mathcal{B}[v_T]}{dv_T} < 0;$$

$$(11) \quad (b) \quad (\ln(\mathcal{A}[v_T]))' \Big|_{v=\underline{v}} \approx 0 \quad \text{and} \quad (\ln(\mathcal{B}[v_T]))' \Big|_{v=\bar{v}} \approx 0; \quad \text{and,}$$

$$(12) \quad (c) \quad \mathcal{A}[v_T] \text{ and } \mathcal{B}[v_T] \text{ are log-convex.}$$

Condition (c) is sufficient for the pricing kernel  $m[v_T]$  to be log-convex. Additionally, (a) and (b) ensure that  $m[v_T]$  will have a minimum in the interior of the interval  $v_T \in [\underline{v}, \bar{v}]$ , and will admit a negatively (positively) sloped region at low (high) volatility levels.

**Proof.** See Appendix A.  $\square$

Part I of Theorem 1 provides a characterization of the pricing kernel, which depends on the assumed change of measure densities. Importantly, the pricing kernel in (8) is determined by the beliefs of both types of investors in the volatility market, since the exponents on  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$  are both positive. Furthermore,  $m[v_T]$  is guaranteed to be positive.

With the pricing kernel determined, the pricing density  $q[v_T]$  can be recovered accordingly, given the physical density  $\Phi[v_T]$ . Even if a relevant question remains about the magnitude of the exponent  $\eta$ , the proportion of each type of investors is assumed to be nonzero.

While, in general, the functional equations from the optimization problem in (4)–(6) can be difficult to solve, the solution for the optimal  $f[v_T]$  offered in (9) is made tractable under two features, namely (i) exponential utility, and (ii) zero net supply for contingent claims for volatility. We will demonstrate that the solution for  $m[v_T]$  in (8) provides flexibility, and could possibly reconcile certain features of the observed pattern of VIX option returns under suitable parameterizations of  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$ .

The setup of Theorem 1 (and also Problem 1) accommodates heterogeneity in beliefs, through  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$ , yet differs from the setup in, for example, Leland (1980), Detemple and Murthy (1994), Basak (2000), Calvet, Grandmont, and Lemaire (2004), Buraschi and Jiltsov (2006), Jouini and Napp (2007), Bakshi, Madan, and Panayotov (2010), and Xiong and Yan (2010). Our setup also differs from the corresponding one for futures on the equity market index (or other traded assets), where the futures are also in zero net supply, but both the spot-futures arbitrage condition *and* the martingale condition for the spot price process must be taken into account when solving for

equilibrium (e.g., Grossman and Stiglitz (1980), Briys, Crouhy, and Schlesinger (1990), and Hong (2000)).

What can be said about the concavity or the convexity of  $g[v_T]$ ? While not much can be inferred analytically without explicitly parameterizing  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$ , we nonetheless note, based on (9), that a sufficient condition for the convexity of  $g[v_T]$  is that  $\mathcal{A}[v]/\mathcal{B}[v]$  be concave, since log preserves concavity.

When  $\mathcal{A}'[v_T] > 0$  ( $\mathcal{B}'[v_T] < 0$ ), as postulated in condition (a) of Theorem 1 (part II), it is provable that the mean of the personalized density  $\Phi_f[v_T]$  ( $\Phi_g[v_T]$ ) is higher (lower) than the mean of the physical density  $\Phi[v_T]$ . Therefore, the pricing kernel in part II of Theorem 1 is designed to reflect the impact of investors with opposite beliefs about volatility outcomes. Moreover, the belief distributions of the investors are restricted to maintain a pricing (risk-neutral) density that dominates the physical density in both tails of the volatility distribution.

What specific change of measure densities  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$  may satisfy the assumptions in part II of Theorem 1? One example would be  $\mathcal{A}[v_T] = \exp(\mu_a v_T^2)$  and  $\mathcal{B}[v_T] = \exp(\mu_b v_T^{-1})$ , for  $\mu_a > 0$  and  $\mu_b > 0$ . In this case,  $m[v_T] = m_0 \exp\left(\eta\mu_a v_T^2 + (1 - \eta)\mu_b v_T^{-1}\right)$ , hence, the pricing kernel is log-convex and has a decreasing (increasing) region at low (high) volatility levels. Alternatively, if  $\mathcal{A}[v_T] = \exp(e^{\mu_a(v_T-1)})$  and  $\mathcal{B}[v_T] = \exp(e^{-\mu_b(v_T-1)})$ , then  $\ln(m[v_T])$  is a generalized hyperbolic cosine function, and has a minimum for  $v_T$  close to unity when  $\eta\mu_a \approx (1 - \eta)\mu_b$ .

Under the assumption on the change of measure densities in equation (10), the optimal  $g[v_T]$  is decreasing, being the logarithm of the ratio of a decreasing change of measure ( $\mathcal{B}'[v_T] < 0$ ) and an increasing change of measure ( $\mathcal{A}'[v_T] > 0$ ), and, by analogy, the optimal  $f[v_T]$  is increasing. Conse-

quently, investors G have short volatility exposure, while investors F have long volatility exposure.

One additional clarification relates to the question: How is the shape of  $m[v_T]$  related to the properties of the cash flows  $f[v_T]$  and  $g[v_T]$ ? At the least, it can be shown that a *necessary condition* for the pricing kernel to be non-monotonic is that  $f[v_T]$  and  $g[v_T]$  are nonlinear in  $v_T$ .

To establish this result, we return to (8)–(9) of Theorem 1 and consider change of measure densities of the form:  $\mathcal{A}[v_T] = e^{\mu_a v_T}$  and  $\mathcal{B}[v_T] = e^{-\mu_b v_T}$ , with  $\mu_a > 0$  and  $\mu_b > 0$ . In this case,

$$(13) \quad f[v_T] = f_0 + \left( \frac{(1-\phi)(\mu_a + \mu_b)}{(1-\phi)\gamma_f + \phi\gamma_g} \right) v_T, \quad g[v_T] = -\frac{\phi}{(1-\phi)} f[v_T], \quad m[v_T] = m_0 e^{\xi v_T},$$

where  $\xi \equiv \eta\mu_a - (1-\eta)\mu_b$  and  $m_0 = (\int e^{\xi v_T} \Phi[v_T] dv_T)^{-1}$ . Equation (13) shows that the linearity of  $f[v_T]$  and  $g[v_T]$  is accompanied by a monotonic pricing kernel, which has two further ramifications. First, contingent claims with payout linear in  $v_T$  may not allow one to fully discern the tail behavior of traded volatility. Second, to obtain non-monotonicity of the pricing kernel in our framework with exponential utility, the change of measure densities,  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$ , must be kept outside the exponential class.

There are two features of the volatility market that distinguish it from other futures markets, for example, those on the S&P 500 equity index. Specifically, the mean-reversion and non-tradability of volatility allow the delinking of pricing kernels at different dates, whereas tradability enforces restrictions on the pricing kernels at different dates. However, these issues are not explored in this paper.

In a setting different from ours, Song and Xiu (2013) estimate non-parametrically a U-shaped pricing kernel of volatility using VIX options, and their finding serves as a direct evidence in support

of the theoretical findings of this paper. In addition, Amengual and Xiu (2013) use variance swaps to study downward jumps in volatility. The models in both these studies are cast in continuous time with jumps, and point to the possibility of obtaining a U-shaped pricing kernel in a different framework.

## D. Model Implications in Example Economies

The implications of part II of Theorem 1, i.e., an asset pricing model where investors exhibit aversion to both low and high volatility states, are illustrated on specific examples that parameterize the belief distribution of investors.

### D.1. Associated Pricing Kernels

**Example 1** Suppose volatility  $v_T$  is distributed lognormally under the physical measure:

$$(14) \quad \Phi[v_T; \mu_p, \sigma_p] = \frac{1}{v_T \sigma_p \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_p^2} \{\ln(v_T) - \mu_p\}^2\right).$$

The personalized densities of investors F and G are also lognormal, with parameters  $\mu_f$  and  $\sigma_f$ , and  $\mu_g$  and  $\sigma_g$ , respectively.

From equation (1), the log change of measure densities are quadratic in  $\ln(v_T)$ :

$$(15) \quad \ln(\mathcal{A}[v_T]) = \ln\left(\frac{\sigma_p}{\sigma_f}\right) - \frac{(\mu_f - \ln(v_T))^2}{2\sigma_f^2} + \frac{(\mu_p - \ln(v_T))^2}{2\sigma_p^2},$$

$$(16) \quad \ln(\mathcal{B}[v_T]) = \ln\left(\frac{\sigma_p}{\sigma_g}\right) - \frac{(\mu_g - \ln(v_T))^2}{2\sigma_g^2} + \frac{(\mu_p - \ln(v_T))^2}{2\sigma_p^2}.$$

We specify that

$$(17) \quad \mu_f > \mu_p > \mu_g, \quad \sigma_f > \sigma_p, \quad \text{and} \quad \sigma_g > \sigma_p,$$

differentiating the beliefs about volatility outcomes of the two types of investors. ♣

From equations (15) and (16), the change of measures for investors F and G satisfy,

$$(18) \quad \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} = \frac{(\sigma_f^2 - \sigma_p^2) \ln(v_T) - \mu_p \sigma_f^2 + \mu_f \sigma_p^2}{v_T \sigma_p^2 \sigma_f^2}, \quad \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]} = \frac{(\sigma_g^2 - \sigma_p^2) \ln(v_T) - \mu_p \sigma_g^2 + \mu_g \sigma_p^2}{v_T \sigma_p^2 \sigma_g^2}.$$

It follows that  $\mathcal{A}[v_T]$  is increasing for  $v_T > \underline{v}$ , and  $\mathcal{B}[v_T]$  is decreasing for  $v_T < \bar{v}$ , where,

$$(19) \quad \underline{v} = \exp \left( - \left( \frac{\sigma_p^2 \sigma_f^2}{\sigma_f^2 - \sigma_p^2} \right) \left( \frac{\mu_f}{\sigma_f^2} - \frac{\mu_p}{\sigma_p^2} \right) \right), \quad \bar{v} = \exp \left( \left( \frac{\sigma_p^2 \sigma_g^2}{\sigma_g^2 - \sigma_p^2} \right) \left( \frac{\mu_p}{\sigma_p^2} - \frac{\mu_g}{\sigma_g^2} \right) \right).$$

Our focus is on change of measure parameterizations, for which  $[\underline{v}, \bar{v}]$  is a suitably wide interval over which  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$  in (15) and (16) are bounded.

Invoking equation (8) of Theorem 1,

$$(20) \quad \ln(m[v_T]) = \frac{\alpha}{2} (\ln(v_T))^2 + \beta \ln(v_T) + (\zeta + \ln(m_0)),$$



where  $m_0$  is a constant of integration, ensuring that  $\int m[v_T] \Phi[v_T] dv_T = 1$ , and,

$$(21) \quad \alpha = \eta \left( \frac{1}{\sigma_p^2} - \frac{1}{\sigma_f^2} \right) + (1 - \eta) \left( \frac{1}{\sigma_p^2} - \frac{1}{\sigma_g^2} \right),$$

$$(22) \quad \beta = \eta \left( \frac{\mu_f}{\sigma_f^2} - \frac{\mu_p}{\sigma_p^2} \right) + (1 - \eta) \left( \frac{\mu_g}{\sigma_g^2} - \frac{\mu_p}{\sigma_p^2} \right),$$

$$(23) \quad \varsigma = \eta \left( \ln \left( \frac{\sigma_p^2}{\sigma_f^2} \right) + \frac{\mu_p^2}{2\sigma_p^2} - \frac{\mu_f^2}{2\sigma_f^2} \right) + (1 - \eta) \left( \ln \left( \frac{\sigma_p^2}{\sigma_g^2} \right) + \frac{\mu_p^2}{2\sigma_p^2} - \frac{\mu_g^2}{2\sigma_g^2} \right).$$

Importantly, the belief distributions implied by equation (17) ensure that  $\alpha > 0$ , and consequently the quadratic function in (20) decreases for low values of volatility and increases for high values of volatility. The log pricing kernel is convex in  $\ln(v_T)$  and achieves a minimum at  $\exp(-\beta/\alpha)$ .

A greater dispersion in the personalized volatility distributions, relative to the physical distribution, as in equation (17), may be needed to support a broader class of pricing kernels in the volatility market, since if  $\sigma_f = \sigma_g = \sigma_p$ , then the pricing kernel is monotonic in  $v_T$  (as  $\alpha = 0$ ).

If  $\sigma_f < \sigma_p$  and  $\sigma_g < \sigma_p$ , then a hump-shaped pricing kernel is obtained, which implies that the physical density dominates the pricing density in both tails of the volatility distribution. Such an implication is also counterfactual.

From equation (9), the optimal volatility contingent cash flow for investors F is:

$$(24) \quad f[v_T] = f_0 + \frac{(1 - \phi)}{(1 - \phi)\gamma_f + \phi\gamma_g} \left\{ \ln \left( \frac{\sigma_g}{\sigma_f} \right) + \frac{1}{2\sigma_g^2} (\mu_g - \ln(v_T))^2 - \frac{1}{2\sigma_f^2} (\mu_f - \ln(v_T))^2 \right\}$$

and the optimal volatility contingent cash flow for investors G is  $g[v_T] = -(\phi/(1 - \phi))f[v_T]$ . Furthermore, over  $[\underline{v}, \bar{v}]$ , where  $\mathcal{A}'[v_T] > 0$  and  $\mathcal{B}'[v_T] < 0$ , the optimal cash flow  $f[v_T]$  is increasing and

$g[v_T]$  is decreasing.

It may be observed that  $f[v_T]$  can be concave or convex over various intervals,

$$(25) \quad f''[v_T] = \frac{1}{v_T^2} \left( \frac{(1-\phi)}{(1-\phi)\gamma_f + \phi\gamma_g} \right) \left( \left( \frac{1}{\sigma_f^2} - \frac{1}{\sigma_g^2} \right) \ln(v_T) - \frac{1}{\sigma_f^2}(\mu_f + 1) + \frac{1}{\sigma_g^2}(\mu_g + 1) \right).$$

For instance, when  $\sigma_f < \sigma_g$ , one could get  $f''[v_T] < 0$  for low values of  $v_T$  and  $f''[v_T] > 0$  for high values of  $v_T$ . The converse obtains when  $\sigma_f > \sigma_g$ . When  $\sigma_f = \sigma_g$ , then  $f[v_T]$  is concave over  $[\underline{v}, \bar{v}]$ .

The crucial point to observe is that either investors F, or investors G are holding convex functions of volatility in the aggregate over some intervals.

[FIGURE 1 about here.]

To highlight some additional aspects of our example economy, Figure 1 plots the pricing kernel and the optimal cash flows  $f[v_T]$  and  $g[v_T]$ , based on the following parameters:

$$(26) \quad \mu_p = 0.000, \quad \mu_f = 0.150, \quad \mu_g = -0.150, \quad \sigma_p = 0.171, \quad \sigma_f = 0.184, \quad \sigma_g = 0.184,$$

with  $\gamma_f = 3.5$ ,  $\gamma_g = 3.0$ , and  $\phi = 0.50$ .<sup>2</sup> Under the assumed parameterization,  $v_T \in (0.50, 2.00)$ .

---

<sup>2</sup>In our illustration, the assumed values of  $\mu_p$  and  $\sigma_p$  roughly correspond to the mean and standard deviation of monthly changes in the log of VIX over 1/1990-12/2013. We also note that the monthly log changes in the VIX are uncorrelated according to the Ljung-Box statistic at lag 12 ( $p$ -value of 0.229). The maximum-likelihood estimate of the first-order autoregressive coefficient is  $-0.125$  ( $p$ -value of 0.058). Furthermore, the frequency of observations in certain intervals of monthly log changes in the VIX are tabulated below:

The derived pricing kernel is declining at low levels of volatility, and then increasing. It is this property of the pricing kernel that could impart negative returns of volatility puts. Incorporating the behavior of the two types of investors with  $\mu_f > \mu_p$  and  $\mu_g < \mu_p$  could give rise to a pricing kernel that is consistent with negative returns of both OTM volatility puts and calls.

In general, a relative increase in  $\gamma_f$  steepens the slope of the pricing kernel in the region of low volatility levels, and shifts it to the right, a trait that can broadly induce more negative put returns. When  $\phi\gamma_g = (1 - \phi)\gamma_f$ , then  $\eta = 0.50$  (from equation (8)), and the response of  $m[v_T]$  to  $v_T$  is symmetric.

Recapitulating the implications of Example 1, the optimal cash flows can be concave or convex over various intervals, and the pricing kernel can be asymmetric, displaying both a decreasing region for low levels of volatility and an increasing region for high levels of volatility.

**Example 2** Keep the physical density,  $\Phi[v_T; \mu_p, \sigma_p]$ , to be lognormal, as in (14), but assume an

---

Log change in the VIX (%)	< -0.2	[-0.2, -0.1)	[-0.1, 0)	(0, 0.1]	(0.1, 0.2]	>0.2
Frequency (%)	0.10	0.18	0.25	0.21	0.13	0.12

---

indicating that a sizable fraction of VIX moves exceed 20% in both the left and right tail. The analysis of Amengual and Xiu (2013) indicates that a source of downward movements in volatility is resolution of policy uncertainty.

alternative functional form for the change of measures for the two types of investors:

$$(27) \quad \mathcal{A}[v_T] = \bar{\omega}_a + (1 - \bar{\omega}_a) \left( \frac{1}{\sigma_a \sqrt{2\pi}} \int_{-\infty}^{\ln(v_T)} \exp\left(-\frac{\{x - \mu_a\}^2}{2\sigma_a^2}\right) dx \right),$$

$$(28) \quad \mathcal{B}[v_T] = \bar{\omega}_b + (1 - \bar{\omega}_b) \left( \frac{1}{\sigma_b \sqrt{2\pi}} \int_{-\infty}^{-\ln(v_T)} \exp\left(-\frac{\{x - \mu_b\}^2}{2\sigma_b^2}\right) dx \right),$$

where  $\bar{\omega}_a \in (0, 1)$  and  $\bar{\omega}_b \in (0, 1)$ . The parameters  $\mu_i$  and  $\sigma_i$ , for  $i = \{a, b\}$ , control the mean and standard deviation of the cumulative Normal distribution in (27)–(28), which is chosen as a device to bound the change of measure densities between  $\bar{\omega}$  and 1. ♣

Such a choice of change of measure densities also preserves an increasing  $\mathcal{A}[v_T]$  and a decreasing  $\mathcal{B}[v_T]$  over the entire domain  $v_T \in (0, \infty)$ :

$$(29) \quad \mathcal{A}'[v_T] = \frac{(1 - \bar{\omega}_a) \exp\left(-\frac{\{\ln(v_T) - \mu_a\}^2}{2\sigma_a^2}\right)}{v_T \sigma_a \sqrt{2\pi}} > 0, \quad \mathcal{B}'[v_T] = -\frac{(1 - \bar{\omega}_b) \exp\left(-\frac{\{-\ln(v_T) - \mu_b\}^2}{2\sigma_b^2}\right)}{v_T \sigma_b \sqrt{2\pi}} < 0,$$

bringing certain deviation from the setup of Example 1.

One common feature across different parameter combinations is that the pricing kernel  $m[v_T]$  associated with this example is flat in the tails, reflecting the boundedness of the change of measures. Moreover, when  $\mu_a$  and  $\mu_b$  are both positive, and not too small, one generally obtains  $m[v_T]$  that displays a decreasing region at some low volatility levels, and then an increasing region for high volatility levels.<sup>3</sup>

---

<sup>3</sup>We note that  $f[v_T]$  is increasing and  $g[v_T]$  is decreasing, and explicit expressions for  $f[v_T]$  and  $g[v_T]$ , and their derivatives, can be obtained analytically in terms of the cumulative Normal distribution function.

[FIGURE 2 about here.]

Figure 2 plots the pricing kernel and the optimal  $f[v_T]$  and  $g[v_T]$  for the following parameters:

$$(30) \quad \mu_p = 0.000, \quad \mu_a = 0.150, \quad \mu_b = 0.200, \quad \sigma_p = 0.171, \quad \sigma_a = 0.100, \quad \sigma_b = 0.100,$$

with  $\gamma_f = 3.5$ ,  $\gamma_g = 3.0$ ,  $\phi = 0.50$ ,  $\varpi_a = 0.060$ , and  $\varpi_b = 0.070$ .

This part of the analysis establishes two model traits. First, the derived  $m[v_T]$  in Figure 2 exhibits bending in the tails, mirroring the boundedness of the change of measure densities. In between,  $m[v_T]$  is first declining and then increasing. For our parameter set,  $f[v_T]$  ( $g[v_T]$ ) is convex for low (high) levels of volatility. These regions correspond to states where the respective investors are selling, rather than buying, volatility contingent cash flows. Examples 1 and 2 thus differ in some economic dimensions, but still rest on heterogeneity in investors' beliefs about the distribution of volatility outcomes and their risk attitudes. The shape of the pricing kernel implies that some investors dislike both low and high volatility states.

## D.2. Expected Option Returns

To provide a link to the empirical evidence presented in Section III, we next discuss the implications of our parametric examples for the pricing density  $q[v_T]$ , and for the pattern of expected returns of options on volatility.

Based on equations (7) and (8),

$$(31) \quad q[v_T] = \frac{(\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta} \Phi[v_T]}{\int (\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta} \Phi[v_T] dv_T}.$$

While in Example 1, (31) reduces to  $q[v_T] = (\Phi_f[v_T])^\eta (\Phi_g[v_T])^{1-\eta} / \int (\Phi_f[v_T])^\eta (\Phi_g[v_T])^{1-\eta} dv_T$ , which is multiplicative in the belief densities, Example 2 does not offer a simple analytical expression for the pricing density.

In Panels A through F of Table 1, we report the expected returns of options on volatility obtained from our Examples 1 and 2 under parameterizations of investors' volatility belief distributions and risk aversions. In this illustrative exercise, Panels A and E correspond to Figures 1 and 2, respectively. Both examples exhibit asymmetries in the returns of volatility options, and most notably, expected put returns increase in strike, whereby deep OTM put returns can be the most negative, while expected call returns decrease in strike.<sup>4</sup> Our parameterizations thus illustrate a pricing measure that dominates the physical measure in both tails of the volatility distribution.

Recognize that our model incorporates the opposing views about volatility of only two types of investors, which leads to expected returns in Table 1 that are somewhat structured. On the other hand, observed option returns likely reflect other opinions, for example, some investors may expect no change in volatility. We anticipate that, for instance, a model adding investors with centrist beliefs about volatility outcomes could possibly deliver a different pattern of near-the-money option returns.

---

<sup>4</sup>We have ascertained that it is possible to preserve the core features of the pricing kernel, as in Examples 1 and 2, under Gamma distributed physical and personalized densities. This analysis is omitted to maintain parsimony in the number of examples.

Such an extension could impart a pricing kernel that places greater weight on outcomes in the neck of the volatility distribution, as opposed to our model that reinforces movements in the tails.

### III. Empirical Evaluation

This section formulates the testable hypothesis regarding our model, and compares its implications to a framework, where volatility is disliked. Next, we outline the essential features of VIX options, given that we measure volatility by the VIX index. Then we discuss the evidence from our empirical tests.

#### A. Testable Implications

The salient feature of our model with heterogeneity in beliefs is the non-monotonic kernel, which exhibits a decreasing region at low volatility levels, and an increasing region at high volatility levels. The specific prediction is that expected returns of calls on volatility decrease in strike, while expected returns of puts on volatility increase in strike (Theorem 1, and as demonstrated in the context of our example economies in Table 1).

How does this prediction compare to a conventional modeling approach, where volatility is disliked? Specifically, in Theorem 2 of Appendix C, we show that a model of this type leads to a pricing kernel which is monotonically increasing in the volatility dimension and, hence, predicts that the expected returns of puts (calls) on volatility are decreasing in strike and are positive (negative) for sufficiently low (high) strikes.

We note that both models imply a pricing kernel with an increasing region for high levels of volatility, and, hence, yield identical predictions with respect to the right tail of the volatility distribution or the pattern of expected returns of OTM volatility calls.

However, the models differ dramatically in their predictions about the left tail of the volatility distribution. Therefore, it is the pattern of expected returns of OTM volatility puts that can help discriminate between the two models, and, hence, Theorem 2 offers a useful counterpoint to Theorem 1 that brings expected returns of volatility puts in the focus of our examination.

## B. VIX Option Data

Proceeding to the empirical analysis, we employ VIX options with 28 days to expiration, and calculate the return of a VIX call and put as,<sup>5</sup>

$$(32) \quad 1 + r_{t,T}^{c,vix}[y] = \frac{(\text{VIX}_{t+T} - y\text{FVIX}_{t,T})^+}{C_{t,T}^{vix}[y]} \quad \text{and} \quad 1 + r_{t,T}^{p,vix}[y] = \frac{(y\text{FVIX}_{t,T} - \text{VIX}_{t+T})^+}{P_{t,T}^{vix}[y]},$$

---

<sup>5</sup>While VIX option attributes are standard, we nonetheless note that VIX at time  $t$ , denoted  $\text{VIX}_t$ , is the square-root of the expected quadratic variation of the logarithm of the market index over the next 30 days, under the pricing measure  $\mathbb{Q}$ . In practice,  $\text{VIX}_t$  is approximated from the price of a traded index option portfolio (e.g., Carr and Wu (2009)). The price at time  $t$  of the VIX futures with time to maturity  $T$ , denoted  $\text{FVIX}_{t,T}$ , is  $\mathbb{E}_t^{\mathbb{Q}}(\text{VIX}_{t+T})$ , where  $\mathbb{E}_t^{\mathbb{Q}}$  is the time  $t$  conditional expectation under the pricing measure  $\mathbb{Q}$ . Similarly, the time  $t$  prices of a call and a put on the VIX with expiration date  $t+T$  and strike  $K$  is  $C_{t,T}^{vix}[K] = e^{-r_f} \mathbb{E}_t^{\mathbb{Q}}((\text{VIX}_{t+T} - K)^+)$  and  $P_{t,T}^{vix}[K] = e^{-r_f} \mathbb{E}_t^{\mathbb{Q}}((K - \text{VIX}_{t+T})^+)$ , where  $r_f$  is the  $T$ -period riskfree return.



where  $FVIX_{t,T}$  is the front-month VIX futures price, and  $C_{t,T}^{\text{vix}}[y]$  ( $P_{t,T}^{\text{vix}}[y]$ ) is the price of the VIX call (put) with moneyness  $y \equiv K/FVIX_{t,T}$ . Using *ask quotes* on VIX options from the CBOE, we obtain one return observation per month, corresponding to a fixed moneyness, and avoid overlapping returns.

For the calculation in equation (32), we identify VIX calls and puts, which are closest to 5%, 15%, and 25% OTM, respectively. The specific moneyness levels are chosen, taking into consideration the minimum strike intervals for VIX options at the CBOE. We compute returns to expiration, using the settlement value of VIX to determine the option payoffs. Our data set thus consists of six series of OTM VIX option returns, both calls and puts, and with 28 days to expiration.

Important to our themes, trading in volatility has seen rapid growth, with the VIX options pit being currently the second most active, after the one for S&P 500 index options. Since exchange trading in VIX options started in March 2006, our sample spans the 93-month period from 03/2006 to 12/2013.

Pertinent to our empirical approach, two other points deserve further comment. First, to uncover a pricing kernel with increasing and decreasing regions one needs to rely on instruments that can separate up and down volatility movements. Therefore, variance swaps, which do not offer Arrow-Debreu like payoffs, may be uninformative and may not be suitable for this purpose. Lacking other suitable instruments that offer tail exposures, we employ VIX option data in our model assessments, despite the nascency of the volatility option market. Second, if conditional expected returns are lower for volatility puts deeper OTM, then average returns of volatility puts, which reflect unconditional expectations, should exhibit the same property (e.g., Coval and Shumway (2001), and Bakshi,

Madan, and Panayotov (2010)).

### C. Tests Based on the Moneyness Pattern of VIX Put Returns

Providing an essential link to the theoretical models, our exercise now focuses on the VIX option returns, whose pattern across option types and moneyness is revealing about the pricing density and underlying pricing kernel that are supported in the data. Panels A through C of Table 2 yield potential insights about tail behavior in the volatility market.

Do returns of VIX options get more negative deeper OTM? While addressing the small sample concern, in Panel A of Table 2 we test for differences in average VIX option returns across strikes by drawing 50,000 pairwise bootstrap samples of returns of VIX puts (calls) that are 25% and 15%, 25% and 5%, and 15% and 5% OTM, respectively. Then we calculate the difference between the average returns for each pair of bootstrap samples. Reported are the  $p$ -values for the differences between the average returns of the respective VIX options, calculated as the proportion of bootstrap samples where the first option in a pair has *higher* average return than the second option.

Most importantly, we find that the average return of 25% OTM VIX puts is lower than the return of 15% (5%) OTM VIX puts at a 0.8% (1.7%) significance level. At the same time, these  $p$ -values indicate that the differences across moneyness are statistically insignificant for VIX calls. Therefore, the pairwise differences in average VIX put returns across the three moneyness levels provide statistically significant evidence, contradicting Theorem 2, but possibly consistent with our model of the volatility market. As highlighted before, it is the returns of volatility puts that could help discriminate between the competing hypotheses.

Complementing the above, we follow Patton and Timmermann (2010), and apply the *MR*, *UP*, and *DOWN* bootstrap-based procedures that they advocate to test for a monotonic relation between expected returns of options on volatility and their strikes. The null hypothesis for the *MR* test is that average VIX option returns are flat with respect to moneyness, versus the alternative of an increasing pattern. Furthermore, the *UP* and *DOWN* tests, which combine information on both the number and magnitude of departures from a flat pattern, are designed to detect positive or negative segments in the relation between average option returns and moneyness.

Panel B of Table 2 reports for VIX puts *p*-values of 0.049, 0.005, and 0.723, for the *MR*, *UP*, and *DOWN*, respectively. The *MR* (*UP*) test thus supports a statistically significant increasing pattern (positive segment) in the relation between VIX put returns and moneyness, and provides statistical confirmation that average returns of VIX puts get more negative deeper OTM.

To clarify further our result, Panel C of Table 2 also shows the average returns per month, which are 10.21%, 4.53%, and  $-32.68\%$  for VIX puts that are 5%, 15%, and 25% OTM, respectively. Therefore, VIX put returns *increase* in strike, and are conspicuously low for the farthest OTM puts. On the other hand, the average returns are  $-27.36\%$ ,  $-31.30\%$ , and  $-34.84\%$  for 5%, 15%, and 25% OTM VIX calls, respectively, and are thus consistent with both models. We note here that the average VIX option returns are not statistically different from zero, as indicated by the reported 90% confidence intervals obtained in 50,000 bootstrap draws, whereas the pairwise test is able to furnish statistical significance in support of the model that hinges on heterogeneity in beliefs.

In conclusion, the average VIX put returns in our sample appear inconsistent with a theory postulating a pricing kernel which is upward sloping in the volatility dimension, as per Theorem 2,

and Table 2 indicates that the risk-neutral probability in both tails of the volatility distribution can potentially exceed the corresponding physical probability. In particular, the model of the volatility market that we suggest in Section II appears better suited to reflect the empirical features of observed VIX option returns, and is more consistent with the left tail of the volatility distribution, thus pointing to a direction for further volatility modeling.

## IV. Concluding Remarks

This paper features a model that incorporates heterogeneity in beliefs about volatility and allows for a kernel that is non-monotonic in the volatility dimension.

We provide examples demonstrating that, within the model with heterogeneity in beliefs, expected returns of puts on volatility can be increasing in strike and negative for deep out-of-the-money strikes. Contrasting our model's implications for volatility tail behavior, a model where high volatility states are disliked, predicts that expected returns of puts on volatility are decreasing in strike, and are positive for sufficiently low strikes. Using the two models' opposite predictions about the behavior of volatility in the left tail, we conduct tests indicating that the model with heterogeneity in beliefs appears to better reflect features of the data, both qualitatively and quantitatively.

While we have a discrete-time model, there is some progress in modeling heterogeneity in beliefs in continuous-time, as in Dieckmann (2011) and Chen, Joslin, and Tran (2012). The recent works of Song and Xiu (2013) and Amengual and Xiu (2013) highlight the possible role of jumps in generating realistic pricing kernels of the volatility market.

It bears emphasizing that our model presumes a volatility market that is dominated by investors with zero equity delta. In a more elaborate model setting, the investors would trade the equity index, as well as volatility contingent claims, exploiting their beliefs about both equity returns and volatility outcomes. The relevant optimization problem of the investors would then involve (i) a two-dimensional physical density, (ii) change of measure densities that are functions of equity returns and volatility, and (iii) long and short exposures in equity and volatility cash flows. However, this extension may require a richer market that trades product options of all strike pairs in equity and volatility, and is left to a follow-up study.

Taken together, our theoretical models and empirical results can pave the way for a better appreciation of the motives for trading volatility, and hopefully provide impetus for improved models for pricing and hedging volatility. When VIX options are available to offer exposure to volatility tails, much can be learned in the realm of traded volatility.

## References

- Aït-Sahalia, Y.; M. Karaman; and L. Mancini. “The Term Structure of Variance Swaps and Risk Premia.” Unpublished working paper, Princeton University (2013).
- Aït-Sahalia, Y., and A. Lo. “Nonparametric Risk Management and Implied Risk Aversion.” *Journal of Econometrics*, 94 (2000), 9–51.
- Aït-Sahalia, Y., and L. Mancini. “Out of Sample Forecasts of Quadratic Variation.” *Journal of Econometrics*, 147 (2008), 17–33.
- Amengual, D., and D. Xiu. “Resolution of Policy Uncertainty and Sudden Declines in Volatility.” Unpublished working paper, University of Chicago (2013).
- Andersen, T.; T. Bollerslev; F. Diebold; and P. Labys. “Modeling and Forecasting Realized Volatility.” *Econometrica*, 71 (2003), 529–626.
- Andersen, T.; T. Bollerslev; and N. Meddahi. “Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities.” *Econometrica*, 73 (2005), 279–296.
- Bagnoli, M., and T. Bergstrom. “Log-Concave Probability and its Applications.” *Economic Theory*, 26 (2005), 445–469.
- Bakshi, G.; D. Madan; and G. Panayotov. “Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels.” *Journal of Financial Economics*, 97 (2010), 130–154.
- Basak, S. “A Model of Dynamic Equilibrium Asset Pricing with Heterogeneous Beliefs and Exogenous Risk.” *Journal of Economic Dynamics and Control*, 24 (2000), 63–95.

- Basak, S. “Asset Pricing with Heterogeneous Beliefs.” *Journal of Banking and Finance*, 29 (2005), 2849–2881.
- Black, F. “Studies of Stock Price Volatility Changes.” In “Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section,,” 177–181. American Statistical Association, Alexandria, VA (1976).
- Bollerslev, T., and V. Todorov. “Tails, Fears, and Risk Premia.” *Journal of Finance*, 66 (2011), 2165–2211.
- Brennan, M. “The Pricing of Contingent Claims in Discrete Time Models.” *Journal of Finance*, 34 (1979), 53–68.
- Briys, E.; M. Crouhy; and H. Schlesinger. “Optimal Hedging under Intertemporally Dependent Preferences.” *Journal of Finance*, 45 (1990), 1315–1324.
- Buehler, H. “Consistent Variance Curve Models.” *Finance and Stochastics*, 10 (2006), 178–203.
- Buraschi, A., and A. Jiltsov. “Model Uncertainty and Option Markets with Heterogeneous Beliefs.” *Journal of Finance*, 61 (2006), 2841–2897.
- Calvet, L.; J.-M. Grandmont; and I. Lemaire. “Aggregation of Heterogeneous Beliefs, Asset Pricing and Risk Sharing in Complete Markets.” Working paper, HEC Paris and CREST (2004).
- Carr, P., and R. Lee. “Robust Replication of Volatility Derivatives.” Unpublished working paper, New York University and University of Chicago (2008).
- Carr, P., and L. Wu. “Variance Risk Premia.” *Review of Financial Studies*, 22 (2009), 1311–1341.

- CBOE. “VIX: Fact & Fiction.” Unpublished working paper, Research Department, Chicago Board of Options Exchange (2009).
- Chen, H.; S. Joslin; and N.-K. Tran. “Rare Disasters and Risk Sharing with Heterogeneous Beliefs.” *Review of Financial Studies*, 25 (2012), 2189–2224.
- Christoffersen, P.; S. Heston; and K. Jacobs. “Capturing Option Anomalies with a Variance-Dependent Pricing Kernel.” *Review of Financial Studies*, 26 (2013), 1963–2006.
- Cont, R., and T. Kokholm. “A Consistent Pricing Model for Index Options and Volatility Derivatives.” Unpublished working paper, Columbia University and Aarhus University (2009).
- Coval, J., and T. Shumway. “Expected Option Returns.” *Journal of Finance*, 56 (2001), 983–1009.
- Cox, J. C., and C. Huang. “Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process.” *Journal of Economic Theory*, 49 (1989), 33–83.
- DeGroot, M., and M. Schervish. *Probability and Statistics*. Addison Wesley, New York, NY (2002).
- Detemple, J., and S. Murthy. “Intertemporal Asset Pricing with Heterogeneous Beliefs.” *Journal of Economic Theory*, 62 (1994), 294–320.
- Dieckmann, S. “Rare Event Risk and Heterogeneous Beliefs: The Case of Incomplete Markets.” *Journal of Financial and Quantitative Analysis*, 46 (2011), 459–488.
- Egloff, D.; M. Leippold; and L. Wu. “The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments.” *Journal of Financial and Quantitative Analysis*, 45 (2010), 1279–1310.



- Föllmer, H., and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, Berlin, NY (2004).
- Gatheral, J. “Consistent Modeling of SPX and VIX Options.” In “The Fifth World Congress of the Bachelier Finance Society,,” John Wiley and Sons, London (2008).
- Grossman, S., and J. Stiglitz. “On the Impossibility of Informationally Efficient Markets.” *American Economic Review*, 70 (1980), 393–408.
- Hansen, L., and R. Jagannathan. “Implications of Security Market Data for Dynamic Economies.” *Journal of Political Economy*, 99 (1991), 225–261.
- Hong, H. “A Model of Returns and Trading in Futures Markets.” *Journal of Finance*, 55 (2000), 959–988.
- Huang, C., and R. Litzenberger. *Foundations for Financial Economics*. Prentice Hall, Upper Saddle River, NJ (1988).
- Jackwerth, J., and G. Vilkov. “Asymmetric Volatility Risk: Evidence from Option Markets.” Unpublished working paper, University of Konstanz (2013).
- Jouini, E., and C. Napp. “Consensus Consumer and Intertemporal Asset Pricing with Heterogeneous Beliefs.” *Review of Economic Studies*, 74 (2007), 1149–1174.
- Karatzas, I., and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, NY (1991).
- Kogan, L.; S. Ross; J. Wang; and M. Westerfield. “The Price Impact and Survival of Irrational Traders.” *Journal of Finance*, 61 (2006), 195–229.

Leland, H. “Who Should Buy Portfolio Insurance?” *Journal of Finance*, 35 (1980), 581–594.

Lin, Y.-N. “VIX Option Valuation.” Unpublished working paper, Department of Finance, National Chung Hsing University (2009).

Neftci, S. *An Introduction to the Mathematics of Financial Derivatives*. Academic Press, New York, NY (2007).

Patton, A., and A. Timmermann. “Monotonicity in Asset Returns: New Tests with Applications to the Term Structure, the CAPM and Portfolio Sorts.” *Journal of Financial Economics*, 98 (2010), 605–625.

Rubinstein, M. “The Valuation of Uncertain Income Streams and the Pricing of Options.” *Bell Journal of Economics*, 7 (1976), 407–425.

Rubinstein, M. “Implied Binomial Trees.” *Journal of Finance*, 49 (1994), 771–818.

Sepp, A. “VIX Option Pricing in a Jump-Diffusion Model.” *Risk*, April (2008), 84–89.

Song, Z., and D. Xiu. “A Tail of Two Option Markets: State-Price Densities and Volatility Risk.” Unpublished working paper, Federal Reserve Board and University of Chicago (2013).

Todorov, V. “Variance Risk Premium Dynamics: The Role of Jumps.” *Review of Financial Studies*, 23 (2010), 345–383.

Xiong, W. “Bubbles, Crises, and Heterogeneous Beliefs.” Unpublished working paper, Princeton University (2013).

Xiong, W., and H. Yan. “Heterogeneous Expectations and Bond Markets.” *Review of Financial Studies*, 23 (2010), 1433–1466.

Zhu, Y., and J. Zhang. “Variance Term Structure and VIX Futures Pricing.” *International Journal of Theoretical and Applied Finance*, 10 (2007), 111–127.

## Appendix A: Proof of Theorem 1

**Proof of Part I of Theorem 1:** The choice problem of the two investors in Problem 1 implies the optimality conditions,

$$(A-1) \quad e^{-\gamma_f f[v_T]} \mathcal{A}[v_T] = \lambda^f \frac{q[v_T]}{\Phi[v_T]} \quad \text{and} \quad e^{-\gamma_g g[v_T]} \mathcal{B}[v_T] = \lambda^g \frac{q[v_T]}{\Phi[v_T]},$$

where  $\lambda^f$  and  $\lambda^g$  are the Lagrange multipliers associated with the budget constraints of the investors.

Now define the pricing kernel as  $m[v_T] \equiv q[v_T]/\Phi[v_T]$ . From equation (A-1), take a logarithmic derivative to obtain the equations:

$$(A-2) \quad f'[v_T] = \frac{1}{\gamma_f} \left( \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} - \frac{m'[v_T]}{m[v_T]} \right) \quad \text{and} \quad g'[v_T] = \frac{1}{\gamma_g} \left( \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]} - \frac{m'[v_T]}{m[v_T]} \right).$$

Since the claims are in zero net supply, we have  $\phi f'[v_T] + (1 - \phi) g'[v_T] = 0$ . This implies,

$$(A-3) \quad \phi f'[v_T] + (1 - \phi) g'[v_T] = 0 = \frac{\phi}{\gamma_f} \left( \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} - \frac{m'[v_T]}{m[v_T]} \right) + \frac{1 - \phi}{\gamma_g} \left( \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]} - \frac{m'[v_T]}{m[v_T]} \right).$$

Rearranging (A-3), we get the differential equation,

$$(A-4) \quad \frac{m'[v_T]}{m[v_T]} = \left( \frac{\phi \gamma_g}{(1 - \phi) \gamma_f + \phi \gamma_g} \right) \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} + \left( \frac{(1 - \phi) \gamma_f}{(1 - \phi) \gamma_f + \phi \gamma_g} \right) \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]}.$$

The solution to (A-4), when expressed in terms of the change of measure densities, and letting

$\eta \equiv (\phi\gamma_g)/((1-\phi)\gamma_f + \phi\gamma_g)$ , is,

$$(A-5) \quad \ln(m[v_T]) = \ln(m_0) + \eta \ln(\mathcal{A}[v_T]) + (1-\eta) \ln(\mathcal{B}[v_T]),$$

where the constant  $m_0$  satisfies  $\int m[v_T]\Phi[v_T]dv_T = 1$ . Hence,  $m[v_T] = m_0 (\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta}$ .

Based on equations (A-3) and (A-4), we arrive at the optimal volatility contingent cash flows,

$$(A-6) \quad f'[v_T] = \left( \frac{(1-\phi)}{(1-\phi)\gamma_f + \phi\gamma_g} \right) \left( \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} - \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]} \right),$$

$$(A-7) \quad g'[v_T] = - \left( \frac{\phi}{(1-\phi)\gamma_f + \phi\gamma_g} \right) \left( \frac{\mathcal{A}'[v_T]}{\mathcal{A}[v_T]} - \frac{\mathcal{B}'[v_T]}{\mathcal{B}[v_T]} \right).$$

Solving the differential equation (A-6) and imposing  $\int f[v_T]q[v_T]dv_T = 0$ , we obtain:

$$(A-8) \quad f[v_T] = - \left( \frac{(1-\phi)}{(1-\phi)\gamma_f + \phi\gamma_g} \right) \int \ln \left( \frac{\mathcal{A}[v_T]}{\mathcal{B}[v_T]} \right) q[v_T]dv_T + \left( \frac{(1-\phi)}{(1-\phi)\gamma_f + \phi\gamma_g} \right) \ln \left( \frac{\mathcal{A}[v_T]}{\mathcal{B}[v_T]} \right).$$

For  $f[v_T]$  to be concave, we need  $f''[v_T] < 0$ . If  $\mathcal{A}[v_T]/\mathcal{B}[v_T]$  is posited to be concave, then  $\ln(\mathcal{A}[v_T]/\mathcal{B}[v_T])$  preserves concavity.  $\square$

**Proof of Part II of Theorem 1:** A function  $\mathcal{L}[v]$  is log-convex on an interval if  $\mathcal{L}[v]$  is positive and  $\ln(\mathcal{L}[v])$  is convex (Bagnoli and Bergstrom (2005)). It follows from equation (A-5) and condition (c) that  $\ln(m[v_T])$  is convex, being a convex combination of two convex functions. Therefore,  $m''[v_T]/m[v_T] - (m'[v_T]/m[v_T])^2 > 0$ , and hence,  $m''[v_T] > 0$ .

Together with conditions (a) and (b), this implies that  $m[v_T]$  is log-convex, decreasing at  $\underline{v}$  and increasing at  $\bar{v}$ , and achieves a minimum in the interior of  $[\underline{v}, \bar{v}]$ .  $\square$

## Appendix B: A Model of Volatility Exposure in Continuous Time

The setup of the continuous-time economy is similar in its information structure and investor beliefs to those adopted in Detemple and Murthy (1994), Basak (2000, 2005), Buraschi and Jiltsov (2006), and Kogan, Ross, Wang, and Westerfield (2006). Hence, we intend to be brief.

An important property of our setup is that instantaneous market volatility, which evolves stochastically over time, is not a traded asset (see, e.g., CBOE (2009)). Therefore, in our analysis we assume that trading opportunities are confined to a position in a futures contract that settles into the instantaneous volatility at maturity. We further assume that the volatility market is dynamically complete, and our aim is to show that the pricing kernel in this dynamic setting is monotonic in volatility.

Other features of the continuous-time model are as follows. First, the current time is 0, the futures contract matures at time  $T$ , and investors consume only at time  $T$ . Second, the economy is inhabited by two investor types,  $F$  and  $G$ , who are endowed with some initial cash,  $Z_0^f > 0$  and  $Z_0^g > 0$ , respectively. Third, the two investor types are present with mass  $0 < \phi < 1$  and  $1 - \phi$ , respectively.

Uncertainty about the evolution of volatility is described by a one-dimensional standard Brownian motion  $\omega_t^{\mathbb{P}}$ , and it is assumed that volatility  $v_t$  follows geometric Brownian motion, under the physical probability measure  $\mathbb{P}$ ,

$$(B-1) \quad dv_t = \mu_p v_t dt + \sigma_p v_t d\omega_t^{\mathbb{P}},$$

where  $\mu_p$  and  $\sigma_p$  are constants.

In choosing their trading policy in the futures contract on volatility, each investor uses a subjective probability measure  $\mathbb{P}^i$ , for  $i \in \{f, g\}$ . We follow, among others, Detemple and Murthy (1994) and Kogan, Ross, Wang, and Westerfield (2006), and define,

$$(B-2) \quad \omega_t^{\mathbb{P}^i} = \omega_t^{\mathbb{P}} + \vartheta_i t, \quad \vartheta_i \equiv \frac{\mu_p - \mu_i}{\sigma_p}, \quad i \in \{f, g\}, \quad 0 \leq t \leq T,$$

which implies that investor  $i$  believes that volatility  $v_t$  evolves as,

$$(B-3) \quad v_t = v_0 \exp\left(\left(\mu_i - \frac{1}{2}\sigma_p^2\right)t + \sigma_p \omega_t^{\mathbb{P}^i}\right), \quad i \in \{f, g\}.$$

It follows from the Radon-Nikodym theorem (e.g., Karatzas and Shreve (1991)) that the change of measure density for investor  $i$  at time  $T$ ,  $\xi_T^i$ , is,

$$(B-4) \quad \xi_T^i = \exp\left(-\frac{1}{2}\vartheta_i^2 T - \vartheta_i \omega_T^{\mathbb{P}}\right).$$

To fix the remaining notation, let the futures price be  $y_t$  and  $r$  be the constant instantaneous interest rate. As investor  $i$ 's position in the futures contract dynamically changes, her wealth evolves as (see also Briys, Crouhy, and Schlesinger (1990) and Hong (2000)),

$$(B-5) \quad dW_t^i = rW_t^i dt + N_t^i dy_t,$$

where  $N_t^i$  is the number of shares held (long or short) in the futures contract at time  $t$  and, hence, equation (B-5) reflects marking to market. The futures contract is in zero net supply, and so  $\phi N_t^f +$

$(1 - \phi)N_t^g = 0$  for each  $t$ .

At maturity  $y_T = v_T$ . The futures price at time  $t$  is determined as  $y_t = \mathbb{E}_t^{\mathbb{P}} \left( \frac{m_T}{m_t} v_T \right)$ , where  $m_T$  denotes the pricing kernel process at time  $T$ , and we normalize  $\mathbb{E}_0^{\mathbb{P}}(m_T) = 1$ .

Following Cox and Huang (1989), one can convert the investor's maximization problem into a static problem, as stated next.

**Problem 2** Choose wealth  $W_T^i$  to maximize  $\left( \frac{-1}{\gamma_i} \right) \mathbb{E}_0^{\mathbb{P}}(\xi_T^i \exp(-\gamma_i W_T^i))$ , subject to  $\mathbb{E}_0^{\mathbb{P}}(m_T W_T^i) = W_0^i$ . The market clearing condition is  $\phi W_T^f + (1 - \phi)W_T^g = \phi Z_0^f + (1 - \phi)Z_0^g = \bar{Z}$ .

We now state a theoretical result from our continuous-time economy.

**Proposition 1** In the continuous-time economy, the pricing kernel  $m_T$  is related to the change of measure densities  $\xi_T^f$  and  $\xi_T^g$  as,

$$(B-6) \quad m_T = m_* (\xi_T^f)^\eta (\xi_T^g)^{1-\eta}, \quad \eta \equiv \frac{\phi \gamma_g}{(1 - \phi)\gamma_f + \phi \gamma_g},$$

where  $\xi_T^f$  and  $\xi_T^g$  are as shown in equation (B-4), and  $m_* = 1/\mathbb{E}_0^{\mathbb{P}} \left( (\xi_T^f)^\eta (\xi_T^g)^{1-\eta} \right)$ . Furthermore, the pricing kernel  $m_T$  is monotonic in volatility  $v_T$ .

**Proof:** The first-order optimality conditions for the two investors are,

$$(B-7) \quad \xi_T^f \exp(-\gamma_f W_T^f) = \lambda^f m_T, \quad \text{and} \quad \xi_T^g \exp(-\gamma_g W_T^g) = \lambda^g m_T.$$

Imposing the market clearing condition, and normalizing  $\mathbb{E}_0^{\mathbb{P}}(m_T) = 1$  to eliminate the two Lagrangian multipliers  $\lambda^f$  and  $\lambda^g$ , yields equation (B-6).



Note that  $\omega_T^{\mathbb{P}} = \frac{1}{\sigma_p} \log(v_T/v_0) + \frac{1}{\sigma_p} (\sigma_p^2/2 - \mu_p) T$ , and the pricing kernel is multiplicative in  $\xi_T^f$  and  $\xi_T^g$ . As a result,  $\xi_T^f$  and  $\xi_T^g$  are exponential in  $\log(v_T/v_0)$ .

Using these features, along with the fact that  $\mu_p, \mu_f, \mu_g$ , and  $\sigma_p$  enter as constants in the derived solution (B-6) implies that the pricing kernel is monotonic in  $\log(v_T/v_0)$ . Such a pricing kernel can only admit an upward or downward sloping region in  $v_T/v_0$ .  $\square$

## Appendix C: Implications of a Model where Volatility is Disliked

The model presented here provides one set of predictions about the tails of the volatility distribution, which could be explored using VIX option data. It also allows us to place in perspective the model of the volatility market that we suggest in Section II.

Let  $S_{t+T} \in (0, \infty)$  be the level of the market index at time  $t + T$  and  $v_{t+T} \in (0, \infty)$  be the market return volatility.

Claims, contingent on  $S_{t+T}$  and/or  $v_{t+T}$ , can be generically represented by their payoff  $h[S_{t+T}, v_{t+T}]$ .

When considering generic contingent claims dependent on the two-dimensional vector  $(S_{t+T}, v_{t+T})$ , for which there could be data from financial markets, it suffices to work with the projection of the pricing kernel onto the space, generated by the market index and return volatility, even if the pricing kernel may admit higher-dimensional state dependencies.<sup>6</sup> Hence, we consider here a pricing kernel

---

<sup>6</sup>The time- $t$  price  $S_t^h$  of the claim with payoff  $h[S_{t+T}, v_{t+T}]$  at  $t + T$  is  $S_t^h = (1 + r_f)^{-1} \mathbb{E}^{\mathbb{P}}(h[S_{t+T}, v_{t+T}] \times \mathbf{m}) = (1 + r_f)^{-1} \mathbb{E}^{\mathbb{P}}\left(h[S_{t+T}, v_{t+T}] \times \mathbb{E}^{\mathbb{P}}\left(\mathbf{m} \mid S_{t+T}, v_{t+T}\right)\right) = \mathbb{E}^{\mathbb{P}}(h[S_{t+T}, v_{t+T}] \times \tilde{\mathbf{m}}[S_{t+T}, v_{t+T}]/1 + r_f)$ , where  $\mathbb{E}^{\mathbb{P}}(\cdot)$  is expectation under the physical probability measure, and  $r_f \geq 0$  is the  $T$ -period riskfree return. Moreover,  $\mathbf{m}$  represents the probability change of measure with the bond price as numeraire, and  $\tilde{\mathbf{m}}[S_{t+T}, v_{t+T}] \equiv \mathbb{E}^{\mathbb{P}}(\mathbf{m} \mid S_{t+T}, v_{t+T})$ . Accordingly, all that is

that is dependent on  $S_{t+T}$  and  $v_{t+T}$ , and is denoted as  $m[S_{t+T}, v_{t+T}]$ .

If we normalize  $S_t = 1$ , then  $z_{t+T} \equiv \ln(S_{t+T}) \in (-\infty, \infty)$  represents the logarithmic return of the market. One can write the pricing kernel as  $m[z_{t+T}, v_{t+T}]$ , or, for brevity, as  $m[z, v]$ . Let  $\Phi[z, v]$  represent the joint density function corresponding to the physical probability measure  $\mathbb{P}$ . For the discussion to come, we denote the conditional density of  $z$  given  $v$  as  $\widehat{\Phi}_1[z|v] = \Phi[z, v] / \int_{-\infty}^{\infty} \Phi[z, v] dz$ , and the marginal density of  $v$  as  $\Phi_2[v] = \int_{-\infty}^{\infty} \Phi[z, v] dz$ .

Define the expected return of the put on volatility with strike price  $K$  as,

$$(C-1) \quad 1 + \mu_v^{\text{put}}[K] = \frac{\int_0^{\infty} \int_{-\infty}^{\infty} (K - v)^+ \Phi[z, v] dz dv}{\int_0^{\infty} \int_{-\infty}^{\infty} (K - v)^+ m[z, v] \Phi[z, v] dz dv},$$

and analogously for the call on volatility, where  $\ell^+ \equiv \max(\ell, 0)$ .

Three assumptions underpin our analysis concerning expected returns of options on volatility.

**Assumption C.1** *The pricing kernel  $m[z, v]$  is declining and convex in market index returns  $z$  for each volatility  $v$  of market index returns.*

Assumption C.1 is broadly consistent with asset pricing theory that often postulates a pricing kernel that is declining and convex in market index returns (e.g., Rubinstein (1976), Brennan (1979), Hansen and Jagannathan (1991), and Ait-Sahalia and Lo (2000)).

**Assumption C.2**  *$m[z, v]$  is increasing in the volatility  $v$  of market index returns for each  $z$ .*

Assumption C.2 reflects the economic intuition that high volatility states are generally disliked, required to be modeled is  $\mathbb{E}^{\mathbb{P}}(m|S_{t+T}, v_{t+T})$ .

and hence, may be associated with a high level of the pricing kernel.

**Assumption C.3** *The conditional density of market index returns  $z$  given volatility  $v$ ,  $\widehat{\Phi}_1[z|v]$ , with lower volatility dominates, in the sense of second-order stochastic dominance, those with higher volatility.*

While Assumptions C.1 and C.2 relate to the pricing kernel  $m[z, v]$ , Assumption C.3 is about the physical probability measure. Assumption C.3 is based on the argument that increases in volatility that are not compensated by an increase in the expected market return would be disliked by concave utilities, and hence the second-order stochastic dominance (e.g., Huang and Litzenberger (1988)).

**Theorem 2** *Under Assumptions C.1, C.2, and C.3, expected returns of puts (calls) on volatility decrease in strike and are positive (negative) for sufficiently low (high) strikes.*

**Proof.** Follows from Lemma 1, Lemma 2, and Lemma 3 below.

Recall that in our notation  $z \in (-\infty, \infty)$  is market index return,  $v \in (0, \infty)$  is volatility of market index return, and  $\Phi[z, v]$  is their joint probability density.

The expected return  $\mu_v[K]$  of an option on volatility with strike  $K$  is,

$$(C-2) \quad 1 + \mu_v[K] = \frac{\int_0^\infty \int_{-\infty}^\infty h[v; K] \Phi[z, v] dz dv}{\int_0^\infty \int_{-\infty}^\infty h[v; K] m[z, v] \Phi[z, v] dz dv},$$

where  $h[v; K] = (K - v)^+$  or  $h[v; K] = (v - K)^+$ , respectively, for volatility puts and calls.

Upon integrating the market index return  $z$ , we may write (C-2) as,

$$(C-3) \quad 1 + \mu_v[K] = \frac{\int_0^\infty h[v; K] \Phi_2[v] dv}{\int_0^\infty h[v; K] \widehat{m}[v] \Phi_2[v] dv},$$

where  $\widehat{m}[v]$  and the marginal density  $\Phi_2[v]$  of  $v$  are defined as,

$$(C-4) \quad \widehat{m}[v] = \int_{-\infty}^{\infty} m[z, v] \widehat{\Phi}_1[z|v] dz \quad \text{and} \quad \Phi_2[v] = \int_{-\infty}^{\infty} \Phi[z, v] dz \quad \text{for } v \in (0, \infty),$$

and  $\widehat{\Phi}_1[z|v] = \Phi[z, v]/\Phi_2[v]$  is the conditional density of  $z$ , with  $\int_{-\infty}^{\infty} \widehat{\Phi}_1[z|v] dz = 1$  (see, for instance, DeGroot and Schervish (2002)).

According to Lemma 1, the expectation of  $m[z, v]$  conditional on volatility is increasing in volatility, i.e.,  $d\widehat{m}[v]/dv > 0$ . Besides,  $\int_0^\infty \widehat{m}[v] \Phi_2[v] dv = 1$ . Therefore, the conditions of Lemma 3 are satisfied for  $\widehat{m}[v]$  and  $\Phi_2[v]$  in equation (C-3), and the statements in Theorem 2 follow.  $\square$

**Lemma 1** *For a function  $m[z, v] > 0$ , let  $\frac{\partial m[z, v]}{\partial z} < 0$  (Assumption C.1), and  $\frac{\partial m[z, v]}{\partial v} > 0$  (Assumption C.2). Then, by Assumption C.3 and Lemma 2, we have  $\frac{d\widehat{m}[v]}{dv} > 0$ , where  $\widehat{m}[v]$  is as defined in equation (C-4).*

**Proof.** Observe that

$$(C-5) \quad \frac{d\widehat{m}[v]}{dv} = \int_{-\infty}^{\infty} \frac{\partial m[z, v]}{\partial v} \widehat{\Phi}_1[z|v] dz + \int_{-\infty}^{\infty} m[z, v] \frac{\partial \widehat{\Phi}_1[z|v]}{\partial v} dz.$$

The first term in equation (C-5) is positive, given our condition on  $\partial m[z, v]/\partial v > 0$ . The statement of the lemma follows, since the second term is also positive given (C-6) of Lemma 2, for all  $v$ .  $\square$

**Lemma 2** Suppose  $v^c \leq v^d$ . Further, suppose lower volatility conditional densities second-order stochastically dominate the higher volatility conditional densities with  $\widehat{\Phi}_1[z|v^c] \succeq^{ssd} \widehat{\Phi}_1[z|v^d]$  (our Assumption C.3). Then,

$$(C-6) \quad \int \mathcal{M}[z] \widehat{\Phi}_1[z|v^c] dz \leq \int \mathcal{M}[z] \widehat{\Phi}_1[z|v^d] dz, \quad \text{for } v^c \leq v^d,$$

for a decreasing convex function  $\mathcal{M}[z]$ , given some fixed volatility.

**Proof.** From Theorem 2.93 of Föllmer and Schied (2004, page 103), we have,

$$(C-7) \quad \int (-\mathcal{M}[z]) \widehat{\Phi}_1[z|v^c] dz \geq \int (-\mathcal{M}[z]) \widehat{\Phi}_1[z|v^d] dz, \quad \text{for } v^c \leq v^d,$$

for an increasing concave function  $(-\mathcal{M}[z])$ . Reversing the inequality gives equation (C-6).  $\square$

**Lemma 3** Suppose  $m[v]$  is a pricing kernel,  $\int_0^\infty m[v] \Phi[v] dv = 1$ , and  $\Phi[v]$  is the density of  $v$ . When  $\frac{dm[v]}{dv} > 0$ , the following statements are true: (i) expected put returns decrease in strike and are positive for sufficiently low strikes, and (ii) expected call returns decrease with strike and are negative for sufficiently high strikes.

**Proof.** For the proof of the signs and strike dependence of expected option returns in (i) and (ii), see, e.g., Coval and Shumway (2001), with the exception that our results pertain to a pricing kernel that is monotonically increasing in volatility (i.e.,  $dm[v]/dv > 0$ ).  $\square$

The proof of Theorem 2 hinges on the result that the expectation of  $m[z, v]$ , conditional on volatility, is increasing in volatility (see Lemma 1). Such a result can also be derived under the alternative

stronger assumption that the conditional density of  $z$  given  $v$ ,  $\widehat{\Phi}_1[z|v]$ , with lower volatility first-order stochastically dominates the higher volatility counterparts, in which case one can relax the assumption of convexity. First-order stochastic dominance could be obtained, for instance, when there is a strong leverage effect (e.g., Black (1976)) and high volatility states map to reduced expected market returns.

One may view Theorem 2 as extending the theoretical results on expected returns of options written on the market index (Coval and Shumway (2001)).

Theorem 2 and, hence, the pattern of expected returns of volatility options across moneyness provides testable implications for the change of measures, or pricing kernels, that can be supported in the volatility market.

TABLE 1  
**Implications of the Model with Heterogeneity in Beliefs for Expected VIX Option Returns**

	Example 1					Example 2	
	Panel A	Panel B	Panel C	Panel D		Panel E	Panel F
<i>Model parameters</i>							
$\mu_p$	0.000	0.000	0.000	0.000	$\mu_p$	0.000	0.000
$\sigma_p$	0.171	0.171	0.171	0.171	$\sigma_p$	0.171	0.171
$\gamma_f$	3.500	3.500	3.000	3.000	$\gamma_f$	3.500	3.500
$\gamma_g$	3.000	3.000	3.500	3.500	$\gamma_g$	3.000	3.000
$\mu_f$	0.150	0.100	0.150	0.100	$\mu_a$	0.150	0.150
$\mu_g$	-0.150	-0.100	-0.150	-0.100	$\mu_b$	0.200	0.200
$\sigma_f$	0.184	0.179	0.184	0.179	$\sigma_a$	0.100	0.100
$\sigma_g$	0.184	0.179	0.184	0.179	$\sigma_b$	0.100	0.100
					$\omega_a$	0.060	0.160
					$\omega_b$	0.070	0.150
<i>Means of investors' volatility belief distributions</i>							
Investor type F	15%	10%	15%	10%		15%	10%
Investor type G	-15%	-10%	-15%	-10%		-15%	-10%
<i>Expected VIX option returns</i>							
25% OTM put	-46%	-35%	-28%	-20%	25% OTM put	-41%	-38%
15% OTM put	-31%	-23%	-13%	-9%	15% OTM put	-35%	-32%
5% OTM put	-19%	-14%	-4%	-3%	5% OTM put	-23%	-22%
5% OTM call	-5%	-4%	-21%	-15%	5% OTM call	-24%	-18%
15% OTM call	-13%	-9%	-30%	-22%	15% OTM call	-30%	-24%
25% OTM call	-21%	-15%	-39%	-29%	25% OTM call	-33%	-27%

### Note to Table 1

The pricing density for Example 1 and Example 2 can be computed as,

$$q[v_T] = \begin{cases} \frac{(\Phi_f[v_T])^\eta (\Phi_g[v_T])^{1-\eta}}{\int (\Phi_f[v_T])^\eta (\Phi_g[v_T])^{1-\eta} dv_T} & \text{Example 1} \\ \frac{(\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta} \Phi[v_T]}{\int (\mathcal{A}[v_T])^\eta (\mathcal{B}[v_T])^{1-\eta} \Phi[v_T] dv_T} & \text{Example 2,} \end{cases}$$

where  $\Phi[v_T] = \Phi[v_T; \mu_p, \sigma_p]$  is the lognormal physical density of volatility, as displayed in (14), and  $\Phi_f[v_T] = \Phi[v_T; \mu_f, \sigma_f]$  and  $\Phi_g[v_T] = \Phi[v_T; \mu_g, \sigma_g]$  are the personalized densities of the two types of investors. The change of measure densities  $\mathcal{A}[v_T]$  and  $\mathcal{B}[v_T]$  are presented in equations (27)–(28). The expected returns of the respective options are calculated using the ratios of their expected payoffs under the physical and the pricing densities, that is, for puts,  $(\int (K - v_T)^+ \Phi[v_T] dv_T / \int (K - v_T)^+ q[v_T] dv_T) - 1$ . The parameters employed in Panels A and E correspond to those used in Figure 1 and Figure 2.



TABLE 2  
**Assessing Differences in Returns of VIX Puts and Calls Across Strikes**

<i>Panel A: Bootstrap p-values for differences in average VIX option returns</i>				
	Puts	Calls		
25% OTM versus 15% OTM	0.008	0.359		
25% OTM versus 5% OTM	0.017	0.272		
15% OTM versus 5% OTM	0.246	0.242		

<i>Panel B: Bootstrap-based tests of monotonic relation (p-values)</i>				
	<i>MR</i>	<i>UP</i>	<i>DOWN</i>	
VIX puts	0.049	0.005	0.723	
VIX calls	0.639	0.668	0.272	

<i>Panel C: Average VIX option returns (28 days, in %), open interest, and traded volume</i>				
	Average returns (%)		Open interest	
	[90% CI]		{Volume}	
	Puts	Calls	Puts	Calls
25% OTM	-32.68	-34.84	34,621	70,684
	[-68, 8]	[-83, 26]	{2,493}	{8,660}
15% OTM	4.53	-31.30	39,943	55,047
	[-24, 35]	[-72, 20]	{6,560}	{6,716}
5% OTM	10.21	-27.36	33,657	49,199
	[-11, 32]	[-63, 16]	{5,672}	{6,024}

## Note to Table 2

Returns to expiration of 28-day VIX options are constructed as,

$$1 + r_{t,T}^{c,\text{vix}}[y] = \frac{(\text{VIX}_{t+T} - y\text{FVIX}_{t,T})^+}{C_{t,T}^{\text{vix}}[y]} \quad \text{and} \quad 1 + r_{t,T}^{p,\text{vix}}[y] = \frac{(y\text{FVIX}_{t,T} - \text{VIX}_{t+T})^+}{P_{t,T}^{\text{vix}}[y]}$$

for each month from 03/2006 to 12/2013 (93 months), using *ask quotes*. The moneyness  $y \equiv K/\text{FVIX}_{t,T}$  of a VIX option at time  $t$ , which expires at time  $t + T$ , for  $T = 28$  days, is determined relative to the time- $t$  price  $\text{FVIX}_{t,T}$  of the VIX futures contract with settlement date  $t + 28$  days.

In Panel A, we test for differences in average VIX option returns across strikes and option types. We consider differences between puts (calls) that are 25% and 15%, 25% and 5%, and 15% and 5% OTM, respectively. For each such pair of moneyness levels, we draw 50,000 pairwise bootstrap samples of the respective option returns, and for each pair of bootstrap samples we calculate the difference between their average returns. A negative difference indicates that, in the respective bootstrap sample, the first option in a pair has higher average return than the second option. The reported  $p$ -values in Panel A are the proportion of negative differences in the 50,000 bootstrap samples for each pair of option moneyness levels.

Panel B presents the results from the bootstrap-based *MR*, *UP*, and *DOWN* monotonicity tests of Patton and Timmermann (2010). The null hypothesis for the *MR* test is that the pattern of VIX option returns with respect to moneyness is flat, versus the alternative that the pattern is monotonic. The *UP* and *DOWN* tests detect positive or negative segments in the pattern.  $p$ -values lower than 5% imply rejection of the null. Panel C shows (i) the average option returns (28 days, in %), (ii) bootstrapped 90% confidence intervals (in square brackets) for average returns, (iii) open interest (average daily number of contracts), and (iv) traded volume (in curly brackets, average daily number of contracts).

FIGURE 1  
Pricing kernel and optimal volatility contingent cash flows from Example 1

Plotted are (i) the pricing kernel  $m[v_T]$ , and (ii) the optimal volatility contingent cash flows,  $f[v_T]$  and  $g[v_T]$ , for the two types of investors. The physical density (as shown in equation (14)) and the personalized densities are assumed to be lognormal, and the personalized change of measure densities are as shown in equations (15)–(16). We use  $\mu_p = 0.000$ ,  $\mu_f = 0.150$ ,  $\mu_g = -0.150$ ,  $\sigma_p = 0.171$ ,  $\sigma_f = 0.184$ , and  $\sigma_g = 0.184$ . The coefficients of absolute risk aversion are set to  $\gamma_f = 3.5$ ,  $\gamma_g = 3.0$ , and  $\phi = 0.50$ . Panels A and B are based on equation (20), while Panels C and D, which plot the optimal volatility contingent cash flows, are based on equation (24).

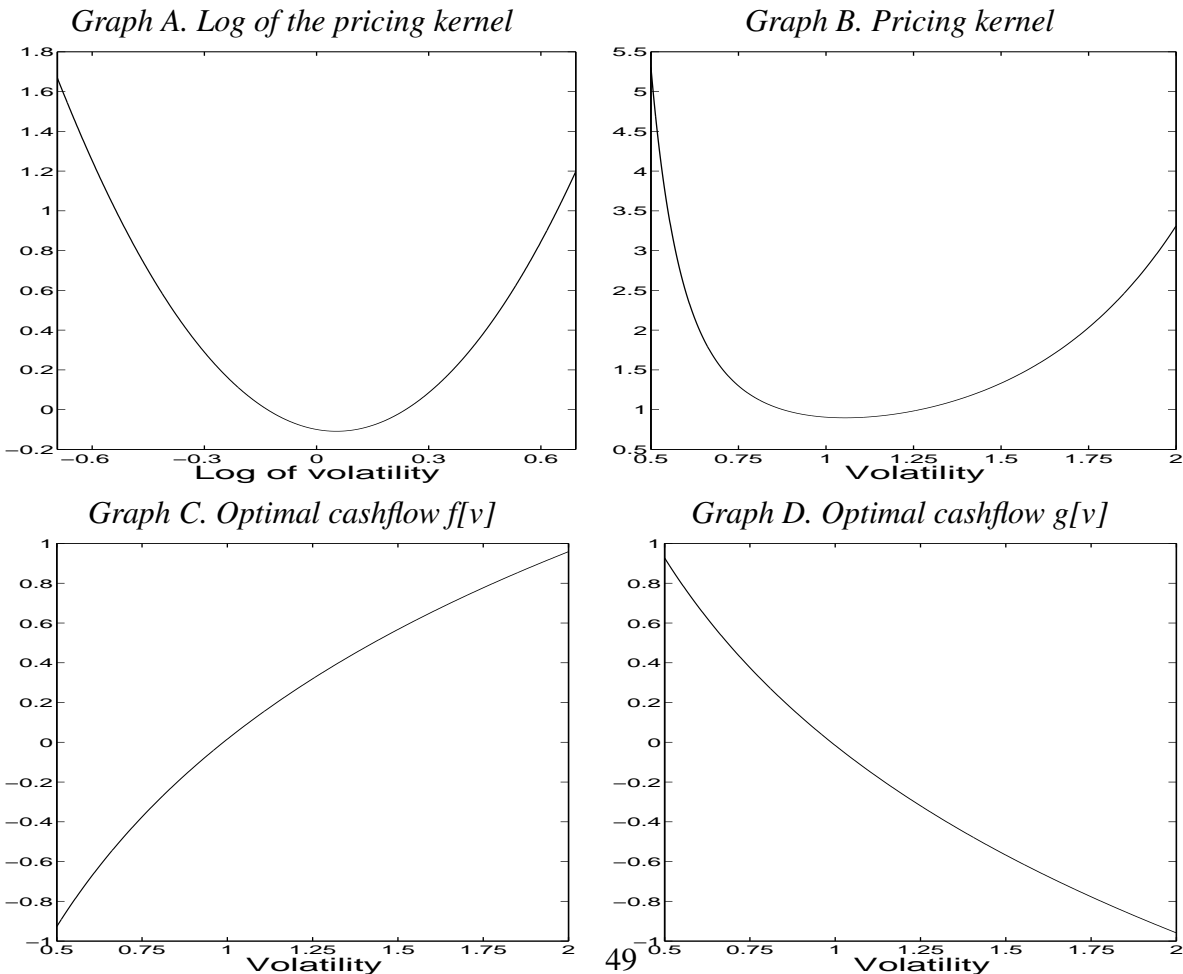


FIGURE 2  
Pricing kernel and optimal volatility contingent cash flows from Example 2

Plotted are (i) the pricing kernel  $m[v_T]$ , and (ii) the optimal volatility contingent cash flows,  $f[v_T]$  and  $g[v_T]$ , for the two types of investors. The physical density (as shown in equation (14)) is assumed to be lognormal, the personalized change of measures are as displayed in equations (27)-(28), and the personalized densities are  $\Phi_f[v_T] = \mathcal{A}[v_T]\Phi[v_T] / \int \mathcal{A}[v_T]\Phi[v_T]dv_T$  and  $\Phi_g[v_T] = \mathcal{B}[v_T]\Phi[v_T] / \int \mathcal{B}[v_T]\Phi[v_T]dv_T$ . We use  $\mu_p = 0.000$ ,  $\mu_a = 0.150$ ,  $\mu_b = 0.200$ ,  $\sigma_p = 0.171$ ,  $\sigma_a = 0.100$ , and  $\sigma_b = 0.100$ . The coefficients of absolute risk aversion are set to  $\gamma_f = 3.5$ ,  $\gamma_g = 3.0$ , and  $\phi = 0.50$ ,  $\varpi_a = 0.060$ , and  $\varpi_b = 0.070$ . Our parameterization implies means of 15% and -15% for the personalized densities of investors F and G, respectively.

