The Price of Higher Order Catastrophe Insurance: The Case of VIX Options

Bjørn Eraker

Aoxiang Yang *

January 15, 2020

Abstract

We develop an equilibrium pricing model aimed at explaining observed characteristics in equity returns, VIX futures and VIX options data. To derive our model we first specify a general framework based on affine jump-diffusive state-dynamics and representative agent endowed with Duffie-Epstein recursive utility. This allows us to derive moments of equity returns under the objective and risk-neutral measures, and subsequently semi-closed form solutions to prices of equity options, VIX futures, and VIX options. We calibrate this model to fit the salient features of the data, including moments of consumption and equity returns, variance premium, and various features of VIX derivatives data. The model matches the extremely right-skewed volatility smiles seen in VIX options, a downward-sloping term structure of implied Black'76 volatilities, large negative rates of return on VIX futures, and large VIX option risk premia. It also matches other characteristics of VIX options data, including time-variation in the shape of implied volatilities.

*Both authors are from the Wisconsin School of Business, Department of Finance. Corresponding author: Bjørn Eraker (bjorn.eraker@wisc.edu).

We thank Ing-Haw Cheng, Sang Byung Seo, Ivan Shaliastovich, Julian Thimme (MFA discussant), Jessica Wachter (AFA discussant), and seminar participants at American Finance Association Meeting, San Diego 2020, Midwest Finance Association Meeting, Chicago 2019, 5th Annual Financial Econometrics and Risk Management Conference and Workshop (CFIRM) 2019, University of Copenhagen, Tilburg University, Maastricht University, Rotterdam University, and the University of Wisconsin for helpful comments. The remaining errors are our own.

1 Introduction

While most derivatives trading now takes place electronically there were two major open outcry pits left at the CBOE in 2019: S&P 500 (SPX) options and VIX options. The VIX index, which is itself computed from SPX option prices, has in recent years become an increasingly important underlying for derivatives traders. It is not hard to see the appeal of VIX derivatives for someone who is seeking to hedge against market turmoil: VIX returns are strongly negatively correlated with SPX returns, and thus, a long VIX position is a negative β investment that offers diversification/ hedging benefits to investors with positive market exposure. Since the VIX index itself is not directly investable, investors rely on derivatives to obtain VIX exposure.

The main objective of this paper is to try to understand the pricing of VIX derivatives from the viewpoint of an equilibrium model. In particular, we are interested to see if it is possible to design an equilibrium model that reproduces the salient features of the VIX futures and options data, SPX returns data, and consumption data. To do so we derive a model where a representative agent is endowed with Duffie-Epstein recursive utility and faces an endowment process with time-varying volatility (σ_t) and jumping volatility to volatility with time-varying intensity (λ_t). The exogenous shocks to consumption and its higher order moments drive asset prices. Specifically, the aggregative stock market value obtains as the present value of a levered claim to consumption, as in Bansal and Yaron (2004). In equilibrium, shocks that lead to higher uncertainty lower stock-market valuations, as to generate a higher conditional expected rate of return. This volatility-feedback effect endogenizes the negative contemporaneous return-volatility correlation (sometimes referred to as the leverage effect) that is observed to be very strong in the data. The model also endogenizes the stock market volatility itself, and by extension, the forward-looking expected stock market volatility. Since the VIX index is interpretable as a conditional risk-neutral 30 day forward-looking estimate of market volatility, the model is interpretable as an equilibrium model of VIX. We use the property of the conditional cumulant generating function for log stock price to obtain an explicit expression (up to ODEs) for equilibrium VIX, and then use a Fourier-type payoff transform analysis to derive a semi-closed form (up to a single integral) formula for the value of VIX options.

While there are countless studies of equity options market data, relatively fewer papers study VIX options. Mencía and Sentana (2013) use a panel of VIX futures and options to fit a no-arbitrage based time-series model. Park (2015) uses SPX and VIX options information to predict market returns (SPX), VIX futures returns, SPX and VIX options returns. Huang, Schlag, Shaliastovich, and Thimme (2019) derive a diffusion-based no-arbitrage model to explain negative delta-hedged VIX options returns. Both papers conclude that volatility of volatility risk is priced with a negative market risk price. Park (2016) specifies a reduced form model for VIX directly in order to price derivatives.

Compared with previous studies, this paper provides three contributions. First, we present reduced-form empirical evidence on VIX options prices. Second, we develop a general, perfectly tractable framework for pricing assets, especially derivatives with arbitrary state-dependent payoffs, in a continuous-time affine economy where the representative agent has Duffie-Epstein preferences. Third, conveniently applying the results in the general framework we develop an equilibrium pricing model for VIX options and calibrate it to fit and reconcile salient features of the data on macro quantities as well as on a wide spectrum of assets markets. We describe these contributions in turn.

Our reduced form empirical analysis has two parts. We first seek to understand some basic properties of *ex-ante* pricing information, including the patterns of implied Black ('76) volatility surfaces. Among the interesting features of implied volatility data are the facts that they imply a severe rightly skewed risk-neutral distribution of VIX "returns". The right skewed distribution contrasts equity return distributions which tend to be negatively skewed, as with the SPX. It is much more heavily skewed to the right than SPX returns are skewed to the left. Secondly, unlike equity options, VIX options display a downward sloping term structure: longer term VIX options have lower implied Black volatility than do short maturity ones. This persuasive feature persists irrespective of strikes and market conditions (i.e., high or low VIX), and is very pronounced. As we will show, the pattern is related to mean-reversion in VIX and lack thereof in the distributional assumptions underlying (Black '76) implied volatility computation. Third, shocks to implied volatility of volatility is positively, but imperfectly correlated with the level of VIX itself. This actually rules out single factor representations such as the Heston (1993) CIR diffusion for conditional variance.

The second element of our reduced-form empirical evidence is a look at *ex-post* realized VIX options returns. Huang, Schlag, Shaliastovich, and Thimme (2019) find that delta-hedged VIX options returns are statistically significantly negative on average. Their interpretation is that after controlling for directional volatility risk, volatility-of-volatility risk is priced. We compute average rates of return on VIX calls and find them to be negative, although not statistically significantly so. The average returns on puts are mostly statistically significantly positive. A long call position gives the buyer a positive volatility exposure. We can think of the underlying for the options as being the VIX futures and thus, since VIX futures yield average rates of return that are somewhere in the -30% to -40% range per annum (see Eraker and Wu (2017)), calls (puts) should have negative (positive) expected return. Our analysis confirms this.

Regarding the paper's second contribution, we develope a general consumption-based, affine equilibrium pricing model coupled with recursive preferences. In it, a group of broadly defined state variables, always including consumption endowment, follow an jump-diffusive affine structure, and the representative agent has Duffie-Epstein preferences with the IES being set at one for tractability. In this respect, it is similar to the models of Eraker and Shaliastovich (2008), Benzoni, Collin-Dufresne, and Goldstein (2011) and Tsai and Wachter (2018). In fact, we show that the state-price density approximately solved in Eraker and Shaliastovich (2008) under the help of log-linearization techniques exactly converges to that in our model as the IES approaches one. For our purposes, very important is the recursive preference assumption, which guarantees that state variables as higher-order conditional moments of consumption, such as its growth volatility and volatility of volatility, etc, are priced in equilibrium, which is further essential for generating non-zero premia on VIX derivatives. This is because as a measure of market volatility VIX naturally in equilibrium is a function of those higher-order state variables, none of which, however, could be priced under CRRA preferences, implying assets only having VIX exposure would simply be priced according to the objective measure, further implying zero premia on them. In light of its above features, we expect our framework is valuable in various future equilibrium derivatives pricing research.

Our third contribution is the specific equilibrium VIX model, which, to our best knowledge, is the first structural, consumption-based attempt at VIX option pricing in the literature. It is comprised of a theoretical part and an empirical part. Since the economic mechanism has already been detailed in the second paragraph, we briefly descibe the model's empirical fit. Our calibrated model seeks to match moments from a number of observed macro quantities, starting with the first two moments of aggregate consumption growth, interest rates, and stock returns. Our model matches the consumption growth and interest rate data up to negligible differences, matches equity market volatility, as well as risk premia associated with VIX derivatives data, and generates an equity premium and a variance risk premium that are pretty close to those seen in the data.

As a comparison, previous consumption-based models such as Bansal and Yaron (2004) and Wachter (2013) are mostly unable to correctly address VIX derivatives premia. The rare disaster mechanism in Wachter (2013) implies an unrealistically large VIX and variance premium. Bansal and Yaron (2004) do have a time-varying volatility process, but it suffers the problem that the process is too persistent to account for the observed sharply downward-sloping term structures in VIX derivative premia and implied volatilities. Moreover, a persistent long-run expected consumption growth is completely unnecessary for explaining VIX derivatives data.[1]

[1] VIX in equilibrium is a function in only higher-order moments of consumption. In long-run risk type models, equilibrium VIX doesn't depend on x_t at all.

In a single factor volatility model, say Heston's (1993) model, squared VIX is proportional to spot variance, and thus inherits its statistical properties. Since a square root process has local variance that is proportional to its level, the Heston model will produce an implied VIX volatility that is perfectly correlated with VIX itself. In the data, however, the contemporaneous correlation between VIX and implied (Black '76 or Black-Scholes) volatility of VIX options is only about 0.5. Our two factor (σ_t^2 and λ_t) model replicates this positive but imperfect correlation. The model also matches higher order moments of VIX option return distributions, including variance, skewness and kurtosis. It matches implied VIX volatility in a number of dimensions: the average ATM implied VIX volatility (i.e., VVIX) is almost matched identically. The average implied volatility surface, meaning implied VIX volatility as a function of maturity and strike, is similar to what we observe: it is vastly skewed to the right (as would be consistent with a right-skewed underlying VIX distribution), and it has a sharply downward sloping term structure similar to what we observe in the data. Another phenomenon we document is that during normal times, VIX implied volatilities are concave over (most) strikes but during high VIX (financial crisis) regimes, the implied volatility is convex. Surprisingly, our model reproduces this change from concavity to convexity.

The rest of the paper is organized as follows. Section 2 and 3 respectively describes our sample of VIX options and presents reduced-form evidence. Section 4 presents our general theory, while 5 presents the specifics as it applies to our joint consumption/returns/VIX options model - the equilibrium VIX model. Section 6 presents results from our model calibration exercise and Section 7 summarizes our findings.

2 Data

The sample was collected from the CBOE[2] and consists of data sampled at the one-minute interval over the period 2005 until the end of 2018. The data set consists of best bids, best asks, bid/ask quantities, and open high/low in addition to contact characteristics over the one-minute intervals. The fact that the data are time-stamped down to the minute interval mitigates the problem of non-synchronous quotes that are often problematic in end-of-day data.

VIX options and futures are cash-settled to a special VIX computation known as the VRO. The VRO is computed from prices of constituent SPX options that are compiled through a special auction that is held pre-market on the VIX expiration day, typically the third or fourth Wednesday of the month. This contrasts the VIX itself, which is computed from midpoints. While in theory VRO should

[2]See `https://datashop.cboe.com` for details.

differ little from the open value of the VIX on the settlement day, in practice it may. Griffin and Shams (2018) present evidence suggesting that, since far OTM SPX options can be traded very cheaply and have a comparably large impact on the computation of the SOQ, the market is prone to manipulation.

Some remarks regarding the relationship between VIX futures and options are in order. VIX futures market prices have no direct effect on VIX options - both are settled to VRO. However, the fact that the underlying VIX index is not a marketable asset has implications for both futures prices and options. The most important impact on the prices of futures contracts is they do not adhere to a standard futures-spot no-arbitage parity condition. For example, for a stock index value S_t, a τ period futures price $F_t(\tau)$ will satisfy

$$F_t(\tau) = S_t e^{(r-q)(T-t)} \tag{1}$$

where r and q are the continuously compounding risk free rate and dividend yield, respectively. This implies that $F_t(\tau)$ and S_t do not deviate by a substantial amount.

For VIX futures with long maturities however, the deviation between spot VIX and VIX futures prices can be very large. Mechanically, this happens because there is no way to arbitrage the devations. Fundamentally, futures prices should incorporate market participants' expectations of mean reversion in VIX. Prices can also reflect a risk premium. Whaley (2013) and Eraker and Wu (2017) present evidence suggesting that expected returns on VIX futures are substantially negative.

VIX options do not satisfy Put-Call parity with respect to the underlying VIX index. They do however satisfy a version of Put-Call parity that includes the same-maturity futures, namely

$$C_t = P_t + (F_t - K)e^{-r(T-t)} \tag{2}$$

where C_t and P_t are respectively prices of calls and puts with strike K and T maturity and F_t a T maturity futures price. Keeping in mind that mean reversion will imply that F_t is below spot VIX when spot VIX is high (and vice versa), an ATM option ($K = F_t$) will have a strike that is below spot VIX when spot VIX is high, and above spot VIX when spot VIX is low. The fact that the VIX itself is not a Martingale, but futures are, suggests that we should apply Black (1976)'s pricing formula for options on futures rather than Black-Scholes in this context, particularly, in computing implied volatilities of VIX options.

3 Exploratory Data Analysis

3.1 Option Implied Volatilities

To characterize the pricing of VIX options we first study Implied Volatilities. Figure 1 plot implied volatility for VIX options on two different days. On December 12, 2008, the VIX was high at 65.48 and on March 26, 2017 the VIX was low at 10.78. These days are typical of what we observe in high and low VIX states in our sample.

[Figure 1 about here.]

There are a number of features of the data that are worth commenting on.

First, in both cases, for a given strike, the implied volatility is greater for short-maturity options. That is, the term structure of implied volatility is downward sloping irrespective of the level of VIX. To understand why this happens, it is important to remember that Implied Volatility, in this case computed from the Black '76 model for pricing options on futures, assumes that the underlying is a random walk. If a time-series follows a random walk, its forecasted variance increases linearly with the forecast horizon. The downward-sloping term structure we observe in VIX options implied volatility, therefore, is evidence that the market does not think that VIX variance increases proportionally with the forecast horizon. Rather, it suggests that the market knows and understands that the VIX is not a Martingale.

Second, the shapes of the implied volatility functions are mildly concave in the high-VIX/short-maturity case, seen in the red six-day maturity case in the top graph. In the low VIX case, the implied volatility functions are uniformly forming a concave "frown" rather than the usual convex "smile" seen in most equity data, including the SPX.

Third, and perhaps most surprising, if we compare maturities in the 70-90 day range with relatively high strikes (say 40), we see that they were in some sense more expensive in the 2017 low volatility state than they were in the 2008 high volatility state. For example, both 69 and 97-day maturities in the 40-50 strike range were trading at implied volatilities below 100% in November of 2008, but 83-day maturity 40-50 strike range options traded at above 100% implied volatilities in 2017.

[Table 1 about here.]

Table 1 shows the average VIX implied volatility surface over strike and maturity. As seen, the two predominant patterns discussed above are visually evident: the term-structure is sharply downward sloping and the volatility surface is increasing and concave in the strike levels.

[Figure 2 about here.]

Figure 2 shows the relationship between VIX level, as measured by one-month futures prices, and ATM VIX option implied Black volatility. The color coding shows data by year. As seen in the plot, there is generally a positive relationship and the unconditional correlation is 0.48. The strength of the relation between the futures level and the implied VIX volatility is however time-varying. By running a regression year-by-year we find that the slope coefficients vary from a low of 0.01 in 2009 to 0.1 in 2014. This is not to be interpreted as a causal relation: we do not believe that vol-of-vol, as measured by ATM VIX volatility, is varying deterministically over the calendar. Rather, the evidence suggests that vol-of-vol, and thereby VIX ATM implied, is related to some persistent factor that is imperfectly correlated with volatility itself. In our structural model, therefore, we specify a structure in which aggregate consumption growth volatility (σ_t) is driven by exogenous shocks with two components. The first is a regular CIR-style diffusion term. Second, aggregate volatility is also discontinuous, with jump in it arriving at a rate (λ_t) which follows an independent self-exciting diffusion process. In equilibrium, both VIX and vol-of-vol are non-linear functions of σ_t and λ_t. This modeling specification allows us to match the positive, yet imperfect, time-varying correlation between vol-level and vol-of-vol seen in Figure 2.

[Table 2 about here.]

3.2 Option Returns

Much like traditional asset pricing research, recent developments in option research emphasize risk premia associated with factors-shocks. Coval and Shumway (2001) show that average returns to SPX options are statistically significantly negative. Their paper shows that even delta-neutral straddles that are immune to tail-risk experience large negative returns. Bondarenko (2003) reports Sharpe ratios of -0.38 and -3.93 for 4% and 6% OTM SPX puts, respectively, while Eraker (2012) find Sharpe ratios of about -1/2 for ATM straddles. It is well known that implied volatility exceeds realized volatility by some considerable amount (i.e, Jackwerth and Rubenstein (1996), Bollerslev, Tauchen, and Zhou (2009) among others), which is interpreted as a volatility risk premium. Whaley (2013), Eraker and Wu

(2017), and Dew-Becker, Giglio, Le, and Rodriguez (2017) show that VIX futures and variance swaps yield large negative average rates of return.

Table 2 presents summary statistics on returns to VIX options positions using data from 2006 until the end of 2018. The average returns are buy-and-hold returns over the maturity of the option. For example, an ATM one-month call had an average one-month return of -0.26 (-26%) in our sample. Likewise, an ATM call with six-month maturity had an average six-month return of -.23 (-23%). Reconciling these two numbers is actually not trivial. First off, the one-month -26% return is in fact reasonably consistent with evidence from the VIX futures and variance swap market cited above: maintaining a positive volatility exposure is very costly when using short-maturity derivative instruments. Comparably, long-maturity calls appear to have been quite cheap, as the average return is "only" -23%, thus seemingly allowing a trader to keep a long position in a longer-maturity call and then sell it a month before expiration with minimal loss. It also suggests that a long-short position with the short leg consisting of short-maturity options and the long of long maturities could be profitable. One immediate caveat to note here is that the expected rates of return are difficult to estimate. In fact, it is difficult to conclude that even the average return rates are statistically significantly different from zero: with the exception of ITM short maturities, the confidence intervals for mean call returns overlap with zero which is to say - we cannot reject the null that the average rate of return on these contracts is different from zero. The reason for this is the large standard deviation of the returns which is determined by the non-linearity of the option payoffs along with the very high volatility of the underlying VIX index.

Put options generally have statistically significantly positive rates of return. Since put options give negative volatility exposure, we should expect puts to generate a positive risk premium, and they do: Table 2 shows that puts on average have positive rates of return, and in the case of the long maturity contracts, statistically significantly so. Overall, negative volatility risk premium evidenced in the average positive put returns and Sharpe ratios is consistent with aformentioned empirical evidence from VIX futures, VIX futures ETNs and underlying SPX options.

[Figure 3 about here.]

In order to illustrate further the properties of payoffs to VIX options, Figure 3 depicts the marked-to-market value of short maturity ATM options. The graphs depict the value of a trading strategy that invests in one-month maturity VIX calls and puts, respectively, and hold those options until maturity. We plot the market value of trades initiated at the bids and asks respectively, as to illustrate the performance of an investor who trades by crossing the market or uses market orders. As seen, call options systematically lose while put options on average yield positive rates of return. Noteworthy is the difference in value based on whether the trades are initiated at the bid or the ask, signifying the large bid ask spreads in VIX options. Indeed, someone who were cross the market to buy puts at posted asking prices would almost be back to breakeven after eleven years of trading.

The figure also illustrates how the difference in return characteristics is manifested: since VIX itself has a very right-skewed return distribution, calls occasionally generate large positive returns reminiscent of crash-insurance. The returns on put options are bounded from below by -100%, and also do not experience large value increases similar to what we see for calls because VIX tends to decrease at a much slower speed than it increases. In our model we model the asymmetric volatility shocks by a mixture of a self-exciting volatility jump process and a CIR style diffusion.

It is finally worth noting that the largest asymmetric increase in the value of the VIX calls did not happen in the Fall of 2008, but rather on February 5th, 2018, a day otherwise not associated with a negative macro event. On this day the SPX fell 4.6%, but it was the move in volatility derivatives that captured financial press headlines. Notably, while VIX itself increased dramatically, VIX futures, which normally would increase by a fraction of about $\frac{1}{2}$ of VIX, itself increased dramatically around the close of the market. The sharp increase was accelerated by the need for both 2x levered and short VIX futures ETFs and ETNs to *purchase* VIX futures at the close[3]. The short VIX ETN "XIV" was terminated with a 96% loss while the SVXY - a short VIX ETF lost about 91% and changed its investment objective to -$\frac{1}{2}$ of its capital.

4 A General Model

In this section, we present a general equilibrium asset pricing model featuring recursive preferences for a representative agent in an economy with affine jumping-diffusive states. Exact solutions are characterized for the value function, risk-free rate, the state-price density, and its induced risk-neutral measure. Those results can be conveniently applied to solve our specific stochastic volatility and time-varying volatility jump risk model for VIX option pricing in the next section.

[3]While it seems counterintuitive that both levered long and short ETPs need to purchase futures in the case of an increase in futures values, they do: short funds need to buy-to-cover to maintain a target portfolio of -100% of their capital while long levered funds too need to purchase because their objective is to maintain a (in this case 2:1) leverage.

4.1 Preferences

Consider a continuous-time formulation of the economy where the representative agent's preferences over the uncertain consumption stream C_t can be described by a recursive utility function developed in Duffie and Epstein (1992). For tractability purpose, we assume a limiting case of their model that sets the intertemporal elasticity of substitution (IES) equal to one:

$$V_t = E_t \int_t^{\infty} f(C_s, V_s) ds \tag{3}$$

$$f(C, V) = \beta(1-\gamma)V(\ln C - \frac{1}{1-\gamma}\ln((1-\gamma)V)), \tag{4}$$

where V_t represents the continuation value. The parameter β is the rate of time preference, and γ is the relative risk aversion. The logarithm form in (4) indicates that we have set IES equal to one. It is well known that (4) is equivalent to the $\psi \to 1$ limiting case of the more general formulation:

$$f(C, V) = \frac{\beta}{1-\frac{1}{\psi}} \frac{C^{1-\frac{1}{\psi}} - ((1-\gamma)V)^{\frac{1}{\theta}}}{((1-\gamma)V)^{\frac{1}{\theta}-1}}, \tag{5}$$

where $\theta = (1-\gamma)/(1-\frac{1}{\psi})$. The novel and appealing characteristic of the generalized preferences described in (5) is that they break the tight link between intratemporal risk-aversion (γ) and intertemporal substitutability (ψ) and allow to capture the agent's preference for the timing of the resolution of uncertainty. For example, risk aversion is usually assumed larger than the reciprocal of the IES, $\gamma > 1/\psi$, in which case the agent prefers early resolution of uncertainty. As we will show, this ensures that the compensations for various risks in the economy are of the right sign and quantitatively important. Note that although the preferences in (5) collapse into the familiar power utility with $\gamma = \frac{1}{\psi}$, in which case only risks to current consumption are priced, setting $\psi = 1$ together with $\gamma > 1$ (i.e., (4)) still totally retains us the desired property of the recursive preferences. From an empirical perspective, the relative size of the IES and one is also a source of debate. A number of studies conclude that reasonable values for this parameter lie in a range close to one, or slightly lower than one (Vissing-Jørgensen (2002); Thimme (2017)), while the long-run risks literature (Bansal and Yaron (2004)) often relies on an IES greater than one.

4.2 State Variables

We follow Duffie, Pan, and Singleton (2000) and Eraker and Shaliastovich (2008) and assume that there is a set of n state variables in the economy which follow an affine diffusion-jump process. Obviously, for any an endowment-economy model to make sense, (log) consumption is always one state variable. The other state variables are very broadly defined. For example, one needs to include as a state variable the intensity of the jumps to consumption in a time-varying disaster risk framework (Wachter (2013)), and the price level or the expected inflation when pricing a nominal bond (Bansal and Shaliastovich (2013)). Specifically, we fix the probability space $\{\Omega, \mathcal{F}, \mathcal{P}\}$ and the information filtration $\mathcal{F}_t$, and suppose that X_t is a Markov process in some state space $\mathcal{D} \subseteq \mathbb{R}^n$ with a stochastic differential equation representation

$$dX_t = \mu(X_t)dt + \Sigma(X_t)dB_t + \xi_t \cdot dN_t, \tag{6}$$

where B_t is an $\mathcal{F}_t$ adapted standard Brownian motion in $\mathbb{R}^n$. The term $\xi_t \cdot dN_t$ (element-by-element multiplication) captures n mutually conditionally independent jumps arriving with intensities respectively equal to the n elements of the vector $l(X_t)$ and jump sizes respectively equal to the n elements of the random vector ξ_t defined on $\mathcal{D}$.[4] Formally, each ith element of N_t is a Poisson process with time-varying intensity equal to the ith element of $l(X_t)$. We further assume that jump sizes ξ are i.i.d. over time but not necessarily in cross-section. Their distributions are together specified through the vector moment generating function $\varrho : \mathbb{C}^n \to \mathbb{C}^n$ (also called the "jump transform")

$$\varrho(u) = E[e^{u\xi}]. \tag{7}$$

We assume that all the n moment-generating functions exist such that each $\varrho_i(\cdot)$ is well defined for both complex and real arguments on some region of the complex plane. We further impose an affine structure on the drift, diffusion and intensity functions

$$\mu(X_t) = \mathcal{M} + \mathcal{K}X_t \tag{8}$$

$$\Sigma(X_t)\Sigma(X_t)' = h + \sum_i H_i X_{t,i} \tag{9}$$

$$l(X_t) = l + LX_t, \tag{10}$$

for $(\mathcal{M}, \mathcal{K}) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$, $(h, H) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$, $(l, L) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$. For X to be well defined, there are additional joint restrictions on the parameters of the model, which are addressed in Duffie and Kan (1996). To facilitate above matrix manipulations, note that $H = [H_1, H_2, ..., H_n]$ and $X_t = (X_{t,1}, X_{t,2}, ..., X_{t,n})'$.

[4] Subsequently, a "$\cdot$" always represents an element-by-element multiplication; a "$\cdot/$" represents an element-by-element division.

We assume an endowment economy and that the log consumption supply is always the first state variable of the economy. With a selection vector $\delta_c = (1, 0, 0, ..., 0)'$, this means

$$\ln C_t = \delta_c' X_t. \tag{11}$$

We also assume that the market is complete[5] and, particularly, the risk-free rate exists.

4.3 Value Function and Risk-Free Rate

Let W denote the wealth of the representative agent and $J(W, X)$ the value function. Because X_t is a Markov process and nothing depends on t explicitly in the specifications of both preferences and state variable dynamics, we conjecture that J is not explicitly t-dependent. Conjecture that the equilibrium price-dividend ratio for the consumption claim (i.e., a perpetual claim that exactly delivers the aggregate consumption as its dividend each period) is constant. In particular, let S_t denote the price of the consumption claim. Then

$$\frac{S_t}{C_t} = A \tag{12}$$

for some constant A.[6] (11), (12) and a vector version of Ito's Lemma for jump-diffusion processes together imply

$$\frac{dS_t}{S_{t^-}} = \left(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c\right)dt + \delta_c'\Sigma(X_t)dB_t + \delta_c'\left[(e^{\xi} - \mathbf{1}) \cdot dN_t\right], \tag{13}$$

where e^{ξ} is an $n \times 1$ vector which ith element is e^{ξ_i} and $\mathbf{1}$ is an $n \times 1$ vector of ones. Recall that the instantaneous net return on the consumption claim is

$$\frac{dS_t + C_{t^-}dt}{S_{t^-}} = \frac{dS_t}{S_{t^-}} + \frac{1}{A}dt. \tag{14}$$

Let r_t denote the instantaneous net risk-free rate. To solve for the value function, consider the Hamilton-Jacobi-Bellman equation for an investor who allocates wealth W_t between S_t and the risk-free asset.

[5] For our purposes of VIX option pricing, and, particularly, VIX option implied volatility computation, in our specific model of Section 5 we will assume the VIX index itself is not directly tradable, as in reality. All other markets, however, are complete.

[6] Indeed, as pointed out in Wachter (2013), the fact that S_t/C_t is constant (and equal to $1/\beta$ actually, which we will verify) arises directly from the assumption of unit IES, and is independent of the details of the model. See, e.g., Weil (1990)

Let α_t be the fraction of wealth invested in the (risky) consumption claim S_t, and (with some abuse of notation) let C_t be the agent's consumption choice. Wealth then follows the process

$$dW_t = \big[W_{t^-}\alpha_t(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{A} - r_t) + W_{t^-}r_t - C_{t^-}\big]dt$$

$$+W_{t^-}\alpha_t\delta_c'\Sigma(X_t)dB_t + W_{t^-}\alpha_t\delta_c'\big[(e^{\xi} - \mathbf{1})\cdot dN_t\big]. \tag{15}$$

Optimal consumption and portfolio choice must satisfy the following Hamilton-Jacobi-Bellman equation

$$\sup_{\alpha_t, C_t}\Bigg\{J_W\big(W_t\alpha_t\big(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{A} - r_t\big) + W_t r_t - C_t\big) + J_X'\mu(X_t)$$

$$+\frac{1}{2}tr\left(\begin{bmatrix} J_{XX} & J_{XW} \\ J_{XW}' & J_{WW}\end{bmatrix}\begin{bmatrix} \Sigma(X_t)\Sigma(X_t)' & W_t\alpha_t\Sigma(X_t)\Sigma(X_t)'\delta_c \\ W_t\alpha_t\delta_c'\Sigma(X_t)\Sigma(X_t)' & W_t^2\alpha_t^2\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c\end{bmatrix}\right)$$

$$+E_{\xi_1}\big[J(W_t + W_t\alpha_t\big(e^{\xi_1} - 1\big), X_{t,1} + \xi_1, X_{t,-1})\big]l_1(X_t)$$

$$+\sum_{i=2}^{n}E_{\xi_i}\big[J(W_t, X_{t,i} + \xi_i, X_{t,-i})\big]l_i(X_t) - J(W_t, X_t)\mathbf{1}'l(X_t) + f(C_t, J(W_t, X_t))\Bigg\} = 0, \tag{16}$$

where J_i denotes the first derivative of J with respect to i, for i equal to W or X, and J_{ij} the second derivative of J with respect to first i and then j. The dimensions of all resulting matrices are well understood. For example, J_X is an $n \times 1$ vector. The operator $tr()$ represents the trace of the operated matrix. Note that the instantaneous expected change in the value function is given by the continuous drift plus the expected change due to jumps to state variables. The effects of jumps are not symmetric: jump to consumption affects J through both W and X, whereas jumps to other state variables affect J only through X. In Appendix A, we show that the form of the utility function and the envelop condition $f_C = J_W$ imply that the wealth-consumption ratio $A = \beta^{-1}$. Moreover, the value function takes the form

$$J(W, X) = \frac{W^{1-\gamma}}{1-\gamma}I(X). \tag{17}$$

The function $I(X)$ is given by

$$I(X) = e^{a+b'X}, \tag{18}$$

where a is given by

$$a = \frac{1}{\beta}\Big\{\big((1-\gamma)\delta_c + b\big)'\mathcal{M} + \frac{1}{2}\big((1-\gamma)\delta_c + b\big)'h\big((1-\gamma)\delta_c + b\big)$$

$$+\Big(\big(\varrho_1(1-\gamma+b_1) - \varrho_1(b_1)\big)\delta_c' + \varrho(b)' - \mathbf{1}'\Big)l + (1-\gamma)\beta ln(\beta)\Big\} \tag{19}$$

and b solves the n-equation system

$$\frac{1}{2}b'Hb + \big(\mathcal{K}' + (1-\gamma)\delta_c'H - \mathrm{diag}(\beta)\big)b + L'\varrho(b)$$

$$+\big(\varrho_1(1-\gamma+b_1) - \varrho_1(b_1)\big)L'\delta_c + \frac{(1-\gamma)^2}{2}\delta_c'H\delta_c + (1-\gamma)\mathcal{K}'\delta_c - L'\mathbf{1} = 0. \tag{20}$$

Here (20) is a system of n equations for n unknowns $(b_1, b_2, ..., b_n)$. It depends on the specification of $\varrho(\cdot)$ and generally has no closed-form solution. The system is quadratic and can be readily solved with a low dimension of n in several special cases, for example, when there are no jumps with state-dependent intensities, or when there is jump only in consumption while the jump intensity doesn't depend on consumption itself, and so forth. In many other cases, including our VIX model considered in the next section, even if the system is not entirely quadratic the equation for some b_i is quadratic so that b_i can be easily solved out which then can be treated as a constant in equations for remaining $b_i s$ and helps reduce those remaining equations into a quadratic system.

Because in most settings $\gamma > 1$, the state variable X_i with a positive (negative) associated coefficient b_i would be negatively (positively) correlated with the value function, i.e., negatively priced. As will be shown in the next subsection, the equilibrium prices of risks to X_t can be summarized by the following $n \times 1$ vector

$$\lambda = \gamma\delta_c - b. \tag{21}$$

Intuitively, except for log consumption, any state variable positively (negatively) correlated with the value function naturally commands a positive (negative) market price of risk. Because the log consumption affects the value function additionally through W, its market price of risk has an extra term γ. Now if $\gamma = 1 \equiv 1/\psi$, the Duffie-Epstein preferences collapse into the familiar CRRA preferences, and thus, as one can easily verify, (20) admits $b = (0, 0, ..., 0)'$ as a solution and from (21) $\lambda = (\gamma, 0, 0, ..., 0)$, which means only innovations to consumption will be priced. Therefore, while consumption is the only priced factor in CRRA utility models, Duffie-Epstein preferences imply that all state variables are potentially priced.[7]

Appendix A also shows that the instantaneous risk-free rate is given by

$$r_t = \underbrace{\beta + \delta_c'\mu(X_t) + (\frac{1}{2} - \gamma)\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \delta_c'\big[\big(\varrho(b+1-\gamma) - \varrho(b-\gamma)\big)\cdot l(X_t)\big]}_{\text{CRRA Preferences}} + \underbrace{\delta_c'\Sigma(X_t)\Sigma(X_t)'b}_{\text{Duffie-Epstein Preferences}}. \tag{22}$$

[7]Of course for any state variable to be priced, it ultimately has to influence consumption in some systematic way.

The terms above the first bracket in (22) is what would still arise even if we had assumed CRRA preferences for the representative agent. β represents the role of discounting, $\delta_c'\mu(X_t)$ intertemporal smoothing, and $(\frac{1}{2}-\gamma)\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c$ [8] precautionary savings due to diffusion risks in consumption. $\delta_c'\big[\big(\varrho(b+1-\gamma)-\varrho(b-\gamma)\big)\cdot l(X_t)\big]$ represents the representative agent's response to jump risks in consumption. Suppose the jump size for consumption is always negative, $\varrho_1'(\cdot)<0$, then $\varrho_1(b+1-\gamma)-\varrho_1(b-\gamma)$ is negative regardless of b. Intuitively, an increase in the probability of a downward jump in consumption, $l_1(X_t)$, increases the representative agent's desire to save, and thus lowers the risk-free rate. The term above the second bracket in (22) represents the representative agent's saving motive response to risks in the economy that would only arise under Duffie-Epstein preferences. $\delta_c'\Sigma(X_t)\Sigma(X_t)$ captures the comovement between the diffusion in consumption and that in each state variable, while b determines the sign of the influence of each state variable on marginal utility of consumption. Multiplication of them, if positive (negative), summarizes an additional feature the representative agent likes (dislikes) about the diffusion risks in the economy. To better understand this point, think about volatility as the second state variable. Assume that the comovement between consumption and volatility is negative, i.e., $(\delta_c'\Sigma(X_t)\Sigma(X_t)')_2<0$, and that volatility positively (negatively) affects marginal utility (utility), i.e., $b_2>0$. Multiplying them together would yield a negative push on the risk-free rate. Intuitively, because the times at which consumption is low are also times at which volatility and thus marginal utility is high, the representative agent dislikes and wants to avoid being impacted by this source of diffusion risk. He will thus have another precautionary saving motive, which pushes down the risk-free rate.

4.4 State-Price Density

Calculation of prices and rates of returns in the economy is simplified considerably by making use of the state-price density and the induced risk-neutral measure, which reflect the equilibrium compensation investors require for bearing various risks in the economy. Unlike time-additive preferences, recursive preferences imply that the state-price density depends explicitly on the value function. In particular, Duffie and Skiadas (1994) show that the state-price density associated with the preferences in (3) and (4) is equal to

$$\pi_t = \exp\Big\{\int_0^t f_V(C_s,V_s)ds\Big\} f_C(C_t,V_t), \tag{23}$$

[8] Here, $\frac{1}{2}$ aries from applying Ito's Lemma because we are working with log consumption, which is quantitatively not important.

where f_C and f_V denote the derivatives of f with respect to the first and second argument, respectively. In Appendix A, we show that π_t can be expressed in equilibrium as

$$\pi_t = \beta^\gamma \exp\Big\{\eta t - \beta b' \int_0^t X_s ds + a + (b' - \gamma\delta_c')X_t\Big\}, \tag{24}$$

where $\eta = \beta(1-\gamma)ln\beta - \beta a - \beta$. Applying Ito's Lemma on (24) then implies

$$\frac{d\pi_t}{\pi_{t^-}} = \mu_{\pi,t}dt + (b' - \gamma\delta_c')\Sigma(X_t)dB_t + \Big(e^{(b-\gamma\delta_c)\cdot\xi} - \mathbf{1}\Big)dN_t, \tag{25}$$

where

$$\mu_{\pi,t} = \eta - \beta b' X_t + (b' - \gamma\delta_c')\mu(X_t) + \frac{1}{2}(b' - \gamma\delta_c')\Sigma(X_t)\Sigma(X_t)'(b - \gamma\delta_c). \tag{26}$$

Alternatively, given the form of the risk-free rate in (22), we can solve for $\mu_{\pi,t}$ following a familiar no-arbitrage argument. As no-arbitrage implies that $\pi_t e^{\int_0^t r_s ds}$ is a martingale, $E_t[d(\pi_t e^{\int_0^t r_s ds})]$ must be zero. Using Ito's Lemma, we can obtain

$$\mu_{\pi,t} = -r_t - \big(\varrho(b - \gamma\delta_c) - \mathbf{1}\big)' l(X_t). \tag{27}$$

Making use of (19), (20) and (22), one can verify that (27) is equivalent to (26). We now define

$$\Lambda_t = \Sigma(X_t)'\lambda, \tag{28}$$

where, recall (21), $\lambda = \gamma\delta_c - b$ determines the market prices of risks in the different components of X_t such that if $\lambda_i = 0$ then innovations to $X_{t,i}$ are not priced. λ_i is related, through $\varrho_i(\cdot)$, to the price of jump risk with jump size ξ_i in the ith state variable. $\Lambda_{t,i}$ is literally the total price of the Brownian motion risks associated with $X_{t,i}$.

In Appendix B we rigorously prove a convergence result: the equilibrium state-price density in (24) is actually an $IES \to 1$ limit of the more general state-price density approximately solved in Eraker and Shaliastovich (2008), thanks to the similarity between the state variable dynamics used in the two papers. Equivalently, the error in the log-linear approximation in that paper vanishes as the IES approaches one.

4.5 Risk-Neutral Measure

Although with the state-price density at hand we can already start to price assets with any state-variable-dependent payoffs, it is sometimes more convenient, or even necessary, to work with the evolution of the state variables under the risk-neutral measure induced by the state-price density π_t. The following theorem is a generalization of Proposition 5 in Duffie, Pan, and Singleton (2000).

Theorem 1. *Under the risk-neutral measure Q induced by the state-price density π_t the state variables follow*

$$dX_t = (\mathcal{M}^Q + \mathcal{K}^Q X_t)dt + \Sigma(X_t)dB_t^Q + \xi_t^Q \cdot dN_t^Q, \tag{29}$$

where

$$\mathcal{M}^Q = \mathcal{M} - h\lambda \tag{30}$$

$$\mathcal{K}^Q = \mathcal{K} - H\lambda \tag{31}$$

$$dB_t^Q = dB_t + \Lambda_t dt \tag{32}$$

defines a Brownian motion under the risk-neutral measure.

The Q jump-arrival intensities are given by

$$l^Q(X_t) = l(X_t) \cdot \varrho(-\lambda). \tag{33}$$

The Q jump-size densities are characterized by the vector moment generating function $\varrho^Q : \mathbb{C}^n \rightarrow \mathbb{C}^n$

$$\varrho^Q(u) = E^Q[e^{u\xi}] = \varrho(u - \lambda)./\varrho(-\lambda), \tag{34}$$

where element-by-element multiplication and division are respectively performed in (33) and (34).

Notice that if $\lambda_i = 0$, then there is no difference in the P vs. Q jump measures and market prices of both diffusion and jump risks associated with X_i are zero. This is another intuitive demonstration of why λ summarizes the market prices of risks associated with the diffusions and jumps in the economy.

The jump intensity is greater (smaller) under the equivalent measure Q whenever λ is negative (positive). Usually, we cannot say much about the relationship between the jump size under the equivalent measure Q and that under the objective measure P, except for a special case in which the jump sizes are exponentially distributed. Suppose ξ is exponentially distributed with mean denoted by $E^P(\xi)$. Intuitively, one can construct the following reward-to-risk ratio vector, which, though a bit too simple, can be an illustration of the equilibrium compensations for jump size risks

$$\Lambda^{Jump} \equiv \left(E^P(\xi) - E^Q(\xi)\right)./Std^P(\xi) = \mathbf{1} - \varrho(-\lambda). \tag{35}$$

It's easy to see Λ_i^{Jump} and λ_i have the same sign, implying that the mean of ξ is adjusted upward (downward) under Q measure for negatively (positively) priced state variables. However, it is somewhat misleading, although tempting, to coin this measure a market price of jump risk. Jump risks are characterized, and thus priced, not only according to their means and standard deviations, but also higher moments.

5 A Structural Approach to VIX Option Pricing

In this section, we present our model framework for pricing VIX options, which is a special case of the general affine pricing framework developed in the previous section. A simple affine state dynamics is first employed to price VIX in explicit form in the general equilibrium of the economy and a generalized Fourier payoff transform analysis is then performed to derive a pricing formula for VIX options as a single integral.

5.1 The Model

Consider an endowment economy where there exists a representative agent who has recursive preferences as described in (3) and (4), and consumption, dividends and in the end asset prices and returns, are influenced by a key variable, which is the conditional volatility of consumption growth, σ_t, which itself is exposed to potential diffusion and jump risks. Specifically, we assume the following affine structure for the evolutions of the state variables

$$d\ln C_t = (\mu - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t^C \tag{36}$$

$$d\sigma_t^2 = \kappa^V(\theta^V - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t \tag{37}$$

$$d\lambda_t = \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda. \tag{38}$$

Here $\ln C_t$ is the log consumption supply. σ_t^2 is the instantaneous conditional variance of consumption growth, which is driven by the continuous Brownian shock dB_t^V as well as the discontinuous shock $\xi_V dN_t$. Here N_t is a compounded Poission process which instantaneous arrival intensity is λ_t which itself follows a mean-reverting diffusion process. ξ_V is a time-invariantly distributed random variable representing the jump size with a moment generating function $\varrho(\cdot)$. In line with Eraker and Shaliastovich (2008) we assume $\xi_V > 0$, implying only upward jumps in volatility are possible. Here all three standard Brownian motions B_t^C, B_t^V and B_t^λ and the jump size ξ_V are mutually independent. Our specification of the consumption process abstracts from the important mechanisms in leading asset pricing models such as the long-run productivity risks and the rare disasters occurred to consumption. Instead, we focus on potential jumps to consumption growth volatility, which specification is natural given our concentration on pricing VIX and VIX derivatives. In a nutshell, we are pursuing the simplest framework that, nevertheless, allows capturing as many aspects of VIX derivatives data as possible.

As Cox, Ingersoll, and Ross (1985) discuss, the solution to (38) has a stationary distribution provided that $\kappa^\lambda > 0$ and $\theta^\lambda > 0$. This stationary distribution is Gamma with shape parameter $2\kappa^\lambda\theta^\lambda/\sigma_\lambda^2$ and scale parameter $\sigma_\lambda^2/(2\kappa^\lambda)$. If $2\kappa^\lambda\theta^\lambda > \sigma_\lambda^2$, the Feller condition (from Feller (1951)) is satisfied, implying a finite density at zero. The stationary distribution of λ_t is highly right-skewed, arising from the square root term multiplying the Brownian shock in (38): the square root term implies that high realizations of λ_t make the process more volatile, and thus further high realizations more likely than they would be under a standard AR process. The model therefore implies that there are times when jumps to volatility can occur with high probability (akin to financial crisis times), but these times are themselves very rare. For similar reasons, there is a σ_t term multiplying the Brownian shock in (37), helping both prevent σ_t^2 from falling below zero and correctly replicate the right-skewness in its distribution.

5.2 State-Price Density

Applying the results derived for the general framework, Appendix C shows that the value function of the representative agent in this model is given by

$$J(W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \exp\left(a + b_2\sigma_t^2 + b_3\lambda_t\right) \tag{39}$$

where

$$a = \frac{1}{\beta}\Big((1-\gamma)(\mu + \beta\ln\beta) + b_2\kappa^V\theta^V + b_3\kappa^\lambda\theta^\lambda\Big) \tag{40}$$

$$b_2 = \frac{(\kappa^V + \beta)}{\sigma_V^2} - \frac{\sqrt{(\kappa^V+\beta)^2 - \sigma_V^2\gamma(\gamma-1)}}{\sigma_V^2} \tag{41}$$

$$b_3 = \frac{\kappa^\lambda + \beta}{\sigma_\lambda^2} - \frac{\sqrt{(\kappa^\lambda+\beta)^2 - 2\sigma_\lambda^2(\varrho(b_2)-1)}}{\sigma_\lambda^2}. \tag{42}$$

Assume parameter values are such that b_2 and b_3 are both well defined. When this is the case, note that (41) implies $b_2 > 0$ if we assume $\gamma > 1$. (42) then implies $b_3 > 0$ since by definition $\varrho(b_2) - 1 = E[e^{b_2\xi_V} - 1] > 0$ which is due to the positivities of both b_2 and ξ_V. Hence from (39) the value function (marginal utility) is decreasing (increasing) in both σ_t^2 and λ_t. This means an increase in consumption growth volatility reduces utility (increases marginal utility) for the representative agent. Similarly, an increase in the probability of a volatility jump also reduces utility (increases marginal utility) for the representative agent. Both results are intuitive. Because under recursive preferences the marginal utility depends on, besides consumption, the value function, which is explicitly affected by σ_t^2 and λ_t, the agent requires compensation for bearing the risks in both σ_t^2 and λ_t.

The instantaneous risk-free rate is given by

$$r_t = \beta + \mu - \gamma\sigma_t^2, \tag{43}$$

where β represents the role of discounting, μ intertemporal smoothing and $\gamma\sigma_t^2$ precautionary savings. Appendix C also shows that the state-price density is given by

$$\frac{\pi_t}{\pi_{t^-}} = -r_t dt - \Lambda_t' dB_t + \left(e^{b_2\xi_V} - 1\right)dN_t - \lambda_t E[e^{b_2\xi_V} - 1]dt \tag{44}$$

$$\Lambda_t = \Sigma(X_t)'\lambda \tag{45}$$

$$\lambda = (\gamma, -b_2, -b_3)'. \tag{46}$$

Remember that the vector λ determines the market prices of risks in the different components of X_t such that innovations to $X_{t,i}$ are positively (negatively/not) priced if and only if $\lambda_i > 0$ $(< 0/ = 0)$. Therefore, in the present model log consumption $\ln C_t$ has a positive market price of risk while consumption growth volatility σ_t^2 and volatility jump intensity λ_t each warrants a negative market price of risk. The fact that innovations to all three state variables are priced is in sharp contrast with the CRRA utility model in which only innovations to consumption are priced, and important for our model to generate non-zero VIX derivative premia akin to those seen in the data.

Using the results in Theorem 1, one can show that the evolution of the state variables under the risk-neutral measure Q induced by the state-price density is given by (also see Appendix C for other risk-neutral parameters)

$$d\ln C_t = \left(\mu - (\frac{1}{2} + \gamma)\sigma_t^2\right)dt + \sigma_t dB_t^{C,Q} \tag{47}$$

$$d\sigma_t^2 = \kappa^{V,Q}(\theta^{V,Q} - \sigma_t^2)dt + \sigma_V\sigma_t dB_t^{V,Q} + \xi_V^Q \cdot dN_t^Q \tag{48}$$

$$d\lambda_t = \kappa^{\lambda,Q}(\theta^{\lambda,Q} - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^{\lambda,Q}, \tag{49}$$

where

$$\kappa^{V,Q} = \kappa^V - b_2\sigma_V^2; \quad \kappa^{\lambda,Q} = \kappa^\lambda - b_3\sigma_\lambda^2 \tag{50}$$

$$\theta^{V,Q} = \frac{\kappa^V}{\kappa^V - b_2\sigma_V^2}\theta_V; \quad \theta^{\lambda,Q} = \frac{\kappa^\lambda}{\kappa^\lambda - b_3\sigma_\lambda^2}\theta^\lambda. \tag{51}$$

From (47), the drift of consumption growth is adjusted downward by $\gamma\sigma_t^2$ under Q measure. Equations (48) through (51) show that for both σ_t^2 and λ_t processes the mean reversion becomes slower and the long-run mean becomes higher under Q measure. Moreover, Appendix C shows that the jump-arrival intensity is magnified under the Q measure by a percentage $\varrho(b_2) - 1$: λ_t under P versus $\varrho(b_2)\lambda_t$ under Q. And as analyzed previously the jump size may be adjusted upward or downward under the Q measure, with a moment generating function $\varrho(u)$ under P versus $\varrho(u + b_2)/\varrho(b_2)$ under Q. But in the special case that ξ_V is exponentially distributed, the jump size is adjusted upward in the sense that its mean is increased under Q. Specifically, let $\xi_V \sim \exp(\mu_\xi)$ under P, then Appendix C shows that $\xi_V^Q \sim \exp(\frac{\mu_\xi}{1-\mu_\xi b_2})$ under Q.

5.3 Equity Price

To price VIX derivatives, let's first price VIX. We start with specifying an aggregate dividend process, which is denoted D_t and modeled as leveraged consumption as in Abel (1999) and Campbell (2003), multiplied by a positive idiosyncratic noise term. Specifically, assume $D_t = C_t^\phi e^{\sigma_D B_t^D}$. Then

$$d\ln D_t = \phi d\ln C_t + \sigma_D dB_t^D, \tag{52}$$

where B_t^D is a standard Brownian motion independent of any other random variable in the model. As a result, this type of risk would not be priced in equilibrium. That is, the state variable $\ln D_t$ is redundant, which point one can confirm by introducing a fourth state variable $\ln D_t$ as specified above in the present framework, solving the model all over again and verifying that $b_4 = 0$, i.e., $\ln D_t$ doesn't enter the agent's value function after $\ln C_t$ has. However, the presence of equity market idiosyncratic risk does affect the level of VIX and give us another free parameter σ_D so as to better match data moments in our later calibration.

Let $X_t = [\ln C_t, \sigma_t^2, \lambda_t]'$ and $P(X_t, D_t)$ denote the price of the claim to all future aggregate dividends. Then this price is just obtained by taking the expectation under the risk-neutral measure of all discounted future dividends:

$$
\begin{aligned}
P(X_t, D_t) &= \int_0^\infty E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} D_{t+\tau}\right) d\tau \\
&= \int_0^\infty E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} e^{\phi \ln C_{t+\tau}} e^{\sigma_D B_{t+\tau}^D}\right) d\tau \\
&= e^{\sigma_D B_t^D} \int_0^\infty e^{\frac{\sigma_D^2}{2}\tau} E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} e^{\phi \ln C_{t+\tau}}\right) d\tau.
\end{aligned}
\tag{53}
$$

To compute the expectation in the integrand, we follow Duffie, Pan, and Singleton (2000) and compute a discounted characteristic function of X_t under the risk-neutral measure:

$$
\varrho_X^Q(u, X_t, \tau) \equiv E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} e^{u' X_{t+\tau}}\right) \tag{54}
$$

defined for $u \in \mathbb{C}^n$. Under appropriate technique conditions (See Duffie, Pan, and Singleton (2000)), ϱ_X^Q is exponential affine in X_t:

$$
\varrho_X^Q(u, X_t, \tau) = e^{\alpha(\tau) + \beta(\tau)' X_t}, \tag{55}
$$

where $\alpha(\tau)$ and $\beta(\tau)$ satisfy the following complex-valued ordinary differential equations

$$
\dot{\beta}(\tau) = -\Phi_1 + \mathcal{K}^{Q'} \beta(\tau) + \frac{1}{2}\beta(\tau)' H \beta(\tau) + L^{Q'}\left(\varrho^Q(\beta(\tau)) - 1\right) \tag{56}
$$

$$
\dot{\alpha}(\tau) = -\Phi_0 + \mathcal{M}^{Q'} \beta(\tau) + \frac{1}{2}\beta(\tau)' h \beta(\tau) + l^{Q'}\left(\varrho^Q(\beta(\tau)) - 1\right), \tag{57}
$$

subject to boundary conditions $\beta(0) = u, \alpha(0) = 0$. A closed-form solution to $\varrho_X^Q(u, X_t, \tau)$ is difficult, if not impossible, to obtain, but solving for $\varrho_X^Q(u, X_t, \tau)$ for a specific u is straightforward. In the case of the equity market, (53) implies we only need to solve for $\left(\alpha(\tau), \beta(\tau)\right)$ for $u = (\phi, 0, 0)'$. It turns out, as shown in Appendix C, that the solutions obey

$$
\beta_1(\tau) = \phi, \forall \tau \tag{58}
$$

$$
\beta_2(\tau) = \frac{2(\phi - 1)(\gamma - \frac{1}{2}\phi)\left(1 - e^{-\eta_\phi \tau}\right)}{(\eta_\phi + b_2\sigma_V^2 - \kappa^V)\left(1 - e^{-\eta_\phi \tau}\right) - 2\eta_\phi} \tag{59}
$$

where

$$
\eta_\phi = \sqrt{(b_2\sigma_V^2 - \kappa^V)^2 + 2(\phi - 1)(\gamma - \frac{1}{2}\phi)\sigma_V^2} \tag{60}
$$

$$
\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2 \beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\left(\beta_2(\tau) + b_2\right) - \varrho(b_2) \tag{61}
$$

$$\dot{\alpha}(\tau) = -\beta + \mu(\phi - 1) + \kappa^V \theta^V \beta_2(\tau) + \kappa^\lambda \theta^\lambda \beta_3(\tau). \tag{62}$$

While $\beta_1(\tau)$ and $\beta_2(\tau)$ have closed-form solutions given by (58) through (60), we need to solve for $\beta_3(\tau)$ and $\alpha(\tau)$ numerically using (61), (62) and the boundary conditions. Given $\big(\alpha(\tau), \beta(\tau)\big)$, (53), (54) and (55) together then recover the price of the claim to future aggregate dividends:

$$P(X_t, D_t) = D_t \int_0^\infty e^{\frac{\sigma_D^2}{2}\tau + \alpha(\tau) + \beta_2(\tau)\sigma_t^2 + \beta_3(\tau)\lambda_t} d\tau. \tag{63}$$

For notational convenience, we define $G(\sigma_t^2, \lambda_t) \equiv \int_0^\infty e^{\frac{\sigma_D^2}{2}\tau + \alpha(\tau) + \beta_2(\tau)\sigma_t^2 + \beta_3(\tau)\lambda_t} d\tau$. Then

$$P(X_t, D_t) = D_t G(\sigma_t^2, \lambda_t). \tag{64}$$

5.4 Equity Premium

The instantaneous equity premium conditional on no jumps occurring in our economy, as shown in Appendix C, is given by

$$\begin{aligned} \mu_{P,t} + \frac{D_{t-}}{P_{t-}} - r_t &= \gamma\phi\sigma_t^2 - b_2 \frac{G_1}{G}\sigma_V^2\sigma_t^2 - b_3\frac{G_2}{G}\sigma_\lambda^2\lambda_t + \lambda_t E\Big[e^{b_2\xi_V}\big(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\big)\Big] \\ &= \sigma_{P,t}'\Lambda_t + \lambda_t E\Big[e^{b_2\xi_V}\big(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\big)\Big] \end{aligned} \tag{65}$$

where

$$\sigma_{P,t} = \big[\phi\sigma_t, \frac{G_1}{G}\sigma_V\sigma_t, \frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\big]' \tag{66}$$

where G_1 and G_2 respectively denotes the derivative of $G(\cdot,\cdot)$ with respect to σ_t^2 and λ_t, both evaluated at (σ_t^2, λ_t). Four components arise in order. The first term, $\gamma\phi\sigma_t^2$, represents a standard CRRA risk premium which arises from the compensation for the diffusion risk in consumption, dB_t^C. The second component, $-b_2\frac{G_1}{G}\sigma_V^2\sigma_t^2$, captures the compensation for the diffusion risk in volatility, dB_t^V. Appendix C shows that $\beta_2(\tau)$ is negative for all τ as long as $1 < \phi < 2\gamma$ which we actually will assume here and in our calibration, which immediately implies $G_1 < 0$ (i.e., the price-dividend ratio is decreasing in σ_t^2). Thus the second component takes a positive value. The third component, $-b_3\frac{G_2}{G}\sigma_\lambda^2\lambda_t$, which has a similar interpretation as the second one, stands for the compensation for the diffusion risk in volatility jump intensity, dB_t^λ. Appendix C shows that given $\beta_2(\tau)$ is negative $\beta_3(\tau)$ is also negative for all τ, which implies $G_2 < 0$ (i.e., the price-dividend ratio is decreasing in λ_t) and thus the third component also takes a positive value. The last term, differently, captures the compensation for the jump risk in volatility, $\xi_V dN_t$. It is positive noting $b_2 > 0$, $\xi_V > 0$ and $G_1 < 0$. Intuitively, at the times volatility jumps upward two things happen simultaneously: first, marginal utility jumps upward by a percentage equal to $e^{b_2\xi_V}$; second, the stock price jumps downward by a percentage equal to $1 - \frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}$. Therefore, equity is a risky, rather than hedging, asset, and thus investors need get compensated for the jump risks in holding it.

Note that (65) is the equity premium conditional on no jumps to volatility occurring. The instantaneous equity premium in population is given by (65) plus the expected percentage change of the equity price if a jump to volatility occurs. That's to say, the population equity premium in the economy is given by $\mu_{P,t} + \frac{D_{t-}}{P_{t-}} - r_t$ plus a negative term: $\lambda_t E\Big[\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)} - 1\Big]$. Finally, we can write the analytical expression for the population equity premium as

$$r_t^e - r_t = \sigma'_{P,t}\Lambda_t + \lambda_t E\Big[\big(e^{b_2\xi_V} - 1\big)\big(1 - \frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}\big)\Big]. \tag{67}$$

Note that the last term in (67) is still positive, implying that the positive compensation for jump risks dominates the direct negative effect of jumps on equity return.

Proposition 2. *In equilibrium, innovations to σ_t^2 and λ_t are both negatively priced; the price-dividend ratio $G(\sigma_t^2,\lambda_t)$ is strictly decreasing in both σ_t^2 and λ_t. Therefore, all sources of risks (diffusion and jump risks) in σ_t^2 and λ_t help contribute to a positive equity premium.*

5.5 VIX

Let $B_t = [B_t^C, B_t^V, B_t^\lambda]'$ and $B_t^Q = [B_t^{C,Q}, B_t^{V,Q}, B_t^{\lambda,Q}]'$. Appendix C shows that the dynamics of the log equity price under the risk-neutral measure Q is given by

$$d\ln P_t = \Big(\mu_{\ln P,t} - \sigma'_{P,t}\Lambda_t\Big)dt + \sigma'_{P,t}dB_t^Q + \sigma_D dB_t^D + \ln\big[\frac{G(\sigma_t^2+\xi_V^Q,\lambda_t)}{G(\sigma_t^2,\lambda_t)}\big]dN_t^Q, \tag{68}$$

where $\mu_{\ln P,t}$ denotes the drift term in $d\ln P_t$ under P measure. Intuitively, the drift of the growth of equity price is adjusted downward by $\sigma'_{P,t}\Lambda_t$ under Q relative to under P. Moreover, the jumps to volatility also have a greater negative impact on equity price under Q than under P, because, first, ξ_V^Q on average is greater than ξ_V, and second, the arrival intensity of jump is $l^Q = \varrho(b_2)\lambda_t$ under Q compared with just $l^P = \lambda_t$ under P. Also note that there is no difference between B_t^D and $B_t^{D,Q}$ since this source of risk is not priced.

Note that given the model parameters have an annual interpretation, VIX, as a measure of risk-neutral 30-day forward-looking market return volatility, can be expressed as[9]

$$VIX(X_t) = \text{Std}_t^Q\big[\ln P_{t+\frac{1}{12}}\big], \tag{69}$$

In order to express VIX as an explicit function in state variables, we follow Eraker and Wu (2017) in using the property of the conditional cumulant generating function for $\ln P_{t+\frac{1}{12}}$. It's thus convenient to first express $\ln P_t$ as a function affine in state variables. Define the log price-dividend ratio as $g(\sigma_t^2, \lambda_t) = \ln G(\sigma_t^2, \lambda_t)$. It then follows from (64) and a highly accurate log-linear approximation of the price-dividend ratio G that[10]

$$\begin{aligned} \ln P_t &= g(\sigma_t^2, \lambda_t) + \ln D_t \\ &\simeq \big(g^* - g_1^* \sigma^{2*} - g_2^* \lambda^*\big) + g_1^* \sigma_t^2 + g_2^* \lambda_t + \phi \ln C_t + \sigma_D B_t^D, \end{aligned} \tag{70}$$

where g_1 and g_2 denote partial derivatives, and letters with asterisks denote relevant functions or variables evaluated at steady states. It follows that

$$\begin{aligned} VIX^2(X_t) &= \text{Var}_t^Q\big[\ln P_{t+\frac{1}{12}}\big] \\ &= \text{Var}_t^Q\big[g_1^* \sigma_{t+1/12}^2 + g_2^* \lambda_{t+1/12} + \phi \ln C_{t+1/12} + \sigma_D B_{t+1/12}^D\big] \\ &= \text{Var}_t^Q\big[g_1^* \sigma_{t+1/12}^2 + g_2^* \lambda_{t+1/12} + \phi \ln C_{t+1/12}\big] + \frac{1}{12}\sigma_D^2. \end{aligned} \tag{71}$$

To compute the conditional variance above, we rely on the property of cumulant generating functions. Appendix C shows that by doing so we can write VIX-squared as a function affine in σ_t^2 and λ_t:

$$VIX^2(\sigma_t^2, \lambda_t) = a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t, \tag{72}$$

where $a_{1/12}$, $c_{1/12}$ and $d_{1/12}$ are three positive constants. Intuitively, the idiosyncratic noise σ_D only factors its impact on $a_{1/12}$, the constant component of VIX-squared: the equity market idiosyncratic

[9] Here, following Eraker and Wu (2017), we just model VIX^2 as the risk-neutral variance of 30-day log equity return. In Branger and Völkert (2012) and others, $VIX_t^2 = E_t^Q\big[\int_t^{t+1/12} (d\ln P_s)^2\big]$. These two methods should generate very similar level of VIX given the short horizon - one month.

[10] Seo and Wachter (2018) show that this type of log-linear approximation is highly accurate essentially because it is used only after the price-dividend ratio is exactly solved out, which is different from the Campbell-Shiller log-linear approximation usually used prior to solving the model as in many asset pricing papers. Moreover, we have verified that P_t is actually indistinguishably exponential affine.

noise affects the level of VIX in a fashion that is independent of the agent's attitude towards risk. Therefore, the VIX is given by

$$VIX(\sigma_t^2, \lambda_t) = \sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}. \tag{73}$$

Intuitively, equilibrium VIX is positively exposed to the two higher-order state variables: consumption growth volatility and volatility jump intensity. Because both factors command a negative market price of risk, so does VIX. This implies that in principle an asset with positive VIX exposure should earn itself a negative premium. Examples include VIX futures and VIX call options.

5.6 VIX Futures

To provide intuition that VIX futures do command a negative expected return in the model, we follow Eraker and Wu (2017) to consider the futures curve for VIX-squared, a hypothetical futures curve, given by

$$F^{VIX^2}(X_t;\tau) = a_{1/12} + c_{1/12}E^Q[\sigma_{t+\tau}^2|\sigma_t^2, \lambda_t] + d_{1/12}E^Q[\lambda_{t+\tau}|\sigma_t^2, \lambda_t]. \tag{74}$$

Appendix C proves the following proposition:

Proposition 3. *Suppose $\varrho(\cdot)$ is convex.*[11] *Then the VIX-squared futures curve, which expression is given in (C.52) through (C.54) in Appendix C, is upward slopping in steady-state at least for τ not too small.*

Because in steady state the expected change in VIX-squared under the objective measure is always zero for any horizon, proposition 3 implies that the expected return on holding the VIX-squared futures with a maturity not too short is negative, consistent with the evidence in Eraker and Wu (2017).

As shown, the VIX-squared futures price has a closed-form expression, but the VIX futures price does not. To compute $F^{VIX}(X_t;\tau)$, we need to deal with the square root in the expression of VIX and thus have to reply on the Fubini's theorem, the property of the characteristic function and finally a numerical integration. The pricing formula is given in Appendix $C.VII$.

5.7 Valuing Equity Options

Before pricing VIX options, let's first consider equity options, which values in the current framework are homogenous of degree one with respect to the underlying stock price. To facilitate computation, we

[11] For example, this is true if ξ_V is exponentially distributed.

divide the usual no-arbitrage equity option pricing equation through by the underlying stock price for a normalization. Specifically, let $P^E(X_t, \tau, K)$ denote the normalized price of a European put option written on the market index with maturity τ and normalized strike K. No-arbitrage then implies

$$P^E(X_t, \tau, K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(K - P_{t+\tau}/P_t\Big)^+\Big], \tag{75}$$

where P_t and $P_{t+\tau}$ are respectively the equity prices at t and $t+\tau$. Note that (70) reveals $P_{t+\tau}$ is exponential affine in $X_{t+\tau}$, which makes it convenient to use the generalized Fourier transform analysis (see, e.g., Lewis (2001)). Appendix C shows that by doing so we can finally write the put price as

$$P^E(X_t, \tau, K) = -\frac{1}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} e^{-iz(c^*-\ln P_t)-\frac{1}{2}\sigma_D^2 z^2\tau}\varrho_X^Q\big(-iz(\phi, g_1^*, g_2^*)', X_t, \tau\big)\frac{K^{iz+1}}{z^2 - iz}dz, \tag{76}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) < 0$, ϱ_X^Q represents the complex-valued discounted characteristic function defined in (54), and c^* is a constant given in Appendix $C.VIII$.

If we use integral variable substitution $x = z - z_i i$, a numerically implementable pricing formula obtains as

$$-\frac{1}{2\pi}\int_{-\infty}^{+\infty} Re\Big[e^{(z_i - xi)(c^*-\ln P_t)-\frac{1}{2}\sigma_D^2(x+iz_i)^2\tau}\varrho_X^Q\Big((z_i - xi)(\phi, g_1^*, g_2^*)', X_t, \tau\Big)\frac{K^{1-z_i+xi}}{(x+z_i i)(x+(z_i-1)i)}\Big]dx, \tag{77}$$

where we only need to take the real part of the complex-valued integrand because the put price is theoretically guaranteed to be real.[12]

5.8 Valuing VIX Options

A (European) VIX call option renders its holder the right, but not obligation, to obtain the difference between the VIX index[13] at an expiration date $t+\tau$ and a pre-specified strike K. Different from equity option, it is more convenient to directly compute VIX option price without normalization. No-arbitrage implies that the price of the VIX call option is given by

$$C^{VIX}(X_t, \tau, K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(VIX_{t+\tau} - K\Big)^+\Big], \tag{78}$$

[12]For $z_i > 1$, (76) and (77) are also the pricing formulas for an equity call option with maturity τ and normalized strike K.

[13]The standard convergence of futures price to the underlying price as time to maturity approaches zero holds regardless of whether the underlying is tradable or not.

where it follows from (73) that

$$VIX_{t+\tau} = VIX(\sigma^2_{t+\tau}, \lambda_{t+\tau}) = \sqrt{a_{1/12} + c_{1/12}\sigma^2_{t+\tau} + d_{1/12}\lambda_{t+\tau}}. \tag{79}$$

To calculate the expectation in (78), we again use the generalized Fourier transform analysis. Appendix C shows that by doing so we can finally write the VIX call price as

$$C^{VIX}(X_t, \tau, K) = \frac{1}{4\sqrt{\pi}} \int_{iz_i - \infty}^{iz_i + \infty} e^{-iza_{1/12}} \varrho_X^Q\Big(-iz(0, c_{1/12}, d_{1/12})', X_t, \tau\Big) \frac{Ercf\big(K\sqrt{-iz}\big)}{(-iz)^{\frac{3}{2}}} dz, \tag{80}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) > 0$, ϱ_X^Q represents the complex-valued discounted characteristic function defined in (54), and $Ercf(\cdot)$ is the complex-valued complementary error function which expression is given in Appendix $C.IX$.

If we use integral variable substitution $x = z - z_i i$, a numerically implementable pricing formula obtains as

$$\frac{1}{4\sqrt{\pi}} \int_{-\infty}^{+\infty} Re\Big[e^{(z_i - xi)a_{1/12}} \varrho_X^Q\Big((z_i - xi)(0, c_{1/12}, d_{1/12})', X_t, \tau\Big) \frac{Ercf(K\sqrt{z_i - xi})}{(z_i - xi)^{\frac{3}{2}}} \Big] dx, \tag{81}$$

where we only need to take the real part of the complex-valued integrand because the VIX call price is theoretically guaranteed to be real.[14] We then use equation (C.76) to compute VIX put price.

6 Quantitative Analysis

In the following, we perform parameter calibration for our model with the target toward replicating salient features of consumption, equity, and VIX derivatives markets data.

6.1 Calibration

Table 3 displays our choices of the model parameters. To facilitate better comparison with a number of recently developed continuous-time asset pricing models, in our model time is measured in years and parameter values should be interpreted accordingly. A rate of time preference β equal to 2% per annum and an expected consumption growth μ equal to 3% per annum together help give rise to an average real

[14] In the numerical section, we have used both the Riemann rule and the quadrature rule to approximate the integral. They generate the same result. We also have compared the price obtained via integral with that via Monte Carlo simulation. We found the difference is negligible as long as the VIX option is not "too OTM".

yield on one-year Treasury Bill of 1.17%, which is roughly consistent with that documented in Bansal and Yaron (2004), 0.86%. Here μ is set relatively high because we anticipate a relatively large risk aversion or a relatively high mean volatility through calibration, as at least one of them is needed to produce large premia on VIX derivatives as those seen in the data. We set the value of θ^V, the average annualized consumption growth variance without jumps, to be 0.0004, which corresponds to a volatility of 2% per annum, consistent with that used in Wachter (2013). With jumps, the average volatility would be greater than 2%, but cannot be so by too much, in order to be still consistent with the U.S. consumption growth data. A reasonable range of values for the U.S. consumption growth volatility that most historical data agree upon is $1-3\%$. For example, Bansal and Yaron (2004) document a volatility of 2.93% while Wachter (2013) documents a volatility of 1.34%.

Consistent with the literature, the stock market leverage ϕ is calibrated at 2.7, a value between that in Bansal and Yaron (2004), 3, and that in Wachter (2013), 2.6. This value of leverage works well overall in terms of explaining various market data. Implicitly, the IES, which value constitutes a source of debate, is set to one, for tractability. Plus, as discussed in Wachter (2013), a number of studies conclude that the reasonable values for this parameter should be somehow close to one (e.g., Vissing-Jørgensen (2002); Hansen, Heaton, and Li (2008); Thimme (2017)).

θ^λ has the interpretation as the average probability of a jump in consumption volatility per annum. The parameter is hard to identify from monthly consumption data alone. However, studies from equity market data, such as Eraker, Johannes, and Polson (2003), suggest that jumps in equity market return volatility are on average 1.5 times per year, and starting from the mean level of volatility, an average sized jump in volatility increases volatility from 15% to 24% from a posterior perspective. Given that jumps in consumption volatility translate one-to-one into jumps in equity price and return in our model, we are a little more conservative in setting the average jump probability to be once every other year ($\theta^\lambda = 0.5$) with each jump having a larger impact on equity volatility.

We choose μ_ξ such that in equilibrium, starting from the steady-state level of VIX, an average-sized jump in volatility increases VIX from 20.9 to 32.6, which is consistent with an average size of jump in VIX, 11.4, computed from historical VIX data from CBOE during the period 1990-2019.[15] The unconditionally average consumption growth volatility is equal to the square root of $\bar{\sigma}_t^2 = \theta^V + \frac{\mu_\xi \theta^\lambda}{\kappa^V}$. Given θ^V, μ_ξ and θ^λ have been fixed, κ^V is set to 2.5, implying a monthly autocorrelation of 0.8 in VIX which compares to 0.84 in the data. Our chosen consumption volatility parameters imply an average

[15]The number 11.4 is obtained as follows: we take monthly data of the VIX index from CBOE, identify all the months during which VIX rises, sort all those rises in descending order, and then take the mean of the largest 15. Given the time period considered 1990-2019, the number 15 is consistent with our earlier calibration that jumps are on average once every other year.

consumption volatility of 3.06%, which is slightly higher than 2.93% as documented in Bansal and Yaron (2004), and higher than 1.34% in Wachter (2013).

[Table 3 about here.]

In order to calibrate the other model parameters, notably risk-aversion (γ), the diffusion parameter of the volatility (σ_V), the mean reversion of the jump intensity (κ^λ), the diffusion parameter of the jump intensity (σ_λ), and the dividend growth idiosyncratic volatility (σ_D), we design a coarse Simulated Methods of Moments procedure. Specifically, we search the parameter space to overall best match the following six data moments: mean VIX (19.3); standard deviation of VIX (7.4); average monthly return on holding one-month ATM VIX call option (−26%); average one-month ATM VIX call option Black'76 implied volatility (0.69); monthly autocorrelation of one-month ATM VIX option Black'76 implied volatility (0.27); contemporaneous correlation between VIX and one-month ATM VIX option Black'76 implied volatility (0.48). The values and data sources for these moments are summarized in Tables 2 and 4. We are able to match a majority of these moments well.

We calibrate risk aversion at 14, which is slightly higher than that in Bansal and Yaron (2004) (10), Drechsler and Yaron (2011) (9.5), and higher than that in Wachter (2013) (3) and that in Eraker and Wu (2017) (8). Intuitively, the high risk aversion arises from the effort to reconcile sizable premia on VIX derivatives (high risk prices) with a low consumption growth volatility (low risk prices).

σ_V is calibrated at 0.16. Obviously, as a volatility-of-volatility parameter, it heavily influences VIX volatility, VIX derivatives premia, the probability distribution of VIX, and thus the contemporaneous correlation between VIX and one-month ATM VIX option Black'76 implied volatility. The parameter σ_V is again not a substitute for risk aversion since a too large σ_V would make the model behave like a single-factor model. κ^λ is calibrated at a high value, 12, in an effort to match a low monthly autocorrelation of one-month ATM VIX option Black'76 implied volatility. Note that the latter is not monotonically decreasing in the former because as κ^λ increases, the second factor, λ_t, exerts a weaker impact on VIX in equilibrium. Finally, we set $\sigma_\lambda = 2.6$ and $\sigma_D = 0.1$ in order to match mean VIX and various VIX derivative premia and implied volatility. Note that VIX derivative premia are generally decreasing with σ_D as the idiosyncratic risk contained in σ_D is not priced in equilibrium and thus only contributes to the constant component of the VIX index, thereby decreasing VIX return's exposure to σ_t^2 and λ_t.

6.2 Simulation Results

6.2.1 General Moments

Table 4 displays a list of moments from a simulation of the model at calibrated parameters, as well as their counterparts in U.S. data. The model is discretized using an Euler approximation and simulated at a monthly frequency ($dt = 1/12$) for 100,000 months. Simulating the model at higher frequencies produces negligible differences in the results. We then aggregate the data to compute the model moments which are largely reported on a monthly or annual basis. As seen in the table, we match a majority but not all of the key moments that we are interested in. In particular, we match average consumption growth volatility fairly well: 3.06 in the model vs. 2.93 in the data. In terms of equity premium, we overshoot slightly, as our model produces 8.83% per annum. This compares to 8.33% in the CRSP distributed in Ken French's publicly available Mkt-Rf time-series. Our model produces an unconditional stock market volatility of 17.75% which compares to 18.31% in post-1990 S&P 500 data.[16] Our model generates an average (one-year) risk-free rate and risk-free rate volatility on par with what we see in the data, though the volatility of the risk-free rate is a bit higher.

Our model does not match the observed (log) price-dividend ratio very well. Empirically observed p/d ratios vary substantially over time, and display an annualized autocorrelation that exceeds anything we could expect to generate with our model. This is a natural consequence of the fact that our model structure is geared toward explaining derivatives data and calibrated to do so at a relatively high frequency. Price/dividend ratios display annualized persistence that way exceeds that seen in high-frequency derivatives-based variables such as VIX and VVIX. Adding additional state-variables, such as in models by Campbell and Cochrane (1999) (habit) and Bansal and Yaron (2004) (long-run persistent consumption growth) will likely alleviate some of this but is outside the focus of the present paper.[17]

[Table 4 about here.]

Importantly, for the purposes of our study, we match the mean and standard deviation of the VIX index almost exactly. The fact that mean VIX (19.46) is higher than equity return objective volatility (17.75) illustrates the model's ability to generate a large Variance Risk Premium. The monthly

[16] We compare our estimate to SPX volatility using data collected after 1990 as to make the estimate comparable to average VIX. The CRSP value-weighted index return over the risk-free rate has an annual volatility of about 20.30% using data from 1927-2018.

[17] We actually solved and recalibrated an extended version of the present model in which the introduction of persistent long-run growth risks help alleviate the issue on pd substantially while all other moments are largely unharmed. This version is available upon request.

autocorrelation of the simulated VIX index is 0.8, which compares to 0.84 in the data. Turning to the model's ability to match key moments of VIX options data, we see that the average implied volatility for one-month ATM VIX options, denoted $E(imp\_vol_t)$, is estimated at 71.83 in the model simulations which compares to 68.8 in the data. The model produces a volatility of the simulated VIX implied volatility, denoted $\sigma(imp\_vol_t)$, of 12.74 vs. 14.3 in the data - a slight miss on the low side. The model also produces a monthly autocorrelation of one-month ATM VIX implied volatility, denoted $AC_1(imp\_vol_t)$, of 0.49 vs. 0.27 in the data. This is quite a miss even though we have tried to increase κ^λ to make the VIX implied volatility process exhibit weak persistence, while κ^V has to be kept small so as to take care of the persistent VIX index process and other aspects of the model.

Our model matches the observed positive but imperfect correlation between vol-of-vol and VIX at 0.34 vs. 0.48 in the data. It is useful to consider this in relation to a model where time-varying arrival intensity of jumps is constant $\lambda_t = \lambda$. In this case, square VIX would be a linear function of σ_t^2 and thus derive its properties. From this it follows that the local variance of VIX^2 will be a linear function of σ_t^2, or equivalently VIX_t. This again implies that VIX options implied volatility (or simply vol-of-vol) should be (either positively or negatively) perfectly correlated with VIX itself.

Why is a fluctuating λ_t process important in shaping the positive correlation between vol-of-vol and VIX? When λ_t increases, first, it drives VIX up as VIX loads positively on it; second, it also drives up vol-of-vol and thereby increasing the prices of VIX Call options. Huang, Schlag, Shaliastovich, and Thimme (2019) propose a model where stock market spot variance follows a mean-reverting process which volatility is driven by an independent diffusion process. The independence assumption implies that the correlation between VIX^2 and volatility-of-volatility (or VVIX) is zero[18]. Huang, Schlag, Shaliastovich, and Thimme (2019) also present empirical evidence suggesting that VIX and vol-of-vol carry negative risk premia which is true in our model: VIX is an increasing function of σ_t^2 and λ_t, both of which have negative market risk prices, so has VIX. Since vol-of-vol is positively correlated with VIX, it too carries a negative risk price.

6.2.2 VIX Futures Returns

[Table 5 about here.]

Table 5 compare average returns and return standard deviations for VIX Futures prices computed from data (see Eraker and Wu (2017)) and our model. We report average daily arithmetic and loga-

[18]The HSST model implies that VIX^2 is a linear function of stock market variance, V_t. It follows that we can write $dVIX_t^2 = (a + bVIX_t^2)dt + c\sqrt{\eta_t}dW^2$ where η_t is a mean-reverting diffusion independent of V_t and therefore VIX_t^2.

rithmic returns, and return standard deviations.[19] For one-month contracts, both log and arithmetic returns are ballpark the same for the model as in the data. At longer horizons, the model generates a too high (negative) risk premium. This is well known in the variance-risk literature. In fact, Dew-Becker, Giglio, Le, and Rodriguez (2017) report positive returns to long-maturity variance swaps, a finding that cannot be reconciled with a negative volatility risk premium. Our model also matches the observed daily return standard deviations of VIX futures almost exactly, although these moments were never targeted in our parameter calibration.

[Figure 4 about here.]

To better understand how negative average VIX futures returns are generated in the model, Figure 4 shows the expected returns under different market conditions (low vs. high VIX). As in Eraker and Wu (2017), Fig.7, our model generates a consistent positive difference between the Q (risk-neutral) and P (objective) expected path of VIX, irrespective of the initial condition. Since expected returns are given by $E_t^P(VIX_{t+\tau})/E_t^Q(VIX_{t+\tau}) - 1$, this implies that expected VIX futures returns are always negative in our model.

6.2.3 VIX Option Implied Volatilities

[Table 6 about here.]

Table 6 looks at the average VIX option implied volatilities conditioning on options in consideration being ATM, ITM, and OTM. For example, $x\%$ means that we compute the implied volatility of VIX options conditional upon the strike being equal to $x\%$ of the same maturity futures price. This part can be examined together with Table 7, which by contrast averages implied volatilities across the absolute value of strike and is thus directly comparable to our Table 1. The patterns we see are similar across Tables 6 and 7. First, for VIX calls, low (high) strikes, either measured by moneyness or absolute value, generally have lower (higher) implied volatilities, suggesting a positively skewed underlying distribution. Second, shorter (longer) maturities uniformly have higher (lower) implied volatilities, suggesting that in the model the conditional variance of log VIX futures price increases with the horizon of the futures slower than linearly, irrespective of market conditions.

Focusing on the comparison between Tables 1 and 7, at the short maturities and low strikes our model seems to undershoot implied volatility by a bit, as we see for example a strike of 20 averaging 104%

[19] Here the model is re-simulated at a daily frequency, to facilitate comparison with the results in Eraker and Wu (2017).

implied volatility in the data vs. 87% in our model. This is not a large deviation when considering that the size of the bid-ask spread often times exceeds 20 implied volatility points - see Figure 1. However, our model implied volatility gradually catches up with the data as strike increases, as we see a strike of 30 averaging 125% in the data vs. 121% in the model and strike of 40 averaging 132% in the data vs. 134% in the model. At the six month horizon our model slightly overshoots implied volatilities by roughly 10 percentage points across all strikes. This is evidence that mean reversion of VIX in the model is slightly slower than that in the data. On the other hand, the autocorrelation of VIX in the model (0.80) is smaller than that in the data (0.84), suggesting, on the contrary, that VIX mean reversion in the model is slightly faster than that in the data. The current parameter choices reflect a balance struck between those tensions.

[Table 7 about here.]

Figures 5 and 6 further illustrate the implied volatility patterns generated by our model. While Figure 5 shows the steady state implied volatilities and pretty much illustrates the patterns in the previous two tables, Figure 6 shows what happens when we condition upon a high and low initial VIX, respectively. In the low VIX case, where we have set the initial state variables very low so as to generate a VIX of 12.5, we see that the implied volatility curves are almost everywhere increasing and concave. This closely resembles the patterns we saw in the data on March 26, 2017 (Figure 1 bottom). The concave implied VIX-volatility seen in the low-VIX state is related to the fact that in order to generate a low VIX, both state-variables need to be low, implying that the second, the jump arrival intensity, is low. This negates the effect of jumps, as they are rare and do not tremendously fatten the right tail of the conditional distribution, which is therefore dominated by the two diffusive processes. To understand the concavity of VIX implied volatility seen both in the data and in our model in low and normal VIX periods, we note the following. Implied volatility is computed from Black's model for options on futures, which is identical to Black-Scholes with the exception that the underlying is a futures. Thus, the model will generate a concave Implied Volatility function over some strike range, say $[K_1, K_2]$, if the log-normality displays heavier tails than the density of the underlying (i.e., VIX). It follows therefore that if squared-VIX were to be log-normally distributed, options struck on VIX would display a concave Implied Volatility function for strikes that exceed some threshold. This follows from the fact that if x_T is log-normal then a call on its square root, $e^{-r(T-t)}E_t^Q[\max(\sqrt{x_T} - k, 0)]$, will have concave implied volatility for strikes k above a threshold $k > k_0$. As a pure mechanical fact, therefore, VIX options will exhibit a concave Implied Volatility function above some threshold strike price if the risk-neutral conditional density of squared VIX is close to a log-normal one.

The top panel of Figure 6 shows Implied Volatility for VIX options under a high initial VIX. We see that the implied volatility curves have changed to something that looks almost flat and marginally convex especially at the left end. Again, this strikingly resembles the data we see on November 12, 2008 (Figure 1 top). In order to generate a high VIX we must simultaneously have both high spot volatility σ_t and hight jump arrival intensity λ_t. The fact that the probability of a jump arrival is high will fatten right tail of the conditional distribution for VIX. This generates a heavy right tail. The fact that σ_t is high also increases the volatility of σ_t^2 itself through the square root diffusion term. The conditional risk-neutral distribution of VIX-squared is a mixture distribution which first component (σ_t^2) has large variance, and thus, generates heavier tails than a (log) normal. The effect of fattening the tails dissapates with maturity, as the risk-neutral conditional distribution converges toward the stationary stationary distribution.

[Figure 5 about here.]

[Figure 6 about here.]

6.2.4 VIX Option Returns

Figure 7 reports average returns on holding VIX call and put options to maturities. All returns are normalized monthly. Some striking patterns are as follows. First, the model generates a negative (positive) premium for VIX call (put) options, intuitively because the payoff of VIX call (put) is a positive (negative) bet on σ_t^2 and λ_t both of which are negatively priced in equilibrium. In other words, for market participants VIX call options are insurances against possible spikes in σ_t^2 and λ_t and thus a negative premium is generated. Second, the model implies that, ceteris paribus, shorter maturity VIX options always carry a greater premium than longer maturity ones, showing that the shorter the maturity is, the more excessively expensive (cheap) the call (put) is. This is in principle consistent with the downward sloping term structure of VIX option implied volatility shown in Table 6, and is true in the data as we will analyze shortly. Third, the premia for both call and put are decreasing with moneyness, showing that the more out-the-money the VIX call (put) is, the more pronounced its role as a bet on (against) volatility or volatility-of-volatility. This is in principle consistent with the upward sloping VIX option implied volatility curve across moneyness shown in Table 6.

[Figure 7 about here.]

Table 8 further reports returns to VIX options in our model with greater details. ITM (OTM) here indicates 15% in-the-money (out-the-money). Given VIX futures price is most of the time close to 20, 15% corresponds to 3 points, so that this table is directly comparable with Table 2. A few comments on the similarities and dissimilarities between the data averages and the model are in order. First, the model generates negative returns to call options, and positive returns to put options. This is consistent with the data. The model generates negative short maturity (1 month) call returns ranging from -14% (ITM) to -35% (OTM) which compares to -23% for both ITM and OTM's in the data. We need to keep in mind here that over the course of our sample, VIX spiked dramatically on several occasions, including the financial crisis of 2008 but also the events of February 5th, 2018 discussed in Section 3.2, leading to positive returns for OTM VIX calls. For the longer maturity six-month calls we see that the average returns in the model are somewhat larger than those in the data. Another commonality between the model and the data is, the term structure of (absolute) returns on VIX call is downward sloping. For example, for ATM call, the monthly return is -0.24 for one-month vs. $-0.52/6 \approx -0.09$ for six-month.

[Table 8 about here.]

The model generates average returns on ITM, ATM and OTM short maturity put options that closely resemble those computed from data in Table 2. For longer maturity 6-month put options, our model does generate positive average returns, but they are significantly smaller than those in the data. In a nutshell, our model tends to overshoot premia on long-maturity calls while undershooting premia on long-maturity puts.

Turning to an examination of higher-order return moments, we see that our model generates patterns that are strikingly similar to what we estimate from data. For example, one-month maturity call returns have an estimated standard deviation of 116% in the data vs. 122% in the model for in-the-moneys, 199% vs. 209% (ATM), and 337% vs. 309% (OTM). The return standard deviations are equally similar for the short maturity puts. At the longer maturities, our model slightly undershoots return standard deviations for calls as well as for puts. We also match the estimated skewness and kurtosis coefficients fairly closely. One exception here is the large kurtosis coefficient computed from OTM one-month maturity calls and puts: for the call, this coefficient is 127.73 in the model vs. 82.61 in the data; for the put, it is 28.57 in the model vs. 9.68 in the data. In interpreting these deviations the reader should keep in mind that higher order moments such as skewness and kurtosis are difficult to accurately estimate from the relatively short sample of options returns.

6.3 Comparative Static Analysis

[Figure 8 about here.]

In order to gain some additional insights into the workings of our model, we report the results of some comparative statics. In doing so, we also emphasize the necessity of recursive preferences ($\gamma \neq 1/\psi$) for the model to generate non-zero VIX derivatives premia. The left subplot of Figure 8 illustrates the co-movements of six important conditional model moments with risk aversion. As shown, the equity premium is sensitive to risk aversion universally: the equity premium increases with risk aversion almost linearly when the latter is relatively low; and increasingly fast when the latter becomes higher. Recall from equation (67) that equity premium reflects compensations for three sources of risks corresponding to the model's three state variables. The pattern of the equity premium's variation with risk aversion reflects the fact that market price of risk for consumption growth increases linearly with γ, whereas the market prices of risks for volatility and its jumping risk only increase slowly with γ at the beginning and increasingly faster afterwards.

The right subplot of Figure 8 shows the impact of risk-aversion, γ, on the market risk prices associated with the three state-variables, $\lambda = (\gamma, -b_2, -b_3)$. Since the representative agent has recursive preferences, he is concerned about variation in his value function in the future which is affected by risks in σ_t^2 and λ_t. Both state variables therefore enter the agent's pricing kernel and are priced in equilibrium. However, these two state variables are by nature higher-order. Specifically, σ_t^2 measures the spot variance of consumption growth and thus is a second-order moment in terms of its relation with consumption, while λ_t governs the arrival intensity of jump in σ_t^2 (i.e., jumping vol-of-vol) and has third order effects on consumption. Accordingly, market risk prices associated with σ_t^2 and λ_t increase relatively slowly with risk aversion.

Turning back onto the left subplot of Figure 8, it remains to check how other moments, besides the equity premium, vary with risk aversion. Steady-state VIX is increasing with γ, manifesting the former's dependence on state variables that are priced in equilibrium. VIX is a risk neutral measure of market volatility. A larger risk aversion implies a higher market price of risk associated with σ_t^2 and λ_t, and thus a higher risk-neutral persistence and mean of σ_t^2 and λ_t. This implies that the average value of VIX increases relative to objective variance, or put differently the Variance Risk Premium increases[20].

The left subplot of Figure 8 also shows the steady-state return premia on one-month ATM VIX Put and Call options. As VIX call (put) option is a (negative) volatility claim, it earns a negative (positive)

[20]Note that while both risk neutral and physical stock market variance increase with γ, risk-neutral variance increases faster.

premium. Both premia increase in absolute value with γ asymmetrically: Puts increase less than the calls decrease. Thus, more risk aversion leads investors to have to pay a comparably higher premium for the crash insurance offered by VIX calls than the positive premium seen for VIX puts.

Figure 8 also speaks to the necessity of the recursive preference assumption in generating non-zero risk premia on VIX derivatives, since all premia are exactly zero when $\gamma = 1$, in which case the Duffie-Epstein recursive preferences collapse into the CRRA preferences. With the latter neither σ_t^2 nor λ_t would be priced in equilibrium, implying that any claim with mere exposure to σ_t^2 and λ_t would be priced by market investors simply according to the objective measure and thus earns a zero premium. All VIX derivative assets are examples of such a claim.

7 Concluding Remarks

This paper studies the properties of VIX derivatives prices, including the returns to buy-and-hold VIX options positions. We document a negative return premia consistent with a negative price of volatility and volatility jump risk. Our paper follows the by now well established literature on consumption-based asset pricing models where persistent state dynamics generate risk premia that exceed those seen under time separable preferences by separating risk-aversion from intertemporal elasticity of substitution, as in Bansal and Yaron (2004), Eraker and Shaliastovich (2008), Drechsler and Yaron (2011), Wachter (2013) and many others. Our theoretical formulation mirrors the general framework outlined in Eraker and Shaliastovich (2008), but has the advantage that it does not require any linearization approximations.

We use this modeling framework to specify a model in which consumption volatility follows a jump-diffusion. This is different from the consumption disaster literature, as for example Barro (2006) or Wachter (2013), where disasters occur in consumption itself. Our model features a time-varying consumption volatility and also a time-varying intensity of jumps in that volatility process. The fact that there are two state variables that drive the aggregate consumption volatility implies that stock market variance is actually a linear combination of these two. This is also the case for (squared) VIX, defined here as forward-looking 30-day risk-neutral stock market volatility. Although we do not derive a closed form expression for implied VIX volatility, it is obvious that this too is a two factor process that depends on the two fundamental factors, σ_t^2 (consumption volatility) and λ_t (jump arrival intensity). The fact that both physical stock market volatility and VIX depend on these two factors implies that they are imperfectly correlated with consumption volatility. This is obviously also true in the data. Moreover, consistent with the finding in Huang, Schlag, Shaliastovich, and Thimme (2019), VIX and implied VIX volatility (i.e., vol-of-vol) are two imperfectly correlated and negatively priced factors in our model.

Our model replicates many of the observed characteristics of asset market data: it is within striking distance of the equity premium, stock market volatility, the variance risk premium, the correlation between VIX and VVIX, but most importantly for our purposes, it appears to replicate some of the features we observe in the VIX derivative markets data with surprising accuracy. First off, it replicates large negative average returns to VIX futures. Second, it replicates with acceptable degree of accuracy the return premia seen in VIX options data. This includes the higher order moments. Third, we replicate the general shape of VIX option implied volatility functions, including the positive skewness and downward sloping term structure.

In equity and variance swap options it is well known that implied volatilities exhibit convexity (i.e. smile) over strikes. In our VIX option data, the smile is actually a concave frown for the most part of our sample, and particularly so when VIX is low. When VIX is high, it surprisingly changes to a convex smile. Even more surprisingly our model actually replicates this empirical phenomenon.

More work is needed for us to understand exactly what the forces are that drive derivative prices during periods of extreme market stress. Our model obviously endogenizes returns to equity and with it, returns to equity options. There are numerous studies of equity options returns; we know that OTM put options are expensive and that their buyers experience high negative expected returns. We leave the issues of whether the model proposed here can address this, and interestingly, whether the prices of catastrophe insurance implied by prices of OTM equity options are more or less comparable to the price of catastrophe insurance offered by VIX derivatives, in particular VIX call options, as future research topics.

References

Abel, Andrew B, 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

Bansal, Ravi, and Ivan Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *The Review of Financial Studies* 26, 1–33.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481–1509.

Barro, Robert J, 2006, Rare disasters and asset markets in the twentieth century, *The Quarterly Journal of Economics* 121, 823–866.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, *Journal of Financial Economics* 101, 552–573.

Black, Fischer, 1976, The pricing of commodity contracts, *Journal of financial economics* 3, 167–179.

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

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

Branger, Nicole, and Clemens Völkert, 2012, What is the Equilibrium Price of Variance Risk? A Long-Run Risks Model with Two Volatility Factors, *A Long-Run Risks Model with Two Volatility Factors (September 9, 2012).*

Campbell, John Y, 2003, Consumption-based asset pricing, *Handbook of the Economics of Finance* 1, 803–887.

Campbell, John Y, and John H Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of political Economy* 107, 205–251.

Campbell, John Y., and Robert J. Shiller, 1988, Stock Prices, Earnings and Expected Dividends, *Journal of Finance* 43, 661–676.

Coval, D. J., and T. Shumway, 2001, Expected Option Returns, *Journal of Finance* 56, 983–1010.

Cox, J. C, J. E. Ingersoll, and S. A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385–407.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123(2), 225?250.

Drechsler, Itamar, and Amir Yaron, 2011, What's Vol got to do with it, *Review of Financial Studies* 24, 1–45.

Duffie, Darrell, and Larry Epstein, 1992, Stochastic Differential Utility, *Econometrica* 60, 353–394.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical finance* 6, 379–406.

Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Duffie, Darrell, and Costis Skiadas, 1994, Continuous-time security pricing: A utility gradient approach, *Journal of Mathematical Economics* 23, 107–131.

Eraker, Bjørn, 2012, The Volatility Premium, *working paper*.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *The Journal of Finance* 58, 1269–1300.

Eraker, Bjørn, and Ivan Shaliastovich, 2008, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* 18-4, 519–543.

Eraker, Bjørn, and Yue Wu, 2017, Explaining the Negative Returns to VIX Futures and ETNs: An Equilibrium Approach, *Journal of Financial Economics* 125, 72–98.

Feller, William, 1951, Two singular diffusion problems, *Annals of mathematics* pp. 173–182.

Griffin, John, and Amin Shams, 2018, Manipulation in the VIX?, *Review of Financial Studies* 4(1), 1377–1417.

Hansen, Lars Peter, John C Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political economy* 116, 260–302.

Heston, Steve, 1993, Closed-Form Solution of Options with Stochastic Volatility with Application to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Huang, Darien, Christian Schlag, Ivan Shaliastovich, and Julian Thimme, 2019, Volatility-of-Volatility Risk, *Journal of Financial and Quantitative Analysis* forthcoming.

Jackwerth, J. C., and M. Rubenstein, 1996, Recovering Probability Distributions from Option Prices, *Journal of Finance* 51, 1611–1631.

Lewis, Alan, 2001, A simple option formula for general jump-diffusion and other exponential Lévy processes, *Working Paper*.

Mencía, Javier, and Enrique Sentana, 2013, Valuation of VIX derivatives, *Journal of Financial Economics* 108, 367–391.

Park, Yang-Ho, 2015, Volatility-of-Volatility and Tail Risk Hedging Returns, *Journal of Financial Markets* 26, 38–69.

Park, Yang-Ho, 2016, The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives, *Journal of Econometrics* 192, 313–328.

Seo, Sang Byung, and Jessica A Wachter, 2018, Option prices in a model with stochastic disaster risk, *Management Science*.

Tauchen, George, 2011, Stochastic volatility in general equilibrium, *The Quarterly Journal of Finance* 1, 707–731.

Thimme, Julian, 2017, Intertemporal substitution in consumption: A literature review, *Journal of Economic Surveys* 31, 226–257.

Tsai, Jerry, and Jessica A Wachter, 2018, Pricing long-lived securities in dynamic endowment economies, *Journal of Economic Theory* 177, 848–878.

Vissing-Jørgensen, Annette, 2002, Limited Stock Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy* 110, 825–853.

Wachter, Jessica A, 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *The Journal of Finance* 68, 987–1035.

Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *The Quarterly Journal of Economics* 105, 29–42.

Whaley, Robert E, 2013, Trading volatility: At what cost, *Journal of Portfolio Management* 40, 95–108.

Internet Appendix

Internet Appendix A: Solutions to the General Model

A.I. Value Function

The value function $J(W, X)$ satisfies

$$\sup_{\alpha_t, C_t} \Bigg\{ J_W \big(W_t \alpha_t \big(\delta_c' \mu(X_t) + \frac{1}{2} \delta_c' \Sigma(X_t) \Sigma(X_t)' \delta_c + \frac{1}{A} - r_t \big) + W_t r_t - C_t \big) + J_X' \mu(X_t)$$

$$+ \frac{1}{2} tr \left(\begin{bmatrix} J_{XX} & J_{XW} \\ J_{XW}' & J_{WW} \end{bmatrix} \begin{bmatrix} \Sigma(X_t)\Sigma(X_t)' & W_t \alpha_t \Sigma(X_t) \Sigma(X_t)' \delta_c \\ W_t \alpha_t \delta_c' \Sigma(X_t) \Sigma(X_t)' & W_t^2 \alpha_t^2 \delta_c' \Sigma(X_t) \Sigma(X_t)' \delta_c \end{bmatrix} \right)$$

$$+ E_{\xi_1} \big[J(W_t + W_t \alpha_t (e^{\xi_1} - 1), X_{t,1} + \xi_1, X_{t,-1}) \big] l_1(X_t)$$

$$+ \sum_{i=2}^{n} E_{\xi_i} \big[J(W_t, X_{t,i} + \xi_i, X_{t,-i}) \big] l_i(X_t) - J(W_t, X_t) \mathbf{1}' l(X_t) + f(C_t, J(W_t, X_t)) \Bigg\} = 0. \qquad \text{(A.1)}$$

In equilibrium, risk-free asset market clears: $\alpha_t = 1$, and the consumption claim market clears: $C_t = A^{-1} S_t = A^{-1} W_t$. Substituting these policy functions into (A.1) implies

$$J_W W_t \Big(\delta_c' \mu(X_t) + \frac{1}{2} \delta_c' \Sigma(X_t) \Sigma(X_t)' \delta_c \Big) + J_X' \mu(X_t) + \frac{1}{2} tr \Big(J_{XX} \Sigma(X_t) \Sigma(X_t)' + W_t J_{XW} \delta_c' \Sigma(X_t) \Sigma(X_t)' \Big)$$

$$+ \frac{1}{2} J_{XW}' \Sigma(X_t) \Sigma(X_t)' \delta_c + \frac{1}{2} W_t^2 J_{WW} \delta_c' \Sigma(X_t) \Sigma(X_t)' \delta_c + E_{\xi_1} \big[J(W_t e^{\xi_1}, X_{t,1} + \xi_1, X_{t,-1}) \big] l_1(X_t)$$

$$+ \sum_{i=2}^{n} E_{\xi_i} \big[J(W_t, X_{t,i} + \xi_i, X_{t,-i}) \big] l_1(X_t) - J(W_t, X_t) \mathbf{1}' l(X_t) + f(A^{-1} W_t, J(W_t, X_t)) = 0. \qquad \text{(A.2)}$$

Conjecture that the solution to this equation takes the form

$$J(W, X) = \frac{W^{1-\gamma}}{1-\gamma} I(X). \qquad \text{(A.3)}$$

It is helpful to first solve for the wealth-consumption ratio prior to solving for $I(X)$. By definition

$$f(C, V) = \beta(1-\gamma) V (\ln C - \frac{1}{1-\gamma} \ln((1-\gamma)V)). \qquad \text{(A.4)}$$

Therefore,

$$f_C(C, V) = \beta(1-\gamma) \frac{V}{C}. \qquad \text{(A.5)}$$

The envelop condition $f_C = J_W$, together with the derivative (A.5), the conjecture (A.3) and that in equilibrium $J = V$, imply

$$\beta(1-\gamma)\frac{W^{1-\gamma}}{1-\gamma}I(X)\frac{1}{A^{-1}W} = W^{-\gamma}I(X).$$

Solving for A yields $A = \beta^{-1}$. Given the wealth-consumption ratio, it follows that

$$f(\beta W, J(W,X)) = \beta W^{1-\gamma}I(X)\Big(\ln(\beta W) - \frac{1}{1-\gamma}\ln(W^{1-\gamma}I(X))\Big)$$

$$= \beta W^{1-\gamma}I(X)\Big(\ln\beta - \frac{\ln(I(X))}{1-\gamma}\Big). \tag{A.6}$$

Now substituting (A.3) and (A.6) into (A.2) yields

$$I(X_t)\big(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c\big) + \frac{1}{1-\gamma}I_X(X_t)\mu(X_t)$$

$$+tr\Big(\frac{1}{1-\gamma}I_{XX}(X_t)\Sigma(X_t)\Sigma(X_t)' + I_X(X_t)\delta_c'\Sigma(X_t)\Sigma(X_t)'\Big)$$

$$+\frac{1}{2}I_X(X_t)'\Sigma(X_t)\Sigma(X_t)'\delta_c - \frac{1}{2}\gamma I(X_t)\delta_c'\Sigma(X_t)\Sigma(X_t)\delta_c$$

$$+\frac{1}{1-\gamma}E_{\xi_1}\big[e^{(1-\gamma)\xi_1}I(X_t+\delta_c\cdot\xi)\big]l_1(X_t) + \frac{1}{1-\gamma}\sum_{i=2}^{n}E_{\xi_i}\big[I(X_t+\delta_i\cdot\xi)\big]l_i(X_t)$$

$$-\frac{1}{1-\gamma}I(X_t)\mathbf{1}'l(X_t) + \beta I(X_t)\Big(\ln\beta - \frac{\ln(I(X_t))}{1-\gamma}\Big) = 0, \tag{A.7}$$

where $\delta_i \equiv (0,...,0,1,0,...,0)'$ denotes a selection vector for the ith state variable. Conjecture that a function of the form

$$I(X) = e^{a+b'X} \tag{A.8}$$

solves (A.7). Then

$$I_X(X) = bI(X) \tag{A.9}$$

$$I_{XX}(X) = bb'I(X). \tag{A.10}$$

Substituting (A.8) through (A.10) into (A.7) implies

$$\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{1-\gamma}b'\mu(X_t) + \frac{1}{2}tr\Big(\frac{1}{1-\gamma}bb'\Sigma(X_t)\Sigma(X_t)' + b\delta_c'\Sigma(X_t)\Sigma(X_t)\Big)$$

$$+\frac{1}{2}b'\Sigma(X_t)\Sigma(X_t)'\delta_c - \frac{1}{2}\gamma\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{1-\gamma}\varrho_1(1-\gamma+b_1)l_1(X_t) + \frac{1}{1-\gamma}\sum_{i=2}^{n}\varrho_i(b_i)l_i(X_t)$$

$$-\frac{1}{1-\gamma}\mathbf{1}'l(X_t)+\beta\Big(\ln\beta-\frac{a+b'X_t}{1-\gamma}\Big)=0. \tag{A.11}$$

Using equations (8) through (10) to rewrite $\mu(X_t)$, $\Sigma(X_t)\Sigma(X_t)'$ and $l(X_t)$, and making use of the property of the trace of matrices that $tr(AB)=tr(BA)$ whenever both AB and BA are defined, yield an equation linear in X_t:

$$\delta_c'\big(\mathcal{M}+\mathcal{K}X_t\big)+\frac{1}{2}\Big(\delta_c'h\delta_c+(\delta_cH\delta_c)'X_t\Big)+\frac{1}{1-\gamma}b'\big(\mathcal{M}+\mathcal{K}X_t\big)$$

$$+\frac{1}{2}\frac{1}{1-\gamma}\Big(b'hb+(b'Hb)'X_t\Big)+\frac{1}{2}\Big(\delta_c'hb+(\delta_c'Hb)'X_t\Big)+\frac{1}{2}\Big(\delta_c'hb+(\delta_c'Hb)'X_t\Big)$$

$$-\frac{\gamma}{2}\Big(\delta_c'h\delta_c+(\delta_cH\delta_c)'X_t\Big)+\frac{1}{1-\gamma}\Big(\varrho_1(1-\gamma+b_1)-\varrho_1(b_1)\Big)\delta_c'(l+LX_t)$$

$$+\frac{1}{1-\gamma}\varrho(b)'\big(l+LX_t\big)-\frac{1}{1-\gamma}\mathbf{1}'\big(l+LX_t\big)+\beta\big(\ln\beta-\frac{a}{1-\gamma}\big)-\frac{\beta}{1-\gamma}b'X_t=0. \tag{A.12}$$

Collecting terms in X_t results in the following equation system for b:

$$\frac{1}{2}b'Hb+\Big(\mathcal{K}'+(1-\gamma)\delta_c'H-\mathrm{diag}(\beta)\Big)b+L'\varrho(b)$$

$$+\Big(\varrho_1(1-\gamma+b_1)-\varrho_1(b_1)\Big)L'\delta_c+\frac{(1-\gamma)^2}{2}\delta_c'H\delta_c+(1-\gamma)\mathcal{K}'\delta_c-L'\mathbf{1}=0, \tag{A.13}$$

which is a system of n equations for n unknowns $(b_1,b_2,...,b_n)$. The system depends on the moment generating functions of the jump sizes and admits an explicit solution only in special cases. There are at least two cases in which (A.13) collapses into a quadratic equation system in X_t and can be easily solved with a relatively low dimension of n. First, if there are no jumps with state-dependent intensities, then $L=0$. Second, if there is jump only in consumption while the jump intensity doesn't depend on consumption itself, then one can also verify that (A.13) becomes a quadratic system.

While multiple solutions to (A.13) possibly exist, there are at least two ways to make a choice among them. First, Tauchen (2011) suggests choosing the solution which approaches a non-explosive limit as certain coefficient associated with X_t in $\Sigma(X_t)\Sigma(X_t)'$ approaches zero. Second, Wachter (2013) suggests one can choose the solution that, under a simple thought experiment, makes economic sense. One such strategy can be to consider the case that jump size is identically equal to zero while the jump intensity is not. Because essentially those jumps should have no economic consequence, the value function should somehow reduce to its counterpart under the standard diffusion model. Collecting constant terms results in the following characterization of a in terms of b:

$$a = \frac{1}{\beta}\Big\{\big((1-\gamma)\delta_c + b\big)'\mathcal{M} + \frac{1}{2}\big((1-\gamma)\delta_c + b\big)'h\big((1-\gamma)\delta_c + b\big)$$

$$+\Big(\big(\varrho_1(1-\gamma+b_1) - \varrho_1(b_1)\big)\delta_c + \varrho(b) - \mathbf{1}\Big)' l + (1-\gamma)\beta\ln\beta\Big\}. \tag{A.14}$$

A.II. Risk-Free Rate

Taking the derivative of (A.1) with respect to portfolio choice α_t and setting it to zero imply

$$\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \beta - r_t + b'\Sigma(X_t)\Sigma(X_t)'\delta_c - \gamma\alpha_t\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c$$

$$+E_{\xi_1}\big[\big(1+\alpha_t(e^{\xi_1}-1)\big)^{-\gamma}\big(e^{\xi_1}-1\big)e^{b_1\xi_1}\big]l_1(X_t) = 0. \tag{A.15}$$

Evaluating the above equation at $\alpha_t = 1$ and rearranging yield

$$r_t = \beta + \delta_c'\mu(X_t) + (\frac{1}{2} - \gamma)\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c$$

$$+\delta_c'\Big(\big(\varrho(b+1-\gamma) - \varrho(b-\gamma)\big)\cdot l(X_t)\Big) + \delta_c'\Sigma(X_t)\Sigma(X_t)'b. \tag{A.16}$$

A.III. State-Price Density

As discussed in Section 4.4, the state-price density is given by

$$\pi_t = \exp\Big\{\int_0^t f_V(C_s, V_s)ds\Big\} f_C(C_t, V_t). \tag{A.17}$$

Now substitute $W_t = \beta^{-1}C_t$ and (A.8) into (A.3). Then taking (A.3) into (A.5) and taking the partial derivative of (A.6) with respect to V yield, respectively,

$$f_C(C_t, J(W(C_t), X_t)) = \beta^{\gamma}C_t^{-\gamma}e^{a+b'X_t} \tag{A.18}$$

$$f_V(C_t, J(W(C_t), X_t)) = \beta(1-\gamma)\ln\beta - \beta(a + b'X_t) - \beta. \tag{A.19}$$

Finally, substituting (A.18) and (A.19) into (A.17) and noting $\ln C_t = \delta_c' X_t$ yield

$$\pi_t = \beta^\gamma \exp\Big\{\eta t - \beta b' \int_0^t X_s ds + a + (b' - \gamma\delta_c')X_t\Big\}, \tag{A.20}$$

where $\eta = \beta(1-\gamma)\ln\beta - \beta a - \beta$. Finally, applying Ito's Lemma on (A.20) obtains

$$\frac{d\pi_t}{\pi_{t^-}} = \mu_{\pi,t}dt + (b' - \gamma\delta_c')\Sigma(X_t)dB_t + \Big(e^{(b-\gamma\delta_c)\cdot\xi} - \mathbf{1}\Big)dN_t, \tag{A.21}$$

where

$$\mu_{\pi,t} = \eta - \beta b' X_t + (b' - \gamma\delta_c')\mu(X_t) + \frac{1}{2}(b' - \gamma\delta_c')\Sigma(X_t)\Sigma(X_t)'(b - \gamma\delta_c). \tag{A.22}$$

Appendix B: Proof of Convergence of State-Price Density

Eraker and Shaliastovich (2008) solve a general equilibrium pricing model with an affine jump-diffusion structure for the state variable dynamics identical to those we use and a continuous-time extension of the discrete-time Epstein-Zin preferences. Because there is generally no closed-form solution to that model when the IES parameter is different from one, they also use a continuous-time extension of the log-linear approximation (see Campbell and Shiller (1988)) to maintain model tractability. We in this section prove that the state-price density exactly solved in this paper is actually an $IES \to 1$ limit of the state-price density approximately solved in Eraker and Shaliastovich (2008).[21]

Proof. To this end, first note the relation between the time preference parameter β in this paper and δ in their paper

$$\beta = \frac{1-\delta}{\delta}. \tag{B.1}$$

Because the affine structures for state variables are identical between both papers, the difference in the state-price densities can only arise from the difference in λ, which completely determines the market prices of various risks in the economy for both papers. The λ in Eraker and Shaliastovich (2008) is given by

$$\lambda = \gamma\delta_c + (1-\theta)\kappa_1 B, \tag{B.2}$$

[21] Eraker and Shaliastovich (2008) start with the discrete-time model and try to see what the log-linearization equation and the pricing kernel become as the time interval shrinks from one to infinitesimal. Although their such two extensions have been widely used, people do not know theoretically how precisely they work in continuous time. Our convergence result and continuity show that it is accurate if the IES parameter is close to one.

where γ, δ_c and θ have identical interpretations as in our paper, κ_1 is the slope coefficient in log-linearization with an expression given by

$$\kappa_1 = \frac{e^{E \ln(S_t/C_t)}}{1 + e^{E \ln(S_t/C_t)}} \tag{B.3}$$

and B is the coefficient associated with X_t in the approximately solved equilibrium wealth-consumption ratio:

$$\frac{S_t}{C_t} = A + B' X_t. \tag{B.4}$$

Because it has been shown that $\frac{S_t}{C_t} = \frac{1}{\beta}$ when $\psi = 1$, to prove convergence the only possibility is $\frac{S_t}{C_t} = \frac{1}{\beta} + o(1)$ as $\psi \to 1$. It follows from (B.3) that

$$\kappa_1 = \frac{1}{1+\beta} + o(1) = \delta + o(1) \tag{B.5}$$

and from (B.4) that $B = o(1)$. It then follows that as $\psi \to 1$:

$$\begin{aligned}
\lambda &= \gamma\delta_c - \theta\kappa_1 B + \kappa_1 B \\
&= \gamma\delta_c - \theta\kappa_1 B + o(\theta\kappa_1 B) \\
&= \gamma\delta_c - \theta\kappa_1 B + o(\theta\delta B) \\
&= \gamma\delta_c - \big(\theta\delta B + o(\theta\delta B)\big) + o(\theta\delta B) \\
&= \gamma\delta_c - \theta\delta B + o(\theta\delta B),
\end{aligned} \tag{B.6}$$

where the first equality follows from (B.2), the second from the fact that $\theta \equiv \frac{1-\gamma}{1-1/\psi} \to \infty$ as $\psi \to 1$, and the third and the fourth from (B.5).

Note because $B \to 0$ and $\theta \to \infty$ as $\psi \to 1$, it is possible that λ approaches a well-defined limit, but when it does, from (B.6) the limit can only be

$$\begin{aligned}
&\gamma\delta_c - \lim_{\psi \to 1} \theta\delta B \\
&= \gamma\delta_c - \lim_{\psi \to 1} \big(\chi - (1-\gamma)\delta_c\big) \\
&= \delta_c - \lim_{\psi \to 1} \chi,
\end{aligned} \tag{B.7}$$

where the second equality follows from the definition of the vector χ in Eraker and Shaliastovich (2008). Therefore, it follows from comparing (21) with (B.7) that what is only left to show becomes

$$\delta_c - \lim_{\psi \to 1} \chi = \gamma \delta_c - b$$

or

$$\lim_{\psi \to 1} \chi = (1-\gamma)\delta_c + b. \tag{B.8}$$

Let's reproduce the equation system that χ solves, which is equation (2.12) in Eraker and Shaliastovich (2008):

$$\mathcal{K}'\chi - \theta(1-\kappa_1)B + \frac{1}{2}\chi' H \chi + L'\big(\varrho(\chi) - 1\big) = 0, \tag{B.9}$$

which after we substituting in a rearrangement of the definition of χ, $\theta B = \frac{\chi-(1-\gamma)\delta_c}{\kappa_1}$, becomes

$$\mathcal{K}'\chi - \frac{1-\kappa_1}{\kappa_1}\big(\chi - (1-\gamma)\delta_c\big) + \frac{1}{2}\chi' H \chi + L'\big(\varrho(\chi) - 1\big) = 0. \tag{B.10}$$

Therefore, to show (B.8), by continuity we only need to show that $\chi = (1-\gamma)\delta_c + b$ solves the $\psi \to 1$ limit of equation (B.10), which is

$$\mathcal{K}'\chi - \beta\big(\chi - (1-\gamma)\delta_c\big) + \frac{1}{2}\chi' H \chi + L'\big(\varrho(\chi) - 1\big) = 0. \tag{B.11}$$

This is actually true since we exactly reproduce (A.13) after substituting $\chi = (1-\gamma)\delta_c + b$ into (B.11).

The above argument together establishes that whenever there is a solution $\lambda(\psi)$ in Eraker and Shaliastovich (2008) one can find a solution λ of our model which it converges to as $\psi \to 1$. □

Appendix C: Solutions to the VIX Model

C.I. Value Function and State-Price Density

Define $X_t = (lnC_t, \sigma_t^2, \lambda_t)'$ and $B_t = (B_t^C, B_t^V, B_t^\lambda)'$. It follows that the process for X_t is equivalent to

$$dX_t = (\mathcal{M} + \mathcal{K}X_t)dt + \Sigma(X_t)dB_t + \xi \cdot dN_t; \quad \mathcal{M} = (\mu, \kappa^V\theta^V, \kappa^\lambda\theta^\lambda)' \tag{C.1}$$

$$\mathcal{K} = \begin{bmatrix} 0 & -\frac{1}{2} & 0 \\ 0 & -\kappa^V & 0 \\ 0 & 0 & -\kappa^\lambda \end{bmatrix}; \quad \Sigma(X_t) = \begin{bmatrix} \sigma_t & 0 & 0 \\ 0 & \sigma_V\sigma_t & 0 \\ 0 & 0 & \sigma_\lambda\sqrt{\lambda_t} \end{bmatrix}; \quad \Sigma(X_t)\Sigma(X_t)' = h + \sum_{i=1}^{3} H_i X_{t,i} \tag{C.2}$$

$$h = H_1 = \begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix}; \quad H_2 = \begin{bmatrix} 1 & & \\ & \sigma_V^2 & \\ & & 0 \end{bmatrix}; \quad H_3 = \begin{bmatrix} 0 & & \\ & 0 & \\ & & \sigma_\lambda^2 \end{bmatrix} \tag{C.3}$$

$$\xi = (0, \xi_V, 0)'; \quad dN_t = (0, dN_t, 0)'; \quad \varrho(u) = (0, \varrho(u_2), 0)', \tag{C.4}$$

and the jump intensities are summarized by

$$l(X_t) = l + LX_t; \quad l = (0,0,0)'; \quad L = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{C.5}$$

Define $b = (b_1, b_2, b_3)'$. Then (A.13) implies that b should solve

$$-\beta b_1 = 0 \tag{C.6}$$

$$\frac{1}{2}\sigma_V^2 b_2^2 - (\kappa^V + \beta) b_2 + \frac{1}{2}\gamma(\gamma - 1) = 0 \tag{C.7}$$

$$\frac{1}{2}\sigma_\lambda^2 b_3^2 - (\kappa^\lambda + \beta) b_3 + \varrho(b_2) - 1 = 0. \tag{C.8}$$

It follows from (C.6) that $b_1 = 0$, and from (C.7) that

$$b_2 = \frac{(\kappa^V + \beta) \pm \sqrt{(\kappa^V + \beta)^2 - \sigma_V^2 \gamma(\gamma - 1)}}{\sigma_V^2}. \tag{C.9}$$

Here we follow Tauchen (2011) to choose the negative root in (C.9), since otherwise b_2 would explode as $\sigma_V \to 0$. Then (C.8) implies

$$b_3 = \frac{(\kappa^\lambda + \beta) \pm \sqrt{(\kappa^\lambda + \beta)^2 - 2\sigma_\lambda^2(\varrho(b_2) - 1)}}{\sigma_\lambda^2}. \tag{C.10}$$

Once again we choose the negative root in (C.10). After solving out b we apply (A.14) to obtain

$$a = \frac{1}{\beta}\Big\{(1 - \gamma)(\mu + \beta \ln \beta) + b_2 \kappa^V \theta^V + b_3 \kappa^\lambda \theta^\lambda\Big\}. \tag{C.11}$$

Then apply (A.16) to obtain the risk-free rate

$$r_t = \Phi_0 + \Phi_1' X_t \tag{C.12}$$

where

$$\Phi_0 = \mu + \beta; \quad \Phi_1 = (0, -\gamma, 0)'. \tag{C.13}$$

Applying (A.21) then implies the state-price density obeys

$$\frac{d\pi_t}{\pi_{t^-}} = -r_t dt - \Lambda_t' dB_t + \left(e^{b_2 \xi_V} - 1\right) dN_t - \lambda_t E[e^{b_2 \xi_V} - 1] dt \tag{C.14}$$

$$\Lambda_t = [\gamma \sigma_t, -b_2 \sigma_V \sigma_t, -b_3 \sigma_\lambda \sqrt{\lambda_t}]'. \tag{C.15}$$

Using Theorem 1, the evolution of the state variables under the equivalent risk-neutral Q measure induced by the state-price density is given by

$$dX_t = (\mathcal{M}^Q + \mathcal{K}^Q X_t)dt + \Sigma(X_t) dB_t^Q + \xi_V^Q \cdot dN_t^Q; \quad \mathcal{M}^Q = (\mu, \kappa^V \theta^V, \kappa^\lambda \theta^\lambda)' \tag{C.16}$$

$$\mathcal{K}^Q = \begin{bmatrix} 0 & -\frac{1}{2} - \gamma & 0 \\ 0 & -\kappa^V + b_2 \sigma_V^2 & 0 \\ 0 & 0 & -\kappa^\lambda + b_3 \sigma_\lambda^2 \end{bmatrix}; \quad dB_t^Q \equiv \begin{bmatrix} dB_t^{C,Q} \\ dB_t^{V,Q} \\ dB_t^{\lambda,Q} \end{bmatrix} = \begin{bmatrix} dB_t^C \\ dB_t^V \\ dB_t^\lambda \end{bmatrix} + \Lambda_t dt \tag{C.17}$$

$$l^Q(X_t) = l^Q + L^Q X_t = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \varrho(b_2) \\ 0 & 0 & 0 \end{bmatrix} X_t; \quad \varrho^Q(u) = \begin{bmatrix} 0 \\ \frac{\varrho(u_2 + b_2)}{\varrho(b_2)} \\ 0 \end{bmatrix}. \tag{C.18}$$

C.II. Equity Price

In order to price the equity claim, let's compute the discounted characteristic function $\varrho_X^Q(u, X_t, \tau)$ evaluated at $u = (\phi, 0, 0)'$, which, as shown by Duffie, Pan, and Singleton (2000), is equal to $e^{\alpha(\tau) + \beta(\tau)' X_t}$ with $\alpha(\tau)$ and $\beta(\tau)$ solving the following ODEs:

Riccati Equations for Discounted Characteristic Function:

$$\dot{\beta}(\tau) = -\Phi_1 + \mathcal{K}^{Q'} \beta(\tau) + \frac{1}{2} \beta(\tau)' H \beta(\tau) + L^{Q'} \left(\varrho^Q(\beta(\tau)) - 1\right) \tag{C.19}$$

$$\dot{\alpha}(\tau) = -\Phi_0 + \mathcal{M}^{Q'} \beta(\tau) + \frac{1}{2} \beta(\tau)' h \beta(\tau) + l^{Q'} \left(\varrho^Q(\beta(\tau)) - 1\right) \tag{C.20}$$

with boundary conditions $\alpha(0) = 0, \beta(0) = (\phi, 0, 0)$. Given the risk-neutral parameters above, the ODEs become

$$\dot{\beta}_1(\tau) = 0 \tag{C.21}$$

$$\dot{\beta}_2(\tau) = \frac{1}{2}\sigma_V^2\beta_2^2(\tau) + (b_2\sigma_V^2 - \kappa^V)\beta_2(\tau) + \frac{1}{2}\beta_1^2(\tau) - (\frac{1}{2} + \gamma)\beta_1(\tau) + \gamma \tag{C.22}$$

$$\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2\beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\big(\beta_2(\tau) + b_2\big) - \varrho(b_2) \tag{C.23}$$

$$\dot{\alpha}(\tau) = -\beta - \mu + \mu\beta_1(\tau) + \kappa^V\theta^V\beta_2(\tau) + \kappa^\lambda\theta^\lambda\beta_3(\tau). \tag{C.24}$$

Together with boundary conditions, (C.21) implies $\beta_1(\tau) = \phi, \forall\tau$. Then as long as $1 < \phi < 2\gamma$ the solution to (C.22) is given by the following closed form:

$$\beta_2(\tau) = \frac{2(\phi - 1)(\gamma - \frac{1}{2}\phi)\big(1 - e^{-\eta_\phi\tau}\big)}{(\eta_\phi + b_2\sigma_V^2 - \kappa^V)\big(1 - e^{-\eta_\phi\tau}\big) - 2\eta_\phi} \tag{C.25}$$

where

$$\eta_\phi = \sqrt{(b_2\sigma_V^2 - \kappa^V)^2 + 2(\phi - 1)(\gamma - \frac{1}{2}\phi)\sigma_V^2}. \tag{C.26}$$

Note that since $1 < \phi < 2\gamma$ the term inside the square root is guaranteed to be positive. Moreover, $\eta_\phi > |b_2\sigma_V^2 - \kappa^V| \geq b_2\sigma_V^2 - \kappa^V$, implying that the denominator $(\eta_\phi + b_2\sigma_V^2 - \kappa^V)\big(1 - e^{-\eta_\phi\tau}\big) - 2\eta_\phi$ is strictly negative for all τ. Therefore, this argument establishes that $\beta_2(\tau) < 0$ for all τ. Noting the similarity between the forms of equations (C.22) and (C.23), it then follows that a similar argument as above can establish that $\beta_3(\tau) < 0$ for all τ, as long as $\varrho\big(\beta_2(\tau)+b_2\big) - \varrho(b_2) < -\epsilon$ for all τ for some $\epsilon > 0$. But the latter is actually satisfied since $\beta_2(\tau) < 0, \forall\tau$ and $\varrho(\cdot)$ is an increasing function. We've shown that $\beta_2(\tau) < 0, \beta_3(\tau) < 0$ for all τ. This is important because the sign of $\beta_2(\tau)$ and $\beta_3(\tau)$ respectively finally determines how equity price responds to σ_t^2 and λ_t. The fact that $\beta_2(\tau) < 0, \beta_3(\tau) < 0$ for all τ implies that the equity price is decreasing in both σ_t^2 and λ_t (recall (63) and (64)). This completes the proof of proposition 2. Generally (C.23) and (C.24) do not admit closed-form solutions and turn out to become (61) and (62), which we can solve numerically

C.III. Equity Premium

We solve for the equity premium analytically. A familiar no-arbitrage condition on the equity market is that the discounted gains process $\pi_t P_t + \int_0^t \pi_s D_s ds$ is a P-martingale (for its derivation, see e.g., Appendix *A.IV.* of Wachter (2013)). It follows that the drift term in $E_t\big[\frac{d\left(\pi_t P_t + \int_0^t \pi_s D_s ds\right)}{\pi_{t-}P_{t-}}\big]$ is zero. Using Ito's Lemma, this implies

$$\mu_{\pi,t} + \mu_{D,t} - [\phi\sigma_t, \frac{G_1}{G}\sigma_V\sigma_t, \frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}]\Lambda_t + \frac{G_1}{G}\kappa^V(\theta^V - \sigma_t^2) + \frac{G_2}{G}\kappa^\lambda(\theta^\lambda - \lambda_t)$$

$$+\frac{1}{2}\frac{G_{11}}{G}\sigma_V^2\sigma_t^2 + \frac{1}{2}\frac{G_{22}}{G}\sigma_\lambda^2\lambda_t + \frac{D_{t-}}{P_{t-}} + \lambda_t E\Big[e^{b_2\xi_V}\frac{G(\sigma_t^2 + \xi_V, \lambda)}{G(\sigma_t^2, \lambda)} - 1\Big] = 0, \tag{C.27}$$

where $\mu_{\pi,t}$ and $\mu_{D,t}$ represent respectively the drift term in $\frac{d\pi_t}{\pi_{t^-}}$ and $\frac{dD_t}{D_{t^-}}$. Use $\mu_{P,t}$ to denote the drift term in $\frac{dP_t}{P_{t^-}}$, which, when Ito's Lemma applied upon, implies

$$\mu_{P,t} + \frac{D_{t^-}}{P_{t^-}} - r_t = \mu_{D,t} + \mu_{\pi,t} + \frac{G_1}{G}\kappa^V(\theta^V - \sigma_t^2) + \frac{G_2}{G}\kappa^\lambda(\theta^\lambda - \lambda_t)$$

$$+\frac{1}{2}\frac{G_{11}}{G}\sigma_V^2\sigma_t^2 + \frac{1}{2}\frac{G_{22}}{G}\sigma_\lambda^2\lambda_t + \frac{D_{t^-}}{P_{t^-}} + \lambda_t E\big[e^{b_2\xi_V} - 1\big], \tag{C.28}$$

where we have used (27) to substitute r_t. Rearranging (C.27) properly and then substituting into (C.28) give rise to the expression for the equity premium conditional on no jumps occurring:

$$\mu_{P,t} + \frac{D_{t^-}}{P_{t^-}} - r_t = \big[\phi\sigma_t, \frac{G_1}{G}\sigma_V\sigma_t, \frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\big]\Lambda_t + \lambda_t E\Big[e^{b_2\xi_V}(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)})\Big]. \tag{C.29}$$

After accounting for the expected percentage change in equity price if a jump to volatility occurs, we obtain the population equity premium:

$$r_t^e - r_t = \sigma_{P,t}'\Lambda_t + \lambda_t E\Big[\big(e^{b_2\xi_V} - 1\big)\big(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)})\Big] \tag{C.30}$$

where

$$\sigma_{P,t} = \big[\phi\sigma_t, \frac{G_1}{G}\sigma_V\sigma_t, \frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\big]'. \tag{C.31}$$

C.IV. Equity Price Dynamics under Q

We solve for the dynamics of the log equity price under the Q measure. Ito's Lemma implies that under P measure:

$$\frac{dP_t}{P_{t^-}} = \Big(\mu_{D,t} + \frac{G_1}{G}\kappa^V(\theta^V - \sigma_t^2) + \frac{G_2}{G}\kappa^\lambda(\theta^\lambda - \lambda_t) + \frac{1}{2}\frac{G_{11}}{G}\sigma_V^2\sigma_t^2 + \frac{1}{2}\frac{G_{22}}{G}\sigma_\lambda^2\lambda_t\Big)dt$$

$$+\sigma_{P,t}'dB_t + \sigma_D dB_t^D + \big[\frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)} - 1\big]dN_t, \tag{C.32}$$

where $\mu_{D,t}$ denotes the drift term in $\frac{dD_t}{D_{t^-}}$. It follows again from Ito's Lemma that (after some algebra)

$$d\ln P_t = \Big(\phi(\mu - \frac{1}{2}\sigma_t^2) + \frac{G_1}{G}\kappa^V(\theta^V - \sigma_t^2) + \frac{G_2}{G}\kappa^\lambda(\theta^\lambda - \lambda_t) + \frac{1}{2}(\frac{G_{11}}{G} - \frac{G_1^2}{G^2})\sigma_V^2\sigma_t^2 + \frac{1}{2}(\frac{G_{22}}{G} - \frac{G_2^2}{G^2})\sigma_\lambda^2\lambda_t\Big)dt$$

$$+\sigma_{P,t}'dB_t + \sigma_D dB_t^D + \ln\big[\frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\big]dN_t. \tag{C.33}$$

By plugging in the expressions for the diffusions under Q measure in equation (C.17) into (C.33) and replacing ξ_V and N_t with their counterparts under Q, we recover the dynamics of the log equity price under Q as

$$d\ln P_t = \Big(\phi(\mu - \frac{1}{2}\sigma_t^2) + \frac{G_1}{G}\kappa^V(\theta^V - \sigma_t^2) + \frac{G_2}{G}\kappa^\lambda(\theta^\lambda - \lambda_t) - \sigma_{P,t}'\Lambda_t$$

$$+\frac{1}{2}(\frac{G_{11}}{G} - \frac{G_1^2}{G^2})\sigma_V^2\sigma_t^2 + \frac{1}{2}(\frac{G_{22}}{G} - \frac{G_2^2}{G^2})\sigma_\lambda^2\lambda_t\Big)dt + \sigma_{P,t}'dB_t^Q + \sigma_D dB_t^D + \ln\big[\frac{G(\sigma_t^2 + \xi_V^Q, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\big]dN_t^Q. \quad \text{(C.34)}$$

C.V. VIX

By definition, $VIX^2(X_t) = \mathrm{Var}_t^Q\big[\ln P_{t+\frac{1}{12}}\big] = \mathrm{Var}_t^Q\big[\ln \tilde{P}_{t+\frac{1}{12}}\big] + \frac{1}{12}\sigma_D^2$, where we have separated the idiosyncratic noise out and $\ln\tilde{P}$ denotes the portion of the log equity price that only involves systematic risk. The conditional cumulant generating function for $\ln\tilde{P}_{t+\frac{1}{12}}$ is given by

$$\Phi(u) = \ln E_t^Q e^{u\ln\tilde{P}_{t+\frac{1}{12}}} \quad \text{(C.35)}$$

$$= \ln E_t^Q e^{u\lambda_X' X_{t+\frac{1}{12}}} \quad \text{(C.36)}$$

$$= \alpha(u\lambda_X, t, t + \frac{1}{12}) + \beta'(u\lambda_X, t, t+\frac{1}{12})X_t \quad \text{(C.37)}$$

where

$$\lambda_X \equiv (\phi, g_1^*, g_2^*)'. \quad \text{(C.38)}$$

Therefore, using the property of the cumulant generating function, we see that $\mathrm{Var}_t^Q\big[\ln\tilde{P}_{t+\frac{1}{12}}\big] = \tilde{a}_{1/12} + b_{1/12}\ln C_t + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t$, where $\tilde{a}_{1/12}, b_{1/12}, c_{1/12}$ and $d_{1/12}$ are the second derivatives w.r.t. u of $\alpha(u\lambda_X, t, t+\frac{1}{12})$, $\beta_1(u\lambda_X, t, t+\frac{1}{12})$, $\beta_2(u\lambda_X, t, t+\frac{1}{12})$ and $\beta_3(u\lambda_X, t, t+\frac{1}{12})$ evaluated at $u = 0$, respectively. Under appropriate technique conditions (see Duffie, Pan, and Singleton (2000)), let $\alpha(\tau), \beta(\tau)$ solve:

Riccati Equations for Cumulant Generating Function:

$$\dot{\beta}(\tau) = \mathcal{K}^{Q'}\beta(\tau) + \frac{1}{2}\beta(\tau)'H\beta(\tau) + L^{Q'}\big(\varrho^Q(\beta(\tau)) - 1\big) \quad \text{(C.39)}$$

$$\dot{\alpha}(\tau) = \mathcal{M}^{Q'}\beta(\tau) + \frac{1}{2}\beta(\tau)'h\beta(\tau) + l^{Q'}\big(\varrho^Q(\beta(\tau)) - 1\big) \quad \text{(C.40)}$$

with boundary conditions $\alpha(0) = 0, \beta(0) = u\lambda_X$. Then $\alpha(u\lambda_X, t, t+\frac{1}{12}) = \alpha(1/12)$ and $\beta(u\lambda_X, t, t+\frac{1}{12}) = \beta(1/12)$. It turns out those ODEs are

$$\dot{\beta}_1(\tau) = 0 \tag{C.41}$$

$$\dot{\beta}_2(\tau) = \frac{1}{2}\sigma_V^2\beta_2^2(\tau) + (b_2\sigma_V^2 - \kappa^V)\beta_2(\tau) + \frac{1}{2}\beta_1^2(\tau) - (\frac{1}{2}+\gamma)\beta_1(\tau) \tag{C.42}$$

$$\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2\beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\big(\beta_2(\tau)+b_2\big) - \varrho(b_2) \tag{C.43}$$

$$\dot{\alpha}(\tau) = \mu\beta_1(\tau) + \kappa^V\theta^V\beta_2(\tau) + \kappa^\lambda\theta^\lambda\beta_3(\tau). \tag{C.44}$$

Together with boundary conditions, the solutions are

$$\beta_1(\tau) = u\phi, \forall\tau \tag{C.45}$$

$$\dot{\beta}_2(\tau) = \frac{1}{2}\sigma_V^2\beta_2^2(\tau) + (b_2\sigma_V^2 - \kappa^V)\beta_2(\tau) - \frac{1}{2}u\phi\big(2\gamma+1-u\phi\big) \tag{C.46}$$

$$\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2\beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\big(\beta_2(\tau)+b_2\big) - \varrho(b_2) \tag{C.47}$$

$$\dot{\alpha}(\tau) = \mu\phi u + \kappa^V\theta^V\beta_2(\tau) + \kappa^\lambda\theta^\lambda\beta_3(\tau). \tag{C.48}$$

Here $\beta_1(\tau)$ has a closed-form solution, while $\beta_2(\tau), \beta_3(\tau), \alpha(\tau)$ do not.[22] It follows immediately from (C.45) that $b_{1/12} = 0$, i.e., VIX does not explicitly depend on current consumption. Finally, we can write $Var_t^Q\big[\ln\tilde{P}_{t+\frac{1}{12}}\big]$ as an affine function in σ_t^2 and λ_t:

$$\text{Var}_t^Q\big[\ln\tilde{P}_{t+\frac{1}{12}}\big] = \tilde{a}_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t, \tag{C.49}$$

where both $c_{1/12}$ and $d_{1/12}$ can be shown to be positive coefficients. And

$$VIX^2(X_t) = \frac{1}{12}\sigma_D^2 + \tilde{a}_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t. \tag{C.50}$$

For notational convenience, we compound the first two constant terms in (C.50) and denote it as $a_{1/12}$. Then

$$VIX(X_t) = \sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}. \tag{C.51}$$

[22] $\beta_2(\tau)$ does actually admit a closed-form solution. But due to its complication, we numerically solve it

C.VI. Proof of Proposition 3

Proof. The steady-state levels of σ_t^2 and λ_t are $\theta^V + \frac{\mu_\xi \theta^\lambda}{\kappa^V}$ and θ^λ, respectively. Recall that the risk-neutral dynamics of the state variables are summarized in (48) through (51). One can easily verify that (49) implies

$$E^Q[\lambda_{t+\tau}|\sigma_t^2, \lambda_t] = \lambda_t e^{-\kappa^{\lambda,Q}\tau} + \theta^{\lambda,Q}\big(1 - e^{-\kappa^{\lambda,Q}\tau}\big). \tag{C.52}$$

Then (C.52), an application of Ito's Lemma under Q to obtain $d\big(\sigma_t^2 e^{\kappa^{V,Q}t}\big)$ and then integrating, and an application of the law of iterated expectations together imply

$$E^Q[\sigma_{t+\tau}^2|\sigma_t^2, \lambda_t] = \sigma_t^2 e^{-\kappa^{V,Q}\tau} + \Big(\theta^{V,Q} + \frac{\theta^{\lambda,Q}\mu_\xi^Q \varrho(b_2)}{\kappa^{V,Q}}\Big)\Big(1 - e^{-\kappa^{V,Q}\tau}\Big)$$

$$+\mu_\xi^Q \varrho(b_2)\frac{\lambda_t - \theta^{\lambda,Q}}{\kappa^{V,Q} - \kappa^{\lambda,Q}}\Big(e^{-\kappa^{\lambda,Q}\tau} - e^{-\kappa^{V,Q}\tau}\Big). \tag{C.53}$$

Note that the VIX-squared futures curve in steady-state is given by

$$F_t^{VIX^2}(\tau) = a_{1/12} + c_{1/12}E^Q[\sigma_{t+\tau}^2|\sigma_t^2 = \theta^V + \frac{\mu_\xi\theta^\lambda}{\kappa^V}, \lambda_t = \theta^\lambda]$$

$$+d_{1/12}E^Q[\lambda_{t+\tau}|\sigma_t^2 = \theta^V + \frac{\mu_\xi\theta^\lambda}{\kappa^V}, \lambda_t = \theta^\lambda]. \tag{C.54}$$

Since $c_{1/12}$ and $d_{1/12}$ are positive constants, to show the futures curve is upward sloping, we only need to show $E^Q[\sigma_{t+\tau}^2|\theta^V + \frac{\mu_\xi\theta^\lambda}{\kappa^V}, \theta^\lambda]$ and $E^Q[\lambda_{t+\tau}|\theta^V + \frac{\mu_\xi\theta^\lambda}{\kappa^V}, \theta^\lambda]$ are both increasing in τ. That the latter is increasing in τ for all τ follows from $\theta^\lambda < \theta^{\lambda,Q}$. We only need to show the former is increasing in τ. Now it follows from (C.53) that taking the derivative of $E^Q[\sigma_{t+\tau}^2|\theta^V + \frac{\mu_\xi\theta^\lambda}{\kappa^V}, \theta^\lambda]$ with respect to τ results in two components.

The first one is

$$e^{-\kappa^{V,Q}\tau}\Big\{\kappa^{V,Q}(\theta^{V,Q} - \theta^V) + \big(\theta^{\lambda,Q}\mu_\xi^Q \varrho(b_2) - \frac{\kappa^{V,Q}\mu_\xi\theta^\lambda}{\kappa^V}\big)\Big\}. \tag{C.55}$$

There are again two components. The first one is obviously positive since $\theta^{V,Q} > \theta^V$. The second one can be shown proportional to $\kappa^V\theta^{\lambda,Q}\mu_\xi^Q\varrho(b_2) - \kappa^{V,Q}\theta^\lambda\mu_\xi$, where $\kappa^V\theta^{\lambda,Q} > \kappa^{V,Q}\theta^\lambda$ since $\kappa^V > \kappa^{V,Q}$ while $\theta^\lambda < \theta^{\lambda,Q}$. We only need to show $\mu_\xi^Q\varrho(b_2) > \mu_\xi$ to show the second component is positive. Using the property of moment generating functions, it is equivalent to $\varrho'(b_2) > \varrho'(0)$, which is guaranteed by the assumption that $\varrho(\cdot)$ is convex. Therefore, (C.55) is positive.

The second one is proportional to

$$\frac{1}{\kappa^{V,Q}-\kappa^{\lambda,Q}}\Big\{\kappa^{\lambda,Q}e^{-\kappa^{\lambda,Q}\tau}-\kappa^{V,Q}e^{-\kappa^{V,Q}\tau}\Big\}, \tag{C.56}$$

which is positive if and only if $xe^{-\tau x}$ is decreasing in x. Taking the derivative of the former with respect to x yields $e^{-\tau x}(1-\tau x)$, which is negative when τ is not too small. □

C.VII. VIX Futures Pricing

To deal with the square root in the expression of VIX, we adopt the numerical integration method in Appendix A.4. of Eraker and Wu (2017) to compute $F_t^{VIX}(\tau)$:

$$\begin{aligned}
F_t^{VIX}(\tau) &= E_t^Q[\sqrt{VIX_{t+\tau}^2}] \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-E_t^Q[e^{-sVIX_{t+\tau}^2}]}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-e^{-a_{1/12}s}E_t^Q[e^{-s(0,c_{1/12},d_{1/12})'X_{t+\tau}}]}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-e^{-a_{1/12}s}e^{\alpha(s,\tau)+\beta(s,\tau)'X_t}}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_{-\infty}^\infty e^{-s/2}\big(1-e^{-a_{1/12}e^s}e^{\alpha(e^s,\tau)+\beta(e^s,\tau)'X_t}\big)ds.
\end{aligned} \tag{C.57}$$

The second equality is a mathematical result using Fubini's theorem. The fourth equality follows from the definition of the (undiscounted) characteristic function, where $\alpha(s,\tau)$ and $\beta(s,\tau)$ are the solutions (evaluated at τ) to the ODE system (C.41) through (C.44) with boundary conditions $\alpha(0)=0;\beta(0)=(0,-c_{1/12}s,-d_{1/12}s)'$. The last equality follows from a change of variable to make the integrand bell shaped for easier numerical computation.

C.VIII. Equity Option Pricing

The normalized price of an equity put option is

$$P^E(X_t,\tau,K) = E_t^Q\Big[e^{-\int_t^{t+\tau}r_u du}\Big(K-P_{t+\tau}/P_t\Big)^+\Big] \tag{C.58}$$

$$= E_t^Q\Big[e^{-\int_t^{t+\tau}r_u du}\Big(K-e^{\ln P_{t+\tau}-\ln P_t}\Big)^+\Big]. \tag{C.59}$$

Using the Parseval identity (a theorem saying that the payoff function for an option stays unchanged under first a generalized Fourier transform and then a reverse generalized Fourier transform; see e.g., Lewis (2001)), we have

$$P^E(X_t, \tau, K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(K - e^{\ln P_{t+\tau} - \ln P_t}\Big)^+\Big] \tag{C.60}$$

$$= \frac{1}{2\pi} E_t^Q\Big[\int_{iz_i-\infty}^{iz_i+\infty} e^{-\int_t^{t+\tau} r_u du} e^{-iz(\ln P_{t+\tau} - \ln P_t)}\hat{\omega}(z)dz\Big], \tag{C.61}$$

where the generalized Fourier transform of the payoff function of the put option $(K - e^x)^+$ is given by

$$\hat{\omega}(z) \equiv \int_{-\infty}^{+\infty} e^{izx}(K - e^x)^+ dx \tag{C.62}$$

$$= -\frac{K^{iz+1}}{z^2 - iz}, \tag{C.63}$$

for $z_i \equiv Im(z) < 0$ (we restrict the imaginary part of z to be smaller than zero because, as one can easily verify, the integral in (C.62) exists if and only if $Im(z) < 0$). Then, taking the expectation operator inside the integral in (C.61) yields

$$P^E(X_t, \tau, K) = -\frac{1}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du} e^{-iz(\ln P_{t+\tau} - \ln P_t)}\Big]\frac{K^{iz+1}}{z^2 - iz}dz \tag{C.64}$$

$$= -\frac{1}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} e^{-iz(c^* - \ln P_t) - \frac{1}{2}\sigma_D^2 z^2 \tau} \varrho_X^Q\big(-iz(\phi, g_1^*, g_2^*)', X_t, \tau\big)\frac{K^{iz+1}}{z^2 - iz}dz, \tag{C.65}$$

where the second line follows from the definition of the discounted characteristic function and equation (70). Here $c^* \equiv g^* - g_1^*\sigma^{2*} - g_2^*\lambda^*$.

C.IX. VIX Call Option Pricing

The VIX call price can be written as

$$C^{VIX}(X_t, \tau, K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(VIX_{t+\tau} - K\Big)^+\Big] \tag{C.66}$$

$$= E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(\sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}} - K\Big)^+\Big]. \tag{C.67}$$

Using the Parseval identity, we obtain

$$C^{VIX}(X_t,\tau,K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(\sqrt{a_{1/12}+c_{1/12}\sigma_{t+\tau}^2+d_{1/12}\lambda_{t+\tau}}-K\Big)^+\Big] \tag{C.68}$$

$$= \frac{1}{2\pi}E_t^Q\Big[\int_{iz_i-\infty}^{iz_i+\infty} e^{-\int_t^{t+\tau} r_u du}e^{-iz(a_{1/12}+c_{1/12}\sigma_{t+\tau}^2+d_{1/12}\lambda_{t+\tau})}\hat{\omega}(z)dz\Big], \tag{C.69}$$

where the generalized Fourier transform of the payoff function of the call option $\big(\sqrt{x}-K\big)^+$ is given by

$$\hat{\omega}(z) \equiv \int_{-\infty}^{+\infty} e^{izx}(\sqrt{x}-K)^+dx \tag{C.70}$$

$$= \frac{\sqrt{\pi}Ercf\big(K\sqrt{-iz}\big)}{2(-iz)^{\frac{3}{2}}} \tag{C.71}$$

for $z_i \equiv Im(z) > 0$ (we restrict the imaginary part of z to be greater than zero because, as one can easily verify, the integral in (C.70) exists if and only if $Im(z) > 0$). Here $Ercf(\cdot)$ is the complex-valued complementary error function with an expression given by

$$Ercf(z) = 1 - \frac{2}{\sqrt{\pi}}\int_0^z e^{-u^2}du, \tag{C.72}$$

for any complex number z. Then, taking the expectation operator inside the integral in (C.69) and using the definition of the discounted characteristic function under risk-neutral measure as defined in (54), we can rewrite the expression for the call price in the following way:

$$C^{VIX}(X_t,\tau,K) = \frac{1}{4\sqrt{\pi}}\int_{iz_i-\infty}^{iz_i+\infty} e^{-iza_{1/12}}\varrho_X^Q\Big(-iz(0,c_{1/12},d_{1/12})',X_t,\tau\Big)\frac{Ercf\big(K\sqrt{-iz}\big)}{(-iz)^{\frac{3}{2}}}dz, \tag{C.73}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) > 0$.

C.X. VIX Put Option Pricing

For VIX put, the Fourier transform of its payoff function $\big(K-\sqrt{x}\big)^+$ is given by

$$\hat{\omega}(z) = -\frac{\sqrt{\pi}}{2}\frac{1}{(-iz)^{3/2}}\Big(1-Ercf(K\sqrt{-iz})\Big) \tag{C.74}$$

for $z_i \equiv Im(z) < 0$. Thus, to derive its pricing formula we only need to insert equation (C.74) into (C.69), which gives:

$$P^{VIX}(X_t,\tau,K) = -\frac{1}{4\sqrt{\pi}}\int_{iz_i-\infty}^{iz_i+\infty} e^{-iza_{1/12}}\varrho_X^Q\Big(-iz(0,c_{1/12},d_{1/12})',X_t,\tau\Big)\frac{1}{(-iz)^{3/2}}\Big(1-Ercf(K\sqrt{-iz})\Big)dz. \tag{C.75}$$

If we use integral variable substitution $x = z - z_i i$, a numerically implementable pricing formula obtains as

$$-\frac{1}{4\sqrt{\pi}}\int_{-\infty}^{+\infty} Re\Big[e^{(z_i-xi)a_{1/12}}\varrho_X^Q\Big((z_i-xi)(0,c_{1/12},d_{1/12})',X_t,\tau\Big)\frac{1-Ercf(K\sqrt{z_i-xi})}{(z_i-xi)^{\frac{3}{2}}}\Big]dx. \tag{C.76}$$

C.XI. Black (1976) Implied Volatility

We are interested in investigating the implied volatilities for VIX call options, which, compared with the usual Black-Scholes implied volatilities for stock options, warrant some extra comments.

The underlying asset of a VIX option is not the VIX index itself. Instead, it is a VIX futures contract with the same maturity as the option. Therefore, we should equate the pricing formula of options on futures in Black (1976) with the VIX option price either in the data or from the model and then invert the equation to obtain the implied volatility. Specifically, consider a VIX call option with strike K, time to maturity τ, and underlying price $F_t(\tau)$. Let the corresponding continuously compounded yield be $r_t(\tau)$. Under the assumptions of Black (1976), the price of the VIX call option should be given by[23]

$$BC(F_t^{VIX}(\tau),K,\tau,r_t(\tau),\sigma) = e^{-r_t(\tau)\tau}\big[F_t^{VIX}(\tau)\cdot N(d_1) - K\cdot N(d_2)\big], \tag{C.77}$$

where $N(\cdot)$ is the standard normal cumulative distribution function and

$$d_1 = \frac{\ln(F_t^{VIX}(\tau)/K)+\sigma^2\tau/2}{\sigma\sqrt{\tau}} \tag{C.78}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{C.79}$$

Then the model-implied implied volatility $\sigma_t^{imp} = \sigma^{imp}(X_t,\tau,K)$ should solve

$$\begin{aligned} C^{VIX}(X_t,\tau,K) &= BC(F_t^{VIX}(\tau),K,\tau,r_t(\tau),\sigma_t^{imp}) \\ &= BC(F^{VIX}(X_t,\tau),K,\tau,r(X_t,\tau),\sigma^{imp}(X_t,\tau,K)), \end{aligned} \tag{C.80}$$

[23]Note that no-arbitrage implies $VIX_t = e^{-r_t(\tau)\tau}F_t^{VIX}(\tau)$ if VIX were tradable. In this case, the Black (1976) formula reduces to the Black and Scholes (1973) formula.

where $C^{VIX}(X_t, \tau, K)$ is given by (C.73), $r_t(\tau) = -\ln\left(\varrho_X^Q(0, X_t, \tau)\right)/\tau$ by the definition of the discounted characteristic function, and $F_t^{VIX}(\tau)$ is given by (C.57).

Figures

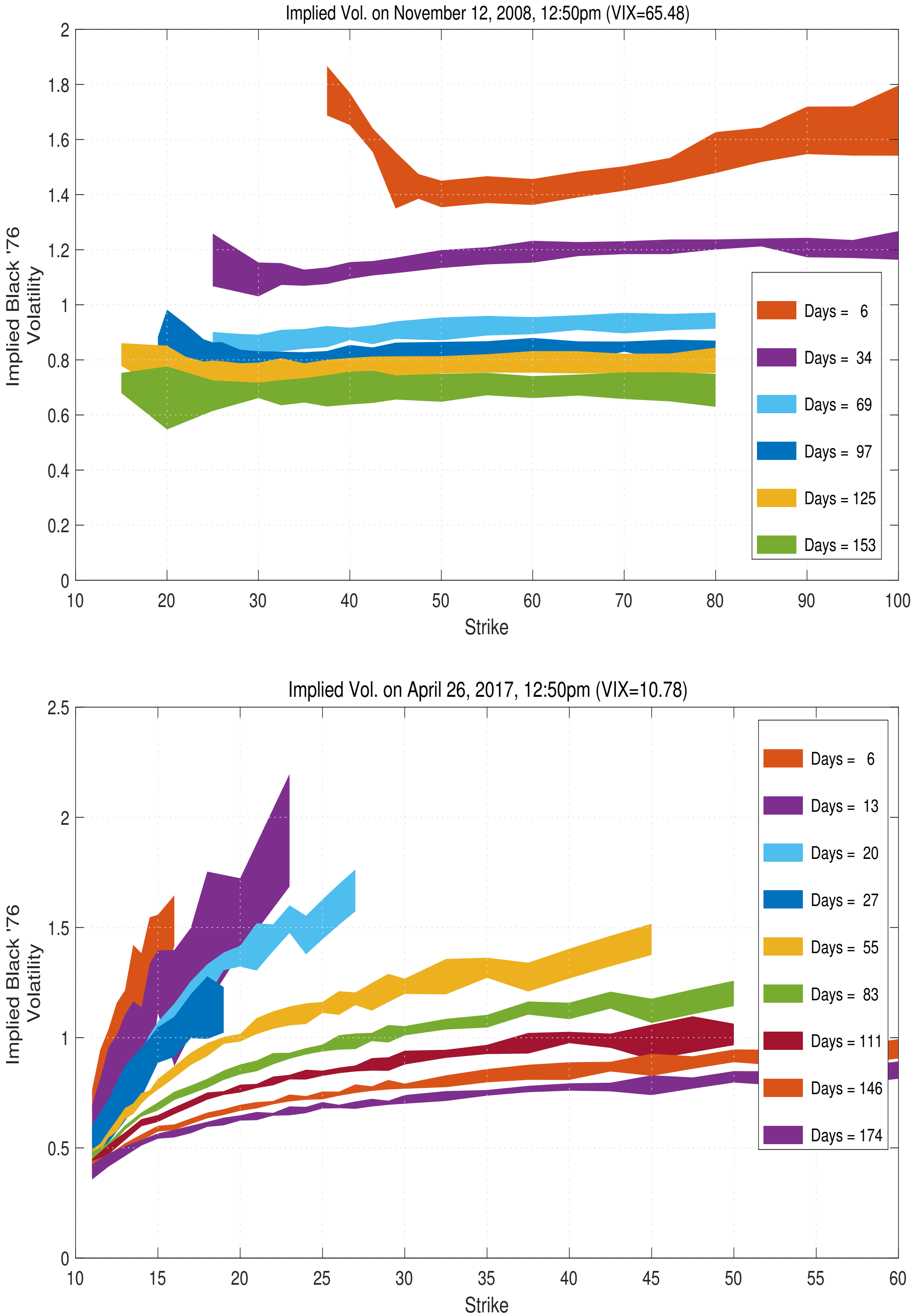

Figure 1: Implied VIX volatility on November 12, 2008 and March 26, 2017. The shaded areas represent the IV computed from bids and asks.

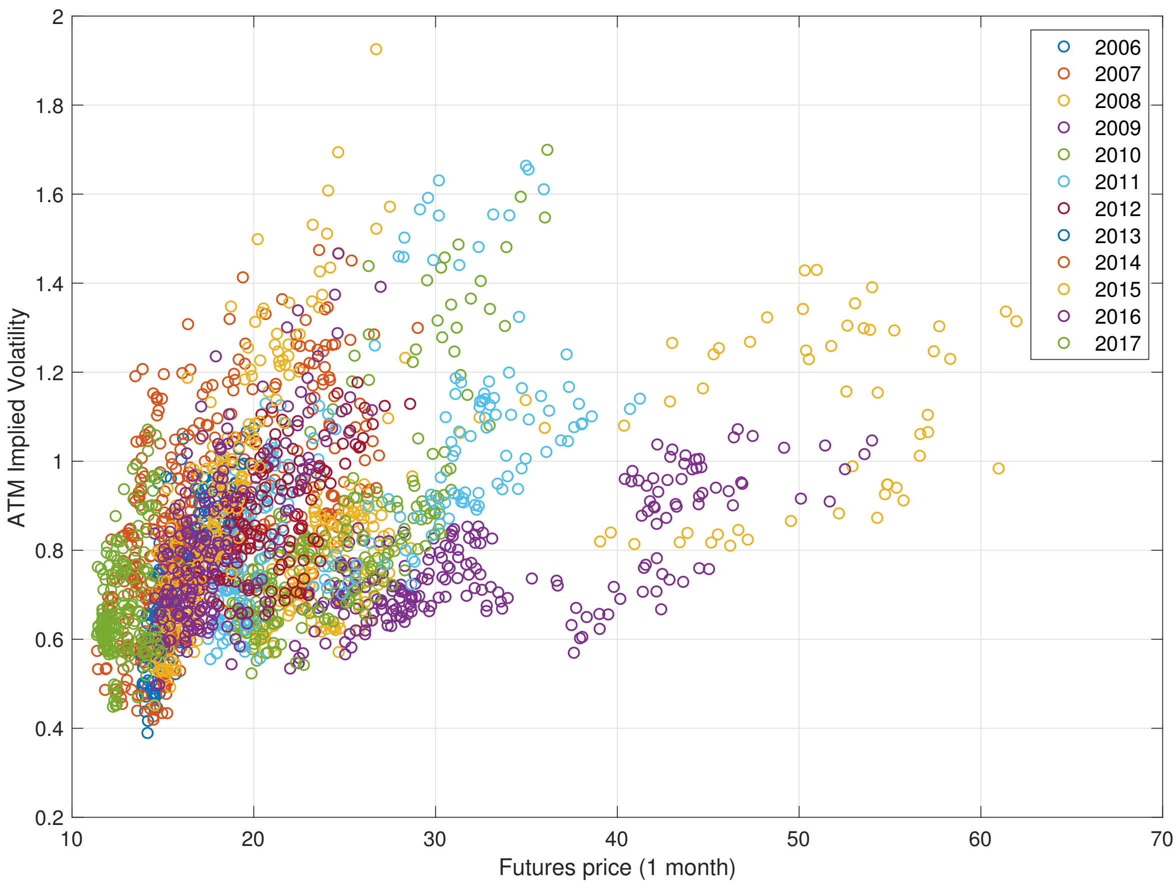

Figure 2: Scatter plot of one-month futures prices vs. one-month ATM implied VIX volatility.

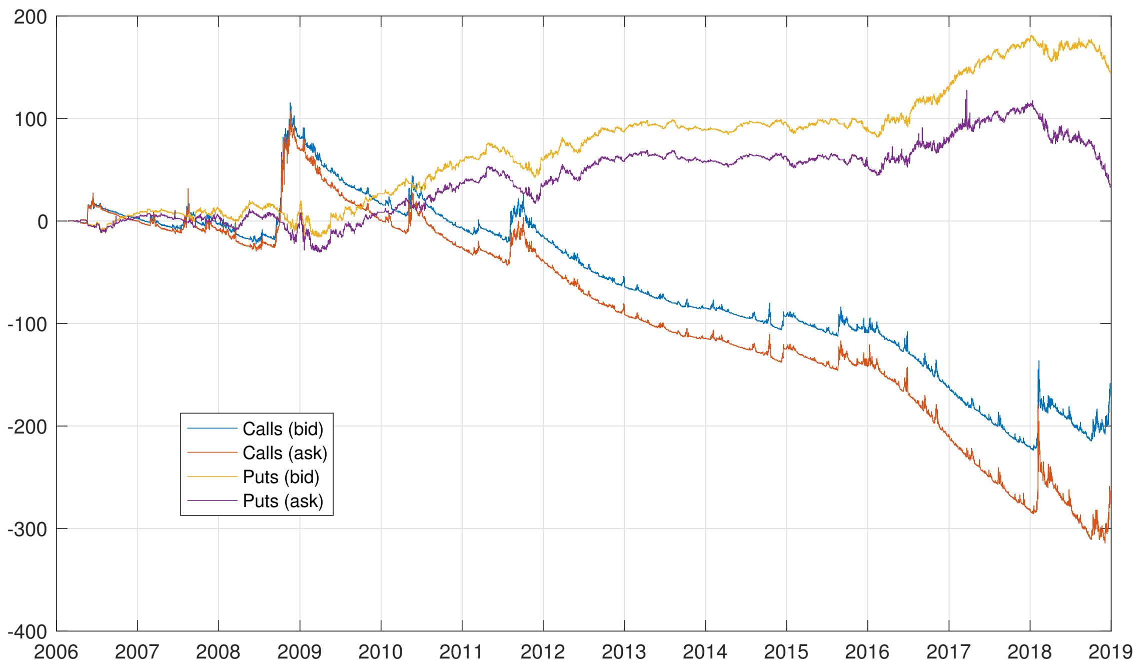

Figure 3: Marked-to-market value of 30 day options.

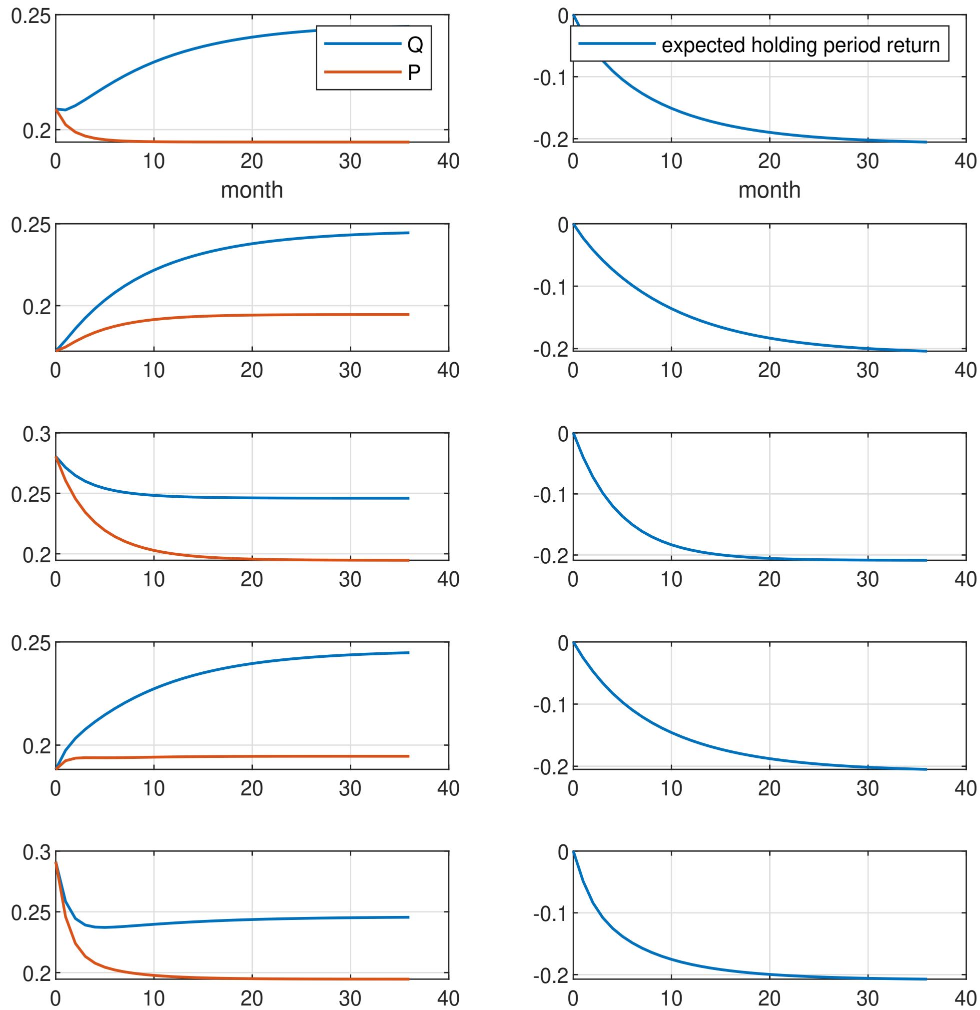

Figure 4: VIX futures curves and holding period returns

The figure illustrates conditional VIX futures term structures and conditional expected holding period returns on VIX futures. Left: VIX futures curves (Q) and the objective-measure expected payoffs (P). Right: expected holding period return, $E_t^P(VIX_{t+\tau})/E_t^Q(VIX_{t+\tau}) - 1$, to a long VIX futures position. State variables conditioned upon for each row are the following. First row: steady state σ_t^2 and λ_t; second row: low σ_t^2 and steady state λ_t; third row: high σ_t^2 and steady state λ_t; forth row: steady state σ_t^2 and low λ_t; last row: steady state σ_t^2 and high λ_t.

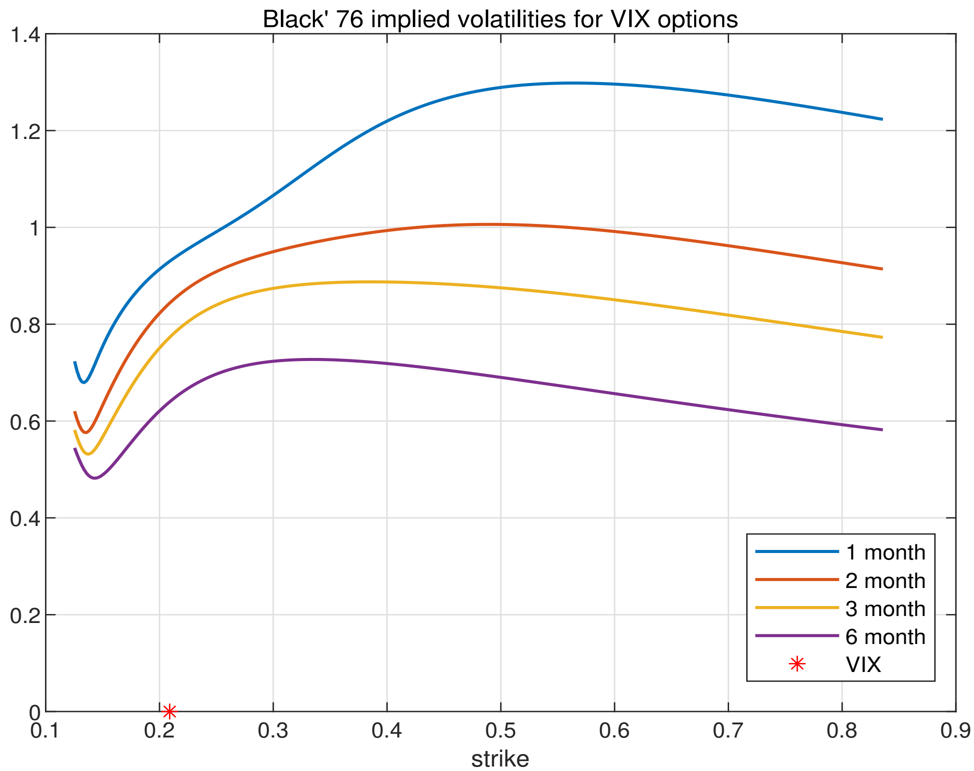

Figure 5: Black'76 implied volatility curves for VIX options in the model

This figure plots a full spectrum of implied volatilities computed from equating the Black(1976) futures option pricing formula with the VIX option price in the model at steady state. The horizontal axis denotes the absolute value of the strike. Implied volatilities are computed for VIX options with four maturities: one-month, two-month, three-month and six-month. There are three pronounced characteristics. First, there is clearly a timing premium across all strikes (given a strike, shorter-maturity option is always more "overpriced"), which reflects the mean reversion of the VIX index. Second, irrespective of maturity, the implied volatility curve is sloping upwards concavely to the right across most strikes. Third, the implied volatility curve eventually slopes downwards when strike exceeds a threshold.

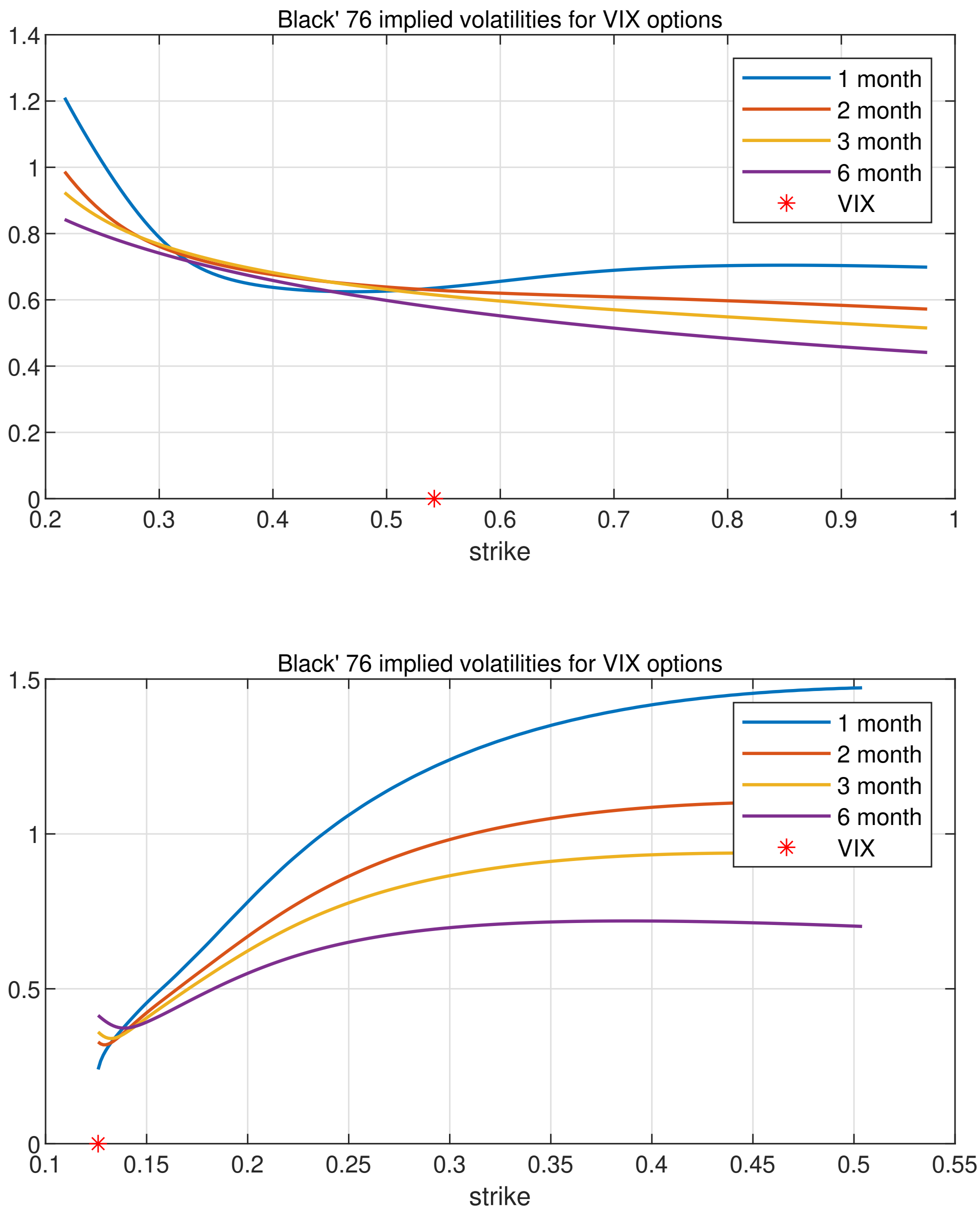

Figure 6: Black'76 implied volatility curves for VIX options in the model: conditional analysis

The figure plots implied volatility curves for VIX options in the model, respectively conditional on today's VIX being high and low. In the upper case, we set both state variables very high: $\sigma_t^2 = 10\sigma_{ss}^2$ and $\lambda_t = 10\lambda_{ss}$, implying a very high VIX, 54. In the lower case, we set both state variables at minimum values: $\sigma_t^2 = \lambda_t = 0$, implying a small value of VIX, 12.5.

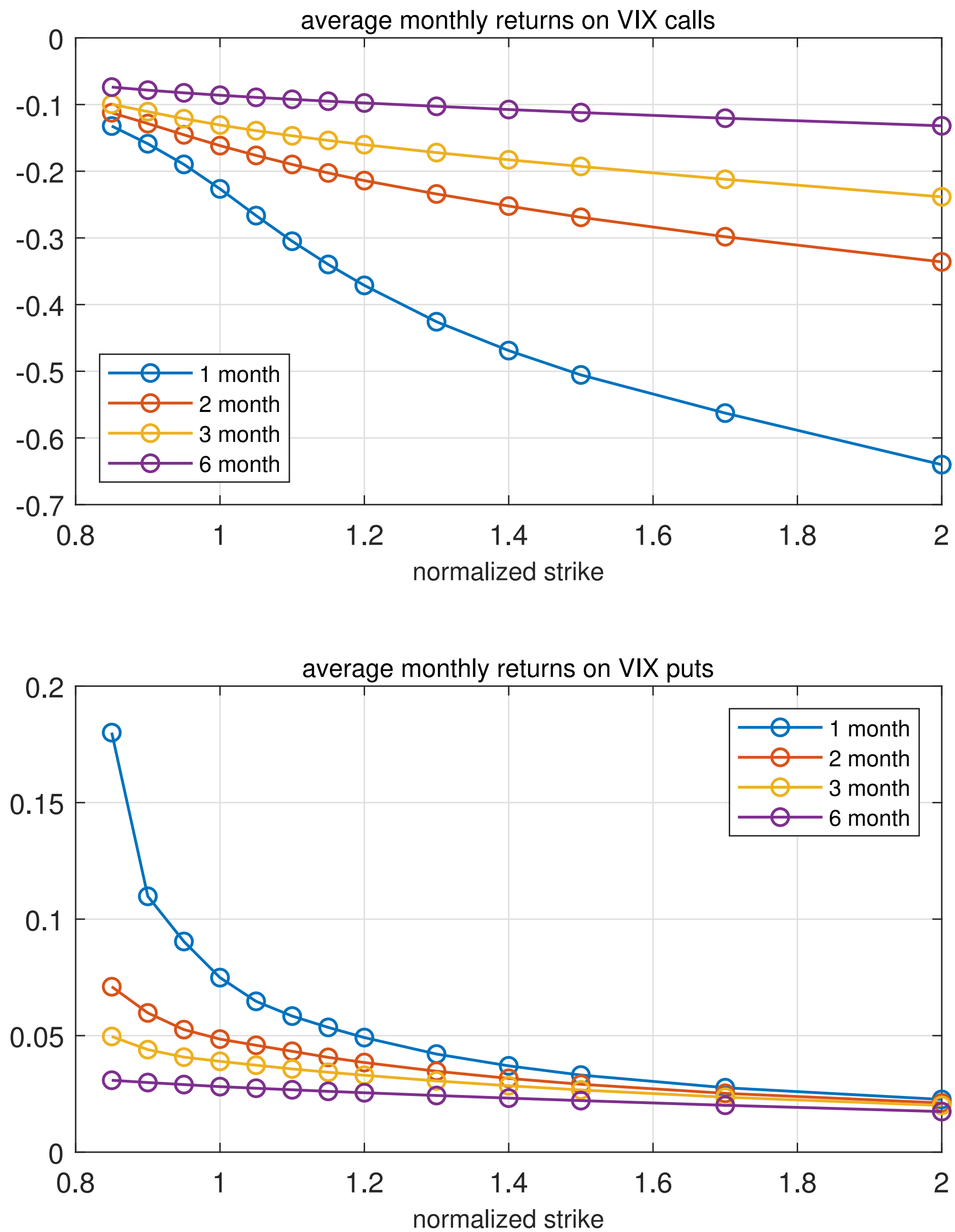

Figure 7: Average returns on VIX options

This upper (lower) figure plots average monthly returns on VIX calls (puts) in the model. In each case, we consider four maturities: 1,2,3, and 6 month, and the horizontal axis denotes the strike of relevant option normalized by its underlying asset price.

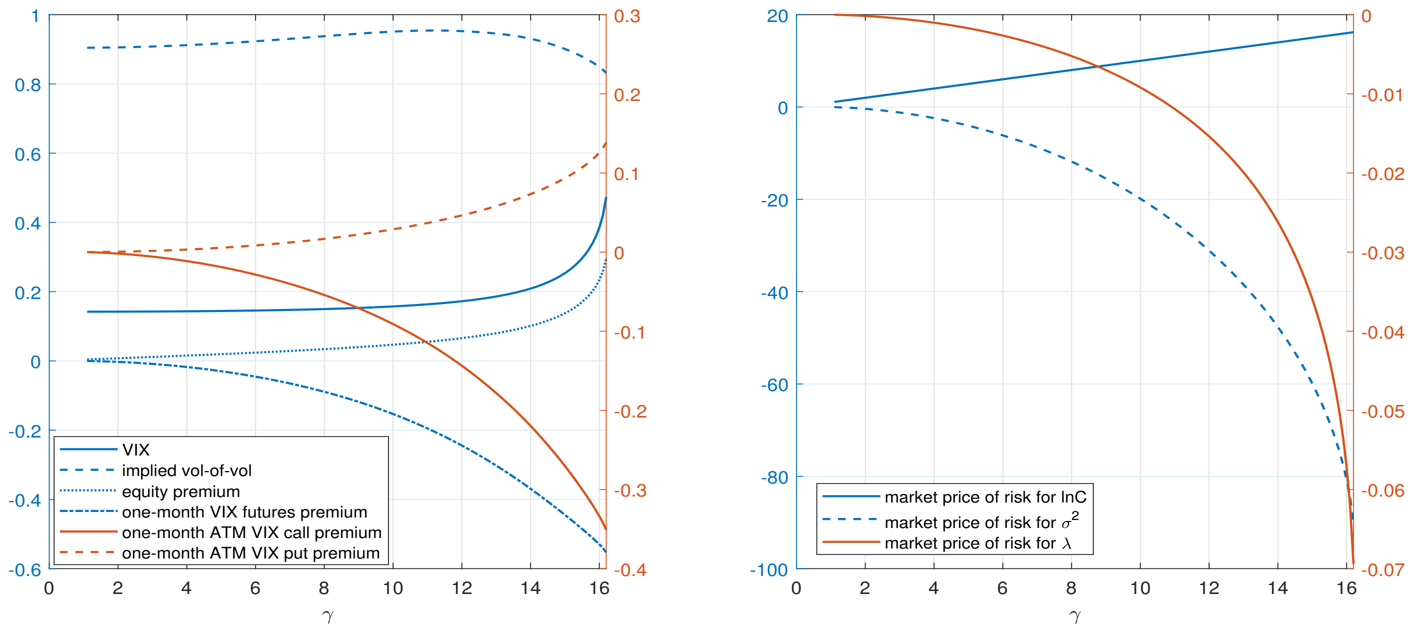

Figure 8: Comparative statics w.r.t. risk aversion

The figure illustrates conditional model moments (left) and market risk prices (right) as a function of the risk aversion, γ, conditional on state variables being at their respective steady state values. The left subplot reports steady-state VIX (VIX), implied volatility of ATM VIX options (implied vol-of-vol), instantaneous steady-state equity premium (equity premium), steady-state one-month VIX futures returns (one-month VIX futures premium), and average returns to ATM VIX calls and VIX puts. The right subplot reports the dependence of the market prices of risks for the model's three state variables, as represented by $(\gamma,-b_2,-b_3)$, on risk aversion, γ. The plot uses γ in the range from 1 to 16.2 - the upper limit for the existence of a model solution.

Tables

Table 1: Average Implied Black '76 Volatility

The table reports average implied Black '76 volatility for VIX options over the 2006-2018 period by maturity and strike.

	Maturity (months)			
	1	2	3	6
strike				
12	0.70	0.59	0.54	0.45
14	0.77	0.66	0.60	0.50
16	0.88	0.74	0.66	0.50
18	0.97	0.79	0.70	0.52
20	1.04	0.83	0.73	0.54
22	1.11	0.88	0.76	0.56
24	1.14	0.92	0.80	0.58
26	1.18	0.95	0.82	0.59
28	1.22	0.98	0.84	0.60
30	1.25	1.00	0.86	0.60
32	1.26	1.02	0.88	0.61
34	1.28	1.04	0.90	0.62
36	1.28	1.06	0.91	0.62
38	1.30	1.08	0.93	0.62
40	1.32	1.09	0.94	0.63

Table 2: VIX Option Returns

The table reports sample statistics on returns to option positions in VIX. Returns are defined as $\text{payoff}_T/p_0 - 1$ where T is the expiration and p_0 is the price (midpoint) of the option one month or six month prior to expiration. ATM is defined as the option with strike closest to the option-implied Futures price of the same maturity as the option. An OTM (ITM) call is defined as a call option with a strike that is 3 points higher(lower) than ATM. OTM/ITM conversely defined for put options. Confidence intervals (CI) for the expected returns are computed by bootstrapping the return distribution.

	CALLS			PUTS		
	ITM	ATM	OTM	ITM	ATM	OTM
			One month maturity			
mean	-0.23	-0.26	-0.23	0.07	0.13	0.20
95% CI	[-0.34,-0.10]	[-0.44,-0.02]	[-0.51,0.31]	[-0.00,0.14]	[0.02,0.24]	[-0.01,0.46]
std	1.16	1.99	3.37	0.65	1.00	2.22
Sharpe1	-0.70	-0.46	-0.23	0.35	0.43	0.31
Sharpe2	-0.62	-0.42	-0.37	0.23	0.32	0.18
skew	2.96	4.44	7.79	-0.15	0.31	2.47
kurt	14.76	29.14	82.61	2.31	1.87	9.68
			Six month maturity			
mean	-0.26	-0.34	-0.17	0.31	0.40	0.51
95% CI	[-0.48,0.03]	[-0.79,1.17]	[-0.50,0.44]	[0.19,0.43]	[0.24,0.55]	[0.30,0.74]
std	1.74	2.08	2.93	0.81	1.02	1.48
Sharpe1	-0.22	-0.23	-0.08	0.52	0.54	0.48
Sharpe2	-0.92	-0.37	-0.78	0.47	0.07	0.28
skew	3.53	3.80	4.92	-0.30	-0.02	0.48
kurt	17.05	17.18	28.79	2.16	1.84	2.00

Table 3: Parameters for the VIX Model

The table reports parameter values for the VIX model in Section 5. The processes for log consumption, consumption growth volatility, volatility jump arrival intensity, and log dividend are respectively given by

$$d\ln C_t = (\mu - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t^C$$

$$d\sigma_t^2 = \kappa^V(\theta^V - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t$$

$$d\lambda_t = \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda$$

$$d\ln D_t = \phi d\ln C_t + \sigma_D dB_t^D$$

where N_t is a Poisson process with instantaneous arrival intensity λ_t, and the jump size ξ_V is exponentially distributed with mean μ_ξ. The representative agent has recursive utility given by

$$V_t = E_t \int_t^\infty f(C_s, V_s)ds$$

$$f(C,V) = \beta(1-\gamma)V(\ln C - \frac{1}{1-\gamma}\ln((1-\gamma)V))$$

Parameters values are interpreted in annual terms.

Rate of time preference β	0.02
Relative risk aversion γ	14
Average growth in consumption μ	0.03
Mean reversion of volatility process κ^V	2.5
Average volatility-squared without jumps θ^V	0.0004
Diffusion scale parameter of volatility process σ_V	0.16
Average volatility jump size μ_ξ	0.005
Mean reversion of jump arrival intensity process κ^λ	12
Average intensity of a jump in volatility θ^λ	0.5
Diffusion scale parameter of jump arrival intensity process σ_λ	2.6
Stock market leverage ϕ	2.7
Idiosyncratic noise in dividend growth σ_D	0.1

Table 4: Simulation: Selected Model Moments

The table reports a list of model moments and their comparison with U.S. data. The model is simulated at a monthly frequency (dt=1/12) and simulated data are then aggregated to an annual frequency. All the moments in the first panel are on an annual basis. Δc denotes log consumption growth rate, Δd log dividend growth rate, pd log price-dividend ratio, r_t^e log return on the dividend claim, and r_t^f yield on one-year riskless bond. All the moments in the second panel are on a monthly basis, but the two variables VIX_t (risk-neutral log equity return volatility index) and imp_vol_t (Black'76 implied volatility for one-month ATM VIX option) are themselves annualized.

	Model	U.S. Data	Data Source
$E[\Delta c]$	2.91	1.80	BY2004
$\sigma(\Delta c)$	3.06	2.93	BY2004
$AC_1(\Delta c)$	0.25	0.49	BY2004
$E[\Delta d]$	7.99	4.61	CRSP
$\sigma(\Delta d)$	11.58	11.49	BY2004
$AC_1(\Delta d)$	0.23	0.21	BY2004
$E[\exp(pd)]$	28.08	26.56	BY2004
$\sigma(pd)$	9.10	29.00	BY2004
$AC_1(pd)$	0.03	0.81	BY2004
$E[r_t^e - r_t^f]$	8.83	8.33	Ken French
$\sigma(r_t^e)$	17.75	18.31	CRSP
$E[r_t^f]$	1.17	0.86	BY2004
$\sigma(r_t^f)$	2.87	0.97	BY2004
$E[VIX_t]$	19.46	19.28	CBOE
$\sigma(VIX_t)$	7.57	7.42	CBOE
$AC_1(VIX_t)$	0.80	0.84	CBOE
$E[imp\_vol_t]$	71.83	68.80	CBOE
$\sigma(imp\_vol_t)$	12.74	14.30	CBOE
$AC_1(imp\_vol_t)$	0.49	0.27	CBOE
$corr(VIX_t, imp\_vol_t)$	0.34	0.48	CBOE

Table 5: Simulation: VIX Futures Returns

This table reports descriptive statistics of the model simulated VIX futures returns. R^1 is the daily average arithmetic constant-maturity return and R^2 is the daily average logarithmic constant-maturity return, *Std* is the standard deviation of daily logarithmic constant-maturity returns. Data moments are from Eraker and Wu (2017). All numbers are in percentages.

Maturity	R^1	R^2	Std
	Model		
1 month	-0.10	-0.18	3.71
2 month	-0.09	-0.14	3.06
3 month	-0.08	-0.12	2.65
4 month	-0.07	-0.10	2.33
5 month	-0.06	-0.08	2.07
	Data		
1 month	-0.12	-0.20	3.98
2 month	-0.07	-0.11	3.00
3 month	-0.01	-0.04	2.47
4 month	-0.03	-0.05	2.21
5 month	-0.01	-0.03	2.01

Table 6: Simulation: Average VIX Option Implied Volatilities (I)

This table reports the model simulated average VIX option Black'76 (annualized) implied volatilities in percentages. Here, implied volatilities are reported across the option's moneyness, i.e., normalized strike. For example, the combination of 6-month and 120% means the option in consideration has a maturity of 6 month and a strike that is equal to 120% of its underlying asset price, i.e., the 6-month VIX futures price.

	80%	90%	100%	110%	120%	150%	200%
			Implied Volatility				
1 month	58.51	64.39	71.83	79.51	87.53	109.47	132.06
2 month	54.01	61.32	68.52	75.09	81.09	93.99	104.28
3 month	52.70	59.83	66.34	72.05	76.74	85.62	90.76
4 month	52.02	58.73	64.57	69.44	73.23	79.64	82.07
5 month	51.55	57.71	62.95	67.09	70.18	74.92	75.75
6 month	51.11	56.71	61.40	64.94	67.48	71.04	70.83

Table 7: Simulation: Average VIX Option Implied Volatilities (II)

The table reports the model simulated average implied Black '76 volatilities for VIX options by maturity and strike.

	Maturity (months)					
	1	2	3	4	5	6
strike						
12	0.69	0.61	0.58	0.57	0.56	0.56
14	0.63	0.54	0.50	0.48	0.47	0.47
16	0.71	0.62	0.57	0.53	0.51	0.50
18	0.79	0.69	0.64	0.60	0.58	0.55
20	0.87	0.77	0.71	0.67	0.63	0.60
22	0.96	0.84	0.76	0.71	0.68	0.64
24	1.04	0.89	0.81	0.75	0.71	0.67
26	1.11	0.93	0.84	0.78	0.73	0.69
28	1.16	0.97	0.87	0.80	0.75	0.71
30	1.21	0.99	0.89	0.81	0.76	0.72
32	1.24	1.01	0.90	0.82	0.77	0.72
34	1.27	1.03	0.91	0.83	0.77	0.73
36	1.30	1.04	0.91	0.83	0.77	0.73
38	1.32	1.05	0.92	0.83	0.77	0.72
40	1.34	1.06	0.92	0.83	0.77	0.72
42	1.35	1.06	0.92	0.83	0.77	0.72
44	1.36	1.06	0.92	0.83	0.77	0.71
46	1.37	1.06	0.91	0.82	0.76	0.71
48	1.37	1.06	0.91	0.82	0.75	0.70
50	1.38	1.06	0.91	0.81	0.75	0.70

Table 8: Simulation: VIX Option Returns

The table reports model moments of returns on holding VIX options, respectively for puts and calls, for ITM, ATM and OTM and for maturities of one and six months. For each maturity, in the first rows are mean returns; in the second rows are standard deviations of returns; in the third rows are Sharpe ratios; in the fourth and fifth rows are respectively skewness and kurtosis of returns. All numbers, except Sharpe ratios, are simply based on holding period total returns. Sharpe ratios are annualized. ITM and OTM respectively stands for 15% in-the-money and 15% out-the-money.

	CALLS			PUTS		
	ITM	ATM	OTM	ITM	ATM	OTM
			One month maturity			
mean	-0.14	-0.24	-0.35	0.05	0.07	0.17
std	1.22	2.09	3.09	0.69	1.11	2.89
Sharpe	-0.39	-0.39	-0.40	0.24	0.22	0.20
skew	4.13	6.54	9.41	0.06	0.77	4.07
kurt	33.31	68.43	127.73	2.41	2.93	28.57
			Six month maturity			
mean	-0.45	-0.52	-0.58	0.15	0.16	0.18
std	1.25	1.42	1.52	0.66	0.84	1.23
Sharpe	-0.52	-0.53	-0.54	0.30	0.26	0.19
skew	3.46	4.18	4.96	-0.48	-0.06	0.57
kurt	17.81	24.51	33.28	2.26	1.85	1.98