Explaining the Negative Returns to Volatility Claims: An Equilibrium Approach

Bjørn Eraker

Yue Wu *

Abstract

We study the returns to investing in VIX futures and VIX Exchange Traded Notes (ETNs). We document a substantial negative return premium for both ETNs and the futures. For example, the a constant maturity portfolio of one-month VIX futures loses about 30% per year over our sample period (2006-2013). We propose an equilibrium model to explain these negative returns. In this model, increases in volatility endogenously lead to decreasing stock prices. Our model explains the negative expected returns to VIX futures and ETNs as well as several other stylized facts about the returns to VIX futures and VIX futures ETNs.

JEL classification: G12, G13, C22, C58.

Keywords: Variance Risk Premium, VIX futures, VIX ETN, Dynamic Equilibrium, Jump-Diffusion

*Both authors are from the Wisconsin School of Business, Department of Finance. Corresponding author: Bjørn Eraker. 975 University Ave, Madison WI, 53706. (608) 890-2594, beraker@bus.wisc.edu. We are grateful for comments from Grigory Vilkov, Mark Ready, Kjell Nyborg, and seminar participants at the European Winter Finance Summit and the University of Wisconsin.
1 Introduction

In 2004 the CBOE Futures Exchange introduced cash settled futures contracts on the CBOE VIX volatility index. While initially sparsely traded, the VIX futures market has become very liquid in recent years. In addition to the futures market itself, since 2009, more than a dozen VIX futures Exchange Traded Notes (ETNs) have been introduced, allowing retail investors to trade VIX futures through regular brokerage accounts. The ETNs follow simple, pre-specified, dynamic trading programs, and in most cases offer constant maturity exposure to n-month futures positions.

The interest in VIX futures and ETNs trading is due at least in part to the perceived positive diversification benefits of the contracts. The CBOE notes through various marketing materials that the VIX correlates negatively with the S&P 500 returns and therefore provides diversification benefits. The CBOE’s own estimates of the VIX-return correlation range from -75% to -86%. Additionally, since the VIX is significantly more volatile than the S&P 500 itself, the VIX, and thus VIX futures, have substantial negative market betas.

The first objective of our paper is to provide descriptive statistics on the average returns to VIX futures positions and the associated ETNs. Szado (2009), Alexander and Korovilas (2012) and Whaley (2013) report negative annualized VIX futures returns. We collect futures data from January 2006 to April 2013 and confirm these findings. For example, if an investor were to invest in VIX futures in January 2006 and roll the position at end-of-day futures prices reported by the CBOE, she would have lost more than 97% of the initial investment by the end of March 2013. This corresponds to an annualized return of about negative 30%. This number is staggering considering that during the first part of the sample period the investor would have more than doubled the initial investment through the peak of the 2008 financial crisis. Not surprisingly, the VIX ETNs perform as badly, if not worse, than the underlying futures. In fact, since the first two VIX ETNs were introduced on January 30, 2009, the VXX and VXZ, which offer exposure to short and medium term futures respectively, have lost an average of 34 and 14 basis points per day (simple returns).
Our second and major objective is to understand what causes this large negative return premium in VIX futures and ETNs. We are interested in testing whether the negative returns are consistent with some notion of equilibrium. Since VIX futures have negative market betas, one might expect that they earn negative returns in a CAPM equilibrium. However, we show that standard linear factor models including the CAPM and the Fama-French three factor models cannot explain the returns. We use the equilibrium model of Eraker and Wang (2011) to derive equilibrium VIX futures prices. This model is based on a dynamic present-value framework where investors worry about downside jump-risk differently from what is the case in the traditional, static CAPM. We show that the model produces a very sizable volatility risk premium. Importantly, the model generates an upward sloping equilibrium futures curve (contango) in steady state. This means that, ceteris paribus, investors who purchase VIX futures pay more than the value of the spot VIX at expiration of the futures contract, on average. The equilibrium model produces a negative premium in all states of the world, whether or not the VIX is above or below its steady state value. Even if the futures curve is in backwardation (downward sloping), the futures may imply a negative risk premium because the physical speed of mean reversion will be faster than the \( Q \) measure speed of mean reversion implicit in the futures prices. These pricing implications are entirely equilibrium outcomes. If the representative agent in the model is risk-neutral, none of these pricing implications hold. In particular, there is no volatility risk premium, the steady state futures curve is essentially flat, and the expected return on VIX futures is zero. We elaborate further below.

Our paper is connected to the extant literature in several ways. Our theoretical model is related to long-run risk models (Bansal and Yaron (2004)) that deliver large volatility risk premia such as those of Eraker and Shaliastovich (2008), and Drechsler and Yaron (2011). Other theoretical justifications for large volatility risk premia include the heterogeneous beliefs model of Buraschi, Trojani, and Vedolin (2011). Bollerslev, Tauchen, and Zhou (2009), Andersen and Todorov (2013) among others show that the volatility risk premium can predict stock market returns. Eraker (2011) shows that a large volatility risk premium is consistent with large negative equity options returns such as found empirically in Bondarenko (2003), Bakshi and Kapadia (2003) and Eraker (2013), among others. Broadie, Chernov, and Johannes (2007) conclude that
jump risk premium, not volatility risk premium, is the primary driver of risk premia in option returns. Recently, Andersen and Todorov (2013) propose a model with a self-exciting jump process but find that this “tail factor” has no incremental power in predicting equity return above the level of volatility itself. This empirical finding lends support to the specification of models in which jump-risks are not disentangled from the diffusive variance, as in our model.

In our model, jump and volatility risk premia obtain endogenously and both are increasing in the level of risk-aversion exerted by the agents. Simplified, if agents are risk averse, they care about the volatility of future cash flows. Their aversion toward high volatility is essentially similar across diffusive and jump driven increments to volatility. Yet, the equilibrium price process we use has characteristics that are similar to existing reduced-form, no-arbitrage models. Our most general model has a two factor volatility specification where jumps in the volatility endogenously lead to negative jumps in the equilibrium price. This is similar, but not identical to, the volatility co-jump models in Duffie, Pan, and Singleton (2000), Eraker, Johannes, and Polson (2003), Bandi and Reno (2012) and Andersen and Todorov (2013) where the correlations between volatility and prices are assumed.

While our paper is the first to attempt a structural explanation for the negative VIX futures return premium, several papers fit statistical models to futures prices and judge the resulting empirical model fit using RMSE or other distance metrics based on the difference between the model and market prices. For example, Zhang and Zhu (2006) analyze the model fit based on Heston (1993), while Lin (2007) and Zhu and Lian (2012) analyze models in the more general class of Duffie, Pan, and Singleton (2000). These papers generally conclude that models with more complicated volatility dynamics (i.e., jumps) are preferable. Egloff, Leippold, and Wu (2010) find that the two factor model of Bates (2000) outperforms Heston’s model. Some studies from related markets include Song (2012) who studies returns on VIX options. He finds that both diffusive volatility-of-volatility and volatility jumps are important in capturing VIX option returns. Carr and Wu (2006) study a sample of returns to variance-swap contracts. Their sample, collected from 1990-2005, contains strikingly large positive (negative) returns to sellers (buyers) of variance swaps.
In our empirical examination we first confirm the large negative returns to VIX futures reported elsewhere. We verify that the negative returns to futures translate into correspondingly negative returns to VIX ETNs that invest in long positions. The returns are particularly bad for short maturity futures and VIX ETNs. Yet, we show that our equilibrium model is generating returns that are almost identical to those we observe in our sample. Our empirical implementation can be summarized as follows: we first estimate our return equilibrium model. This is done using Bayesian MCMC sampler, extending the method in Eraker, Johannes, and Polson (2003) to a structural setting. The advantage of the structural model is that we estimate the risk aversion of the representative agent from returns data alone. We therefore recover the pricing kernel without the use of additional data from derivatives markets. This contrasts empirical studies of reduced form, no-arbitrage models that simultaneously use derivatives and returns data to back out market risk prices, as in Pan (2002), Chernov and Ghysels (2000), Eraker (2004), Andersen and Todorov (2013), and Jackwerth and Vilkov (2014).

We show that the equilibrium model can explain the negative returns to VIX futures and VIX ETNs almost exactly. In particular, the model that includes volatility jumps fits all moments of maturities that are less than four months. For four and five month futures, the model underestimates the variability (standard deviation) of the futures returns. We argue that the model’s inability to account for the return standard deviation of longer maturity contracts is consistent with similar model failures in the affine term structure literature in capturing low frequency movements. We demonstrate that a generalized model that allows for a second volatility factor (i.e, a “central tendency” factor) can be calibrated such that all the moments of the VIX futures data can be matched.

The rest of the paper is organized as follows: in the next section we present basic descriptive evidence on the returns to VIX futures contracts and VIX ETNs. Section 3 presents the equilibrium framework and structural parameter estimates. Section 4 presents our empirical evaluation of the model and compares it to data on VIX futures and VIX ETN options. Section 5 concludes.
2 VIX Futures and ETN Returns

In the following section we provide descriptive evidence of the statistical behavior of VIX futures and ETNs. We start with the futures. Before presenting the evidence, it is worthwhile noting that there are some subtle issues involved in measuring the average returns of VIX futures. In particular, the futures price, like the VIX itself, is extremely volatile. The return distribution also displays right skewness - as the VIX occasionally jumps, a long futures position provides a large positive return. High frequency estimates of average arithmetic returns are upward biased estimates of long horizon buy and hold returns (see Blume (1974)). We therefore report, in addition to the arithmetic returns, the annualized mean log returns as well as the geometric returns. These are defined as follows: let $V_t$ represent a time $t$ value of a portfolio that rolls a futures position at daily closing prices. We compute $V_t = V_{t-1}(1 + r_p(t))$ where

$$1 + r_p(t) = w_{t-1} \frac{F_t(T_1)}{F_{t-1}(T_1)} + (1 - w_{t-1}) \frac{F_t(T_2)}{F_{t-1}(T_2)}$$

is the return to a portfolio of futures with maturities $T_1$ and $T_2$, and $w_{t-1}$ is the weight in the front month futures. We report the average daily arithmetic return,

$$R^1 := \frac{1}{T} \sum_t r_p(t),$$

the average daily log-return,

$$R^2 := \frac{1}{T} \sum_t \ln(1 + r_p(t))$$

as well as the annualized geometric return

$$R^3 := \left( \frac{V_T}{V_0} \right)^{\frac{252}{T}} - 1.$$

While the geometric return is known to be biased for the expected annual return (Blume (1974)), it represents a monotonic transform of the total return over the sample period and we include it for this reason.

[Table 1 about here.]
Table 1 presents measures of return and higher order return moments. Though introduced in 2004, low liquidity in the first two years leads us to start the sample in January 2006. As can be seen from the table, VIX futures averaged negative returns over the sample period. Largest were the losses for the short-term maturity futures, with one month contracts losing an average of 12 basis points per day which roughly annualizes to $252 \times -0.12 = -30\%$. In terms of log-returns the one-month futures averaged negative 20 basis points per day, or -50.4\% annually. The geometric average return was -39.4\%. The returns tend to increase with maturity, and the five-month contract loses a comparably small amount, with “only” -7.33\% percent per annum geometric average loss. Both the average loss for the short maturity contracts, as well as the comparably smaller loss for the long maturities are interesting features of the data. These features of the data are seen in Figure 1 which plots the value and log-value of a dollar invested in rolled positions in VIX futures on January 3rd, 2006.

In order to get a first pass at whether or not the returns are consistent with an equilibrium story, we compute $\alpha$’s using several factor model specifications. As the futures prices are first order dependent on the underlying VIX, we use log changes in VIX ($\Delta VIX_t := \ln(VIX_t/VIX_{t-1})$) in addition to standard Fama-French risk factors. The results are reported in Table 2. As can be seen, the various specifications give large negative $\alpha$’s. The short maturity one to two month $\alpha$’s are statistically significantly negative while the longer maturities are not. We also report tests of the null hypothesis $\alpha_i = 0 \forall i = 1, \ldots, 5$ which are rejected for all model specifications. We conclude accordingly that neither of these asset pricing models can explain the average returns to VIX futures.

To understand where the negative returns come from, we present the average values of the VIX spot and the various maturity futures over the sample period in Table 3. This table shows
that on average, the futures curve is in contango and prices monotonically increase with maturity of the contract. This, mechanically, is the reason why long positions in VIX futures lose money on average. Consider, for example, an investor who buys a one-month contract and holds it until it expires. Her average return would be \(\frac{20.57}{21.48} - 1 = -4.24\%\) per month, or \(-40.52\%\) per year when compounded. This is close to the annually average compounding geometric return in Table 1. Similarly, the annualized, average one-month holding period returns for two to seven month contracts are reported in the row labelled “Implied Return” in Table 3. While not identical, the numbers are on the same order of magnitude as the actual returns reported in Table 1. This shows that in order to understand why the returns to VIX futures are so low, we must understand why the futures curve is on average severely upward sloping in the data. An equilibrium explanation for the negative returns, therefore, will need to generate a sharply upward sloping steady-state futures curve.

### 2.1 Returns to ETNs

The first Exchange Traded Notes (ETNs) linked to VIX futures were introduced in January 2009. The VXX and VXZ offer long exposures to the one and five month futures, respectively. The large negative returns earned on the VIX ETNs are typically thought to be a result of the unfortunate timing of their inception, either during the height of the financial crisis as with the VXX or VXZ, or in its aftermath. It is not true, however, that the decimation of these securities’ value is solely a consequence of the directional move in the VIX over the sample period. The simple evidence of this is the fact that the securities also lose value during periods for which there is no change in the underlying VIX index. For example, during the period March 1, 2010 to June 21, 2012, the VIX went from 19.26 to 20.08, a marginal positive change, but the VXX lost 82.74\% of its value over this sample period while the VXZ lost 29.21\%. Clearly, the directional move in the VIX was not the reason why the VXX lost almost 83\% of its value over this period!
The performance of the VIX ETNs is closely tied, if not identical, to the performance of the synthetic VIX futures portfolios that we have analyzed above. To see this, Figure 2 shows the log value of our synthetic one-month portfolio alongside VXX’s log net asset value. As can be seen, the two are highly correlated and essentially identical with the exception that the VXX depreciates at a slightly higher rate over the sample period. Specifically, the annualized geometric return to the VXX was -65.33% vs. -64.78% for the synthetic portfolio. The difference is only about 0.55% per year, of which 0.89% comprises its management fee. The remaining -34 basis points are presumably due to differences in execution between the VXX and our synthetic portfolio.

[Table 5 about here.]

It is also interesting to study the relative performance of the three ETNs, VXX, TVIX and XIV which span the one-month space with single long, double long, and single short positions, respectively. Since the TVIX and XIV can be replicated by trading in the VXX, we compute the terminal values of the replicating portfolios and compare them to the respective terminal values of the XIV and TVIX. These results are reported in Table 5. We use two performance measures, $G_1$ and $G_2$. $G_1$ denotes the total return difference (annualized) between actual trade price of TVIX and XIV and the value of synthetic securities. $G_2$ denotes the corresponding differential based on the net asset value. As can be seen, the TVIX does about 161 basis points better than the synthetic security of VXX per annum based on the net asset value. Based on the actual trade price, it gains 336 basis points on the synthetic counterpart. The XIV loses 180 and 190 basis points.

[Figure 3 about here.]

Table 5 reveals that the market value of the TVIX tends to deviate from its net asset value. Most VIX ETNs tend to track their net asset values (NAV) relatively closely. The TVIX is an exception to this, and this stock has at times traded significantly above its net asset value. In particular, on March 21, 2012, the TVIX was traded 89% above its net asset value after the
issuer, Credit Suisse, had announced it would stop share issuance. On March 22, Credit Suisse announced that it would issue more shares, resulting in a collapse in the spread between the price and the net asset value. The stock lost almost 60% of its value over the next two days. The stock has traded about 8% above fair value on average since then.

The sharp increase in the price of the TVIX from July 2011 seen in Figure 3 was caused by the second European sovereign debt crisis during which the VIX increased sharply from about 20 to 47 while the S&P 500 dropped about 28%. Anecdotal evidence suggests that there is a high retail investors demand for long ETNs as they provide negative market betas and therefore act as a hedge against financial crisis. VIX futures, in particular the double long TVIX and UVXY ETNs, provide hedges against financial crisis regimes. This mechanism is essential in the equilibrium model, to be discussed next.

3 An Equilibrium Model of VIX Futures

In the following section we outline an equilibrium framework to understand and analyze the returns to VIX futures, and implicitly, ETNs. Our model aims to explain how VIX futures earn high negative expected returns. We take a first pass at this by estimating the (negative) beta of the futures returns and see if the CAPM can explain the returns. With a beta of -2 and a market risk premium of 5% to 7.5%, we end up at a return premium of -10% to -15% if we assume a zero risk free rate, which is about right at least for the latter part of the sample period. The negative beta thus gives us the right sign, but the CAPM still explains only half or less of the magnitude of the negative returns to one month VIX futures. In the next section we develop a model that endogenizes the negative volatility beta.

3.1 Model

To motivate the model, consider a simple one-period model first. At date 0 agents trade claims to a single cash flow, $\tilde{x}_1$, paid one period later, at date 1. The price of the claim is $P_0 = E(\tilde{x}_1)/k$ where by definition the discount rate $k$ is the expected rate of return. Since there
is only one period, risks associated with the terminal payoff \( \tilde{x}_1 \) map one-to-one to risks associated with the return \( R_m = \tilde{x}_1/P_0 \). This means that we can equivalently derive a model based on risks embedded in \( \tilde{x} \) or in \( R \). For example, if we endow a representative agent with quadratic utility we can derive the Sharpe-Lintner single period CAPM and the expected rate of return can be equivalently represented as a function of the variance of \( R_m \) or \( \tilde{x} \).

We extend this simple equilibrium construction in the following way: First, let the terminal time period be denoted \( T \) such that the terminal cash flow is \( \tilde{x}_T \). Second, imagine that claims to \( \tilde{x}_T \) can be traded in the capital market any time prior to and including the terminal date \( T \). Let’s assume that \( P_0 \) represents an equilibrium price of the claim to \( \tilde{x}_T \). Let’s assume that the agents can trade the claim to \( \tilde{x}_T \) again an instant later, at time \( \Delta t \). Why would the price be different from what it was at time 0? In a dynamic economy it is natural to think that there are two types of shocks that impact the price of the risky asset: cash flow shocks and discount rate shocks.

To facilitate both types of shocks we imagine investors learn over time what the size and risk of the terminal cash flow \( \tilde{x}_T \) will be. Specifically, assume that the terminal cash flow is a sum of independent increments, \( \ln x_T = \sum_{i=1}^{N} \Delta \epsilon_i \) where \( \epsilon_i \) represents the \( i \)th cash flow increment \( t_i = i \Delta t \) for \( \Delta t = T/N \). A positive shock will have a positive impact on the price, and vice versa.

In order to incorporate a time-varying discount rate, or expected rate of return, we assume that the cash flow increments have persistent time varying stochastic volatility. Previewing the results of our model, the expected rate of return, or discount rate for the terminal cash flow, will be an increasing function of the volatility of the \( \epsilon_i \)s. Thus, if the volatility of these cash flow shocks increases, the volatility of \( \tilde{x}_T \) is increased as well. If we think about today’s price as the certainty equivalent of \( \tilde{x}_T \), it is clear that by increasing its volatility we 1) increase the volatility of the current price, and 2) the price responds negatively. Both effects are important: The former introduces stochastic, time-varying volatility into the stock price and the latter introduces an endogenous volatility feedback or leverage effect.
We now leave the discrete time setup above in favor of a continuous time economy. As before, agents trade claims to a terminal cash flow \( \tilde{x}_T \). We assume that \( \tilde{x}_T \) is a the terminal value \( x_T = \tilde{x}_T \) of a stochastic process \( x \) which is exogenous. We call a claim to \( \tilde{x}_T \) “the stock market” and assume a unit net supply of stock and zero net supply of bonds.

The cash flow is assumed to be the terminal value of

\[
\frac{dx_t}{x_t} = \mu dt + \sigma_t dB^x_t \tag{2}
\]

\[
d\sigma_t^2 = \kappa(\theta - \sigma_t^2) dt + \sigma_t \sigma_t dB^\nu_t + \xi_t dN_t \tag{3}
\]

\[\xi_t \sim \text{Exp}(\mu_\xi) \tag{4}\]

\[\text{Corr}(dB^x_t, dB^\nu_t) = 0. \tag{5}\]

where \( N_t \) is a Poisson process with arrival intensity \( l_0 \) and where \( \tilde{x}_T = x_T \). \( \sigma_t \) is the volatility of \( x_t \) and is driven by a diffusion, \( B^\nu_t \), and a compound Poisson process, \( \xi_t dN_t \), where the counting process, \( N_t \), has Poisson arrivals with intensity, \( l(\sigma_t^2) = l_0 + l_1 \sigma_t^2 \) where again \( l_0, l_1 \) are parameters such that \( l_0 > 0, l_1 = 0 \) implies constant arrivals and \( l_0 = 0, l_1 > 0 \) implies that arrivals are proportional to \( \sigma_t^2 \). We shall assume that the parameters are chosen so that \( \sigma_t^2 \) is stationary and positive. Under this setup, the (relative) cash flow shocks \( dx_t/x_t \) are uncorrelated and have persistent stochastic volatility \( \sigma_t \). Note that even when the planning horizon \( T - t \) is large, \( \sigma_t \) impacts the variance of the terminal claim \( \tilde{x}_T \). This follows because the variance of the terminal claim equals the integrated variance of the instantaneous shocks, \( dx_t \).

\section{3.2 Equilibrium Stock Prices}

The appendix section 5.2 derives the equilibrium price

\[P_t = \frac{\mathbb{E}_t \{ u'(\tilde{x}_T)\tilde{x}_T \}}{\mathbb{E}_t \{ u'(\tilde{x}_T) \}} e^{-(T-t)r}. \tag{6}\]

Subject to regularity conditions this equation applies generally. It is analytically tractable in the case of exponential affine processes for \( x_T \) in the class of Duffie, Pan, and Singleton (2000), and power utility, \( u'(x) = \beta x^{-\gamma} \). To see this, recall the main insight of Duffie, Pan, and Singleton
(2000). Let $X_t$ be an affine process with domain $D \subseteq \mathbb{R}^N$ and let $u$ be an $N$ dimensional real vector. Then

$$\mathbb{E}_t e^{u'X_{t+\tau}} = e^{\alpha(u,t,t+\tau)+\beta(u,t,t+\tau)'X_t}$$

(7)

where the functions $\alpha : \mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ and $\beta : \mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^N$ solve ordinary, first order differential equations (see Duffie, Pan, and Singleton (2000)).

It now follows by defining $X_t = \{\ln x_t, \sigma_t^2\}$ that by setting $u_{1-\gamma} = (1-\gamma, 0)$ and $u_{-\gamma} = (-\gamma, 0)$ in the numerator and denominator of (6), along with (7), gives the price

$$P_t = \frac{\mathbb{E}_t \left\{ e^{(1-\gamma)\ln \tilde{x}_T} \right\} e^{-(T-t)r_f}}{\mathbb{E}_t \left\{ e^{-\gamma \ln \tilde{x}_T} \right\} e^{-(T-t)r_f}}
= \frac{e^{\alpha(u_{1-\gamma},t,T)+\beta(u_{1-\gamma},t,T)X_t}}{e^{\alpha(u_{-\gamma},t,T)+\beta(u_{-\gamma},t,T)X_t}} e^{-(T-t)r_f}
= x_t e^{-r_f(T-t)+\lambda_0(t,T)+\lambda_\sigma(t,T)\sigma_t^2}$$

(8)

where $\lambda_\sigma(t,T)$ is the second component of the vector $\beta(u_{1-\gamma}, t, T) - \beta(u_{-\gamma}, t, T)$ and $\lambda_0(t,T) = \alpha(u_{1-\gamma}, t, T) - \alpha(u_{-\gamma}, t, T)$. The stock return accordingly is

$$d\ln P_t = d\ln x_t + r_f dt + \frac{\partial \lambda_0(t,T)}{\partial t} dt + \frac{\partial \lambda_\sigma(t,T)}{\partial t} \sigma_t^2 dt + \lambda_\sigma d\sigma_t^2.$$

(9)

### 3.2.1 Infinite horizon limit

In order to avoid the effects of time lapsing, we consider the infinite horizon limit of our model. The equilibrium stock price is given by (see Appendix 5.2)

$$\frac{dP_t}{P_t} = r_f dt + \lambda_0(\sigma_t^2) dt + \sigma_t dB_t^x + \lambda_\sigma d\sigma_t^2 + (e^{\lambda_\sigma \xi_t} - 1 - \lambda_\sigma \xi_t) dN_t.$$

(10)
where

\[
\beta_2(u_{1-\gamma}) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_v^2(\gamma^2 - \gamma)}}{\sigma_v} 
\]

(11)

\[
\beta_2(u_{-\gamma}) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_v^2(\gamma^2 + \gamma)}}{\sigma_v^2} 
\]

(12)

\[
\lambda_0(\sigma_t^2) = -\kappa \theta \lambda_\sigma + (\kappa \lambda_\sigma + \gamma - \sigma_v^2 \lambda_\sigma \beta_2(u_{-\gamma}) ) \sigma_t^2 - l_0 (\varrho (\beta_2(u_{1-\gamma})) - \varrho (\beta_2(u_{-\gamma}))) 
\]

(13)

\[
\lambda_\sigma = \beta_2(u_{1-\gamma}) - \beta_2(u_{-\gamma}) < 0 
\]

(14)

and \(\lambda_0(\sigma_t^2)\) does not depend on \(\mu\). The unconditional equity premium in the economy defined as \(\lambda_0\) is

\[
\lambda_0 = \mathbb{E}(\lambda_0(\sigma_t^2))dt + \mathbb{E}((e^{\lambda_0\xi_t} - 1 - \lambda_\sigma \xi_t) dN_t) \\
= \lambda_0(\mathbb{E} \sigma_t^2)dt + l_0 (\varrho (\lambda_\sigma) - 1 - \lambda_\sigma \mathbb{E} \xi_t) dt 
\]

(15)

where \(\varrho(h) = E(e^{h\xi_t}) = 1/(1 - \mu \xi h)\) is the moment generating function of \(\xi\).

There are two immediate properties of \(\lambda_\sigma\) that are important. First, there is an upper limit on the amount of risk aversion, \(\gamma\) that produces a well defined equilibrium (such that \(\lambda_\sigma\) is real).

Second, when the equilibrium is well defined, \(\lambda_\sigma\) is negative for \(\gamma < 0\) and exactly zero for \(\gamma = 0\). The sign follows from the fact that \(\gamma^2 + \gamma > \gamma^2 - \gamma\). The fact that \(\lambda_\sigma\) is negative is important as we will discuss below.

### 3.3 Risk-Neutral Dynamics

**Proposition 1.** If equilibrium exists, the pricing kernel follows

\[
\frac{dM_t}{M_t} = -r_f dt - \gamma \sigma_t dB_t^\varepsilon - \eta \sigma_t dB_t^v + (e^{-\eta \xi_t} - 1) dN_t - l_0 (\varrho (-\eta) - 1) dt 
\]

(16)

Under the equivalent measure \(Q\), the variance process follows

\[
d\sigma_t^2 = \kappa Q(\theta Q - \sigma_t^2)dt + \sigma_v \sigma_t dB_t^{v,Q} + \xi_t^Q dN_t^Q 
\]

(17)
The risk neutral parameters are

\[ \kappa^Q = \kappa + \eta \sigma_v^2 \]  
\[ \theta^Q = \frac{\theta \kappa}{\kappa^Q} \]  
\[ l_0^Q = l_0 \varrho(-\eta) \]  
\[ \mu_\xi^Q = \mu_\xi \varrho(-\eta) \]  
\[ \eta = -\beta_2 (u_\gamma) < 0 \]

Under the risk-neutral measure, the stock price follows

\[ \frac{dP_t}{P_t} = r_f dt + \sigma_t dB_t^x + \lambda_\sigma \sigma_t dB_t^{\nu,Q} + \lambda_\theta \sigma_t \sqrt{\theta_t} dB_t^\theta,Q + (e^{\lambda_\sigma \xi_t^Q} - 1) dN_t^Q - l_0^Q (\theta^Q(\lambda_\sigma) - 1)dt \]  

A few notes on the functional form of the transformation of the probability measure are in order. First, note that it follows immediately from the expression for \( \eta \) that this parameter is negative for \( \gamma > 0 \). The market price of variance risk is \( \eta \sigma_v \). Accordingly, the model features a negative variance risk premium when the representative agent is risk averse. We explore this in some depth below in connection with variance futures (see Proposition 1). It is also clear from expressions (20) and (21) that jump arrivals and jump sizes are larger under the risk neutral than objective measure. Our model also contrasts reduced form models in that there is a tightly specified relationship between the risk-adjustments for jumps and diffusive risk premia.

The stock price follows

\[ d\ln P_t = r_f dt + (\lambda_0(\sigma_t^2) - \gamma \sigma_t^2)dt + \lambda_\sigma d\sigma_t^2 + \sigma_t dB_t^Q \]  

under the risk-neutral measure where \( d\sigma_t^2 \) is given by Equation (17).

Of essential interest is the \( \tau \) period ahead conditional variance of the log-return, which we can show to be a linear function of the spot variance \( \sigma_t^2 \) under both measures \( i = \{P, Q\} \), as

\[ \text{Var}_i^i(\ln P_{t+\tau}) = a_i(\tau) + b_i(\tau) \sigma_t^2. \]
Notice that VIX is defined as

\[ VIX_t = \text{Std}_t^Q(\ln P_{t+21}) = \sqrt{a_Q(21) + b_Q(21)\sigma_t^2} \]  

(26)

where we have defined a unit of time to be one business day.

### 3.4 Dynamic Equilibrium Effects

[Figure 4 about here.]

The parameter \( \lambda_\sigma \) is an important endogenous parameter, as it determines the impact of economic uncertainty on asset prices. \( \lambda_\sigma \) is negative and increasing (in absolute value) as a function of risk aversion, \( \gamma \), and also the inverse of the persistence parameter \( \kappa \). We show this in Figure 4 (right panel). The fact that \( \lambda_\sigma \) is negative and decreasing in \( \gamma \) has important consequences for the variance risk premium. From equation (19) - (21) we see that average volatility is higher under the risk neutral measure.

By taking unconditional expectations on both sides of Equation(6) we see that \( Ed\ln P/dt - r = \lambda_0 \) is the unconditional equity premium. The left plot in Figure (4) shows that the equity premium is increasing in \( \gamma \) and reaches realistically large values for values of \( \gamma \) that are plausibly small, suggesting a resolution the equity premium puzzle. The equity premium is larger when persistence in volatility is high (\( \kappa \) close to zero). This is analogous to long-run-risk models. In fact, the model can be seen as a limiting case of long-run-risk when intertemporal elasticity of substitution is high.

[Figure 5 about here.]

It is clear from equation (10) that \( \lambda_\sigma \) determines the endogenous impact of volatility shocks on stock prices. Since \( \lambda_\sigma \) increases (in absolute value) in \( \gamma \), the correlation between volatility shocks and stock prices is negative (since \( \lambda_\sigma \) is negative) and increasing in absolute value as \( \gamma \) increases. We illustrate this through scatter plots of daily volatility changes and daily returns simulated...
from the model using $\gamma = 3$ and $\gamma = 8$ in Figure 5. The correlations are $-0.26$ and $-0.63$ respectively.

The equilibrium impact of shocks to volatility on stock prices, as captured through $\lambda_\sigma$, is an important distinguishing feature of our model relative to standard reduced form models. For example, standard models such as Heston’s 1993 famous model for pricing options specify an exogenous correlation between shocks to volatility and prices. In Heston’s model this correlation is not tied to risk aversion, and the correlation can be large or small independent of volatility risk premium and equity premium. In our model, risk-premia for all assets increase monotonically as a function of risk aversion, through $\lambda_\sigma$. This is seen clearly from the expressions (19) - (21). The expressions restrict volatility risk premium for diffusive shocks as well as jumps directly as a function of $\eta$. This contrasts the reduced form literature where there is no explicit link between jump and diffusive risk premia.

Our model also leaves no room for volatility shocks that are not priced. Reduced form models, such as the SVIJ model estimated in Eraker, Johannes, and Polson (2003) are not nested since in that model jumps in volatility can occur without price impact. Importantly, equilibrium implied not only that volatility shocks are priced, but that the relative impact of small (i.e. Brownian) and large (jumps) shocks to volatility load with a factor $\lambda_\sigma$ on stock returns. Our model also rule out the conditional CAPM as that model does not feature priced volatility shocks (see Eraker and Wang (2011) for a discussion of how the conditional CAPM cannot be reconciled with dynamic present value computation).

Figure 6 illustrates some of the key pricing implications through a comparison of the $P$ and $Q$ measures for the stock price. As can be seen, the difference between the $P$ and $Q$ densities increases as $\gamma$ is increasing. The plot also shows that the two different values of $\gamma$ give widely different Black-Scholes implied volatilities for the underlying stock which suggests that the model can generate a substantial volatility risk premium. This is of course key in explaining the negative returns to VIX futures.
### 3.5 Variance Futures

The non-linearity introduced by the square root in (26) necessitates numerical computation of futures prices. While we can do this through a single one-dimensional numerical integration, for purposes of illustration, we discuss the shape of a hypothetical futures contract written on VIX squared. These can be computed directly, and since the squared VIX is a linear function of $\sigma_t^2$, it inherits the properties of the risk premia embedded into the differences between the objective $P$ and risk-neutral $Q$ probability measures. These hypothetical variance futures are priced

\[
F_t^{VIX^2}(\tau) = E_t^Q \{ VIX^2_{t+\tau} \} = E_t^Q \{ a_Q(21) + b_Q(21)\sigma_{t+\tau}^2 \} \\
= a_Q(21) + b_Q(21) E_t^Q \{ \sigma_{t+\tau}^2 \} 
\]  

(27)

In steady-state, the futures curve is

\[
F_t^{VIX^2}(\tau) = a_Q(21) + b_Q(21) E^Q \{ \sigma_{t+\tau}^2 | \sigma_t^2 = E(\sigma_t^2) \} 
\]  

(28)

The following basic facts about VIX-squared futures are easy to verify:

**Proposition 2.** If $\gamma > 0$, then the volatility risk premium is negative and

1. The futures curve is upward sloping (contango) in steady state.

2. Long positions in VIX squared futures earn negative expected returns irrespective of the state of the economy.

3. VIX square futures have negative market betas. The size of the beta is a function of risk aversion.

It is worth commenting on some of these features. Some believe that if the futures curve is upward sloping, they will capture a negative roll by purchasing the futures. This is true only in relation to the physical drift rate of the underlying VIX or $\sigma_t^2$ processes. For example, if the underlying $\sigma_t^2$ process is in steady state, ($\sigma_t^2 = E(\sigma_t^2)$), the negative roll is indicative of the
actual expected returns because the expected terminal value of the VIX squared is its present value $E_t(VIX_{t+\tau}^2) = VIX_t^2$. A positive slope of the futures curve reflects only risk premium in steady state. The larger the risk aversion in our model, the steeper the slope, and the higher (lower) the risk premium earned by short (long) positions. It is important to understand that the expected return to a long position in VIX futures is determined not by the shape of the futures curve per se, but rather by the shape of the curve relative to the expected value of the VIX at expiration of the futures. Accordingly, in principle, if the futures curve were upward sloping but the objective measure drift was greater than that implied by the futures curve, the expected return to a long VIX futures position would have been positive. This, however, is never true in our model economy and investors earn a strictly negative premium. 1

3.6 VIX Futures

How do the actual VIX futures prices differ from the hypothetical VIX squared futures prices we analyzed in the preceding section? To answer this question, Figure 7 shows the various term structures under different assumptions about initial volatility states, $\sigma_t$. We plot the term structure for the pure diffusive (SV) model as well as for the jump model (SVVJ). We compare two things: the actual, model-implied futures curve labeled $SV_Q$ and $SVVJ_Q$ (signifying expectation under $Q$), and also the corresponding objective measure expectation, labeled $SV_P$ and $SVVJ_P$, respectively. The difference between the $Q$ and the $P$ expectations is due to volatility risk premium. The right hand figures show the expected holding period return, $E_t^P(VIX_{t+\tau})/E_t^Q(VIX_{t+\tau}) - 1$, to a long futures position. The expected returns are negative and more sharply so when volatility is high.

Examining the plots to the left in Figure 7, we see the obvious relations between spot volatility and the shape of the futures curve: when spot volatility is high (low) the curve is in backwardation.

---

1The variance risk premium is a function of the difference in drift rate for $\sigma_t^2$ under the two measures. It is easy to see that the $Q$ minus $P$ drift $-\lambda_\sigma \sigma_t \sigma_t^2$ is a positive number regardless the level of $\sigma_t^2$. $\kappa_Q < \kappa$ and $\theta_Q > \theta$ suggest that the physical mean reversion rate is faster than the risk-neutral and the physical mean reversion level is lower, which gives a negative risk premium for VIX long futures.
(contango). When in steady state the futures curve is first convex and then concave. The convexity created at the short end of the curve is due to a Jensen’s inequality term created by the concavity of the square root function.\(^2\) The total expected holding period returns are negative for all maturities and for all initial values of spot volatility. The jump model (SVVJ) provides a larger short-term risk premium (in absolute value) as can be seen by the steeper slope of the expected return graphs. We expect, therefore, that the jump model will generate a higher average return to short term VIX futures.

4 Empirical Analysis

We estimate the model using structural likelihood-based estimation. The estimation approach is similar to that of Eraker, Johannes, and Polson (2003) (hereby EJP) in that we draw the parameters of the model from the posterior distribution \(\Theta_j \sim p(\Theta | Y_T)\) using MCMC sampling. The latent conditional variances, \(\sigma^2_t\), the jump times, \(dN_t\), and jump sizes, \(\xi_t\), are drawn from the respective conditional posterior distributions in a manner similar to that of EJP. The main difference is that our structural model leads to non-standard posterior distributions. This contrasts EJP and similar papers that draw from conjugate densities such as Normal and Gamma. Since our structural model gives rise to complicated non-linear relationships between the “deep” parameters and the affine model coefficients such as \(\lambda_\sigma\), it is not possible to draw any of the parameters directly from conditional distributions. We therefore draw the entire parameter vector directly from the posterior using a Metropolis Hastings draw.

4.1 MCMC Estimation

Bayesian inference for stochastic volatility models have become rather routine and a substantial literature is devoted to developing algorithms for sampling from the posterior distributions of these models. This literature was initiated by Jacquier, Polson, and Rossi (1994), and has

\(^2\)The Jensen’s inequality term is \(E_t(VIX_{t+\tau}) - \sqrt{E_t(VIX^2_{t+\tau})} < 0\) since the square root function is concave.
since seen numerous refinements (see for example Kim, Shephard, and Chib (1998) and references therein).

The mechanics of our MCMC is straightforward. Let $\xi_t$ denote jump sizes and $dN_t$ a jump time indicator and define $\Sigma = \{\sigma_t\}_0^T$, $\Xi = \{\xi_t\}_0^T$, and $N = \{dN_t\}_0^T$.

We construct an algorithm that samples $(\Theta, \sigma_0:T, \xi_0:T, N_0:T)$ conditional upon the observed returns data $\mathcal{R}_T := \{r_t\}_0^T$ only. By Monte-Carlo sampling from the joint posterior $p(\Theta, \Sigma, \Xi, N | \mathcal{R}_T)$, we also implicitly sample from the marginal posterior distributions (see Tanner and Wong (1987)).

Our method for carrying out MCMC sampling follows the general recipe of Eraker, Johannes, and Polson (2003), but differ in some important ways as the conjugacy offered by the reduced form models in that paper is lost in our structural model. For example, in EJP virtually all the model parameters have known conditional posterior distributions which follows from conditioning on the simulated values of volatilities, jump-sizes and times which leads to conditionally gaussian errors associated with the Brownian increments. For our model, conjugacy is lost as the $\lambda$-parameters that determine the equilibrium unconditional return and volatility-feedback effect are non-linear functions of the statistical parameters that determine the dynamics of volatility, and importantly, risk-aversion. For this reason, the entire parameter simulation step will have to be done using Metropolis-Hastings.

4.2 Estimation Results

[Table 6 about here.]

Table 6 reports structural estimates of the parameters in the model. All parameter estimates should be interpreted to be based on a unit of time being one day. This makes the time-series parameters directly comparable to estimates based on daily returns. Depending on whether the model includes jumps or not, $\kappa$ is estimated to be 0.014 and 0.0091, respectively. These estimates imply daily autocorrelatons of $\exp(-0.014) = 0.9861$ and $\exp(-0.0091) = 0.991$. These numbers are broadly consistent with estimates reported elsewhere (see Singleton (2006) for a review).
The most interesting parameter in Table 6 is the risk-aversion parameter, $\gamma$. We estimate this to be more than 10 posterior standard deviations away from zero for the SVVJ jump model. This contrasts conflicting evidence from the conditional CAPM literature where the risk aversion parameters are typically found to be statistically insignificant.

To obtain estimates of the model’s expected return, standard deviation, skewness, and kurtosis, we simulate data from the underlying equilibrium stock price process using the estimated parameters in Table 6. We simulate 10,000 data sets of length $T = 1816$, the length of our futures time-series and use these to compute theoretical futures prices. We then compute the returns to rolled futures positions similarly to our procedure for the real data. We compute sample moments from the simulated data and compare these to the real data.

Table 9 reports the results. We include the first three estimators of the average returns from the data which we previously reported in Table 1 for convenience. Note that the higher order moments are of logarithmic daily returns. Although slightly higher, the returns produced by the SVVJ model are quite close to what we see in the data for the one-month maturity contracts. For the longer maturities both models overstate the size of the negative return premium. Overall however, both models do surprisingly well in fitting average returns. The models also do seemingly well in fitting the higher order moments, with the exception of the SVVJ model’s too high kurtosis. In general, the simulated model returns have moments in the ballpark of that seen in the real data. We further test the null hypothesis of zero difference between the model and the data below.

To carry out these significance tests, we use a model-based bootstrap. We wish to avoid the use of test-statistics based on asymptotic normality because the higher order moments are very
non-normally distributed in small samples. The evidence is presented in Figures 9 and 8. The figures plot the kernel-smoothed densities of moments implied by the model and the corresponding data moment represented by the vertical bars. The non-parametric density estimates should be interpreted as a model-based bootstrap of the sampling distribution for each respective moment.

In examining Figures 9 and 8 we see that the average returns for the SV model are insignificantly different between the real data and the model generated data. For the standard deviations and higher order moments, the results are very different. The SV model basically fails to fit any moment higher than the first. The standard deviations are all significantly different between the model and the data. The skewness and kurtosis are so far off that the vertical bars are absent from the plots.

On the other hand, for the SVVJ model, the moments are surprisingly well matched. With the exception of the standard deviation for the four and five month maturity contracts, the data and the model are insignificantly different. From Table 9 we see that the SVVJ model generates sample kurtosis coefficients that range from 20 to 27 which compares to coefficients that range from 6.02 to 7.28 in the data (see Table 1). This may seem like a large difference, however, Figure 9 reveals that the sampling distribution under the model has most of its mass below what we find in the data. In fact, the medians of the small sample distributions for the kurtosis coefficients under the SVVJ model range from 4.4 to 5. Thus, while it may seem from Table 1 and 9 that the SVVJ model generates too high kurtosis, the differences are in fact almost non-existent.

4.3 A Two-Factor Volatility Model

As we can see in the previous section, the SVVJ model captures all but one aspect of the observed futures returns - it significantly underestimates the variability of longer term contracts. Since the SVVJ model is a single factor affine volatility model, it cannot generate long-run memory like behavior in conditional volatility, as empirically documented by Bollerslev and Mikkelsen (1996) among others. Bates (2000) and Chernov, Gallant, Ghysels, and Tauchen

---

3This happens when the real data moments are outside of the empirical support of the sampling distributions under the model.
(2003) propose two factor nested conditional variance specification in a no-arbitrage, reduced form model. We propose a similar model,

\[
\begin{align*}
\frac{dx_t}{x_t} & = \mu dt + \sigma_t dB^x_t \\
\sigma_t^2 & = \kappa(\theta_t - \sigma_t^2)dt + \sigma_t \sigma_v dB^v_t + \xi_t dN_t \\
\theta_t & = \kappa(\theta - \theta_t)dt + \sigma_\theta \sqrt{\theta_t} dB^\theta_t \\
\xi_t & \sim \text{Exp}(\mu_\xi) \\
N_t & \sim \text{Poisson}(l_0 t)
\end{align*}
\]

\[\text{Corr}(dB^i_t, dB^j_t) = 0, \; i,j \in \{x,v,\theta\} \text{ and } i \neq j\] (34)

In this specification the conditional variance, \(\sigma_t^2\), mean-reverts to a stochastic mean, \(\theta_t\), which again follows a square root process. If we assume that the persistence in \(\theta_t\) is stronger than for \(\sigma_t^2\) (i.e., \(\kappa_\theta\) is “small”), then \(\theta_t\) will generate low frequency movements in volatilities while \(\sigma_t\) accounts for higher frequency movements.

The equilibrium stock price can now be seen to be given by

\[d \ln P_t = r_{f,t} dt + \lambda_0(\sigma_t^2, \theta_t) dt + \lambda_\sigma d\sigma_t^2 + \lambda_\theta d\theta_t + \sigma_t dB^x_t\] (35)

where \(\lambda_j, j = \{\sigma, \theta\}\) are equilibrium coefficients which again are nonlinear functions of the parameters. The squared VIX index is again a linear function of the state-variables

\[VIX_t^2 = \text{Var}_Q^2(\ln P_{t+21}) = a + b\sigma_t^2 + c\theta_t\] (36)

where we can solve for constants \(a, b\) and \(c\) (see Appendix). Since \(VIX_t^2\) is a function of two processes with different autocorrelation functions, the autocorrelation for the \(VIX_t^2\) itself is a mixture, and thus displays long memory-like behavior.

[Figure 10 about here.]

It’s difficult to take our two-factor volatility model and estimate it using return data alone, as we did with the models in the previous section. The purpose of our exercise here is to demonstrate
that the model is capable of matching the moments of the futures returns data. In Figure 10 we show the sampling distributions for the VIX futures data under the model, using calibrated parameters. The parameters are $\kappa = 0.0165$, $\sigma_v = 0.0012$, $\theta \times 10000 = 0.5416$, $\gamma = 7.6275$, $\lambda_0 = 0.003$, $\mu_v \times 10000 = 1.016$, $\kappa_\theta = 0.005$, $\sigma_\theta \times 10000 = 2.175$. As can be seen, the resulting moments are all well inside the tails of the sampling distributions suggesting that our two factor volatility specification provides a plausible description of the true data generating process. Note that the standard deviation of the longer maturity contracts is matched almost exactly.

### 4.4 Variance Swaps

In this section we consider our models’ ability to explain the returns to variance swaps. A variance swap is a zero-cost contract that pays the difference between an agreed upon price, called the swap rate, $VS$, and physical, realized variance over some time period, say $t, ..., T$. The realized variance leg of the swap is simply the sum of squared daily log-returns\footnote{Note that this convention ignores the squared expected return so that strictly speaking it is not a variance.},

$$RV_{t,T} = \sum_{s=t+1}^{T} \ln R_s^2. \quad (37)$$

The variance swap rate is set such that investors are indifferent between the fixed and floating rate legs of the swap. This implies that

$$VS_{t,T} = \mathbb{E}_t^Q(RV_{t,T}). \quad (38)$$

Variance Swaps trade in the OTC market. There is no public source for the proprietary prices from these OTC transactions. Fortunately, eqn. (38) suggests that any measure of $Q$ variance is an estimator for the variance swap rate. A number of recent papers including (citations here) use S&P 500 option implied variance. A variety of possible estimators of $Q$-variance, including the VIX method can be applied to estimate the variance swap rate. (38). We apply the VIX method with the exception that we ignore the
Given some option implied VS rate, we compute one month holding period returns

\[ R_{t,t+22} = \frac{RV_{t,t+22} + VS_{t+22,T}}{VS_{t,T}} - 1 \] (39)

Note that this formula shows that the payoff, \( RV_{t,t+22} + VS_{t+22,T} \), contains the realized variance component, whereas the returns to VIX futures do not. It’s possible therefore for the returns to the two asset classes to differ markedly even if the theoretical prices are strongly related.

With this in mind, the existing literature on variance swap returns suggests some puzzling features relative to our evidence from the VIX futures markets. In particular, there are several papers that report empirical findings that suggest that short maturity variance swaps exhibit very large negative returns, while long maturity ones have returns close to zero, or even positive.

[Table 8 about here.]

Table 8 presents various return estimates from the extant literature, along with our own estimates. We include these five different data sources in order to convey that the returns are largely in the same ballpark. Three of the data-sets include the financial crisis (DGLR, Johnson, and our own data). Those datasets produce returns that range from -17 to -25% for one-month maturities, and close to zero returns for the longest horizons. The Bloomberg data yield negative returns across the entire maturity spectrum, but the sampling period is very particular as the data starts in October of 2008 - the height of the financial crisis. The data from Egloff, Leippold, and Wu (2010) also produces negative returns across all maturities, but this sample stops before the financial crisis, in 2007.

[Table 9 about here.]

Table 9 presents the moments of the variance swap returns from our sample data alongside the model implied returns. At first glance it would appear that our models do a poor job. In particular, the models are incapable or matching the large negative return for the one-month variance swaps. For example, the SVVJ model generates an average log-return of -22% which
compares to -43% in the data. At the longer maturities the discrepancies are less significant. These findings echo the results in Dew-Becker, Giglio, Le, and Rodriguez (2015) who show that the Drechsler and Yaron (2011) model cannot match the large negative returns to one-month claims while at the same time delivering too negative returns for longer maturity contracts.

The results in Table 9 are puzzling given the positive results on behalf of our model in matching the VIX futures returns. To understand why our model fails to explain one market while almost perfectly explaining another, closely related market, we note two facts. First, payoffs to variance swaps depend on physical, realized variance, whereas VIX futures do not. This means that if somehow the VIX itself is inflated by some constant across time relative to physical variance, this would show up in variance swap returns but not necessarily in VIX futures returns. Second, our variance swap rates are in fact not market rates, but rather swap rates that are computed from the midpoint of the bid and asks from S&P 500 options prices.

We examine whether the use of midpoints between the bids and the asks for the underlying options affects our return computations next. Figure 11 plots the average Variance Swap term structure in our sample data. The yellow line represents the midpoint of the bid and the ask, which is what the CBOE uses to compute the VIX index, while the blue and red lines represent the term structures computed from bids and asks. There is a dramatic difference between the rates computed from bids and asks. At the short end (1M) the difference is 17% (in variance units) while for a 12M contract it is about 10%.

Table 10 presents two pieces of evidence relating the bid ask spreads in the underlying options to the variance swap returns. In Panel A we simply compute the variance swap rates from bids, mid points, and asking prices. The results show that indeed, the returns are higher when computed from bids than for asking prices. In Panel B we examine the returns to a “price taker” who will initiate a trade at either bids or asks, hold the position for one month, and then liquidate
the position at the other side of the market. We refer to a buyer as a trader who is purchasing
the variance swap at the asking price and liquidating at the bid one month later, and vise versa.

A seller realizes a positive return for a one month swap, and negative returns for any other
maturity contract. This is so because the bid-ask are so large that by adding the spread to the
cost of making the roundtrip trade, both long and short positions in long maturity contracts are
unprofitable.

5 Concluding Remarks

In this paper we show that the average returns earned on VIX futures and ETNs are abysmal.
Investors who hold long positions in VIX futures will gradually see their wealth vanish as they
realize negative average rates of returns over the long run. We show that this is a statistically
significant feature of the data.

We argue that the negative returns are consistent with equilibrium. Though the size of
the negative return premium is not consistent with a traditional CAPM, which delivers a risk
premium of -15% per year, we show that the equilibrium model of Eraker and Wang (2011) is
capable of explaining the -36% per year average return to a one-month (rolled) futures position.
The SVVJ model, which includes volatility jumps, is our preferred one-factor model. This model
generates a higher short-term volatility risk premium than does the simpler SV model.

It is important to understand the mechanisms that cause these models to assign high volatility
risk premia and negative market betas to the VIX futures. The stock market pays a single
terminal cash flow such that its current market value is the present value of the cash flow,
essentially discounted at an expected rate of return which is proportional to the current volatility,
\( \sigma_t \). The sensitivity of the expected rate of return with respect to changes in volatility is an
increasing function of risk aversion, \( \gamma \), in the model. Thus, in equilibrium, positive (negative)
shocks to \( \sigma_t \) give rise to negative (positive) stock price shocks. The absolute magnitude of this
negative return/volatility correlation is an endogenous quantity that increases (in absolute value)
with \( \gamma \). If we then think about volatility as an asset class, relative to the CAPM, volatility, as
measured by the VIX, is a negative beta asset. Since VIX futures prices are positively dependent on current spot VIX, they too are negative beta assets.

The overall size of the volatility risk premium in our model depends on risk-aversion, volatility persistence, and the specification of volatility jumps in the model. Investors dislike volatility jumps and large jumps lead to discontinuities in stock prices reminiscent of financial crisis. Overall, this leads to a higher volatility risk premium in the short term. Expected returns, accordingly, are a steeper function of maturity under the SVVJ model than the SV model. The SV model comparably generates a higher volatility risk premium by assigning a slightly lower speed of volatility mean reversion. This leads to a somewhat higher negative return premium for long term futures contracts. It also leads to more volatility in futures returns at longer maturities (i.e. five months), in a manner that is consistent with the real market data.

Appendix

5.1 Solving the SDE system

In our most general model the process $x$ that generates the terminal (log) payoff $x_T = \tilde{x}_T$ to the aggregate stock market is given by

$$\frac{dx_t}{x_t} = \mu dt + \sigma_t dB_t^x$$
(40)

$$d\sigma_t^2 = \kappa (\theta - \sigma_t^2) dt + \sigma_v \sigma_t dB_t^v + \xi_t dN_t$$
(41)

$$d\theta_t = \kappa_\theta (\theta - \theta_t) dt + \sigma_\theta \sqrt{\theta_t} dB_t^\theta$$
(42)

$$\xi_t \sim \text{Exp}(\mu_\xi), \quad f(\xi_t) = \frac{1}{\mu_\xi} e^{-\frac{1}{\mu_\xi} \xi_t} \text{ so the expected jump size is } \mu_\xi$$
(43)

$$N_t \sim \text{Poisson}(l_0 t)$$
(44)

$$\text{Corr}(dB_t^i, dB_t^j) = 0, \quad i, j \in \{x, v, \theta\} \text{ and } i \neq j$$
(45)
The state variable $X_t$ in this economy defined as $(\ln x_t, \sigma_t^2, \theta_t)$ and follows an affine process equivalent to the following

$$dX_t = (K_0 + K_1 X_t)dt + \Sigma(X_t)dB_t + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \xi_t dN_t$$  \quad (46)

$K_0 = \begin{pmatrix} \mu & 0 & \kappa \theta \\ 0 & -\frac{1}{2} & 0 \\ 0 & -\kappa & \kappa \end{pmatrix}$  \quad (47)

$K_1 = \begin{pmatrix} 0 & 0 & -\kappa \theta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  \quad (48)

$\sigma(X_t)\sigma(X_t)' = H_0 + \Sigma_{i=1}^3 H_{(i)} X_t(i)$  \quad (49)

$H_0 = H_{(1)} = 0_{3 \times 1}$  \quad (50)

$H_{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  \quad (51)

$H_{(3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{pmatrix}$  \quad (52)

Therefore, $E_t e^{u'X_t} = e^{\alpha(u,T,T) + \beta'(u,T,T)X_t}$, where $\alpha$ and $\beta$ solve a system of ordinary differential equations (see Duffie, Pan and Singleton (2000)), given by

$$\frac{\partial \alpha}{\partial t} = -K_0'\beta - \ell_0(\varrho(\beta_2) - 1)$$  \quad (53)

$$\frac{\partial \beta}{\partial t} = -K_1'\beta - \frac{1}{2} \beta' H \beta$$  \quad (54)

$\alpha(u,T,T) = 0$  \quad (55)

$\beta(u,T,T) = u$  \quad (56)

$\varrho(\beta_2) = E e^{\beta_2 \xi} = \frac{1}{1 - \beta_2 \mu \xi}$  \quad (57)

The analytical solutions for $\beta_1$ and $\beta_2$ can be obtained as

$$\beta_1(u,t,T) = u(1),$$  \quad (58)

$$\beta_2(u,t,T) = a_2(u) + (a_2(u) - a_1(u)) \frac{C(t,u)}{1 - C(t,u)},$$  \quad (59)
while,

\[
a_2(u) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_u^2(u(1)^2 - u(1))}}{\sigma_u^2},
\]

\[
a_1(u) = \frac{\kappa + \sqrt{\kappa^2 - \sigma_u^2(u(1)^2 - u(1))}}{\sigma_u^2},
\]

\[
c_1(u) = \frac{u(2) - a_2(u)}{u(2) - a_1(u)},
\]

\[
c_2(u) = \frac{(a_2(u) - a_1(u))\sigma_u^2}{2},
\]

\[
C(t, u) = c_1 e^{c_2(u)(T-t)}.
\]

We solve \(\beta_3(u, t, T)\) and \(\alpha(u, t, T)\) numerically. However, when \(T \to \infty\), the analytical solution for \(\beta_3(u)\) is

\[
\beta_3(u) = \frac{\kappa \theta - \sqrt{\kappa^2 \theta - 2\kappa a_2(u)\sigma_u^2}}{\sigma_u^2}. 
\]

In the case of \(T \to \infty\), solution for \(\beta_1(u)\) is the same and \(\beta_2(u) = a_2(u)\).

### 5.2 Equilibrium Stock Returns

Let \(P_t\) denote the price of the risky asset at \(t\) and \(a\) be the number of shares the representative agent holds, who has power utility and enjoys consumption at the final date. The equilibrium asset price can be derived by optimal portfolio problem:

\[
\max_a E_t u(a \tilde{x}_T - (a - 1) P_t e^{r_f(T-t)})
\]

From the first order condition and the fixed supply \(a^* = 1\) of the risky asset we we find that its price is

\[
P_t = \frac{E_t u'(\tilde{x}_T)\tilde{x}_T e^{-r_f T}}{E_t u'(\tilde{x}_T) e^{-r_f T}} = \frac{E_t \tilde{x}_T^{1-\gamma} e^{-r_f T}}{E_t \tilde{x}_T^{1-\gamma} e^{-r_f T}}
\]

\[
= \exp\{\alpha(u_{1-\gamma}, t, T) - \alpha(u_{-\gamma}, t, T) + (\beta'(u_{1-\gamma}, t, T) - \beta'(u_{-\gamma}, t, T))X_t - r_f(T-t)\}
\]

\[
\ln P_t = \alpha(u_{1-\gamma}, t, T) - \alpha(u_{-\gamma}, t, T) - r_f(T-t) + \ln x_t
\]

\[
+ (\beta_2(u_{1-\gamma}, t, T) - \beta_2(u_{-\gamma}, t, T))\sigma_t^2
\]

\[
+ (\beta_3(u_{1-\gamma}, t, T) - \beta_3(u_{-\gamma}, t, T))\theta_t
\]

\[
u_{1-\gamma} = \begin{pmatrix} 1 - \gamma & 0 & 0 \end{pmatrix}^\prime, \quad v_{-\gamma} = \begin{pmatrix} -\gamma & 0 & 0 \end{pmatrix}^\prime
\]
If we take the limit $T \to \infty$, then $\beta$ does not depend on $t$. Define

\[
\lambda_\sigma = \beta_2(u_{1-\gamma}) - \beta_2(u_{-\gamma}),
\]
\[
\lambda_\theta = \beta_3(u_{1-\gamma}) - \beta_3(u_{-\gamma})
\]

(70) (71)

Therefore, the dynamics of the stock price is

\[
d \ln P_t = \left( \frac{\partial \alpha(u_{1-\gamma}, t, T)}{\partial t} - \frac{\partial \alpha(u_{-\gamma}, t, T)}{\partial t} \right) dt + r_f dt + d \ln x_t
\]
\[
+ \lambda_\sigma d \sigma_t^2 + \left( \frac{\partial \beta_2(u_{1-\gamma}, t, T)}{\partial t} - \frac{\partial \beta_2(u_{-\gamma}, t, T)}{\partial t} \right) \sigma_t^2 dt
\]
\[
+ \lambda_\theta d \theta_t + \left( \frac{\partial \beta_3(u_{1-\gamma}, t, T)}{\partial t} - \frac{\partial \beta_3(u_{-\gamma}, t, T)}{\partial t} \right) \theta_t dt
\]
\[
= \left( -\mu - \kappa \theta \lambda_\theta - l_0 (\varphi (\beta_2 (u_{1-\gamma})) - \varphi (\beta_2 (u_{-\gamma}))) \right) dt + r_f dt + \left( \mu - \frac{1}{2} \sigma_t^2 \right) dt + \sigma_t dB_t^x
\]
\[
+ \lambda_\sigma d \sigma_t^2 + \left( \kappa \lambda_\sigma + \gamma - \frac{1}{2} \sigma_v^2 \left( \beta_2^2 (u_{1-\gamma}) - \beta_2^2 (u_{-\gamma}) \right) \right) \sigma_t^2 dt
\]
\[
+ \lambda_\theta d \theta_t + \left( -\kappa \lambda_\sigma + \kappa \theta \lambda_\theta - \frac{1}{2} \sigma_\theta^2 \left( \beta_3^2 (u_{1-\gamma}) - \beta_3^2 (u_{-\gamma}) \right) \right) \theta_t dt
\]
\[
= r_f dt + l_0 \left( \varphi (\beta_2 (u_{1-\gamma})) - \varphi (\beta_2 (u_{-\gamma})) \right) dt + \lambda_\sigma d \sigma_t^2 + \lambda_\theta d \theta_t
\]

(72)

Note that $\mu$ cancels in the expression for the expected return. To further write the dynamics of $\ln P_t$ in terms of the underlying shocks, we plug in the dynamics of the state variables to the above equation. A lot of items cancel out and we have

\[
d \ln P_t = r_f dt + \sigma_t dB_t^x + \lambda_\sigma \sigma_t d B_t^x + \lambda_\theta \sigma_t \sqrt{\theta_t} dB_t^\theta
\]
\[
+ \left( \gamma - \frac{1}{2} - \frac{1}{2} \sigma_v^2 \lambda_\sigma (\beta_2 (u_{1-\gamma}) + \beta_2 (u_{-\gamma})) \right) \sigma_t^2 dt
\]
\[
- \frac{1}{2} \sigma_\theta^2 \lambda_\theta (\beta_3 (u_{1-\gamma}) + \beta_3 (u_{-\gamma})) \theta_t dt
\]
\[
- l_0 \left( \varphi (\beta_2 (u_{1-\gamma})) - \varphi (\beta_2 (u_{-\gamma})) \right) dt + \lambda_\sigma \xi_t dN_t
\]

(73)
5.3 Proof of Proposition 1

From section 5.2, we know
\[
P_t = \frac{E_t \left( X_T^{1-\gamma} e^{-r_f T} \right)}{E_t \left( X_T^{\gamma} e^{-r_f t} \right)}
\]
\[
= E_t \left( \frac{X_T^{1-\gamma} e^{-r_f T}}{X_T^{\gamma} e^{-r_f t}} \right)
\]
\[
= E_t \left( \frac{E_{t+\tau} \left( X_T^{1-\gamma} e^{-r_f T} \right)}{E_{t+\tau} \left( X_T^{\gamma} e^{-r_f t} \right)} \right)
\]
\[
= E_t \left( \frac{E_{t+\tau} \left( X_T^{1-\gamma} e^{-r_f (t+\tau)} \right)}{E_{t+\tau} \left( X_T^{\gamma} e^{-r_f (t+\tau)} \right)} \right)
\]
\[
= E_t \left( \frac{M_{t+\tau} P_{t+\tau}}{M_t P_t} \right)
\]
(74)

Therefore it is easy to see that the stochastic discount factor is
\[
M_t = E_t \left( x_T^{1-\gamma} e^{-r_f t} \right) = e^{\alpha (u_-, t, T) + \beta' (u_-, t, T) X_t - r_f t}
\]
(75)

Follow the similar process for the dynamics of the stock price in section 5.2, we can derive the dynamics of the pricing kernel
\[
d \ln M_t = -r_f dt - \left( \frac{1}{2} \gamma^2 + \frac{1}{2} \sigma_e^2 \beta_2^2 (u_-) \right) \sigma_t^2 dt - \frac{1}{2} \sigma_o^2 \beta_3^2 (u_-) \theta_t dt
\]
\[
- \gamma \sigma_i d B^x_t + \beta_2 (u_-) \sigma_o d B^y_t + \beta_3 (u_-) \sigma_y \sqrt{\theta_t} dB^q_t
\]
\[
- l_0 (q (\beta (u_-)) - 1) dt + \beta_2 (u_-) \xi d N_t
\]
(76)
\[
\Gamma := (\gamma, -\beta_2 (u_-), -\beta_3 (u_-), \gamma, -\beta_2 (u_-), -\beta_3 (u_-))
\]
(77)

while \( \eta \) in the one factor model corresponds to \( \Gamma (2) \).

Define \( d \ln M_t^- \) as all the increment for \( \ln M_t \) except for the jump term, it is easy to see
\[
\frac{dM_t^-}{M_t^-} = -r_f dt - l_0 (q (-\Gamma (2)) - 1) dt - \Gamma' \sigma (X_t) dB_t
\]
(78)

and from the definition of \( d \ln M_t^- \), we know \( \ln M_t - \ln M_t^- = \Gamma (2) \xi d N_t \), therefore
\[
\frac{M_t - M_t^-}{M_t^-} = e^{-\Gamma (2) \xi d N_t} - 1 = \left( e^{-\Gamma (2) \xi} - 1 \right) d N_t
\]
(79)
Adding the above two equation together, we get

\[
\frac{dM_t}{M_t} = -r_f dt - l_0 (\rho (-\Gamma(2)) - 1) dt + \left(e^{-\Gamma(2)\xi_t} - 1\right) dN_t - \Gamma' \sigma(X_t) dB_t
\] (80)

It is easy to see that \(E \frac{dM_t}{M_t} = -r_f dt\).

Using Theorem 2.1 in Eraker and Shaliastovich (2008), the dynamics of diffusions and state variables under the equivalent measure \(Q\) are

\[
dB_t = dB_t^Q - \sigma(X_t) \Gamma dt
\] (81)

\[
dX_t = (K_0^Q + K_1^Q X_t) dt + \Sigma(X_t) dB_t^Q + \xi_t^Q dN_t^Q
\] (82)

\[
K_0^Q = K_0
\] (83)

\[
K_1^Q = K_1 - H \Gamma = \begin{pmatrix} 0 & -\frac{1}{2} - \Gamma(1) & 0 \\ 0 & -\kappa - \Gamma(2) \sigma_v^2 & \kappa \\ 0 & 0 & -\kappa_\theta - \Gamma(3) \sigma_\theta^2 \end{pmatrix}
\] (84)

\[
l_0^Q = l_0 \rho (-\Gamma(2))
\] (85)

\[
\mu_\xi^Q = \mu_\xi \rho (-\Gamma(2))
\] (86)

Under the \(Q\) measure, state variables still follow exponential affine and thus expectations of exponential affine functions of the states can be derived semi-analytically (up to ODE’s) similarly as under the objective measure.

Plug in the diffusions under \(Q\) in equation 81 to Equation 73, we get the the dynamics of the stock price under \(Q\) as

\[
d\ln P_t = r_f dt + \sigma_t dB_t^{x,Q} + \lambda_\sigma \sigma_t \sigma_t dB_t^{v,Q} + \lambda_\theta \sigma_\theta \sqrt{\theta_t} dB_t^{\theta,Q} + \left(- \frac{1}{2} - \frac{1}{2} \lambda_s^2 \sigma_v^2\right) \sigma_v^2 dt - \frac{1}{2} \lambda_\theta^2 \sigma_\theta^2 \theta_t dt
\]

\[-l_0 (\theta (\beta_2 (u_{1-\gamma})) - \theta (\beta_2 (u_{-\gamma}))) dt + \lambda_\xi \xi_t^Q dN_t^Q
\]

(87)
Realize

\[ l_0 \left( \varrho \left( \beta_2 (u_{1-\gamma}) \right) - \varrho \left( \beta_2 (u_{-\gamma}) \right) \right) = l_0 \varrho \left( \beta_2 (u_{1-\gamma}) \right) \left( \frac{\varrho \left( \beta_2 (u_{1-\gamma}) \right)}{\varrho \left( \beta_2 (u_{-\gamma}) \right)} - 1 \right) \]

\[ = l_0 \varrho \left( \beta_2 (u_{1-\gamma}) \right) \left( 1 - \frac{1}{\varrho \left( \beta_2 (u_{-\gamma}) \right)} \right) \]

\[ = l_0 \varrho \left( \beta_2 (u_{1-\gamma}) \right) \left( 1 - \frac{1}{\varrho \left( \beta_2 (u_{1-\gamma}) \right)} \right) \]

\[ = \left( \varrho \left( \lambda_{\sigma} \right) - 1 \right) \]

(88)

Plug it back to equation 87 and also use the similar process as we did to get the dynamics of \( \frac{dM_t}{M_t} \) from \( d\ln M_t \), we get

\[ \frac{dP_t}{P_t} = r_f dt + \sigma_t dB^T_{t, Q} + \lambda_{\sigma} \sigma_t \sigma_t dB^n_{t, Q} + \lambda_{\theta} \sigma_t \sqrt{\theta_t} dB^\theta_{t, Q} + (e^{\lambda_{\sigma} \xi_{t, Q}} - 1) dN_{t, Q} - l_0 \left( \varrho \left( \lambda_{\sigma} \right) - 1 \right) dt \]

(89)

The stock has expected return equal to risk-free rate under measure \( Q \).

### 5.4 Proof of Proposition 2

**Proof.** For the first claim: Note that the unconditional mean \( \sigma_t^2 \) is \( \mathbb{E}(\sigma_t^2) = \theta + l_0 \mu_{\xi} / \kappa \). The slope of the futures curve is determined by the last conditional expectation term in (28), as both \( a_Q(21) \) and \( b_Q(21) \) are positive. This conditional expectation is

\[ F(\tau) = \sigma_t^2 e^{-\kappa_Q \tau} + (\varrho_Q + \frac{\mu_{\xi} l_0^Q}{\kappa_Q})(1 - e^{-\kappa_Q \tau}) \]

(90)

which slope is given by

\[ \frac{dF(\tau)}{d\tau} = -\kappa_Q \sigma_t^2 e^{-\kappa_Q \tau} + (\varrho_Q + \frac{\mu_{\xi} l_0^Q}{\kappa_Q})(1 + \kappa_Q e^{-\kappa_Q \tau}) \]

\[ = -\kappa_Q \sigma_t^2 e^{-\kappa_Q \tau} + (\varrho_Q + \frac{\mu_{\xi} l_0^Q}{\kappa_Q})\kappa_Q e^{-\kappa_Q \tau} \]

(91)
Substituting in the steady state $\sigma_t^2 = \theta + \frac{\mu \xi}{\kappa}$ we find that the sign of $\frac{dF(\tau)}{d\tau}$ is determined by

$$-\theta - \frac{\mu \xi}{\kappa} + \theta^Q + \frac{\mu \xi Q}{\kappa^Q} \geq 0$$

(92)

where the equality holds iff $\gamma = 0$.

For the second claim, we need to prove

$$E^Q(\sigma_{t+\tau}^2 | \sigma_t^2 = x) > E^P(\sigma_{t+\tau}^2 | \sigma_t^2 = x)$$

(93)

for all $x \in \mathbb{R}^+$. Define $\bar{\theta}$ and $\bar{\theta}^Q$ to be the unconditional means of the process under the respective measures. The inequality in (93) can be written

$$\sigma_t^2 e^{-\kappa^Q \tau} + \bar{\theta}^Q (1 - e^{-\kappa^Q \tau}) - (\sigma_t^2 e^{-\kappa \tau} + \bar{\theta}(1 - e^{-\kappa \tau})) > 0.$$  

(94)

Since $\kappa^Q < \kappa$ it follows that $\sigma_t^2 (e^{-\kappa^Q \tau} - e^{-\kappa \tau}) > 0$ for any $\sigma_t^2$. It remains to show

$$\bar{\theta}^Q (1 - e^{-\kappa^Q \tau}) - \bar{\theta}(1 - e^{-\kappa \tau}) > 0$$

(95)

which is equivalent to

$$\bar{\theta}^Q - \bar{\theta} > \bar{\theta}^Q e^{-\kappa^Q \tau} - \bar{\theta} e^{-\kappa \tau}.$$  

(96)

Define $f(\tau) = \bar{\theta}^Q e^{-\kappa^Q \tau} - \bar{\theta} e^{-\kappa \tau}$, then

$$f'(\tau) = -\kappa^Q \bar{\theta}^Q e^{-\kappa^Q \tau} + \kappa \bar{\theta} e^{-\kappa \tau}$$

(97)

Recall

$$\bar{\theta}^Q = \theta^Q + \frac{\mu_0 Q}{\kappa^Q}$$

$$= \frac{\theta \kappa + \mu_0 Q}{\kappa^Q}$$

(98)

$$\bar{\theta} = \theta + \frac{l_0 \mu_0}{\kappa}$$

(99)

Plug it back to the equation for $f'(\tau)$, we get

$$f'(\tau) = -\left(\kappa \theta + \frac{l_0 Q}{\mu_0} \right) e^{-\kappa^Q \tau} + (\kappa \theta + l_0 \mu_0) e^{-\kappa \tau}$$

$$= \kappa \theta (e^{-\kappa \tau} - e^{-\kappa^Q \tau}) + \left(l_0 \mu_0 e^{-\kappa \tau} - \frac{l_0 Q}{\mu_0} e^{-\kappa^Q \tau}\right)$$

(100)
Since $\kappa^Q < \kappa$, $\ell^Q > \ell_0$ and $\mu^Q > \mu_\xi$, it is easy to see that $f'(\tau) < 0$, therefore $f(0) > f(\tau)$ for all $\tau$, so

$$\theta^Q - \theta > \theta^Q e^{-\kappa^Q \tau} - \theta e^{-\kappa \tau} \quad (101)$$

For the third claim, note that since VIX squared futures are linear positive functions of $\sigma_t^2$ for any finite maturity, it follows that the correlation between returns and volatility shocks is determined by the sign of $\lambda_\sigma$ which is negative for any $\gamma > 0$.

\[ \square \]

### 5.5 VIX Futures Prices

By definition we have

$$VIX_t^2 = \text{Var}_t^Q(\ln P_{t+21}). \quad (102)$$

The conditional cumulant generating function for $\ln P_{t+21}$ is given by

$$\Phi(u) = \ln \mathbb{E}_t^Q e^{u \ln P_{t+21}}$$

$$= \ln \mathbb{E}_t^Q e^{u \lambda_X X_{t+21}}$$

$$= \alpha(u \lambda_X, t, t + 21) + \beta'(u \lambda_X, t, t + 21)X_t \quad (103)$$

$$\lambda_X := (1, \lambda_\sigma, \lambda_\theta)' \quad (104)$$

Therefore, using the property of the cumulant generating function, we see $VIX_t^2 = \text{Var}_t^Q(\ln P_{t+21}) = a + b \sigma_t^2 + c\theta_t$, while $a, b$ and $c$ are second derivatives of $\alpha(\lambda_X, t, t + 21)$, $\beta_2(\epsilon \lambda_X, t, t + 21)$ and $\beta_3(\epsilon \lambda_X, t, t + 21)$ evaluated at $\epsilon = 0$. We numerically compute $a, b$ and $c$ since the analytical solution gets very messy. However, in the case of one-factor model, we are able to achieve an analytical solution for $b$ as follows:

$$b(\tau, \gamma) = \frac{1}{\kappa_\gamma} + \frac{(1 + 2\gamma)^2 \sigma_\sigma^2}{4(\kappa_\gamma)^3}$$

$$+ \frac{e^{-\kappa_\gamma \tau} \sigma_\sigma^2}{\kappa_\gamma^2} \lambda_\sigma \left(1 - e^{-\kappa_\gamma \tau}\right) \left(1 + 2\gamma + \lambda_\sigma \kappa_\gamma\right)$$

$$- \frac{e^{-\kappa_\gamma \tau} \sigma_\sigma^2}{\kappa_\gamma^2} \left(1 + 2\gamma\right)^2 \left(\tau^2 + 4\kappa_\gamma^2 \kappa_\gamma + \kappa_\gamma(1 + 2\gamma) \tau \lambda_\sigma + \frac{\kappa_\gamma}{\sigma_\sigma^2}\right) \quad (105)$$

$$
\kappa_\gamma = \sqrt{\kappa^2 - \sigma_\sigma^2(\gamma^2 + \gamma)} \quad (106)
$$

The solution of $b$ under $P$ measure is a special case of letting $\gamma = 0$ and the VIX associated $b$ is a special case of letting $\tau = 21$ in the above expression.
We adopt the analytical formula for VIX futures up to an integral by Zhu and Lian (2011), and price VIX futures by numerical integration:

\[
F_{t}^{VIX}(t + \tau) = \mathbb{E}^{Q}_{t} \sqrt{VIX_{t+\tau}^{2}} = \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} \frac{1 - \mathbb{E}^{Q}_{t} e^{sv_{1}(s)X_{t+\tau}}}{s^{3/2}} ds
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} \frac{1 - e^{-as} \mathbb{E}^{Q}_{t} e^{u_{1}(s)'X_{t+\tau}}}{s^{3/2}} ds
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} \frac{1 - e^{-as} e^{\alpha(u_{1}(s), t, t + \tau) + \beta'(u_{1}(s), t, t + \tau)X_{t}}}{s^{3/2}} ds
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-0.5s} \left(1 - e^{-ae^{\alpha(u_{2}(s), t, t + \tau) + \beta'(u_{2}(s), t, t + \tau)X_{t}}} \right) ds, \quad (107)
\]

where

\[
u_{1}(s) = (0, -bs, -cs)'
\]

\[
u_{2}(s) = (0, -be^{s}, -ce^{s})'
\]

The second equality is a mathematical result using Fubini’s theorem and the last equality follows by a change of variable to make the integrand bell shaped for easier numerical computation.

### 5.6 MCMC details

Our MCMC estimator requires a Gibbs scheme where we cycle through draws from the conditional distributions of conditional variances, \(\sigma_{t}^{2}\), jump-times \(N_{t}\), jump-sizes \(\xi_{t}\) and parameters \(\Theta\). The respective conditional distributions are known to be proportional to the joint posterior, which is given by

\[
p(\Sigma, \Xi, N, \Theta | \mathcal{R}_{T}) \propto \prod_{t=1}^{T} p(r_{t}, \sigma_{t}^{2} | \xi_{t}, \sigma_{t-1}^{2}, \Theta)p(\xi_{t} | \Theta)p(\xi_{t})l(\sigma_{t})p(\Theta). \quad (110)
\]

We can now sample elements of \(\Sigma, \Xi, N\) and \(\Theta\) in a sequence of Metropolis draws. However, rather sampling directly from this posterior we prefer to first transform the volatility processes through a log transform. Define \(v_{t} = \ln \sigma_{t}^{2}\) and \(\eta_{t} = \ln(\sigma_{t}^{2} + \xi_{t}) - \ln \sigma_{t}^{2}\). The joint dynamics of \(r_{t}\) and \(v_{t}\) can be found through Ito’s lemma. The Euler discretization is

\[
r_{t} = \lambda_{0} dt + \lambda_{\sigma} (\exp(v_{t}) - \exp(-v_{t-1})) + \exp(v_{t-1}) \epsilon_{t}^{\gamma}
\]

\[
v_{t} = v_{t-1} + (\kappa \theta_{v} - \frac{1}{2} \sigma_{v}^{2}) \exp(-v_{t-1}) - \kappa \eta_{t} dN_{t} + \exp(-v_{t-1}) \sigma_{v} \epsilon_{t}^{\gamma}
\]

38
A bivariate Gaussian specification for $\epsilon^x_t$ and $\epsilon^\sigma_t$ now provides implies the conditional density $p(r_t, v_t \mid v_{t-1}, \eta_t, dN_t)$. The conditional posterior for $\eta$ is

$$p(\eta_t \mid \Sigma, \mathcal{N}, R_T, \Theta) \propto p(r_t, v_t \mid v_{t-1}, \eta_t, dN_t)p(\eta_t \mid \sigma^2_t)I_{\eta > 0}$$

(113)

where

$$p(\eta_t \mid \sigma^2_t, \Theta) = \exp \left( - (\exp(\eta_t) - 1) \frac{\sigma^2_t}{\mu_v} + \eta_t \right) \frac{\sigma^2_t}{\mu_v}$$

(114)

We draw $\eta$ by proposing from a normal with mean $\mu^2_\eta = (\mu_2 \omega^2_1 + \mu_1 \omega^2_2)/(\omega^2_1 + \omega^2_2)$, where $\mu_1 = \ln(\mu_v/\sigma^2_t)$, $\mu_2 = v_t - v_{t-1} - (\kappa \theta_v - \frac{1}{2}\sigma^2_t)e^{-\kappa v_t} - \kappa)$, $\omega_2 = \sigma_v \exp(-\frac{1}{2}v_{t-1})$ and $\omega_1$ is a constant. This proposal density approximates the target density through a normal approximation of $p(\eta_t \mid \cdot)$ in (113).

The posterior distribution for the jump-indicator, $dN$ is available in closed form. Define

$$G(i) = \phi(v_t \mid v_{t-1}, \eta_t, dN_t = i, \Theta)p(\eta_t \mid \sigma^2_t, \Theta)p(dN_t = i)$$

(115)

where $\phi$ is a Gaussian density, then

$$p(dN_t = 1) = \frac{G(1)}{G(1) + G(0)}$$

(116)

is the Binomial probability of $dN_t = 1$.

Finally, our algorithm requires simulating from

$$p(v_t \mid v_{t-1}, v_{t+1}, \eta_t, dN_t, \Theta) \propto p(r_t, v_t \mid v_{t-1}, \eta_t, dN_t)p(r_{t+1}, v_{t+1} \mid v_t, \eta_{t+1}, dN_{t+1})$$

(117)

a non-standard density which ones again requires a Metropolis step. To facilitate a Metropolis draw from this density, we propose a candidate value $v^*$ from a conditionally Gaussian $y^* \sim N(M_t, S_t)$ where the conditional mean is $\frac{1}{2}(v_{t-1} + v_{t+1})$ which was shown to be the $\Delta t \to 0$ limit in a diffusive setting in Eraker (2001). The variance of the proposal density is $S_t = \sigma^2_v e^{-\kappa t} s$ where $s$ is drawn from $s = \{s_l, s_h\}$ is a Bernoulli random number such that the variance can be scaled to be large.

[Figure 12 about here.]

Figure 12 shows the behavior of the MCMC sampler for 40,000 posterior draws of $\gamma$ using different starting values. The figure is suggestive of the ideal behavior of MCMC chains: as the sample size increases the respective posterior draws appear to settle in a subset of the parameter space that suggest they are drawn from the same stationary distribution as they “forget” the starting values.
References

Alexander, Carol, and Dimitris Korovilas, 2012, Volatility Exchange-Traded Notes: Curse or Cure?, *Available at SSRN 2062854*.


Bondarenko, Oleg, 2003, Why are Put Options So Expensive?, *working paper, University of Illinois*. 


Johnson, Travis L, 2015, Risk Premia and the VIX Term Structure, *Available at SSRN 2548050*.


Szado, Eduardo, 2009, VIX Futures and Options A Case Study of Portfolio Diversification During the 2008 Financial Crisis, *Journal of Alternative Investments*.


Figure 1: Value of portfolios that roll the futures contracts at closing prices. Top: Value of $1 invested in Jan 2006. Bottom: Log-value of $1 invested in Jan 2006.
Figure 2: Log-value of the VXX and a synthetic one-month rolling VIX futures position.
**Figure 3:** Price of the TVIX and its Net Asset Value (top) and the relative difference between price and NAV (bottom).
Figure 4: The equilibrium coefficients, $\lambda_0$ and $\lambda_\sigma$ as functions of risk-aversion, $\gamma$. 
Figure 5: Scatter plots of contemporaneous daily changes in volatility and stock returns. The data are simulated from the model using $\gamma = 3$ and 8 respectively.
Figure 6: Top: Objective ($P$) and Risk Neutral ($Q$) one month conditional transition densities for the aggregate stock market for $\gamma = 3$ and 8. Bottom: Implied Black-Scholes volatility. The “objective measure (blue)” implied Black-Scholes volatility is computed assuming zero-volatility and jump risk premia.
Figure 7: Left: VIX futures curve and the objective measure expected payoff in the case of high, medium (steady state), and low initial values of spot variance, \( \sigma^2_t \). Right: expected holding period return, \( \mathbb{E}^P_t (VIX_{t+\tau})/\mathbb{E}^Q_t (VIX_{t+\tau}) - 1 \), to a long futures position.
Figure 8: Sampling distributions from the SV model vs. data: We compute the sampling distributions, under the SV model, by simulating a sample for futures returns and then computing the sample moments. The small sample densities are estimated by kernel densities and compared to the sample moments in the data, shown in the vertical bars.
Figure 9: Sampling distributions from the SVVJ model vs. data: We compute the sampling distributions under the SVVJ model by simulating a sample for futures returns and then computing the sample moments. The small-sample densities are estimated by kernel smoothing and compared to the sample moments in the data, shown in the vertical bars.
Figure 10: Sampling distributions from the two-factor volatility model vs. data: We compute the sampling distributions, under the model in eqns. (30) - (36), by simulating a sample for futures returns and then computing the sample moments. The small sample densities are estimated by kernel densities and compared to the sample moments in the data, shown in the vertical bars.
Figure 11: Average term structure of Variance Swap rates computed from Optionmetrics using best bid, best ask and the midpoint. All numbers are in annualized standard deviation units.
Figure 12: Posterior MCMC draws of the risk-aversion parameter $\gamma$ using different starting values.
Table 1: VIX Futures Return Descriptive Statistics

This table reports the summary statistics of returns to rolling positions in VIX futures. The sample data is at daily frequency from Jan 2006 to May 2013. $R^1$ is the daily average arithmetic return, $R^2$ is the daily average logarithmic return and $R^3$ is the average, annualized geometric return. Standard deviation, skewness and kurtosis are from daily arithmetic returns.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>$R^1$</th>
<th>$R^2$</th>
<th>$R^3$</th>
<th>Std</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Month</td>
<td>-0.12</td>
<td>-0.20</td>
<td>-39.40</td>
<td>3.98</td>
<td>0.83</td>
<td>6.42</td>
</tr>
<tr>
<td>2 Month</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-24.65</td>
<td>3.00</td>
<td>0.62</td>
<td>6.02</td>
</tr>
<tr>
<td>3 Month</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-10.15</td>
<td>2.47</td>
<td>0.66</td>
<td>6.11</td>
</tr>
<tr>
<td>4 Month</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-12.09</td>
<td>2.21</td>
<td>0.79</td>
<td>7.28</td>
</tr>
<tr>
<td>5 Month</td>
<td>-0.01</td>
<td>-0.03</td>
<td>-7.33</td>
<td>2.01</td>
<td>0.70</td>
<td>6.90</td>
</tr>
</tbody>
</table>
Table 2: Factor Regressions

Using the VIX futures logarithmic returns, this table reports factor regressions. The factors are the daily logarithmic change in the VIX, $\Delta VIX = \ln(VIX_t/VIX_{t-1})$, and the Fama-French factors MKT, SMB and HML. $t(\alpha)$ denotes a standard $t$-test of the null hypothesis $\alpha = 0$. $\theta$ is a test statistic for the null-hypothesis, $\alpha = 0$ (see Campbell, Lo, and Mackinlay (1997)), and $p(\theta)$ is the associated $p$-value.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$t(\alpha)$</th>
<th>$\Delta VIX$</th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Model, $H_0: \alpha = 0, \theta = 6.35, p(\theta) = 0.00$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-0.19</td>
<td>-4.51</td>
<td>0.36</td>
<td>-0.67</td>
<td>-0.41</td>
<td>-0.09</td>
</tr>
<tr>
<td>2 Month</td>
<td>-0.10</td>
<td>-2.87</td>
<td>0.24</td>
<td>-0.63</td>
<td>-0.18</td>
<td>-0.11</td>
</tr>
<tr>
<td>3 Month</td>
<td>-0.04</td>
<td>-1.16</td>
<td>0.18</td>
<td>-0.58</td>
<td>-0.10</td>
<td>-0.11</td>
</tr>
<tr>
<td>4 Month</td>
<td>-0.04</td>
<td>-1.48</td>
<td>0.14</td>
<td>-0.61</td>
<td>-0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>5 Month</td>
<td>-0.02</td>
<td>-0.81</td>
<td>0.13</td>
<td>-0.45</td>
<td>-0.10</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

| **$\Delta VIX$, MKT, $H_0: \alpha = 0, \theta = 6.37, p(\theta) = 0.00$** | | | | | | |
| 1 Month | -0.19    | -4.52       | 0.37         | -0.71 |
| 2 Month | -0.10    | -2.89       | 0.24         | -0.67 |
| 3 Month | -0.04    | -1.17       | 0.18         | -0.62 |
| 4 Month | -0.04    | -1.49       | 0.14         | -0.62 |
| 5 Month | -0.02    | -0.82       | 0.13         | -0.47 |

| **MKT, $H_0: \alpha = 0, \theta = 3.99, p(\theta) = 0.00$** | | | | | | |
| 1 Month | -0.17    | -2.89       |              | -2.08 |
| 2 Month | -0.09    | -1.93       |              | -1.58 |
| 3 Month | -0.02    | -0.64       |              | -1.30 |
| 4 Month | -0.03    | -1.01       |              | -1.16 |
| 5 Month | -0.01    | -0.46       |              | -0.98 |

| **$\Delta VIX$, $H_0: \alpha = 0, \theta = 6.48, p(\theta) = 0.00$** | | | | | | |
| 1 Month | -0.21    | -4.51       | 0.48         |
| 2 Month | -0.11    | -2.96       | 0.35         |
| 3 Month | -0.05    | -1.39       | 0.28         |
| 4 Month | -0.05    | -1.68       | 0.24         |
| 5 Month | -0.03    | -1.03       | 0.21         |
Table 3: Average VIX Futures Prices

This table reports the summary statistics of the extrapolated constant maturity (one to seven month) VIX futures. The sample data is at daily frequency from March 2004 to May 2013.

<table>
<thead>
<tr>
<th>Futures Maturity</th>
<th>Spot</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.57</td>
<td>21.48</td>
<td>22.10</td>
<td>22.45</td>
<td>22.70</td>
<td>22.94</td>
<td>23.14</td>
<td>23.29</td>
</tr>
<tr>
<td>Median</td>
<td>17.61</td>
<td>19.51</td>
<td>20.94</td>
<td>21.80</td>
<td>22.31</td>
<td>22.75</td>
<td>23.19</td>
<td>23.44</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>10.11</td>
<td>8.79</td>
<td>8.05</td>
<td>7.51</td>
<td>7.13</td>
<td>6.87</td>
<td>6.68</td>
<td>6.52</td>
</tr>
<tr>
<td>Implied Return</td>
<td>-40.52%</td>
<td>-28.93%</td>
<td>-17.18%</td>
<td>-12.44%</td>
<td>-11.86%</td>
<td>-9.89%</td>
<td>-7.46%</td>
<td></td>
</tr>
</tbody>
</table>

Eigenvalues of the Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>6.85</th>
<th>0.14</th>
<th>0.01</th>
<th>0.00</th>
<th>0.00</th>
<th>0.00</th>
<th>0.00</th>
</tr>
</thead>
</table>

58
Table 4: VIX ETNs Investment Objectives and Average Returns

This table summarizes investment objectives and performance of VIX Exchange Traded Notes (ETNs). The investment objectives are denoted $1x$, $2x$ and $-1x$, reflecting a single long, 100% levered long, and single short position, respectively. \( \bar{r} \) is the average daily arithmetic return in percent.

<table>
<thead>
<tr>
<th>Ticker</th>
<th>First Date</th>
<th>Leverage</th>
<th>Horizon</th>
<th>Underlying</th>
<th>( \bar{r} )</th>
<th>std(( r ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>UVXY</td>
<td>04-Oct-2011</td>
<td>2x</td>
<td>1M</td>
<td>VIX-F</td>
<td>-1.13</td>
<td>8.57</td>
</tr>
<tr>
<td>TVIX</td>
<td>30-Nov-2010</td>
<td>2x</td>
<td>1M</td>
<td>VIX-F</td>
<td>-0.67</td>
<td>8.06</td>
</tr>
<tr>
<td>CVOL</td>
<td>15-Nov-2010</td>
<td>1x</td>
<td>3-4M</td>
<td>VIX-S&amp;P</td>
<td>-0.47</td>
<td>6.96</td>
</tr>
<tr>
<td>TVIZ</td>
<td>30-Nov-2010</td>
<td>2x</td>
<td>5M</td>
<td>VIX-F</td>
<td>-0.38</td>
<td>4.20</td>
</tr>
<tr>
<td>VXX</td>
<td>29-Jan-2009</td>
<td>1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>-0.34</td>
<td>3.94</td>
</tr>
<tr>
<td>VIX</td>
<td>30-Nov-2010</td>
<td>1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>-0.30</td>
<td>4.27</td>
</tr>
<tr>
<td>VIXY</td>
<td>04-Jan-2011</td>
<td>1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>-0.26</td>
<td>4.34</td>
</tr>
<tr>
<td>VIXM</td>
<td>04-Jan-2011</td>
<td>1x</td>
<td>5M</td>
<td>VIX-F</td>
<td>-0.17</td>
<td>2.12</td>
</tr>
<tr>
<td>VIIZ</td>
<td>30-Nov-2010</td>
<td>1x</td>
<td>5M</td>
<td>VIX-F</td>
<td>-0.19</td>
<td>2.08</td>
</tr>
<tr>
<td>VXZ</td>
<td>29-Jan-2009</td>
<td>1x</td>
<td>5M</td>
<td>VIX-F</td>
<td>-0.14</td>
<td>2.00</td>
</tr>
<tr>
<td>XVIX</td>
<td>01-Dec-2010</td>
<td>S/L</td>
<td>-1&amp;5M</td>
<td>VIX-F</td>
<td>-0.05</td>
<td>0.79</td>
</tr>
<tr>
<td>IVOP</td>
<td>19-Sep-2011</td>
<td>-1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>0.22</td>
<td>2.95</td>
</tr>
<tr>
<td>XIV</td>
<td>30-Nov-2010</td>
<td>-1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>0.25</td>
<td>4.27</td>
</tr>
<tr>
<td>SVXY</td>
<td>04-Oct-2011</td>
<td>-1x</td>
<td>1M</td>
<td>VIX-F</td>
<td>0.49</td>
<td>4.33</td>
</tr>
<tr>
<td>ZIV</td>
<td>30-Nov-2010</td>
<td>-1x</td>
<td>5M</td>
<td>VIX-F</td>
<td>0.15</td>
<td>1.99</td>
</tr>
</tbody>
</table>
Table 5: VIX ETN Relative Performance

This table compares the performance of the XIV and TVIX stocks to their synthetic securities using VXX returns. The returns on TVIX are related to the VXX through $r_{TVIX}^t = 2 \times r_{VXX}^t$ and the returns on XIV are $r_{XIV}^t = -r_{VXX}^t$. We use these relationships to construct synthetic TVIX and XIV and compare the prices and fair values of each security to the corresponding synthetic security. $G_i$ is defined as the annualized geometric gain relative to the synthetic security, $G_i = \left( \frac{P_{T,i}}{P_{s,i}} \right)^{252/T} - 1$ where $i = 1, 2$ representing actual trade price and net asset value respectively and $P_{s,i}$ denotes the ending value of the synthetic security created from VXX returns with $P_{0,i} = P_{0,i}$.

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th>G2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVIX</td>
<td>3.36</td>
<td>1.61</td>
</tr>
<tr>
<td>XIV</td>
<td>-1.80</td>
<td>-1.90</td>
</tr>
</tbody>
</table>
Table 6: Model Parameter Estimates

The table reports parameter estimates of the underlying model for S&P 500. The parameter estimates are obtained using stock returns only. We report posterior means and standard deviation based on a joint MCMC simulation of latent volatility, jump times and sizes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SVVJ</th>
<th>SV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta \times 10000$</td>
<td>0.677</td>
<td>0.758</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.014</td>
<td>0.0091</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\sigma_v \times 10000$</td>
<td>10.788</td>
<td>9.716</td>
</tr>
<tr>
<td></td>
<td>(0.357)</td>
<td>(0.307)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.002</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$\mu_v \times 10000$</td>
<td>0.390</td>
<td>(0.096)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>7.945</td>
<td>5.898</td>
</tr>
<tr>
<td></td>
<td>(0.760)</td>
<td>(0.677)</td>
</tr>
</tbody>
</table>
Table 7: VIX Futures Return Moments: Data vs. Model

This table compares average returns of VIX futures positions in data and model simulations. $R^1$ is the daily average arithmetic return, $R^2$ is the daily average logarithmic return and $R^3$ is the average annual geometric return. Standard deviation, skewness and kurtosis are from daily logarithmic returns.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>$R^1$</th>
<th>$R^2$</th>
<th>$R^3$</th>
<th>Std</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-0.12</td>
<td>-0.20</td>
<td>-39.40</td>
<td>3.92</td>
<td>0.56</td>
<td>5.95</td>
</tr>
<tr>
<td>2 Month</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-24.65</td>
<td>2.97</td>
<td>0.41</td>
<td>5.85</td>
</tr>
<tr>
<td>3 Month</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-10.15</td>
<td>2.45</td>
<td>0.49</td>
<td>5.88</td>
</tr>
<tr>
<td>4 Month</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-12.09</td>
<td>2.18</td>
<td>0.61</td>
<td>6.86</td>
</tr>
<tr>
<td>5 Month</td>
<td>-0.01</td>
<td>-0.03</td>
<td>-7.33</td>
<td>2.00</td>
<td>0.54</td>
<td>6.67</td>
</tr>
<tr>
<td><strong>SVVJ</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-0.10</td>
<td>-0.19</td>
<td>-36.14</td>
<td>4.15</td>
<td>1.00</td>
<td>26.33</td>
</tr>
<tr>
<td>2 Month</td>
<td>-0.08</td>
<td>-0.13</td>
<td>-26.29</td>
<td>3.06</td>
<td>0.94</td>
<td>23.10</td>
</tr>
<tr>
<td>3 Month</td>
<td>-0.06</td>
<td>-0.09</td>
<td>-19.60</td>
<td>2.34</td>
<td>0.91</td>
<td>21.85</td>
</tr>
<tr>
<td>4 Month</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-14.84</td>
<td>1.82</td>
<td>0.89</td>
<td>21.26</td>
</tr>
<tr>
<td>5 Month</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-11.37</td>
<td>1.43</td>
<td>0.88</td>
<td>20.97</td>
</tr>
<tr>
<td><strong>SV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-0.09</td>
<td>-0.19</td>
<td>-36.10</td>
<td>4.50</td>
<td>-0.13</td>
<td>3.29</td>
</tr>
<tr>
<td>2 Month</td>
<td>-0.07</td>
<td>-0.14</td>
<td>-28.15</td>
<td>3.56</td>
<td>-0.10</td>
<td>3.17</td>
</tr>
<tr>
<td>3 Month</td>
<td>-0.06</td>
<td>-0.11</td>
<td>-22.57</td>
<td>2.92</td>
<td>-0.08</td>
<td>3.20</td>
</tr>
<tr>
<td>4 Month</td>
<td>-0.05</td>
<td>-0.08</td>
<td>-18.41</td>
<td>2.44</td>
<td>-0.07</td>
<td>3.27</td>
</tr>
<tr>
<td>5 Month</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-15.19</td>
<td>2.06</td>
<td>-0.06</td>
<td>3.36</td>
</tr>
</tbody>
</table>
Table 8: Returns to Variance Claims


<table>
<thead>
<tr>
<th>Period</th>
<th>TYPE</th>
<th>1M</th>
<th>2M</th>
<th>3M</th>
<th>6M</th>
<th>12M</th>
<th>24M</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGLR</td>
<td>Mean</td>
<td>96-'14 Option</td>
<td>−25.6</td>
<td>−5.6</td>
<td>0.8</td>
<td>0.5</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>St.dev.</td>
<td>69</td>
<td>47.8</td>
<td>34</td>
<td>19.7</td>
<td>17.4</td>
<td></td>
</tr>
<tr>
<td>Johnson</td>
<td>Mean</td>
<td>96-'14 Option</td>
<td>−17.27</td>
<td>−8.66</td>
<td>−5.14</td>
<td>−3.6</td>
<td>−1.09</td>
</tr>
<tr>
<td></td>
<td>St.dev</td>
<td>150</td>
<td>85.66</td>
<td>53.29</td>
<td>29.09</td>
<td>23.03</td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>Mean</td>
<td>08-'15 Option</td>
<td>−33.24</td>
<td>−23.15</td>
<td>−19.07</td>
<td>−12.25</td>
<td>−6.93</td>
</tr>
<tr>
<td></td>
<td>St.dev</td>
<td>63.25</td>
<td>52.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELW</td>
<td>Mean</td>
<td>96-'07 OTC</td>
<td>−</td>
<td>−21.27</td>
<td>−13.38</td>
<td>−7.22</td>
<td>−4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>−</td>
<td>35.14</td>
<td>29.16</td>
<td>19.8</td>
<td>12.85</td>
</tr>
<tr>
<td>This paper</td>
<td>Mean</td>
<td>96-'15 Option</td>
<td>−21.09</td>
<td>−9.68</td>
<td>−5.9</td>
<td>−1.43</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>St.dev</td>
<td>71.78</td>
<td>63.14</td>
<td>50.92</td>
<td>38.52</td>
<td>28.12</td>
<td>−</td>
</tr>
</tbody>
</table>
Table 9: Variance Swap Return Moments: Data vs. Model

This table compares average returns of 1 month holding return of variance swaps in data and model simulations. $R^1$ is the average arithmetic return and $R^2$ is the average logarithmic return. Standard deviation, skewness and kurtosis are from daily logarithmic returns. All numbers are in percentage.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>$R^1$</th>
<th>$R^2$</th>
<th>Std</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-21.09</td>
<td>-43.41</td>
<td>71.78</td>
<td>0.47</td>
<td>4.63</td>
</tr>
<tr>
<td>2 Month</td>
<td>-9.68</td>
<td>-22.18</td>
<td>63.14</td>
<td>1.33</td>
<td>6.82</td>
</tr>
<tr>
<td>3 Month</td>
<td>-5.90</td>
<td>-14.49</td>
<td>50.92</td>
<td>1.50</td>
<td>7.43</td>
</tr>
<tr>
<td>6 Month</td>
<td>-1.43</td>
<td>-6.38</td>
<td>38.52</td>
<td>1.47</td>
<td>7.36</td>
</tr>
<tr>
<td>12 Month</td>
<td>0.39</td>
<td>-2.52</td>
<td>28.12</td>
<td>1.30</td>
<td>6.92</td>
</tr>
<tr>
<td>Two Factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-21.47</td>
<td>-55.46</td>
<td>62.54</td>
<td>0.53</td>
<td>5.24</td>
</tr>
<tr>
<td>2 Month</td>
<td>-13.54</td>
<td>-26.70</td>
<td>38.46</td>
<td>1.26</td>
<td>6.81</td>
</tr>
<tr>
<td>3 Month</td>
<td>-10.45</td>
<td>-18.68</td>
<td>31.52</td>
<td>1.34</td>
<td>7.07</td>
</tr>
<tr>
<td>6 Month</td>
<td>-6.34</td>
<td>-9.70</td>
<td>21.15</td>
<td>1.42</td>
<td>7.58</td>
</tr>
<tr>
<td>12 Month</td>
<td>-3.48</td>
<td>-4.61</td>
<td>12.65</td>
<td>1.47</td>
<td>8.16</td>
</tr>
<tr>
<td>SVVJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-7.56</td>
<td>-22.57</td>
<td>51.44</td>
<td>-0.48</td>
<td>3.86</td>
</tr>
<tr>
<td>2 Month</td>
<td>-7.01</td>
<td>-16.42</td>
<td>40.92</td>
<td>0.11</td>
<td>3.44</td>
</tr>
<tr>
<td>3 Month</td>
<td>-6.24</td>
<td>-13.19</td>
<td>35.30</td>
<td>0.20</td>
<td>3.29</td>
</tr>
<tr>
<td>6 Month</td>
<td>-4.49</td>
<td>-7.85</td>
<td>24.63</td>
<td>0.34</td>
<td>3.35</td>
</tr>
<tr>
<td>12 Month</td>
<td>-2.72</td>
<td>-3.91</td>
<td>14.71</td>
<td>0.45</td>
<td>3.68</td>
</tr>
<tr>
<td>SV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td>-3.41</td>
<td>-14.71</td>
<td>47.24</td>
<td>-0.21</td>
<td>3.36</td>
</tr>
<tr>
<td>2 Month</td>
<td>-4.34</td>
<td>-12.92</td>
<td>40.91</td>
<td>-0.15</td>
<td>3.15</td>
</tr>
<tr>
<td>3 Month</td>
<td>-4.27</td>
<td>-11.34</td>
<td>36.96</td>
<td>-0.07</td>
<td>2.97</td>
</tr>
<tr>
<td>6 Month</td>
<td>-3.57</td>
<td>-7.73</td>
<td>28.18</td>
<td>0.07</td>
<td>2.89</td>
</tr>
<tr>
<td>12 Month</td>
<td>-2.52</td>
<td>-4.35</td>
<td>18.71</td>
<td>0.17</td>
<td>3.03</td>
</tr>
</tbody>
</table>
Table 10: Variance Swap Returns from Bid and Ask Data

The table reports the returns to Variance Swaps computed from Optionmetrics using a modified version of the VIX formula where we use only BID or ASK price data. In Panel A we report returns to Variance Swaps based on Bid, Midpoints, and Asking prices. In Panel B we report returns to price takers, defined as a trader who crosses the market to make trade, selling at the bid and buying at the asking price.

<table>
<thead>
<tr>
<th>Panel A: Returns based on bids, midpoints, and asks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Bids</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
<tr>
<td>Midpoints</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
<tr>
<td>Asks</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Returns to price takers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyers</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
<tr>
<td>Sellers</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std</td>
</tr>
</tbody>
</table>