The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXVII, NO. 6 • DECEMBER 2022

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

BJØRN ERAKER and AOXIANG YANG*

ABSTRACT

We develop a tractable equilibrium pricing model to explain observed characteristics in equity returns, VIX futures, S&P 500 options, and VIX options data based on affine jump-diffusive state dynamics and representative agents endowed with Duffie-Epstein recursive preferences. Our calibrated model replicates consumption, dividends, and asset market data, including VIX futures returns, the average implied volatilities in SPX and VIX options, and first- and higher-order moments of VIX options returns. We document a time variation in the shape of VIX-option-implied volatility and a time-varying hedging relationship between VIX and SPX options that our model both captures.

SINCE THEIR INTRODUCTION IN 2006, options on the VIX index have become the second-most traded contracts on the Chicago Board Options Exchange (CBOE), surpassed only by S&P 500 (SPX) options.[1] The trading volume in VIX calls is about twice that of puts, reflecting a demand for speculative bets on, or hedges against, market turmoil in the form of high volatility. In this respect, VIX calls inherit some of the characteristics of out-of-the-money SPX put options, but with some important differences. We investigate the differences in the pricing of SPX and VIX options in this paper with a primary view toward understanding the factors that drive differences in the pricing of these two types of catastrophe insurance contracts.

*Bjørn Eraker is with Wisconsin School of Business. Aoxiang Yang is with Peking University HSBC Business School. We are grateful for comments from Stefan Nagel (Editor) and three anonymous referees. We also 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 University of Wisconsin for helpful comments. We have read *The Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

Correspondence: Bjørn Eraker, Finance Department, Wisconsin School of Business, 5274a Grainger Hall, 975 University Ave, Madison, WI 53706, USA; e-mail: bjorn.eraker@wisc.edu.

[1] The Option Clearing Corporation cleared a total of 1.5 trillion SPX trades versus 29 billion VIX trades year-to-date as of November 2020. See https://www.theocc.com/Market-Data/Market-Data-Reports/Volume-and-Open-Interest/Monthly-Weekly-Volume-Statistics

DOI: 10.1111/jofi.13182

VIX options prices display interesting features that differ markedly from SPX. For example, while implied Black-Scholes volatility is always a convex function of strike for SPX options, we document that the shape varies from concave in normal times to convex in high-volatility periods for VIX options. VIX options' implied volatilities decrease monotonically with maturity and generally increase in the strike. The opposite is true for SPX.

The main objective of this paper is to try to understand these price characteristics from the viewpoint of an equilibrium model. To this end, we derive an equilibrium model that reproduces salient features of VIX futures and options, SPX returns and options, and consumption and dividend data. The model features a representative agent with Duffie-Epstein recursive utility preferences who 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 aggregate stock market value obtains as the present value of a levered claim to consumption with unpriced risks, 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 for equilibrium VIX, and then apply a novel Fourier-type payoff transform analysis to derive a semiclosed form (up to a single integral) formula for the value of VIX options.

While there are countless studies of equity options market data, relatively few 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. Lin and Chang (2009) conduct a horse race between extant reduced-form models and conclude that jumps in volatility help explain VIX options data. Park (2015) uses information in SPX and VIX options to predict market returns (SPX), VIX futures returns, and SPX and VIX options returns. Huang et al. (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. Bakshi, Madan, and Panayotov (2015) derive a two-period model to price VIX options, attributing heterogeneity in beliefs to empirical evidence suggesting that both high- and low-volatility states carry high-risk premia. Park (2016) specifies a reduced-form model for VIX directly in order to price derivatives. This paper, to our knowledge, is the first to consider the pricing of VIX options, SPX options, the equity premium, the variance risk premium (VRP), and risk-free rates while maintaining the discipline imposed by a fully-fledged consumption-based equilibrium framework.

We start our analysis by first seeking to understand some basic properties of ex-ante pricing information in VIX options, including the patterns of implied Black-76 volatility surfaces. Among the interesting features of implied volatility data are the facts that they imply a severely right-skewed risk-neutral distribution of VIX "returns." The right-skewed distribution contrasts equity return distributions that tend to be negatively skewed, as with the SPX. The VIX returns distribution is skewed to the right much more heavily than SPX returns are skewed to the left. Unlike equity options, VIX options display a downward-sloping term structure: longer-term VIX options have lower implied Black volatility than do short-maturity VIX options. This persuasive feature persists irrespective of strikes and market conditions (i.e., high or low VIX). We show that this feature is related to mean reversion in VIX and lack thereof in the distributional assumptions underlying (Black-76) implied volatility computation. In addition, the shock to implied volatility of volatility (VVIX) is positively but imperfectly correlated with the level of VIX itself, suggesting that VIX options prices contain a component independent of VIX. This actually rules out single-factor conditional variance representations such as the models in Heston (1993), Bates (1996), and Eraker (2004). VIX options prices and implied volatilities can move independent of VIX only if there is a separate pricing factor that drives VIX options valuations. In fact, the essence of our two-factor model is to capture the independent moves in VIX options prices.

A second element of our descriptive empirical evidence is a look at ex-post realized VIX options returns. Huang et al. (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 significantly negative. 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 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.

Both VIX calls and SPX puts constitute crash insurance. During a short window of about 20 days in March 2020, the S&P 500 index fell almost 30% from its high in January of the same year. Intraday VIX peaked at more than 80. Figure 1 shows the holding period returns to an investor who were to buy and hold the respective option contracts over the height of the Covid-19 crisis period. Some interesting features of the data emerge. For SPX, option prices rose throughout March and dramatically so on March 16 and 18, days on which the SPX index dropped 12% and 5.18%, respectively. A fortuitous investor who bought the farthest out-of-the-money (OTM) SPX put with strike 1600 in the beginning of March would have had more than a 200-fold increase in value if she had sold out on either one of these days. Note that over this particular sample window, returns are monotonically decreasing in strikes, so the 1600 strike put had the highest return, although this as well as the other low-strike SPX puts eventually expired worthless. Investors in VIX options fared even

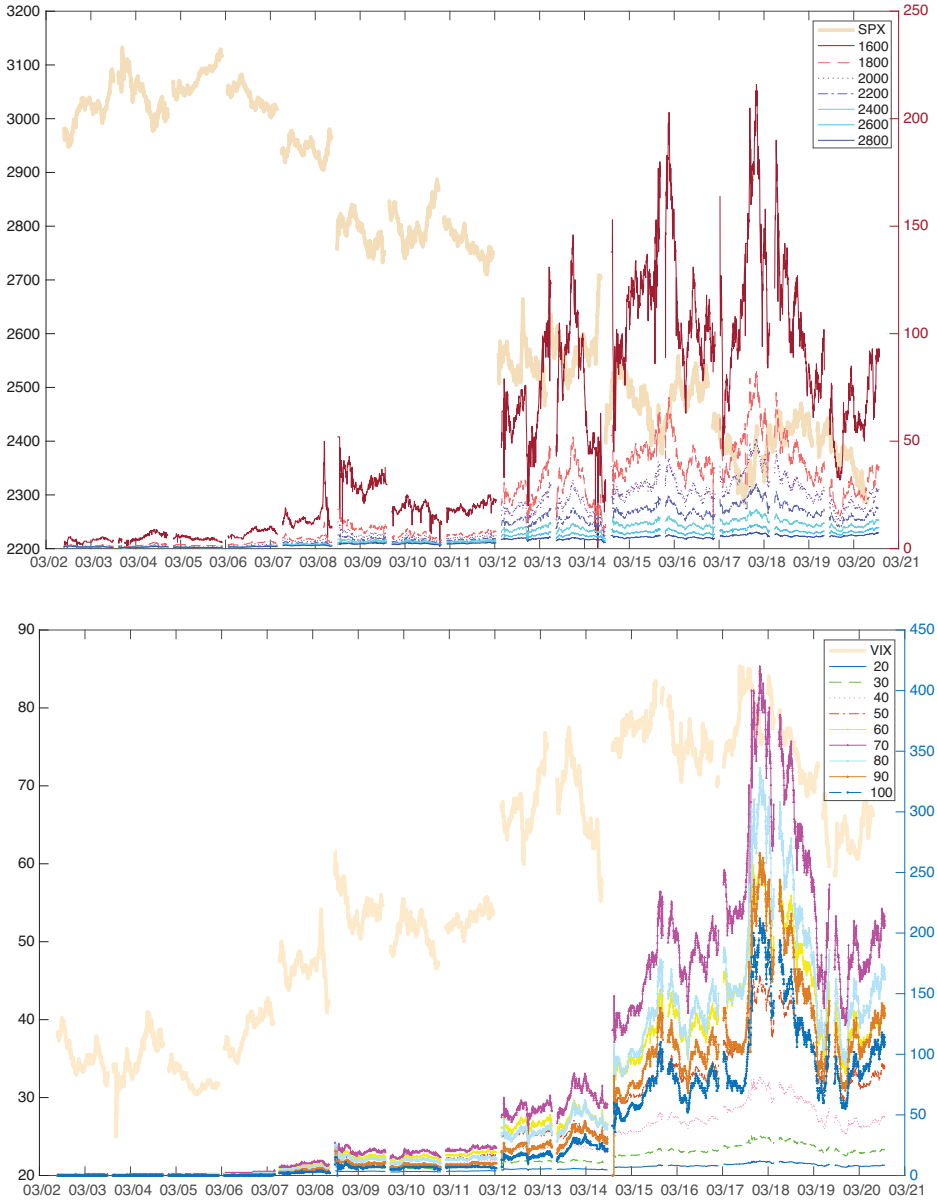

Figure 1. Returns to OTM SPX puts and VIX calls during the height of the Covid-19 crisis, March 2020. (Color figure can be viewed at wileyonlinelibrary.com)

better than SPX. The price of a 70 strike VIX call increased 400-fold from March 2 to March 18, although this option would also expire worthless.

Figure 1 also suggests that the correlation between SPX and VIX options is high, but imperfect. In particular, while prices of SPX puts peaked twice on March 16 and 18, VIX calls showed a much larger spike on March 18. We find that VIX calls correlate even less with SPX puts during calm market periods. Foreshadowing our model's implications, SPX options are impacted by cash flow shocks that do not impact VIX options. The cash flow component is more important than the discount rate component in driving SPX options during periods of low market volatility, whereas the opposite holds during periods of high market volatility. The presence of cash flow shocks breaks the correlation between SPX and VIX options, particularly during low- and normal-volatility periods.

We analyze the factor structure of VIX call and SPX put returns. The results suggest that one distinct common factor drives about 80% of the variation across the two markets. We dub a second factor "SPX skew," a third "extreme tail," and a fourth a pure "VIX" factor. We also study the relationship between the crash insurance offered by OTM SPX puts and OTM VIX calls using reduced-form regressions. Specifically, we study the extent to which SPX options can be hedged with a delta hedge (SPY), a position in VIX futures, as well as positions in VIX calls. We find that during calm market periods, VIX calls do not correlate substantially with SPX puts, and thus, do not improve on hedging performance. By contrast, during turbulent periods, VIX calls substantially reduce hedging errors. We compare these empirical results to results obtained by simulating data from our model and find that the model replicates the results. To understand these relationships better, we link SPX and VIX options returns to exogenous shocks to state variables in our model. In low-VIX regimes, we find that both VIX and SPX options respond linearly to shocks. The variance decompositions show that higher order polynomials of the innovations in the state variables, which proxy for convexity, are important for explaining option returns during high-VIX periods. This explains how VIX options can be useful for hedging SPX positions (or vice versa) in periods of market distress.

Our model matches a number of observed moments of macro quantities and asset prices, starting with the first two moments of aggregate consumption and dividend growth, interest rates, and stock returns. It matches the consumption and dividend growth and interest rate data up to negligible differences, matches both physical and risk-neutral equity market volatility, and generates an equity premium and a VRP both close to those seen in the data. Our model replicates the first two moments of VIX futures ex-post returns with reasonable precision. It also accurately captures the SPX option-implied volatility curves. It further matches various moments of VIX option ex-post return distributions, including mean, variance, skewness, and kurtosis. It matches implied VIX volatility along several dimensions: the average at-the-money (ATM)-implied VIX volatility (i.e., VVIX) is matched almost identically, and the average implied volatility surface over maturity and strike is similar to what we observe—it is concavely and vastly skewed to the right (consistent with

an extremely right-skewed underlying VIX distribution), and it has a sharply downward-sloping term structure. Remarkably, our model also reproduces the change in the implied volatility curve (as a function of the strike) from concave during low- and average-VIX periods to convex during periods of market stress. We argue that this unique model implication is related to the flexibility afforded by our two-factor model.

To derive our model, we first develop a general framework for pricing assets under recursive Duffie-Epstein preferences with intertemporal elasticity of substitution (IES) set to one under the assumption that state variables follow affine jump diffusions, as in Duffie, Pan, and Singleton (2000). The model builds on Duffie and Epstein (1992), Duffie and Lions (1992), and Duffie and Skiadas (1994), shares similarities with the models of Eraker and Shaliastovich (2008), Benzoni, Collin-Dufresne, and Goldstein (2011), and Tsai and Wachter (2018), but has a clear marginal contribution that it is an endowment-based equilibrium model with (i) clearly stated affine state variable dynamics and (ii) precisely characterized equilibrium value function, risk-free rate, prices of risk, and risk-neutral state dynamics. We prove that our state-price density is a precise $IES \to 1$ limit of that approximately solved in Eraker and Shaliastovich (2008). The recursive preference assumption implies that higher order conditional moments of the economic fundamental, such as its growth volatility and volatility-of-volatility, are explicitly priced in equilibrium. Since VIX derivatives depend on these factors, this, in turn, implies that the former carry nonzero risk premia.

The rest of the paper is organized as follows. Sections I and II describe our sample of VIX options and presents reduced-form evidence, respectively. Section III presents our equilibrium model of VIX option pricing. Section IV presents results from our model calibration exercise, and Section V summarizes our findings. The Internet Appendix[2] contains our general theory as well as model derivations and extensions.

I. Data

The sample was collected from the CBOE[3] and consists of data sampled at the one-minute interval over the period January 2006 to June 2020. The data set consists of best bids, best asks, bid/ask quantities, and open high/low in addition to contract characteristics over the one-minute intervals. The fact that the data are time-stamped down to the minute interval mitigates the problem of nonsynchronous quotes, which are often problematic in end-of-day data.

VIX options and futures are cash-settled to a special VIX computation with ticker code VRO. VRO is computed from prices of constituent SPX options that are compiled through a special auction that is held premarket on the VIX expiration day, typically the third or fourth Wednesday of the month. This is in

[2] The Internet Appendix is available in the online version of this article on *The Journal of Finance* website.

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

contrast to the VIX itself, which is computed from midpoints. While in theory VRO should differ little from the open value of the VIX on the settlement day, in practice it may. Griffin and Shams (2018) suggest that the market is prone to manipulation since OTM SPX options can be traded cheaply while having a comparably large impact on VRO.

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-arbitrage 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 deviations. 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 maturity T, and F_t is 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 underlying asset of a VIX option contract is a same-maturity VIX futures contract implies that we should apply Black (1976)'s pricing formula to compute implied volatilities.

II. Exploratory Data Analysis

A. Option-Implied Volatilities

To characterize the pricing of VIX options, we first study implied volatilities. Figure 2 plots implied volatility for VIX options on two different days. On November 12, 2008, the VIX was high at 65.48, and on April 26, 2017, the VIX was low at 10.78. These days are typical of what we observe in high-

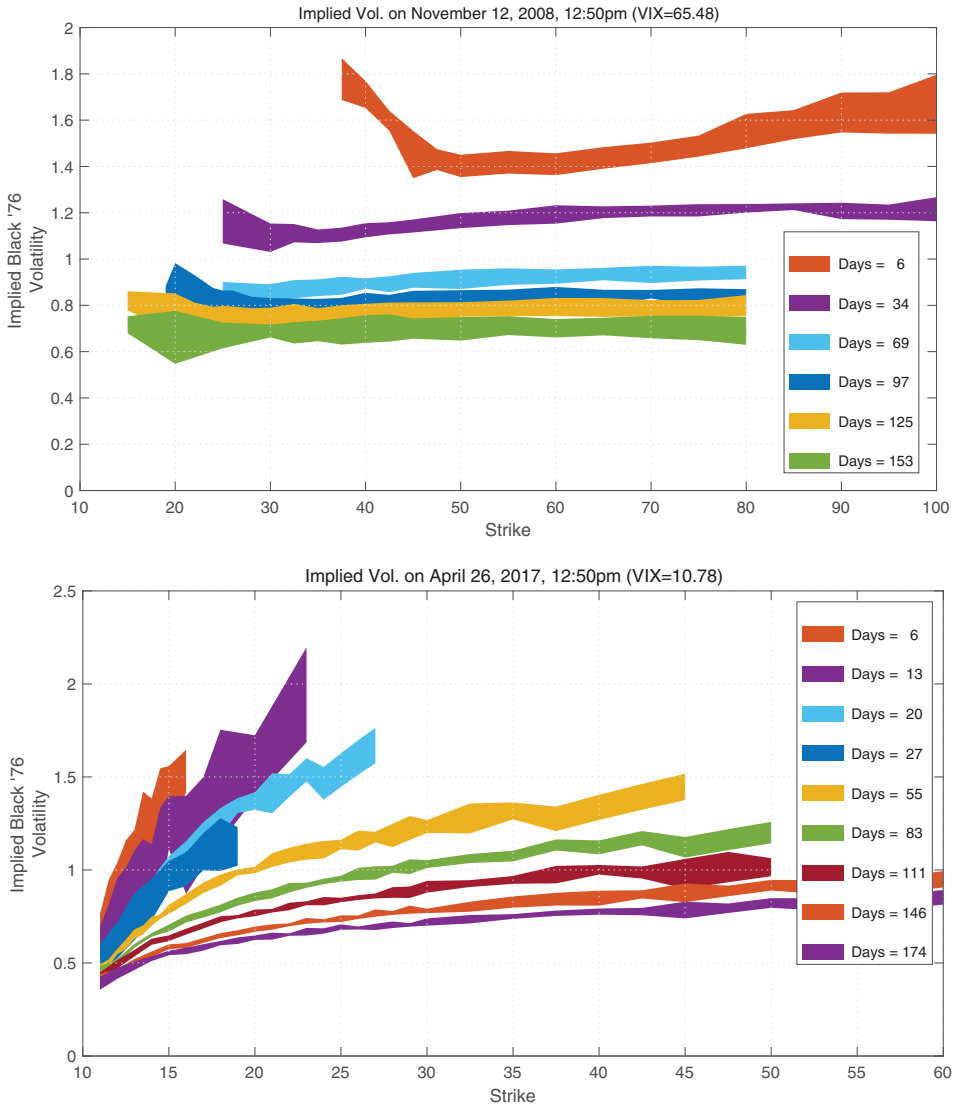

Figure 2. Implied VIX volatility on November 12, 2008 and April 26, 2017. The shaded areas represent the implied volatility computed from bids and asks. (Color figure can be viewed at wileyonlinelibrary.com)

and low-VIX states in our sample. A number of features of the data 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 the VIX. To understand why this is the case, note that Black-76 assumes that the underlying is a random walk.

Table I
Average Implied Black-76 Volatility

The table reports average (annualized) implied Black-76 volatility for VIX options by maturity and strike. The left panel reports results for data over the January 2006 to June 2020 period. The right panel reports results for the benchmark VIX model simulated over 100,000 months.

Maturity (Months)	Data				Model			
Strike	1	2	3	6	1	2	3	6
12	0.72	0.59	0.53	0.48	0.72	0.62	0.59	0.57
14	0.79	0.66	0.60	0.52	0.64	0.54	0.50	0.47
16	0.90	0.74	0.66	0.52	0.71	0.61	0.57	0.50
18	0.98	0.80	0.70	0.54	0.79	0.70	0.64	0.55
20	1.04	0.83	0.73	0.56	0.88	0.77	0.71	0.60
22	1.11	0.88	0.77	0.58	0.96	0.84	0.76	0.64
24	1.15	0.92	0.80	0.59	1.04	0.89	0.81	0.67
26	1.19	0.96	0.82	0.61	1.11	0.93	0.84	0.69
28	1.23	0.99	0.85	0.62	1.16	0.97	0.87	0.71
30	1.26	1.01	0.87	0.63	1.21	1.00	0.89	0.72
32	1.27	1.03	0.89	0.64	1.25	1.02	0.90	0.72
34	1.30	1.05	0.90	0.65	1.28	1.03	0.91	0.73
36	1.30	1.07	0.92	0.66	1.30	1.04	0.92	0.73
38	1.32	1.09	0.93	0.67	1.33	1.05	0.92	0.72
40	1.35	1.11	0.95	0.67	1.34	1.06	0.92	0.72

If a time series follows a random walk, its forecasted variance increases linearly with the forecast horizon. The downward-sloping term structure that we observe in VIX options implied volatility therefore suggests that the market does not think that VIX variance increases proportionally with the forecast horizon.

Second, the shapes of the implied volatility functions are mostly convex in the high-VIX case, especially at the left end of the strike, though they are mildly concave to the right in the high-VIX/short-maturity case, as can be seen in the red six-day maturity case in the top graph. In the low-VIX case, however, the implied volatility functions uniformly form a concave frown rather than the usual convex smile seen in equity options data, including the SPX.

Third, and perhaps most surprising, if we compare maturities in the 70- to 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 to 50 strike range were trading at implied volatilities below 100% in November 2008, while the 83-day maturity in the 40 to 50 strike range traded at implied volatilities above 100% in 2017. Our model successfully replicates all three main characteristics, as shown below.

The left panel of Table I shows the average VIX-implied volatility surface over strike and maturity, which should largely inherit the characteristics of the implied volatility surface in the low-VIX case that occurs a vast majority of the time. As can be seen, the two predominant patterns for the low-VIX case

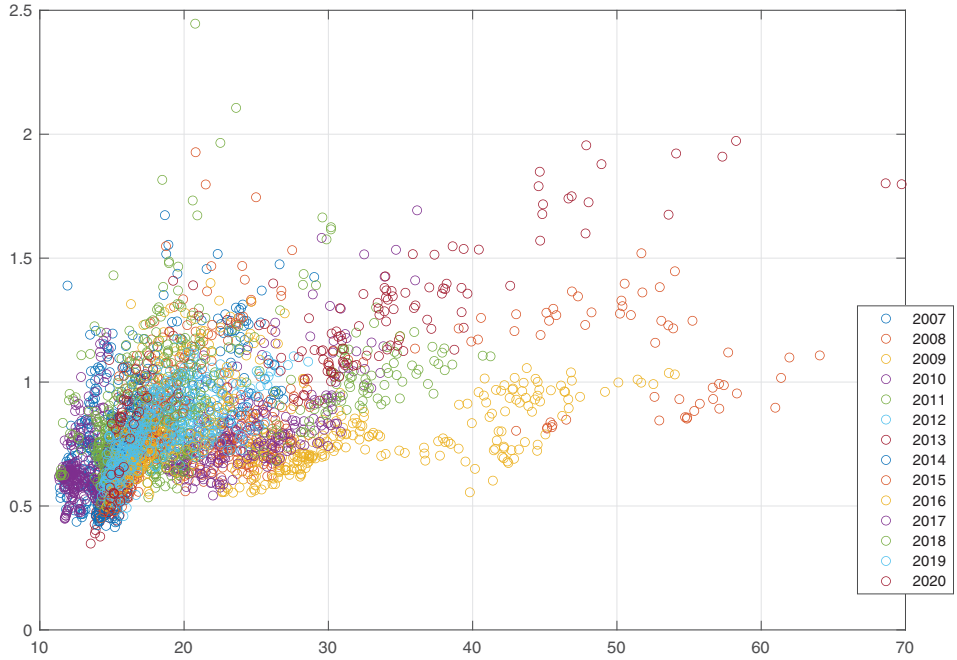

Figure 3. Scatter plot of one-month VIX futures prices versus one-month ATM-implied VIX volatility. (Color figure can be viewed at wileyonlinelibrary.com)

discussed above are visually evident: the term structure is sharply downward sloping and the volatility surface is increasing and concave in the strike levels. In particular, the fact that implied volatility keeps increasing in strike even for very high strike ranges indicates an extremely right-skewed underlying VIX distribution that cannot be rationalized without jumps.

Figure 3 shows the relationship between VIX level, as measured by one-month futures prices, and ATM VIX-option-implied 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. However, the strength of the relation between the futures level and the implied VIX volatility is 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, varies deterministically over the calendar. Rather, the evidence suggests that vol-of-vol, and hence VIX ATM-implied volatility, 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 Cox-Ingersoll-Ross (CIR) style diffusion term. Second, aggregate volatility is also discontinuous, with jumps arriving at a rate λ_t, which follows an independent self-exciting diffusion process. In equilibrium, both VIX and vol-of-vol are nonlinear functions of σ_t and λ_t. This

modeling specification allows us to match not only the positive but imperfect time-varying correlation between vol-level and vol-of-vol seen in Figure 3, but also the implied volatility surface presented in Figure 2 and Table I.

B. Option Returns

B.1. Descriptive Evidence

Much like traditional asset pricing research, recent developments in option research emphasize risk premia associated with factor-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 (2021) finds Sharpe ratios of about −0.5 for ATM straddles. It is well known that implied volatility exceeds realized volatility by some considerable amount (e.g., Jackwerth and Rubinstein (1996), Bollerslev, Tauchen, and Zhou (2009), among others), which is interpreted as a volatility risk premium.

Table II presents summary statistics on returns to VIX options positions using data from January 2006 to June 2020—a period covering both the 2008 financial crisis and the Covid-19 crash in March 2020. For comparison, Table III reports corresponding statistics from the SPX options markets over the same period. Both tables report buy-and-hold returns. The results suggest that VIX calls have negative average rates of return over the sample period. Moreover, the bootstrapped confidence intervals indicate that the returns are statistically significantly negative at a one-sided 5% sized test. This is similar to SPX put options, which also lose between 69% and 29% on average. Like VIX options, OTM SPX returns are statistically significantly negative at the 5% level using a one-sided test.

Long VIX put positions give negative exposure to VIX. In accordance with the negative risk premium associated with VIX futures positions, one might expect VIX puts to earn positive premiums, and they do: Table II shows that short-maturity puts on average have statistically significantly positive rates of return. Longer-maturity VIX puts yield close to zero average returns. The fact that VIX puts yield positive short-term and zero long-term average returns is consistent with extant evidence on average rates of return on variance swaps and VIX futures. Our findings are consistent with those of Eraker and Wu (2017) and Dew-Becker et al. (2017), who document a sharply downward-sloping term structure of risk premia for variance claims.

Figure 1 suggests that the pre-Covid-19 prices of VIX options were cheaper than those of SPX options. Below we seek to add to this anecdotal evidence by presenting a more detailed analysis of returns to the respective option classes over a longer period that includes the 2008 financial crisis as well as the Covid-19 crisis.

Figure 4 presents visual evidence on the performance of option investments in VIX calls and OTM SPX puts. These are not values of self-financing

Table II

VIX Option Returns

The table reports sample statistics on buy-and-hold returns to option positions in VIX. Returns are defined as $payoff_T/p_0 - 1$, where T is the expiration and p_0 is the price (midpoint) of the option one month or six months 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. OTM (ITM) is defined as a call option with a strike that is three points higher (lower) than ATM. Confidence intervals (CIs) for the expected returns are computed by bootstrapping the return distribution. The sample period is January 2006 to June 2020. Sharpe ratios are annualized.

	CALLS			PUTS		
	ITM	ATM	OTM	ITM	ATM	OTM
			One-Month Maturity			
Mean	−0.33	−0.48	−0.60	0.10	0.21	0.54
	[−0.47, −0.19]	[−0.66, −0.29]	[−0.81, −0.39]	[0.01,0.19]	[0.07,0.35]	[0.22,0.86]
Std	1.57	3.05	4.98	0.66	1.01	2.21
Sharpe	−0.73	−0.55	−0.42	0.55	0.73	0.85
Skew	5.60	6.70	7.56	−0.23	0.27	2.39
Kurt	50.51	59.47	68.29	2.24	1.89	9.33
			Six-Month Maturity			
Mean	−0.53	−0.59	−0.61	−0.13	−0.02	0.21
	[−0.79, −0.23]	[−0.86, −0.30]	[−0.92, −0.25]	[−0.27,0.01]	[−0.21,0.16]	[−0.07,0.49]
Std	1.94	2.58	3.28	0.80	1.01	1.45
Sharpe	−0.39	−0.33	−0.27	−0.24	−0.04	0.20
Skew	3.67	4.06	4.75	−0.31	−0.03	0.49
Kurt	18.68	21.70	27.22	2.16	1.84	2.03

Table III
SPX Put Option Returns

The table reports sample statistics on buy-and-hold returns to OTM S&P 500 options. Returns are defined as $payoff_T/p_0 - 1$, where T is the expiration and p_0 is the price of the option one, two, or six months prior to expiration. The sample period is January 2006 to June 2020. The 90% confidence interval for the mean return is computed by bootstrapping. Sharpe ratios are annualized.

	30d			60d			180d		
K/S	0−0.85	0.85−0.90	0.90−0.95	0−0.85	0.85−0.90	0.90−0.95	0−0.85	0.85−0.90	0.90−0.95
Mean	−0.69	−0.57	−0.30	−0.29	−0.44	−0.31	−0.79	−0.67	−0.57
	[−0.95, −0.36]	[−0.83, −0.26]	[−0.53, −0.03]	[−0.82, 0.34]	[−0.75, −0.06]	[−0.55, −0.05]	[−0.96, −0.58]	[−0.87, −0.44]	[−0.78, −0.34]
Std	3.55	3.34	2.41	6.64	4.18	2.52	1.02	1.11	1.09
Sharpe	−0.74	−0.60	−0.50	−0.08	−0.17	−0.26	−1.18	−0.90	−0.69
Skew	19.30	12.79	6.46	11.53	8.56	5.85	7.63	5.01	3.09
Kurt	402.18	202.79	61.79	153.37	92.44	47.63	62.33	30.75	13.78

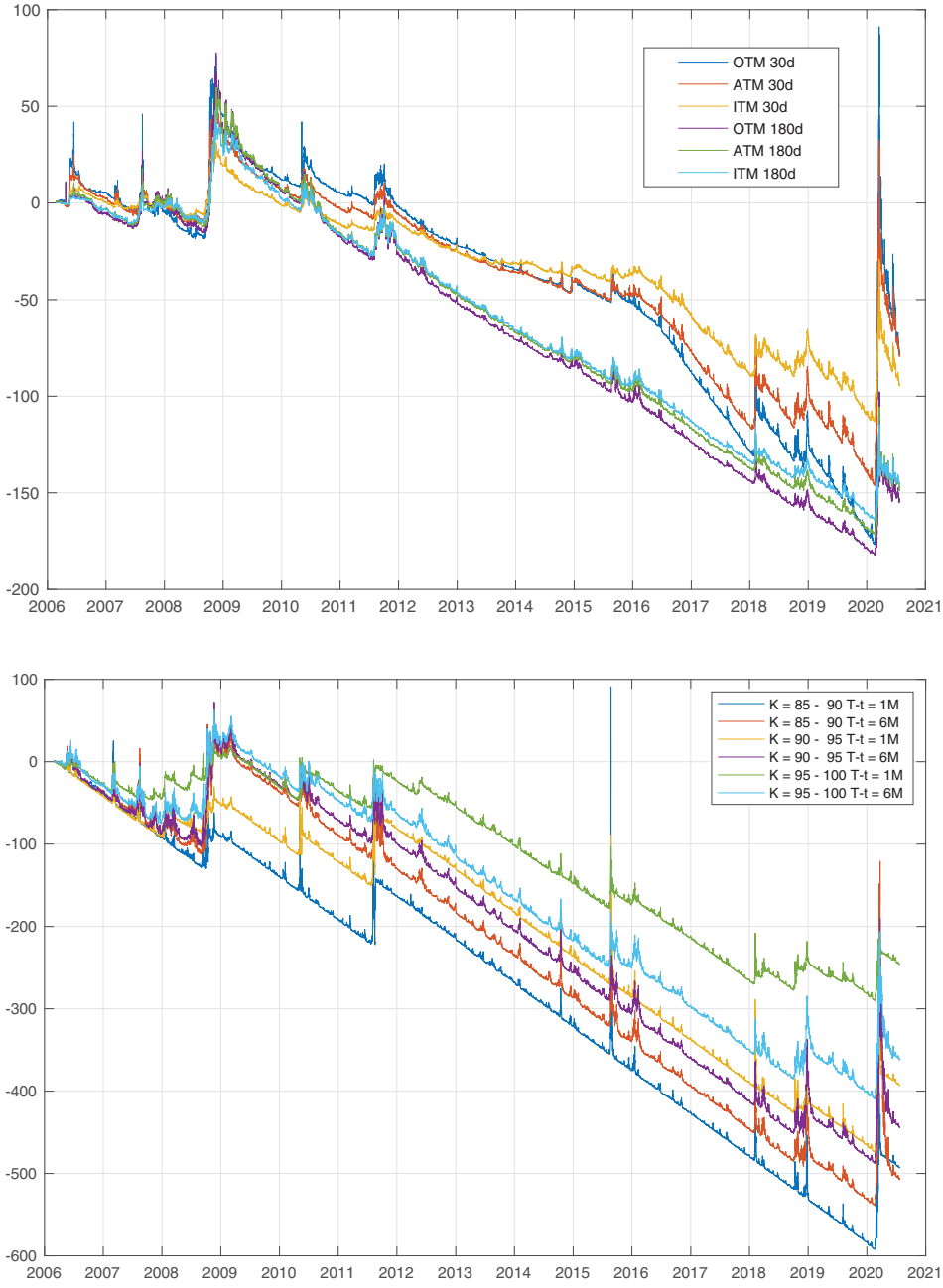

Figure 4. Marked-to-market value of 30- and 180-day VIX call options (top) and SPX put options (bottom). (Color figure can be viewed at wileyonlinelibrary.com)

portfolios, as we would typically look at for bond and stock investments. The reason we cannot compute the value of self-financing portfolios is that the majority of OTM options expire worthless. The natural way to overcome this is for portfolio managers to keep only a small fraction of ones' capital allocated to short or long positions. However, this approach implies that average returns depend on the arbitrary amount of starting capital.

To overcome this challenge, Figure 4 presents portfolio values of investments that are constantly replenished with cash. Specifically, we assume that one invests a single dollar each month in a target security based on moneyness and maturity. The graphs show the cumulative profit and loss on these investments including minute-to-minute marking-to-market. The figures show periods for which there is a constant one-dollar per month loss on average. This is true, for example, for the 2012 to mid-2015 period for all the SPX puts and for six-month maturity VIX calls. This happens as the options have a near-constant time decay plus some random, smaller moves due to fluctuations in market prices.

The graphs also show that the main driver of returns to these constant cash investments is the payoffs of the options. OTM options pay off during periods of financial turmoil. The fall of the 2008 financial crisis led to large payoffs for both VIX calls and SPX puts. Figure 4 shows that for both SPX and VIX options, March 2020 produced the largest payoffs seen in the sample. Short-term VIX options jumped so much that they temporarily erased the entire cumulative loss of the previous 14 years. Confirming evidence from Figure 1, the payoffs to VIX options holders were larger than for SPX.

Other periods, such as the European currency crisis periods in 2010 and 2011 also produce positive option payoffs but to varying degrees depending on the underlying, the strike, and maturities. There are also other noticeable features. For example, on August 24th, 2015, the Dow opened up 1,000 points lower in response to a substantial decline in the Chinese market. This led to an extreme spike in short-term OTM SPX put options, momentarily wiping out losses of the previous 11 years to buyers of these contracts. Prices quickly reverted and the episode had no impact on long-term performance. As seen in Figure 4, the impact on VIX options was much less dramatic. On February 5, 2018, the SPX fell 4.6%, while the VIX index more than doubled. The event, dubbed "volpocalypse" by some VIX market participants, forced the termination or restructuring of several VIX futures-linked exchange-traded-funds (ETF).

B.2. Principal Component Analysis

To further understand the connection between returns to SPX and VIX options, we perform principal component analysis (PCA) of the returns associated with the various maturity and moneyness categories. PCA and latent factor analysis are standard tools to uncover factor structures in returns. In their classic study, Roll and Ross (1982) apply factor analysis for equity returns, Christoffersen, Fournier, and Jacobs (2017) and Horenstein, Vasquez,

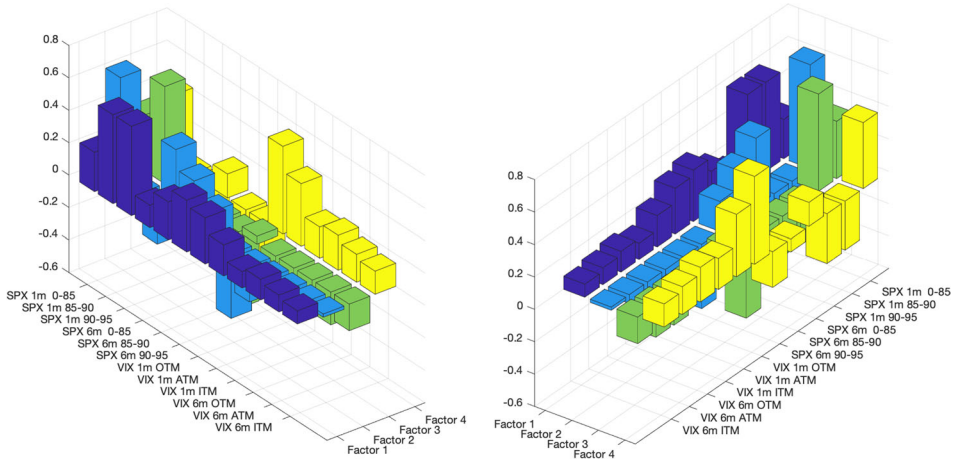

Figure 5. PCA factor loadings. The plots show the PCA factor loadings for the four first factors in the data as seen from two different angles. (Color figure can be viewed at wileyonlinelibrary.com)

and Xiao (2019) analyze factor structures in equity options, and Johnson (2017) performs PCA on the VIX term structure.[4]

Figure 5 shows factor loadings associated with the first four principal components. The first factor (dark blue) can be interpreted as a level factor. It loads highly on short-term SPX options and less on VIX options. However, this factor accounts for 85.5% of the variation in the data and, as such, is important for VIX options as well. The second (light blue) is a short-term SPX skew factor. It loads positively on short-term deep OTM SPX options and negatively on short-term near ATM options. This skew factor accounts for about 5.5% of the variation. The third factor (green) can be interpreted as an extreme tail factor as it loads positively on far OTM SPX options and negatively on everything else. It accounts for 3.8% of the variation. The fourth factor, which accounts for 2.5% of variation, is essentially a VIX factor, as it loads on VIX as well as far OTM SPX short-term options.

Figure 5 reveals that VIX and SPX options contain some common and some idiosyncratic components. In our equilibrium model, first, shocks to consumption variance and variance-of-variance are common to VIX and SPX, whereas cash-flow shocks are specific to SPX. Second, VIX and SPX endogenously obtain different exposures to these shocks. Both facts allow for an imperfect correlation between payoffs to SPX and VIX options. Our model generates a common factor structure, where the factor coefficients depend on "deep parameters" (e.g., persistence in priced risk factors) along with preference parameters.

[4] Other factor structure analyses for individual stock options include Bakshi, Kapadia, and Madan (2003), Bakshi and Kapadia (2003), Serban, Lehoczky, and Seppi (2008), Duan and Wei (2009), Vasquez (2017), and Brooks, Chance, and Shafaati (2018), among others. Different from these studies, we apply PCA on different types of options—SPX and VIX options.

C. VIX Options as Hedges for SPX Options

To understand exactly the relationship between SPX and VIX options, consider what it would take to synthetically recreate an SPX option by dynamically trading in the underlying in addition to other instruments. From Black and Scholes' seminal 1973 paper, we know that we can replicate the option payoff by holding delta number of shares of the SPX index, provided that the index follows a geometric Brownian motion.[5] If we generalize the distributional assumptions of Black and Scholes to include stochastic volatility, we will have to include an additional hedging instrument to hedge the SPX option. For example, under Heston (1993)'s model, it can be shown that the introduction of VIX futures will complete the market such that a dynamically adjusted portfolio of SPX futures and VIX futures will replicate SPX options. Likewise, model economies with additional risk factors need more instruments to complete the market.

Let $P_t = P(t, X_t)$ denote the price of an SPX put option where X_t is an N-dimensional state variable. Assuming that X_t is a continuous time, continuous path process, standard arguments imply that Ito's formula describes the dynamics of P,

$$dP_t = \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial X}' dX_t, \tag{3}$$

where the partial derivatives $\frac{\partial P}{\partial t}$ and $\frac{\partial P}{\partial X}$ are functions of t and X_t and the option's strike and maturity (arguments suppressed). We later make explicit assumptions about the evolution of X_t and preferences to derive an explicit formula for P_t. Absent any such assumptions, we can compute hedge coefficients through a regression

$$dP_t = \alpha_t + \beta_t' dX_t + d\epsilon_t, \tag{4}$$

where α_t and β_t are regression coefficients and $d\epsilon_t$ an error term. The regression coefficients β_t are time-varying as they approximate $\frac{\partial P}{\partial X}$, which depends on (t, X_t), provided that the data used to run the regression are sampled over a small time interval.

Figure 6 shows results from the regression

$$\triangle P_t^i = \alpha_{i,t} + \beta_{i,t}^{\text{SPY}} \triangle SPY_t + \beta_{i,t}^{\text{VIX futures}} \triangle F_t + \beta_{i,t}^{\text{VIX option}} \triangle C_t + error_{i,t}, \tag{5}$$

where $\triangle P_{t,i}$ are changes in SPX put options, $\triangle SPY_t$ is the change in the SPDR S&P 500 ETF, $\triangle F_t$ is the change in the front-month VIX futures contract, and $\triangle C_t$ is the change in the value of a VIX call option index.[6] The regressions

[5] As a matter of practical implementation, a hedger would have to trade the SPX futures or an SPX-linked ETF such as the SPY.

[6] The VIX option index is an equally weighted index of call midpoint prices that include options that are at least 10% out of the money. It implicitly puts more (less) weight on lower (higher) strikes. The averaging is applied to mitigate stale quotes.

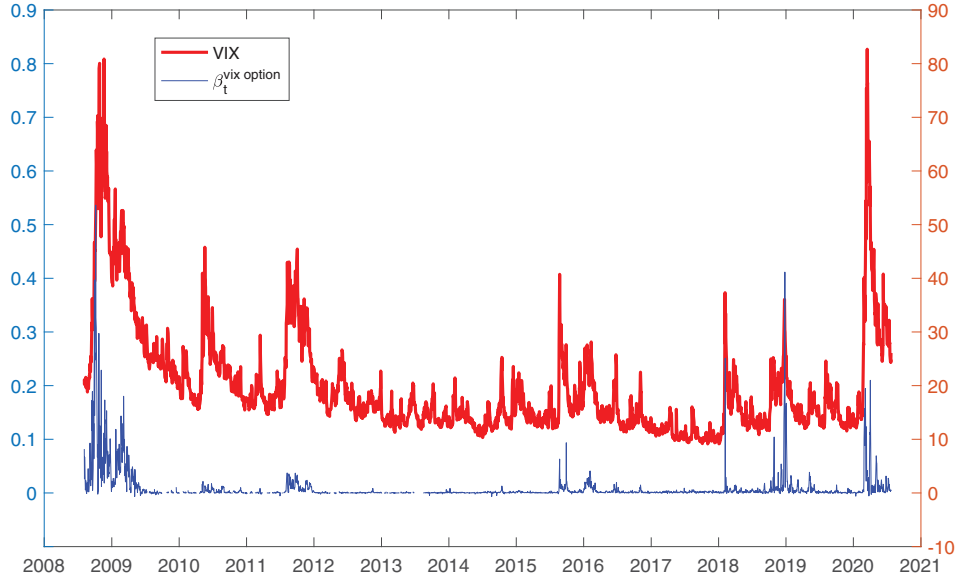

Figure 6. Hedge regressions. The figure shows $\beta_t^{\text{vix option}}$ computed through the regression $\triangle P_t^i = \alpha_{i,t} + \beta_{i,t}^{\text{SPY}} \triangle SPY_t + \beta_{i,t}^{\text{VIX futures}} \triangle F_t + \beta_{i,t}^{\text{VIX option}} \triangle C_t + error_{i,t}$, where P_t is the SPX put option, SPY_t is the S&P 500 index ETF, F_t is front-month VIX futures, and C_t is the OTM VIX call option. The regression is run intraday using 10-minute price changes from overlapping data sampled at the one-minute frequency. The figure shows the average estimated slope coefficient $\beta_t^{\text{VIX option}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \beta_{i,t}^{\text{VIX option}}$, where i indexes SPX put options that are at least 10% out of the money and have less than 40 days to maturity and N_t is the number of SPX options that satisfy these criteria on day t. (Color figure can be viewed at wileyonlinelibrary.com)

are run day-by-day using 10-minute price changes from overlapping data sampled at the one-minute interval. The typical number of observations is 405 one-minute intervals within a day. To average over sampling noise, the figure shows the average slope coefficient $\beta_t^{\text{VIX option}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \beta_{i,t}^{\text{VIX option}}$ for SPX puts with less than 40 days to maturity.

The figure shows the results for $\beta_t^{\text{VIX option}}$, allowing us to study the ability of OTM VIX calls to hedge OTM SPX puts. As can be seen, the estimated slope coefficients are close to zero during periods of market calm, and are positive during periods of high volatility. The results suggest that during high-volatility periods, such as the Great Recession and the COVID-19 crisis periods, VIX options serve as useful hedges for SPX options. Our two-factor model, which replicates the pattern seen in Figure 6, can be used to shine light on the nonlinear relationship between changes in OTM SPX puts and OTM VIX calls.

III. A Structural Approach to VIX Option Pricing

This section presents our model framework for pricing VIX options, which is a special case of the general model developed in Section I of the

Internet Appendix. We first specify a specific economic environment and describe the equilibrium VIX. We then perform a generalized Fourier payoff transform analysis to derive a pricing formula for VIX options as a single integral. Section III of the Internet Appendix derives pricing formulas for VIX futures and SPX options also as single integrals.

A. The Model

Consider an endowment economy with a representative agent who has Duffie and Epstein (1992) recursive preferences described by

$$V_t = E_t \int_t^\infty f(C_s, V_s) ds, \tag{6}$$

$$f(C, V) = \beta(1-\gamma)V\left(\ln C - \frac{1}{1-\gamma}\ln((1-\gamma)V)\right), \tag{7}$$

where V_t represents the continuation value. The parameter β is the rate of time preference, γ is the relative risk aversion, and the IES is implicitly set at one. Consumption, dividends, and ultimately, asset prices and returns are influenced by a key variable, namely, the conditional volatility of consumption growth, σ_t, which itself is exposed to both diffusion and jump risks. Specifically, we assume the following affine structure for the evolutions of state variables $X_t \equiv [\ln C_t, \sigma_t^2, \lambda_t]'$:

$$d\ln C_t = (\mu - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t^C, \tag{8}$$

$$d\sigma_t^2 = \kappa^V(\theta^V - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t, \tag{9}$$

$$d\lambda_t = \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^\lambda, \tag{10}$$

where $\ln C_t$ is the log consumption supply, and σ_t^2 is the instantaneous conditional variance of consumption growth. B_t^C, B_t^V, and B_t^λ are Brownian motions. $\xi_V dN_t$ is a jump term where N_t is a compounded Poisson process with instantaneous arrival intensity λ_t that itself follows a mean-reverting diffusion process, and ξ_V is a time-invariantly distributed random variable that represents the jump size with a moment generating function $\varrho(\cdot)$. Motivated by Eraker and Shaliastovich (2008), Park (2016), and our VIX option data, we assume that $\xi_V > 0$, implying that upward jumps in volatility are emphasized. We assume that 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 endowment dynamics abstract from important mechanisms in leading asset pricing models such as long-run productivity risks and rare disasters that occur to consumption. Instead, we focus on jumps to consumption growth volatility, which is natural given our VIX and VIX derivatives pricing concentration. In short, we pursue the simplest framework that captures as many aspects of VIX derivatives data as possible.

As Cox, Ingersoll Jr, and Ross (1985) discuss, the solution to (10) 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 $2\kappa^{\lambda}\theta^{\lambda} > \sigma_{\lambda}^2$, then the Feller (1951) condition 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 (10). 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(1) process. The model therefore implies that there are times when jumps to volatility can occur with high probability, but such times are themselves rare. For similar reasons, there is a σ_t term multiplying the Brownian shock in (9), which helps both prevent σ_t^2 from falling below zero and correctly replicate the right-skewness in its distribution.

B. State-Price Density

Section III of the Internet Appendix shows that the equilibrium value function of the representative agent is given by[7]

$$J(W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} \exp\left(a + b_2\sigma_t^2 + b_3\lambda_t\right), \tag{11}$$

where

$$a = \frac{1}{\beta}\left((1-\gamma)(\mu + \beta \ln \beta) + b_2\kappa^V\theta^V + b_3\kappa^{\lambda}\theta^{\lambda}\right), \tag{12}$$

$$b_2 = \frac{(\kappa^V + \beta)}{\sigma_V^2} - \frac{\sqrt{(\kappa^V + \beta)^2 - \sigma_V^2\gamma(\gamma-1)}}{\sigma_V^2}, \tag{13}$$

$$b_3 = \frac{\kappa^{\lambda} + \beta}{\sigma_{\lambda}^2} - \frac{\sqrt{(\kappa^{\lambda} + \beta)^2 - 2\sigma_{\lambda}^2(\varrho(b_2) - 1)}}{\sigma_{\lambda}^2}. \tag{14}$$

Assume that parameter values are such that b_2 and b_3 are both well defined. Note then that (13) implies $b_2 > 0$ if we assume $\gamma > 1$. Equation (14) then implies $b_3 > 0$, since by definition $\varrho(b_2) - 1 = E[e^{b_2\xi_V} - 1] > 0$ due to the positivities of both b_2 and ξ_V. Hence, from (11), 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

[7] Because the specific model is a special case of our general model, the IES=1 implication that the wealth-consumption ratio is constant, $W_t/C_t = 1/\beta$, is inherited, which contrasts with the data slightly. This is a sacrifice for precise framework tractability. Importantly for our purposes, the price-dividend ratio is not constant, as we show below.

agent. Both results are intuitive. Under recursive preferences, marginal utility depends on consumption as well as the value function, which is explicitly affected by σ_t^2 and λ_t. The agent thus requires compensation for bearing risks in both σ_t^2 and λ_t.

Section III of the Internet Appendix shows that the instantaneous risk-free rate is given by

$$r_t = \beta + \mu - \gamma\sigma_t^2, \tag{15}$$

where β represents the role of discounting, μ intertemporal smoothing, and $\gamma\sigma_t^2$ precautionary savings.[8] The Internet Appendix also shows that the state-price density is given by

$$\frac{d\pi_t}{\pi_{t^-}} = -r_t dt - \Lambda_t' dB_t + \left(e^{b_2\xi_V} - 1\right)dN_t - \lambda_t E[e^{b_2\xi_V} - 1]dt, \tag{16}$$

$$\Lambda_t = \Sigma(X_t)'\lambda, \tag{17}$$

$$\Sigma(X_t) = Diag\left(\sigma_t, \sigma_V\sigma_t, \sigma_\lambda\sqrt{\lambda_t}\right), \tag{18}$$

$$\lambda = (\gamma, -b_2, -b_3)', \tag{19}$$

where *Diag* represents a diagonal matrix. 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 all three state variables are priced is in sharp contrast with the CRRA utility model, in which only innovations to consumption are priced and VIX derivatives have zero premia in equilibrium.

Section III of the Internet Appendix shows that the evolution of the state variables under the risk-neutral measure Q induced by the state-price density is given by

$$d\ln C_t = \left(\mu - (\frac{1}{2} + \gamma)\sigma_t^2\right)dt + \sigma_t dB_t^{C,Q}, \tag{20}$$

$$d\sigma_t^2 = \kappa^{V,Q}(\theta^{V,Q} - \sigma_t^2)dt + \sigma_V\sigma_t dB_t^{V,Q} + \xi_V^Q \cdot dN_t^Q, \tag{21}$$

$$d\lambda_t = \kappa^{\lambda,Q}(\theta^{\lambda,Q} - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^{\lambda,Q}, \tag{22}$$

[8] Note that r_t can become negative when σ_t^2 is sufficiently high. A standard arbitrage when the real risk-free rate is negative involves borrowing consumption goods at negative rates, storing them until maturity, and then repaying a fraction back. This strategy does not work since no physical storage technology is available in the economy. For the same reason, negative real interest rates are also possible in models such as Bansal and Yaron (2004) (not because of log-linear approximation errors) and Wachter (2013).

where

$$\kappa^{V,Q} = \kappa^V - b_2\sigma_V^2; \quad \kappa^{\lambda,Q} = \kappa^\lambda - b_3\sigma_\lambda^2, \tag{23}$$

$$\theta^{V,Q} = \frac{\kappa^V}{\kappa^V - b_2\sigma_V^2}\theta^V; \quad \theta^{\lambda,Q} = \frac{\kappa^\lambda}{\kappa^\lambda - b_3\sigma_\lambda^2}\theta^\lambda. \tag{24}$$

Equation (20) shows that the drift of consumption growth is adjusted downward by $\gamma\sigma_t^2$ under the Q measure. Equations (21) through (24) show that, for both σ_t^2 and λ_t, mean reversion becomes slower and the long-run mean becomes higher under the Q measure. Moreover, the Internet Appendix 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. As analyzed in the general model, 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. In the special case in which 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 $\xi_V^Q \sim \exp(\frac{\mu_\xi}{1-\mu_\xi b_2})$ under Q.

By now, we have drawn on all of the key results from our general model, which help characterize the equilibrium value function, risk-free rate, pricing kernel, and risk-neutral dynamics. We next apply these results to price dividend strips, SPX, VIX, SPX options, and VIX futures and options.

C. Equity Price

Let us first price SPX. The continuous-time literature (Abel (1999), Campbell (2003), Wachter (2013)) specifies the aggregate dividend process, D_t, as leveraged consumption: $D_t = C_t^\phi$, so that D_t does not introduce a new state variable. However, this assumption has two undesirable consequences: first, ϕ shapes both the exposure of dividend risk to consumption risk and the average growth rate of dividend relative to consumption; second, consumption and dividend are perfectly correlated. To address these shortcomings, we follow the long-run risk literature (Bansal and Yaron (2004)) that models dividend and consumption separately,

$$d\ln D_t = \phi d\ln C_t + \mu_D dt + \sigma_D dB_t^D, \tag{25}$$

where ϕ captures stock market leverage, μ_D allows a flexible dividend growth rate, and B_t^D is a standard Brownian motion independent of any other random process in the model, thus representing the idiosyncratic risk in dividend growth.[9] As a result, the state variable $\ln D_t$ remains redundant: dB_t^D

[9] The long-run risk literature (Bansal and Yaron (2004)) typically assumes $d\ln D_t = \mu_d dt + \phi\sigma_t dB_t^C + \sigma_D dB_t^D$, which is equivalent to our specification (25) with properly chosen μ_d, up to

does not enter the pricing kernel, and $\ln D_t$ does not enter the value function, which one can confirm by including a fourth state variable $\ln D_t$ in X_t, solving the model all over again, and verifying that $b_4 = 0$. We note that the parameter σ_D has dual roles: besides governing the correlation between consumption and dividend, σ_D also affects dividend growth volatility, SPX return volatility, and eventually, the level and composition of VIX, thus affecting VIX derivatives premia. Our choice of σ_D in calibration takes care of both aspects.

Let $P(X_t, D_t)$ denote the price of the claim to all future aggregate dividends (the dividend claim). Then no-arbitrage implies that $P(X_t, D_t)$ is obtained as

$$\begin{aligned} P(X_t, D_t) &= \int_0^\infty E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} D_{t+\tau}\right) d\tau \\ &= e^{\sigma_D B_t^D + \mu_D t} \int_0^\infty e^{\left(\frac{\sigma_D^2}{2} + \mu_D\right)\tau} E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} e^{\phi \ln C_{t+\tau}}\right) d\tau, \end{aligned} \tag{26}$$

where the risk-neutral expectation in the first line represents the price of a dividend strip paid off τ periods ahead. To compute the risk-neutral expectation in the second line as well as other no-arbitrage asset prices such as riskless bond prices and derivatives prices, we follow Duffie, Pan, and Singleton (2000) and define an important function, the discounted characteristic function of X_t under the risk-neutral measure,

$$\varrho_X^Q(u, X_t, \tau) \equiv E_t^Q\left(e^{-\int_t^{t+\tau} r_u du} e^{u' X_{t+\tau}}\right). \tag{27}$$

Section III of the Internet Appendix shows that ϱ_X^Q is exponential affine in X_t for arbitrary $u \in \mathbb{C}^3$, and proves the following proposition.

PROPOSITION 1: *The equilibrium price of the dividend claim (i.e., SPX) is*

$$\begin{aligned} P(X_t, D_t) &= D_t G(\sigma_t^2, \lambda_t) \\ &= D_t \int_0^\infty e^{\left(\frac{\sigma_D^2}{2} + \mu_D\right)\tau + \alpha(\tau) + \beta_2(\tau)\sigma_t^2 + \beta_3(\tau)\lambda_t} d\tau, \end{aligned} \tag{28}$$

where $(\alpha(\tau), \beta_2(\tau), \beta_3(\tau))$ *solve ordinary differential equations (ODEs) in Section III of the Internet Appendix, and* $G(\sigma_t^2, \lambda_t)$ *is the price-dividend ratio function.*

D. Equity Premium

We next discuss the equity premium. No-arbitrage implies that the instantaneous equity premium conditional on no jumps occurring in our economy, as

a Jensen's term that is quantitatively unimportant. We write dividend in the form of (25) for convenience of applying the discounted characteristic function as defined in (27), a very useful tool in continuous-time models, which requires the dividend to be log-linear in consumption: $D_t = C_t^\phi e^{\mu_D t} e^{\sigma_D B_t^D}$.

shown in Section III of the Internet Appendix, is given by

$$\mu_{P,t} + \frac{D_{t^-}}{P_{t^-}} - r_t = \gamma\phi\sigma_t^2 - b_2\frac{G_1}{G}\sigma_V^2\sigma_t^2 - b_3\frac{G_2}{G}\sigma_\lambda^2\lambda_t + \lambda_t E\left[e^{b_2\xi_V}\left(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\right)\right]$$
$$= \sigma_{P,t}'\Lambda_t + \lambda_t E\left[e^{b_2\xi_V}\left(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\right)\right] \tag{29}$$

with

$$\sigma_{P,t} = \left[\phi\sigma_t, \frac{G_1}{G}\sigma_V\sigma_t, \frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\right]', \tag{30}$$

where G_1 and G_2, respectively, denotes the partial derivative of $G(\cdot, \cdot)$ with respect to σ_t^2 and λ_t. Four components arise in order. The first term, $\gamma\phi\sigma_t^2$, represents a standard constant relative risk aversion (CRRA) risk premium that 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. The Internet Appendix shows that $\beta_2(\tau)$ is negative for all τ as long as $1 < \phi < 2\gamma$, which we assume here and in our calibration, which immediately implies $G_1 < 0$ (i.e., the price-dividend ratio decreases 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 compensation for the diffusion risk in volatility jump intensity, dB_t^λ. The Internet Appendix 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 decreases in λ_t), and thus, the third component also takes a positive value. In contrast, the last term captures compensation for the jump risk in volatility, $\xi_V dN_t$. It is positive since $b_2 > 0$, $\xi_V > 0$, and $G_1 < 0$. Intuitively, when volatility jumps upward, two things happen simultaneously: first, marginal utility jumps upward by a percentage equal to $e^{b_2\xi_V}$, and second, the stock price jumps downward by a percentage equal to $1 - \frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}$. Investors therefore demand a jump risk premium for holding equity.

The instantaneous equity premium is given by (29) plus the expected percentage change in the equity price if a jump to volatility occurs. That is, the population equity premium in the economy is given by $\mu_{P,t} + \frac{D_{t^-}}{P_{t^-}} - r_t$ plus a negative term: $\lambda_t E[\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)} - 1]$. Finally, we can write the analytical expression for the population equity premium as

$$r_t^e - r_t = \sigma_{P,t}'\Lambda_t + \lambda_t E\left[\left(e^{b_2\xi_V} - 1\right)\left(1 - \frac{G(\sigma_t^2 + \xi_V, \lambda_t)}{G(\sigma_t^2, \lambda_t)}\right)\right]. \tag{31}$$

Note that the last term in (31) remains positive, meaning that the positive compensation for jump risks dominates the direct negative expected effect of jumps on equity returns. The above analysis establishes the following proposition.

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.*

E. VIX

Having obtained equilibrium SPX, we next define VIX. Given our model parameters have an annual interpretation, VIX, as a measure of risk-neutral 30-day forward-looking market volatility, can be expressed as[10]

$$VIX(X_t) = \mathrm{Std}_t^Q\left[\ln P_{t+1/12}\right]. \tag{32}$$

To express VIX as a function explicitly in state variables, we follow Eraker and Wu (2017) and use the property of the conditional cumulant generating function for $\ln P_{t+1/12}$, which requires expressing $\ln P_{t+1/12}$ as an affine function in state variables. Define the log price-dividend ratio as $g(\sigma_t^2, \lambda_t) = \ln G(\sigma_t^2, \lambda_t)$. It follows from (25), (28), and a highly accurate log-linear approximation of the price-dividend ratio G around steady states that[11]

$$\begin{aligned}\ln P_t &= g(\sigma_t^2, \lambda_t) + \ln D_t \\ &\simeq \left(g^* - g_1^* \sigma^{2*} - g_2^* \lambda^*\right) + g_1^* \sigma_t^2 + g_2^* \lambda_t + \phi \ln C_t + \mu_D t + \sigma_D B_t^D,\end{aligned} \tag{33}$$

where g_1 and g_2, respectively, denotes the partial derivative of $g(\cdot, \cdot)$ with respect to σ_t^2 and λ_t, and letters with asterisks denote relevant functions or variables evaluated at steady states. It follows that

$$VIX^2(X_t) = \mathrm{Var}_t^Q\left[g_1^* \sigma_{t+1/12}^2 + g_2^* \lambda_{t+1/12} + \phi \ln C_{t+1/12}\right] + \frac{1}{12}\sigma_D^2, \tag{34}$$

where to compute the conditional variance, we rely on the property of cumulant generating functions. Section III of the Internet Appendix shows that by doing

[10] Following the convention in the literature, we define VIX^2 as the risk-neutral variance of 30-day log market return. The precise definition of VIX^2 is $VIX_t^2 = -2(E_t^Q[\ln P_{t+1/12}] - \ln E_t^Q[P_{t+1/12}])$ as shown, for instance, in Martin (2011), Result 5. In Section IV of the Internet Appendix, we show that our results change quantitatively negligibly under the precise definition because the third- and higher-order conditional moments of log equity return are relatively unimportant compared with the second-order conditional moment. We thank an anonymous referee for this point.

[11] Our such log-linearization is highly accurate for two reasons. First, as in Seo and Wachter (2019), our log-linearization of the price-dividend ratio is used only after the price-dividend ratio is exactly solved out. This is different from the Campbell-Shiller log-linear approximation used before solving the model in many asset pricing papers. Second, as argued in Pohl, Schmedders, and Wilms (2018) and Lorenz, Schmedders, and Schumacher (2020), a necessary condition for the log-linearization technique generating a nontrivial numerical error is that factors that impact the price-dividend ratio are extremely persistent. This is not the case in our calibration. As we have verified, P_t is actually indistinguishably different from exponential affine, and numerical errors in VIX calculations (due to log-linearization of P_t) across various states never exceed 1%.

so, we can write VIX-squared as a function affine in σ_t^2 and λ_t,

$$VIX^2(\sigma_t^2, \lambda_t) = a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t, \tag{35}$$

where $a_{1/12}$, $c_{1/12}$, and $d_{1/12}$ are three positive constants that in equilibrium depend on investors' preference parameters. For example, all of them are increasing in risk aversion γ because VIX, as risk-neutral variance, implicitly incorporates market investors' attitudes toward risks. The more market investors are risk-averse, the higher the VIX index. In addition, $a_{1/12}$, $c_{1/12}$, and $d_{1/12}$ also depend on endowment dynamics parameters. For example, the more persistent σ_t^2 is, the greater its impact on SPX volatility and thus VIX, that is, the higher $c_{1/12}$ is. However, we emphasize that the dividend-specific risk parameter σ_D only impacts the constant component of VIX, $a_{1/12}$. Since the risk is orthogonal to consumption risks and not priced, it affects VIX in a fashion independent of investors' attitudes towards risks.[12] It follows that in equilibrium, the VIX index has a square-root affine structure:

$$VIX(\sigma_t^2, \lambda_t) = \sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}. \tag{36}$$

Several observations are worth noting. First, in the reduced-form literature, VIX typically takes an affine or exponential affine structure (Mencía and Sentana (2013), Park (2016)) that delivers convenience for VIX option pricing. In our model, however, VIX has a square-root affine structure, which we handle with a novel generalized Fourier transform in order to price VIX options. Importantly, as we will explain, the square-root structure is essential for replicating the concave VIX-option-implied volatility curves seen in the data. Second, as VIX loads positively on state variables σ_t^2 and λ_t, both of which command a negative market price of risk, so does VIX. This implies that, in principle, an asset with positive (negative) VIX exposure should earn a negative (positive) premium with no ambiguity, as in the data. Examples include VIX futures and call options (put options), as we verify quantitatively below.

F. VIX Options

Our key focus is a (European) VIX call option that renders its holder the right, but not the obligation, to obtain the difference between the VIX index at an expiration date $t + \tau$ and a prespecified strike K.[13] No-arbitrage implies

[12] Section III of the Internet Appendix shows that the differential equations pinning down $c_{1/12}$ and $d_{1/12}$ depend on almost all model parameters except σ_D. If they depended on σ_D, then the impact of σ_D on $c_{1/12}$ and $d_{1/12}$ would be γ-dependent.

[13] Note that the standard convergence of the futures price to the underlying price as the time to maturity approaches zero holds regardless of whether the underlying is tradable. As a related issue, just as in reality, the VIX futures-spot parity does not hold in our model. This is because the VIX index is not tradable. That VIX were tradable is equivalent to the existence of an investment technology allowing the agent to transfer $\sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}$ units of consumption goods

that the price of a VIX call option is given by

$$C^{VIX}(X_t, \tau, K) = E_t^Q\left[e^{-\int_t^{t+\tau} r_u du}\left(VIX_{t+\tau} - K\right)^+\right], \tag{37}$$

where it follows from (36) that

$$VIX_{t+\tau} = VIX(\sigma_{t+\tau}^2, \lambda_{t+\tau}) = \sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}}. \tag{38}$$

The challenge in computing the expectation in (37) is to properly transform the nonstandard option payoff function $(\sqrt{x} - K)^+$, to which end we apply a novel generalized Fourier transform analysis. Section III of the Internet Appendix shows that by doing so, we can finally write the VIX call price as

$$C^{VIX}(X_t, \tau, K) = \frac{1}{4\sqrt{\pi}} \int_{iz_i - \infty}^{iz_i + \infty} e^{-iza_{1/12}} \varrho_X^Q\left(-iz(0, c_{1/12}, d_{1/12})', X_t, \tau\right) \times \frac{Ercf\left(K\sqrt{-iz}\right)}{(-iz)^{\frac{3}{2}}} dz, \tag{39}$$

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 (27), and $Ercf(\cdot)$ is the complex-valued complementary error function with an expression given in the Internet Appendix.

A similar generalized Fourier transform analysis on the put's payoff function $(K - \sqrt{x})^+$ gives the VIX put price as

$$P^{VIX}(X_t, \tau, K) = -\frac{1}{4\sqrt{\pi}} \int_{iz_i - \infty}^{iz_i + \infty} e^{-iza_{1/12}} \varrho_X^Q\left(-iz(0, c_{1/12}, d_{1/12})', X_t, \tau\right) \times \frac{1 - Ercf(K\sqrt{-iz})}{(-iz)^{3/2}} dz, \tag{40}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) < 0$.[14] As the payoff structure we are

at t into $\sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}}$ units of consumption goods at $t + \tau$ (for any τ). However, any such intertemporal consumption transfer is ruled out in the model. A τ-maturity VIX future at t is in essence a random consumption strip that pays off $\sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}}$ units of consumption goods at $t + \tau$. The futures price is the equilibrium (time $t + \tau$) price of such a strip. Relatedly, we only consider a VIX futures option and back out its implied volatility using the Black (1976) formula. On the other hand, we back out implied volatility for SPX options using the Black and Scholes (1973) formula, since SPX (the dividend claim) is a tradable asset.

[14] In Section IV, we use both the Riemann rule and the quadrature rule to approximate the integrals. The two methods generate the same result. We also compare the price obtained via integral with that via risk-neutral Monte Carlo simulation. We find that the difference is negligible as long as the VIX option is not "too far OTM," which applies to all of our reported results.

looking at is not common, we prove the existence of relevant Fourier transforms in the Internet Appendix. An important contribution of our paper to the option pricing literature is thus to fully characterize the working of a generalized Fourier transform argument to price a European call and put option with a square-root affine underlying payoff structure, which no previous papers have done to our best knowledge.

IV. Quantitative Analysis

In the following, we calibrate parameters for our model with the target toward replicating salient features of consumption, dividends, equity, VIX, and VIX derivatives data.

A. Calibration

Table IV displays our choices of model parameters. To facilitate comparison with recent 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 1% per annum and an expected consumption growth μ equal to 3% per annum together help give rise to a low average real yield on the one-year Treasury Bill of 0.18%, roughly consistent with that documented in Bansal and Yaron (2004), 0.86%. We set μ relatively high to counter the negative effect of a relatively large risk aversion or a relatively high mean volatility on risk-free rate, since at least one of the latter is needed to produce premia on VIX derivatives as large as those seen in the data. But this results in an excessively high dividend growth through the stock market leverage parameter ϕ. To counter this side effect, we set the adjustment term $\mu_D = -2\%$. These parameters lead to a dividend growth of 5.84% per annum, close to the data.

We set the value of θ^V, the average annualized consumption growth variance without jumps, to 0.0004, which corresponds to a volatility of 2% per annum, consistent with that used in Wachter (2013). A reasonable range of values for the U.S. consumption growth volatility that most previous research agrees on is 1% to 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, 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, the value of which constitutes a source of debate, is set to one for tractability. A number of studies conclude that a reasonable value for this parameter should be somehow close to one (e.g., Vissing-Jørgensen (2002), Hansen, Heaton, and Li (2008), Wachter (2013), Thimme (2017)).

The parameter 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 using equity market data,

Table IV
Parameters for the VIX Model

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

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

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

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

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

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

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

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

Parameters values are interpreted in annual terms.

Rate of time preference β	0.01
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
Adjustment in dividend growth drift μ_D	−0.02
Idiosyncratic risk in dividend growth σ_D	0.1

such as Eraker, Johannes, and Polson (2003), suggest that jumps in equity market return volatility occur 1.5 times per year on average. Starting from the mean level of volatility, an average-sized jump in volatility increases volatility from 15% to 24%. 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, an average-sized jump in volatility increases steady-state VIX from 20.9 to 32.6, which is consistent with the average jump size in VIX, 11.4, computed from historical VIX data from CBOE over

the period 1990 to 2020.[15] The unconditionally average consumption growth volatility is equal to the square root of $\bar{\sigma}_t^2 = \theta^V + \frac{\mu_\xi \theta^\lambda}{\kappa^V}$. With θ^V, μ_ξ, and θ^λ fixed, we then set κ^V at 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 consumption volatility of 3.08%, which is slightly higher than 2.93% documented in Bansal and Yaron (2004) and higher than 1.34% in Wachter (2013).

To calibrate the other model parameters, notably risk-aversion (γ), the diffusion parameter for volatility (σ_V), the mean reversion of the jump intensity (κ^λ), the diffusion parameter for 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 best match the following six data moments: mean VIX (19.3), standard deviation of VIX (7.4), average holding-period return on one-month ATM VIX call option (−48%), average one-month ATM VIX option Black-76-implied volatility (0.69), monthly autocorrelation of one-month ATM VIX option Black-76 implied volatility (0.53), and 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 II and V. We match a majority of these moments well.

We calibrate risk aversion at 14, slightly higher than that in Bansal and Yaron (2004) (10), Bollerslev, Tauchen, and Zhou (2009) (10), and Drechsler and Yaron (2011) (9.5), and higher than that in Wachter (2013) (3) and 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 low consumption growth volatility (low-risk prices), while maintaining a reasonable stock market leverage ϕ.

We calibrate σ_V 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. The parameter 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, becomes shorter-lasting and impacts equilibrium VIX option price less (note that σ_t^2 is also a volatility-of-volatility factor that impacts VIX option price). Finally, we set $\sigma_\lambda = 2.6$ and $\sigma_D = 0.1$ in order to match the mean and standard deviation of VIX and VIX option premia and implied volatility. Note that VIX derivative (futures and option) premia are generally decreasing with

[15] The number 11.4 is obtained as follows: we take monthly data of the VIX index from CBOE, identify all months during which VIX rises, and then take the mean of the largest 15. Given the period 1990 to 2020, the number 15 is consistent with our earlier calibration that jumps are on average once every other year.

Table V
Simulation: Selected Model Moments

The table reports 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 one-month 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
$E[\Delta c]$	2.96	1.80	BY2004
$\sigma(\Delta c)$	3.08	2.93	BY2004
$AC_1(\Delta c)$	0.27	0.49	BY2004
$E[\Delta d]$	5.84	4.61	CRSP
$\sigma(\Delta d)$	11.56	11.49	BY2004
$AC_1(\Delta d)$	0.24	0.21	BY2004
$corr(\Delta c, \Delta d)$	0.69	0.59	DY2011
$E[\exp(pd)]$	21.98	26.56	BY2004
$\sigma(pd)$	9.14	29.00	BY2004
$AC_1(pd)$	0.04	0.81	BY2004
$E[r_t^e - r_t^f]$	8.81	8.33	Ken French
$\sigma(r_t^e)$	17.71	18.31	CRSP
$E[r_t^f]$	0.18	0.86	BY2004
$\sigma(r_t^f)$	2.86	0.97	BY2004
$E[VIX_t]$	19.41	19.28	CBOE
$\sigma(VIX_t)$	7.51	7.42	CBOE
$AC_1(VIX_t)$	0.80	0.84	CBOE
$E[imp\_vol_t]$	71.84	68.80	CBOE
$\sigma(imp\_vol_t)$	12.64	14.30	CBOE
$AC_1(imp\_vol_t)$	0.49	0.53	CBOE
$corr(VIX_t, imp\_vol_t)$	0.32	0.48	CBOE

σ_D, as the dividend idiosyncratic risk contained in σ_D is not priced in equilibrium and thus contributes only to the constant component of the VIX index, thereby decreasing the exposure of VIX (derivative) returns to σ_t^2 and λ_t.

B. Simulation Results

B.1. General Moments

Table V displays 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 simulated data to compute model moments primarily reported on a monthly or annual basis. As seen in the table, we match a majority of the key moments that we are interested in. In particular, we match average consumption growth

volatility fairly well: 3.08% in the model versus 2.93% in the data. Notably, we match the correlation between consumption and dividend growths closely: 0.69 in the model versus 0.59 in the data. This outperforms leading asset pricing models, as, for example, the correlation is 0.31 in Bansal and Yaron (2004), 0.32 in Drechsler and Yaron (2011), and 1 in Wachter (2013), implying that the balance between systematic and idiosyncratic risks in dividend growth is more reasonable in our model. In terms of the equity premium, we overshoot slightly, as our model produces 8.81% 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 17.71%, which compares to 18.31% in post-1990 S&P 500 data.[16] Our model generates an average (one-year) risk-free rate on par with what we see in the data, though the model-implied volatility of the risk-free rate is a bit higher.

Our model does not match the observed (log) price-dividend ratio well. Empirically observed price-dividend 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 far exceeds that seen in high-frequency derivatives-based variables such as VIX and VVIX. Adding additional state variables, such as those in models by Campbell and Cochrane (1999) (habits), Bansal and Yaron (2004) (long-run persistent consumption growth), and Wachter (2013) (persistent disaster risk factor) helps the model fit the price-dividend data better. In Section IV of the Internet Appendix, we solve and calibrate an extended version of our baseline model in which the introduction of persistent long-run growth risks brings the volatility and persistence of the price-dividend ratio substantially closer to the data while leaving all the other moments largely unaffected. In the Internet Appendix, we also report additional model moments as well as return predictability, and discuss how the shortfall in long-term return predictability can be addressed in an extended three-factor model.

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 (19.41) is higher than equity return objective volatility (17.71) illustrates the model's ability to generate a large unconditional VRP close to that in the data. The monthly autocorrelation of the simulated VIX index is 0.8, close 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 as $E(imp\_vol_t)$, is estimated at 71.84 in the model simulations, which compares to 68.8 in the data. The model produces a volatility of the simulated VIX-implied volatility, denoted as $\sigma(imp\_vol_t)$, of 12.64 versus 14.3 in the data—a slight miss on the low side. The model also produces a

[16] We compare our estimate to SPX volatility using data collected after 1990 to make the estimate comparable to the 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 to 2020.

monthly autocorrelation of one-month ATM VIX-implied volatility, denoted as $AC_1(imp\_vol_t)$, of 0.49 versus 0.53 in the data.

Our model matches the observed positive but imperfect correlation between the implied VIX volatility and VIX at 0.32 versus 0.48 in the data. It is useful to consider this in relation to a model in which the arrival intensity of volatility jumps is constant $\lambda_t = \lambda$ (Eraker and Wu (2017)) or volatility-driven $\lambda_t = \lambda_0 + \lambda_1\sigma_t^2$ (Drechsler and Yaron (2011)). In these cases, VIX^2 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 is a linear function of σ_t^2, or equivalently, VIX_t^2. This shows that VIX-option-implied volatility (or simply vol-of-vol) should be (either positively or negatively) perfectly correlated with VIX itself.

Our two-factor model breaks up the otherwise rigid correlation between VIX and vol-of-vol by having an additional self-exciting λ_t factor. The latter typically drives VIX and vol-of-vol in the same direction as follows. When λ_t increases, it first drives VIX up as VIX loads positively on it. It second drives up the prices of VIX call options and thus implied vol-of-vol. Huang et al. (2019) 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 also loads positively on σ_t^2 and λ_t, it too carries a negative risk price. However, Huang et al. (2019) propose a model in which stock market spot variance follows a mean-reverting process in which volatility is driven by an independent diffusion process. The independence assumption counterfactually implies that the correlation between VIX^2 and volatility-of-volatility (or VVIX) is zero.[17]

B.2. VIX Futures Returns

Table VI compares average returns and return standard deviations for VIX futures prices computed from the data (see Eraker and Wu (2017)) and our model. We report average daily arithmetic and logarithmic returns as well as return standard deviations by resimulating our model at a daily frequency to facilitate comparison with the results in Eraker and Wu (2017). For one-month contracts, both log and arithmetic returns are roughly the same between the model and 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 et al. (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. The VRP is always positive in our model, which follows from the fact that VRP is a positive linear function of two positive state variables.

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

Table VI
Simulation: VIX Futures Returns

This table reports descriptive statistics for model simulated VIX futures returns. R^1 is the daily average arithmetic constant-maturity return, R^2 is the daily average logarithmic constant-maturity return, and 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			
One month	−0.10	−0.18	3.75
Two months	−0.09	−0.14	3.12
Three months	−0.08	−0.12	2.71
Four months	−0.07	−0.10	2.40
Five months	−0.06	−0.08	2.14
Data			
One month	−0.12	−0.20	3.98
Two months	−0.07	−0.11	3.00
Three months	−0.01	−0.04	2.47
Four months	−0.03	−0.05	2.21
Five months	−0.01	−0.03	2.01

To better understand how negative average VIX futures returns are generated in the model, Figure 7 shows the expected returns under different market conditions (low versus high VIX). As in Eraker and Wu (2017), figure 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.

B.3. SPX-Option-Implied Volatilities

Before proceeding to our key focus, VIX options, we discuss our model's ability to accurately capture SPX-option-implied volatilities. Figure 8 illustrates the Black and Scholes (1973) implied volatilities for SPX put options in our model's steady states. The implied volatilities resemble those we see in the data well. First, the levels of ATM- and OTM-implied volatilities for various maturities are on par with those in the data. Second, fixing moneyness, the implied volatility term structure is upward sloping for ATM options and gradually transitions to downward sloping for (far) OTM options. Third, the implied volatility curve is decreasing with the strike for most strike ranges, consistent with a highly left-skewed risk-neutral distribution of SPX returns.[18]

[18] See Bates (2000), Broadie, Chernov, and Johannes (2007), Eraker (2004), Pan (2002), Santa-Clara and Yan (2010), Backus, Chernov, and Martin (2011), and Seo and Wachter (2019), among others, for models that generate left skewness. In our model, left skewness is endogenously achieved through the volatility feedback mechanism.

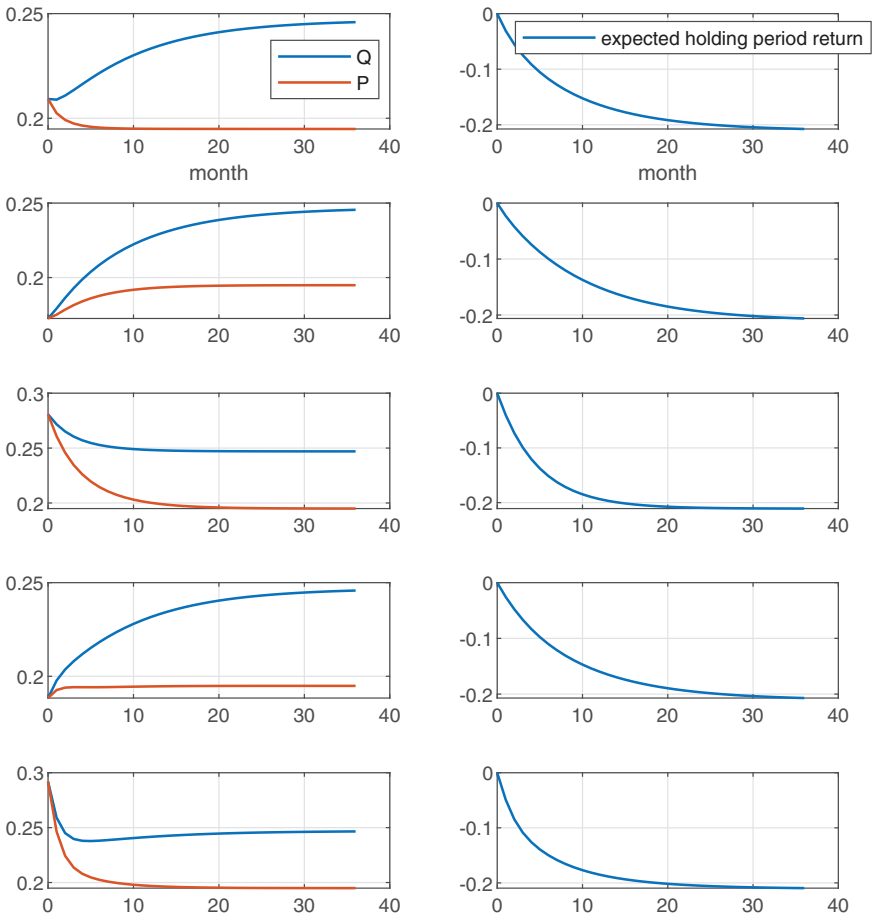

Figure 7. VIX futures curves and holding period returns. The figure illustrates conditional VIX futures term structures and conditional expected holding period returns on VIX futures in the model. Left: VIX futures curves (Q) and the objective-measure expected payoffs (P). Right: expected holding period 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 as follows. 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; fourth row: steady-state σ_t^2 and low λ_t; last row: steady-state σ_t^2 and high λ_t. (Color figure can be viewed at wileyonlinelibrary.com)

B.4. VIX-Option-Implied Volatilities

Table I, right panel, reports VIX-implied volatilities from our model and is thus comparable to Table I, left panel, which uses real data. At short maturities and low strikes, our model mildly undershoots implied volatility, as seen for the 20 strike, which averages 104% implied volatility in the data versus 88% in our model. This is not a significant deviation when considering that the size of the bid-ask spread often exceeds 20 implied volatility points—see Figure 2. At higher strikes, our model generates data that are close to what is reported in

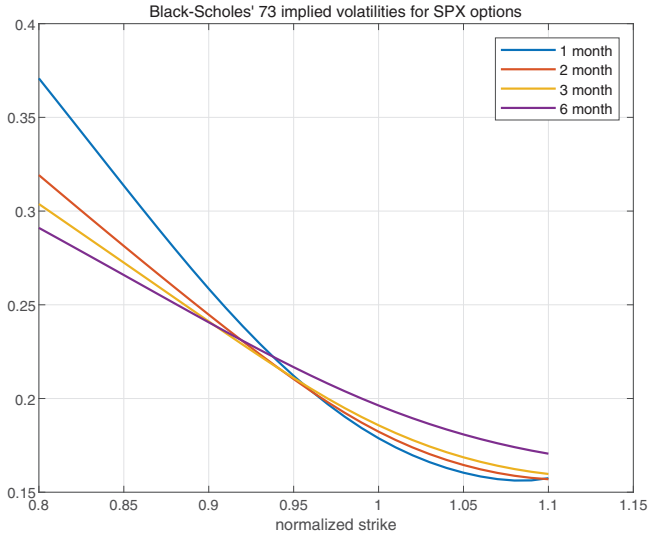

Figure 8. Black-Scholes-implied volatility curves for SPX options. This figure plots (annualized) implied volatility curves computed from equating the Black-Scholes (1973) option pricing formula with the SPX option price in the model at steady state. The horizontal axis denotes strike normalized by SPX. Implied volatilities are computed for SPX options with four maturities: 1, 2, 3, and 6 months. (Color figure can be viewed at wileyonlinelibrary.com)

Table I, left panel. For example, at a 30 strike, the data average 126% implied vol, which compares to 121% in the model. At a strike of 40, the numbers are 135% and 134%, respectively. At the six-month horizon, our model overshoots implied volatilities by roughly 5 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, in contrast, 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.

Figures 9 and 10 further illustrate the implied volatility patterns generated by our model. While Figure 9 shows the steady-state implied volatilities and pretty much illustrates the patterns in Table I, right panel, Figure 10 shows what happens when we condition on a high and low initial VIX. In the low-VIX case, where we set the initial state variables so low as to generate a VIX of 12.6, we see that the implied volatility curves are almost everywhere increasing and concave. This closely resembles the patterns we see in the data on April 26, 2017 (Figure 2 bottom). The top panel shows that under a high initial VIX, 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 2 top).

To understand this contrast, note that two forces shape VIX-option-implied volatility curves. First, equilibrium VIX-squared is linear in σ_t^2 and λ_t, but

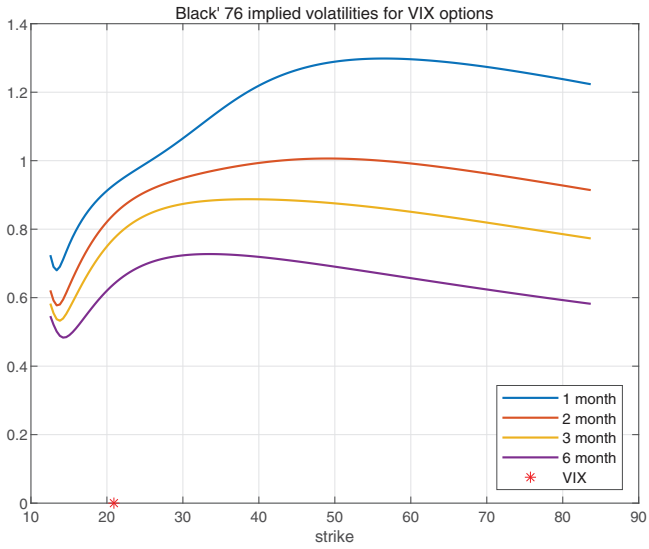

Figure 9. Black-76 implied volatility curves for VIX options. This figure plots (annualized) implied volatility curves computed from equating the Black (1976) futures option pricing formula with the VIX option price in the model at steady state. The horizontal axis denotes the absolute value of the strike. Implied volatilities are computed for VIX options with four maturities: 1, 2, 3, and 6 months. (Color figure can be viewed at wileyonlinelibrary.com)

driven mainly by σ_t^2. Therefore, the implied volatility is shaped by the risk-neutral conditional distribution of σ_t^2. Second, the option is written on the VIX, the square root of VIX-squared. The square-root payoff structure carries a moderate effect on distribution shape, that is, a square-root transform reduces (increases) a random variable's right (left) skewness. Now consider the shapes of implied volatility curves across the two different market conditions in order.

The concavity in implied volatility seen in the low-VIX state is related to the fact that, in order to generate a low VIX, both state variables are set low. When λ_t is low it mean-reverts fast so that during the option's lifetime, it likely experiences a considerable increase. By contrast, σ_t^2 mean-reverts slowly and remains persistently low unless there is a jump. The σ_t^2 dynamics described by equation (21) show that the effect of jumps dominates the distribution of σ_t^2 and thus VIX-squared, inducing a fat right tail in the distribution of squared-VIX. This again leads to an increase in implied volatility across strikes. The moderate effect of the square root transformation from VIX-squared to VIX, however, works in the opposite direction by dampening the right-skewness of VIX's distribution, which is why the implied volatility curve is concave. In other words, without the square-root payoff structure, implied volatility is convexly increasing in the strike, whereas without the possibility of jumps, implied volatility would decrease sharply beyond a certain threshold strike and also would undershoot its data counterpart. Only combining a possibility of jumps and a square-root payoff structure can deliver a concavely increasing

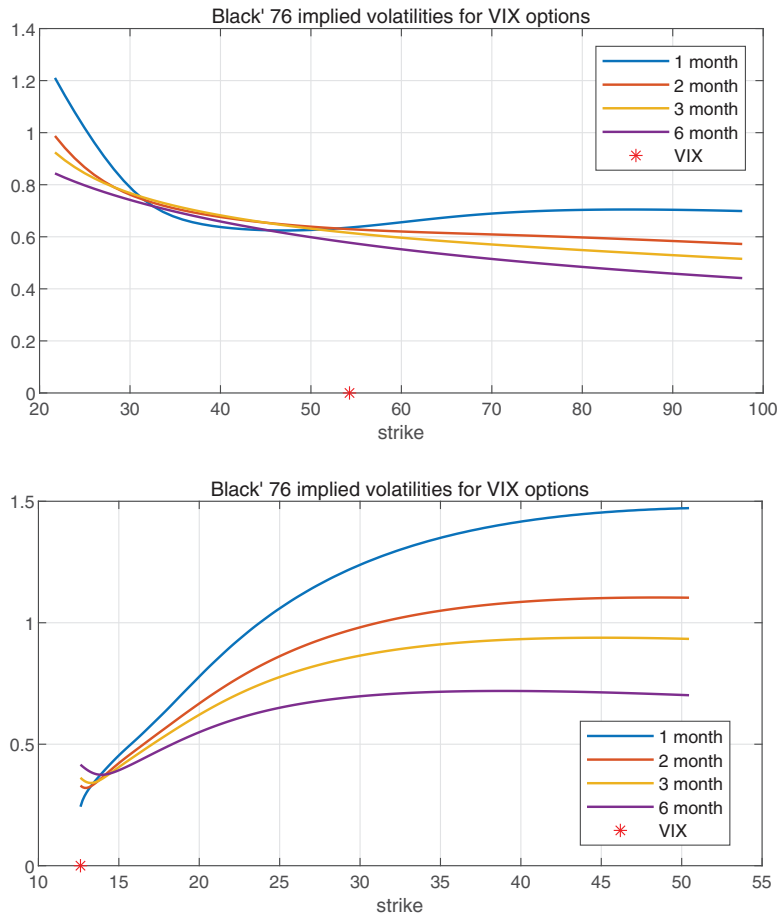

Figure 10. Black-76 implied volatility curves for VIX options: conditional analysis. The figure plots (annualized) implied volatility curves for VIX options in the model conditional on high and low initial VIX. In the top panel, we set both state variables very high: $\sigma_t^2 = 10\sigma_{ss}^2$ and $\lambda_t = 10\lambda_{ss}$, implying a very high VIX, 54.3. In the bottom panel, we set both state variables at minimum values: $\sigma_t^2 = \lambda_t = 0$, implying a small VIX, 12.6. (Color figure can be viewed at wileyonlinelibrary.com)

implied volatility curve, not only in low-volatility times (Figure 10) but also in average times (Figure 9).

In contrast, in order to generate a high VIX, simultaneously we need both high spot volatility σ_t and high jump arrival intensity λ_t. The high probability of a jump arrival fattens the right tail of the conditional distribution for VIX in high-VIX regimes, generating high implied volatilities for VIX options with high strikes. This counters the moderate effect of the option's square-root payoff structure and generates a relatively heavy right tail of VIX's distribution, preventing the implied volatility curve from sharply sloping downward to the

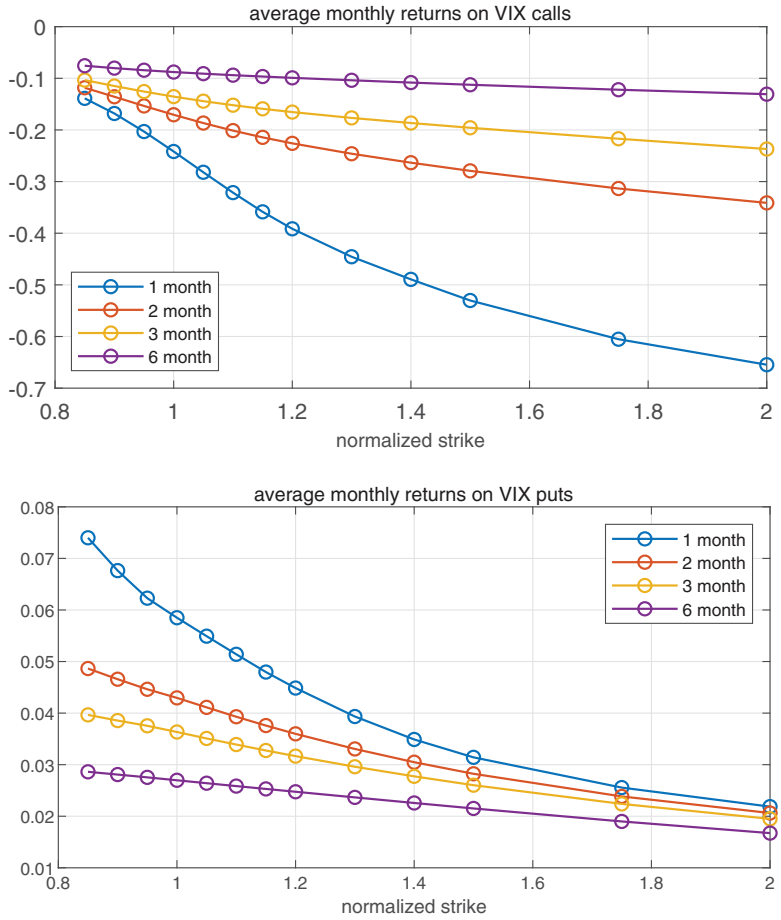

Figure 11. Average returns on VIX options. This top (bottom) panel plots average (monthly) returns on VIX calls (puts) in the model. In each case, we consider options with four maturities: 1, 2, 3, and 6 months, and the horizontal axis denotes the strike of the relevant option normalized by its underlying asset price. (Color figure can be viewed at wileyonlinelibrary.com)

right. The fact that σ_t is high also increases the volatility of σ_t^2 itself through the square-root diffusion term, assigning fat tails to both sides. Reinforced by the moderate effect of the square-root payoff structure that contributes to left skewness, VIX now has a particularly fat tail at the left end, making the implied volatility curve convex there.

B.5. VIX Option Returns

Figure 11 reports average returns on holding VIX call and put options to maturity. All returns are normalized to a monthly frequency. Some striking patterns can be seen. First, the model generates a negative (positive) premium

Table VII
Simulation: VIX Option Returns

The table reports model moments of returns on holding VIX options for maturities of one and six months. ITM and OTM are defined as 15% in-the-money and 15% out-of-the-money. Sharpe ratios are annualized. All other numbers are based on buy-and-hold returns.

	CALLS			PUTS		
	ITM	ATM	OTM	ITM	ATM	OTM
			One-Month Maturity			
Mean	−0.14	−0.24	−0.36	0.05	0.06	0.07
Std	1.23	2.12	3.18	0.68	1.09	2.47
Sharpe	−0.40	−0.40	−0.39	0.24	0.18	0.10
Skew	4.46	7.15	10.42	0.05	0.76	3.10
Kurt	39.49	83.93	162.06	2.39	2.89	15.80
			Six-Month Maturity			
Mean	−0.45	−0.53	−0.58	0.15	0.16	0.17
std	1.25	1.41	1.52	0.65	0.83	1.22
Sharpe	−0.52	−0.53	−0.54	0.32	0.27	0.19
Skew	3.53	4.27	5.06	−0.50	−0.08	0.56
Kurt	18.61	25.65	34.85	2.32	1.88	1.98

for VIX call (put) options, intuitively because the payoff of a 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 represent insurance 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 their longer-maturity counterparts, which implies that the shorter the maturity is, the more excessively expensive (cheap) the call (put) is. This result is, in principle, consistent with the downward-sloping term structure of VIX-option-implied volatility shown in Table I, right panel. Below we show that this is true in the VIX-implied volatility data. Third, the premia for both call and put options are decreasing with moneyness, showing that the more OTM the VIX call (put) is, the more pronounced its role as a bet on (against) volatility and volatility-of-volatility. This result is, in principle, consistent with the upward-sloping VIX-option-implied volatility curve across strike shown in Table I, right panel.

Table VII further reports returns to VIX options in our model in greater detail. Here, ITM (OTM) indicates 15% in (out) -of-the-money. Given the VIX futures price is most of the time close to 20, 15% corresponds to three points, so this table is directly comparable with Table II. 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, consistent with the data. Second, the model generates negative short maturity (one-month) call returns ranging from −14% (ITM) to −36% (OTM), which compare to -33% (ITM) and -60% (OTM) in the data. Keep in

mind, however, that over the course of our sample, VIX spiked dramatically on several occasions, including the financial crisis of 2008, the events of February 5, 2018 discussed in Section II, and the Covid-19 crisis, leading to large positive returns for OTM VIX calls. Turning to the longer-maturity (six-month) call option average returns, our model matches the data well: -45% (ITM) to -58% (OTM) in the model versus -53% (ITM) and -61% (OTM) in the data. Another commonality between the model and the data is that the term structure of (absolute) returns on VIX call is downward sloping. For example, for ATM calls, the monthly return is -0.24 (-0.48) for the one-month maturity versus $-0.53/6 \approx -0.09$ ($-0.59/6 \approx -0.1$) for the six-month maturity in the model (data).

The model generates positive average returns on ITM, ATM, and OTM short- and long-maturity put options, consistent with the data in Table II. The quantities are somehow different, with the average return heavily dependent on moneyness in the data, but not in the model. In particular, the data show that short-maturity OTM puts have large average returns due to the 2008 financial crisis and Covid-19 crisis periods.

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 157% in the data versus 123% in the model for ITM, 305% versus 212% for ATM, and 498% versus 318% for OTM. The return standard deviations are matched even better for short-maturity puts, as one-month-maturity put returns have an estimated standard deviation of 66% in the data versus 68% in the model for ITM, 101% versus 109% for ATM, and 221% versus 247% for OTM. At longer maturities, our model slightly undershoots return standard deviations for both calls and puts. Notably, we also match the estimated skewness and kurtosis coefficients fairly closely, especially for the puts. One exception here is the large model-implied kurtosis for OTM one-month-maturity calls: this coefficient is 162.06 in the model versus 68.29 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 a relatively short sample of option returns.

C. VIX Options as Hedges For SPX Options

To get a clearer insight into the workings of our model economy, we present results of a variance decomposition of data simulated from the model. We condition on the initial values of the state variables σ_t^2 and λ_t so as to generate high and low volatility states or VIX states. In each case, we simulate a large number ($N = 50,000$) of realizations of state variables one day ahead and compute option prices and VIX futures prices. We then ask which of the state variables are important to various states of the world by running regressions of the price changes on changes in state variables as well as squared and cubed state variables. The latter allows us to approximately pinpoint the importance of convexity in option prices.

Table VIII
Variance Decomposition

This table reports decompositions of variance in different assets into fractions attributable to cash flow risk (D_t), volatility risk (σ_t^2), second-order volatility risk ($(\sigma_t^2)^2$), third-order volatility risk ($(\sigma_t^2)^3$), jump intensity risk (λ_t), and second-order jump intensity risk ($(\lambda_t)^2$), conditional on a high-VIX, medium-VIX, and low-VIX regime, respectively. Definitions of OTM, ATM, and VIX futures characteristics are as in Figure 12. All numbers are in percentages. We simulate the model (starting from four different initial states) for two periods (days) to obtain the changes in all relevant variables. We repeat the simulation 50,000 times to obtain 50,000 observations and then obtain variance decomposition from linear regressions.

Asset	D_t	σ_t^2	$(\sigma_t^2)^2$	$(\sigma_t^2)^3$	λ_t	$(\lambda_t)^2$
High-VIX regime I: $\sigma_t^2 = 10\sigma_{ss}^2, \lambda_t = 10\lambda_{ss}, VIX_t = 54.3$						
ATM SPX Put	7.7	3.5	77.6	10.9	0.3	0.0
OTM SPX Put	0.0	10.0	55.3	34.6	0.0	0.0
ATM VIX Call	0.0	11.7	77.5	10.6	0.0	0.3
OTM VIX Call	0.0	14.2	60.9	24.9	0.0	0.0
VIX Futures	0.0	81.5	11.7	0.6	6.1	0.1
High-VIX regime II: $\sigma_t^2 = 5\sigma_{ss}^2, \lambda_t = 5\lambda_{ss}, VIX_t = 39.4$						
ATM SPX Put	18.5	0.2	69.3	11.3	0.7	0.0
OTM SPX Put	0.1	9.3	48.4	42.3	0.0	0.0
ATM VIX Call	0.0	2.2	82.7	14.2	0.4	0.5
OTM VIX Call	0.0	5.6	38.1	56.2	0.0	0.1
VIX Futures	0.0	81.7	10.5	0.7	7.0	0.1
Medium-VIX regime: $\sigma_t^2 = \sigma_{ss}^2, \lambda_t = \lambda_{ss}, VIX_t = 20.9$						
ATM SPX Put	41.4	33.0	19.2	5.3	1.1	0.0
OTM SPX Put	0.3	4.5	37.9	56.0	1.3	0.0
ATM VIX Call	0.0	27.7	47.8	21.2	3.0	0.3
OTM VIX Call	0.0	5.9	89.1	4.5	0.4	0.0
VIX Futures	0.0	79.0	11.6	1.7	7.7	0.0
Low-VIX regime: $\sigma_t^2 = 0.01\sigma_{ss}^2, \lambda_t = 0.01\lambda_{ss}, VIX_t = 12.7$						
ATM SPX Put	98.0	1.9	0.0	0.0	0.1	0.0
OTM SPX Put	35.1	0.4	0.0	0.0	64.5	0.0
ATM VIX Call	0.0	89.5	0.0	0.0	10.5	0.0
OTM VIX Call	0.0	23.5	0.6	0.0	75.9	0.0
VIX Futures	0.0	85.5	0.0	0.0	14.5	0.0

The results are presented in Table VIII. Starting from the bottom, we see that low values of the state variables, consistent with a VIX of 12.7, imply that ATM SPX options returns are driven almost entirely by cash flow risk (D_t). OTM SPX options depend less on cash flow risk (35.1%) but heavily on variation in variance jump risk (λ_t). Intuitively, OTM SPX options depend primarily on the possibility of a crash occurring, the probability of which is λ_t. This is consistent with Bollerslev and Todorov (2011) and Bollerslev, Todorov, and Xu (2015), among others, who derive option-based tail-risk measures from far OTM options and argue that these are priced state variables. The convexity terms, $(\sigma_t^2)^2$, $(\sigma_t^2)^3$, and $(\lambda_t)^2$, do not matter in the low-volatility regime.

VIX options and futures never depend on cash flow risk D_t. In the low-VIX regime, they depend linearly on σ_t^2 and λ_t. In the steady state and higher-VIX

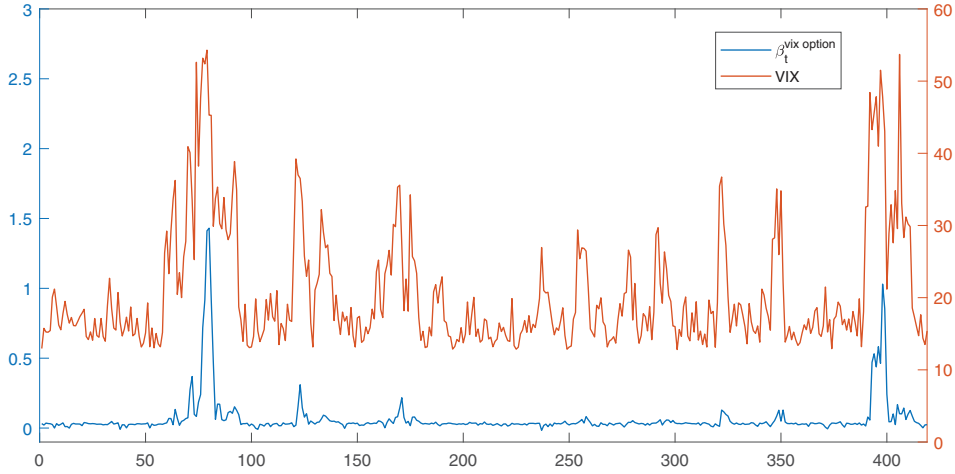

Figure 12. Hedge regressions. The figure shows $\beta_t^{\text{vix option}}$ computed by the following regression using model-simulated data: $\Delta P_t = \alpha_t + \beta_t^{\text{SPX}} \Delta \ln SPX_t + \beta_t^{\text{vix futures}} \Delta F_t + \beta_t^{\text{vix option}} \Delta C_t + error_t$, where P_t is a half-month 30% OTM SPX put normalized by the SPX index (for stationarity in long-sample simulation), $\ln SPX_t$ is the log market index, F_t is the half-month VIX futures, and C_t is a half-month 50% OTM VIX call. Consistent with our data regressions in the empirical section, we select relatively far-OTM VIX calls and SPX puts. The regression is run each day using daily price changes with a rolling window of one month. We then average daily coefficients within each month and plot $\beta_t^{\text{vix option}}$ and VIX_t as a monthly time series. (Color figure can be viewed at wileyonlinelibrary.com)

regimes, we see that both VIX and SPX options depend increasingly on the convexity terms. This is, in part because, as we increase the jump frequency λ_t, jump realizations become more important, leading to the squared and cubed σ_t^2 terms instrumenting for the convexity in the options prices for both VIX and SPX.

Our variance decompositions are based on state variables that are not observable (to econometricians): the state variables cannot be traded, and there are no instruments that load on the state variables only, rendering the market incomplete. Options traders can hedge SPX or VIX options using a standard delta hedge through index futures or ETFs, and they can hedge volatility exposure with variance claims such as VIX futures. Recall that Figure 6 depicts the performance of hedging SPX options using stock prices (SPY), VIX futures, and VIX call options, and shows that during crisis periods, VIX calls significantly improve hedging performance. Figure 12 replicates this exercise using model-simulated data. As seen, our model replicates the essential feature of the data: during periods of low volatility, the estimated factor loading on VIX call options fluctuates around zero, whereas during periods of high volatility, the estimated factor loading becomes positive. The pattern can be traced back to information contained in the variance decomposition (Table VIII). First, during normal times when VIX is low (λ_t and σ_t^2 low), variation in D_t dominates

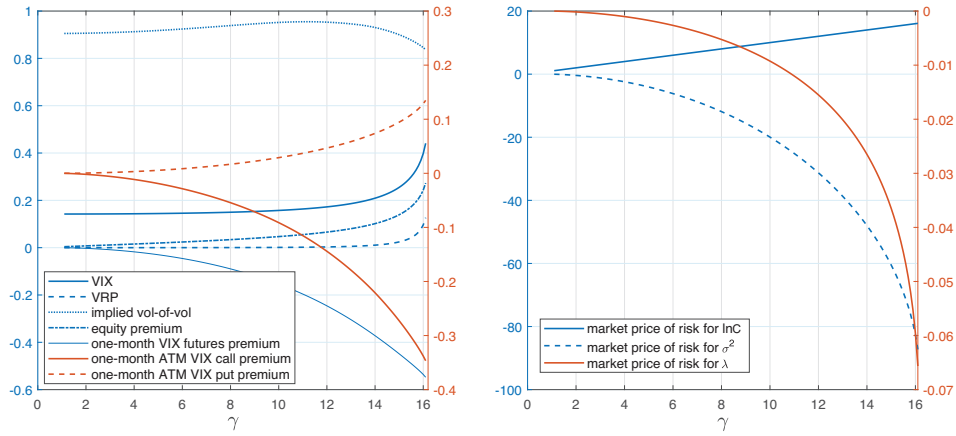

Figure 13. Comparative statics with respect to risk aversion. The figure plots steady-state conditional model moments (left) and market risk prices (right) as a function of risk aversion, γ. The left panel depicts VIX (VIX), variance risk premium (VRP), implied volatility of ATM VIX options (implied vol-of-vol), instantaneous equity premium (equity premium), one-month VIX futures premium, and one-month ATM VIX call and VIX put premiums. The right panel depicts the dependence of the market prices of risk 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 16.1, the upper limit for the existence of a model solution. (Color figure can be viewed at wileyonlinelibrary.com)

returns to SPX and SPX options. This leads to a low correlation between VIX calls and SPX puts. Second, during high-VIX periods when σ_t^2 are high, variation in σ_t^2 and its polynomial terms dominate variation in SPX options. Since λ_t and particularly σ_t^2 drive all of the variation in VIX options, the correlation between SPX and VIX options increases with these variables.

D. Comparative Static Analysis

To gain some additional insights into the workings of our model, we report results of some comparative statics. In doing so, we also emphasize the necessity of recursive preferences ($\gamma > 1/\psi$) for the model to generate nonzero VIX derivatives premia. The left panel of Figure 13 shows the comovements of seven important steady-state conditional model moments with risk aversion. As shown, the equity premium is sensitive to risk aversion at all levels of risk aversion: it increases with risk aversion almost linearly when the latter is relatively low, and increasingly fast when the latter becomes higher. Recall from equation (31) that the equity premium reflects compensation 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 risk for volatility and its jump risk increase slowly with γ at the beginning and increasingly faster afterward.

The right panel of Figure 13 shows the impact of risk aversion on the market prices of risk associated with the three state variables, $\lambda = (\gamma, -b_2, -b_3)$. Since the representative agent has recursive preferences (and prefers early resolution of uncertainties for the current parameter configuration), she is concerned about variation in her value function in the future, which is affected by the risk 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 jumps in σ_t^2 and has third- or even higher-order effects on consumption. Accordingly, market prices of risk associated with σ_t^2 and λ_t increase relatively slowly with risk aversion.

Turning back to the left panel of Figure 13, it remains to check how moments besides the equity premium vary with risk aversion. VIX increases 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, higher risk-neutral persistence, and higher mean σ_t^2 and λ_t, and a higher VIX. But this also implies that the average value of VIX increases relative to objective variance, or put differently, the VRP increases.

The left panel of Figure 13 also shows the steady-state risk premia on one-month ATM VIX put and call options. As the VIX call (put) option is a (negative) volatility claim, it earns a negative (positive) premium. Both premia, however, increase in absolute value with γ asymmetrically: premia in the puts increase more slowly than the calls. Thus, a larger risk aversion leads investors to be willing to pay a comparably higher premium for the crash insurance offered by VIX calls than the positive premium they demand for holding VIX puts.

Finally, Figure 13 speaks to the necessity of the recursive preference assumption in generating nonzero risk premia on VIX derivatives, since all premia are exactly zero when $\gamma = 1$, in which case the Duffie-Epstein recursive preferences collapse into CRRA preferences. With the latter, neither σ_t^2 nor λ_t would be priced in equilibrium. This would imply that a claim with mere exposure to σ_t^2 and λ_t would earn a zero premium.

V. Concluding Remarks

This paper studies the properties of VIX derivatives prices, including returns to buy-and-hold VIX options positions. We document negative return premia consistent with a negative price of volatility and volatility jump risk. Our paper follows the well-established literature on consumption-based asset pricing models in which persistent state dynamics generate risk premia that exceed those seen under time-additive preferences by separating risk-aversion from IES, 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 Shalias-

tovich (2008), but has the advantage that it does not require any linearization approximations in deriving the pricing kernel.

We use this modeling framework to specify a model that features time-varying consumption volatility and time-varying intensity of jumps in that volatility process. This is different from the consumption disaster literature, such as Barro (2006) or Wachter (2013), where disasters occur in consumption itself. Our model produces a smooth aggregate consumption consistent with what we see in U.S. data.

Our model replicates many of the observed characteristics of asset market data: it is within striking distance of the equity premium, unconditional stock market volatility, the VRP, the correlation between VIX and VVIX, and the weak persistence in 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, it replicates large negative average returns to VIX futures. Second, it replicates with an acceptable degree of accuracy the return premia seen in VIX options data. This includes the higher order moments. Third, it replicates 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 replicates this empirical phenomenon.

We show that variation in VIX options is not necessarily spanned by SPX options, as a PCA decomposition shows that VIX options returns contain variation not seen in SPX options. The model also replicates the time-varying nature of the hedging relationship between SPX options, the underlying SPX index, VIX futures, and VIX options. In regressing SPX put option changes on changes in these variables, we find that VIX options are nearly uncorrelated with SPX options in low-volatility periods, while the correlation spikes in high-volatility periods. Our model explains this through essentially time-varying factor loadings: when volatility is low, ATM SPX options depend primarily on cash flow news, while ATM VIX options depend on volatility and jump arrival intensity. In high-volatility periods, the correlations increase, and VIX call options can serve as important hedging instruments for SPX puts.

Initial submission: January 15, 2020; Accepted: August 16, 2021
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

REFERENCES

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

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters implied by equity index options, *Journal of finance* 66, 1969–2012.

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies* 16, 527–566.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bakshi, Gurdip, Delip Madan, and George Panayotov, 2015, Heterogeneity in beliefs and volatility tail behavior, *Journal of Financial and Quantitative Analysis* 50, 1389–1414.

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, *Quarterly Journal of Economics* 121, 823–866.

Bates, David S., 1996, Jump and stochastic volatility: Exchange rate processes implicit in deutsche mark options, *Review of Financial Studies* 9, 69–107.

Bates, David S., 2000, Post-'87 crash fears in S&P 500 futures options, *Journal of Econometrics* 94, 181–238.

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, Fischer, and Myron 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.

Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears, and risk premia, *Journal of Finance* 66, 2165–2211.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113–134.

Bondarenko, Oleg, 2003, Why are put options so expensive? Working paper, University of Illinois.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model specification and risk premia: Evidence from futures options, *Journal of Finance* 62, 1453–1490.

Brooks, Robert, Don M. Chance, and Mobina Shafaati, 2018, The cross-section of individual equity option returns, Working paper, University of Alabama.

Campbell, John Y., 2003, Consumption-based asset pricing, in G. Constantinides, M. Harris, and R. Stulz, eds.: Handbook of the Economics of Finance (Elsevier Science/North-Holland, Amsterdon, The Netherlands).

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.

Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs, 2017, The factor structure in equity options, *Review of Financial Studies* 31, 595–637.

Coval, Joshua D., and Tyler Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983–1010.

Cox, John C., Jonathan E. Ingersoll Jr, and Stephen 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, 225–250.

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

Duan, Jin-Chuan, and Jason Wei, 2009, Systematic risk and the price structure of individual equity options, *Review of Financial Studies* 22, 1981–2006.

Duffie, Darrell, and Larry Epstein, 1992, Stochastic differential utility, *Econometrica* 60, 353–394.

Duffie, Darrell, and Pierre-Louis Lions, 1992, PDE solutions of stochastic differential utility, *Journal of Mathematical Economics* 21, 577–606.

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, 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 59, 1367–1403.
Eraker, Bjørn, 2021, The volatility premium, *Quarterly Journal of Finance* 11, 1–35.
Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58, 1269–1300.
Eraker, Bjørn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 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* 54, 173–182.
Griffin, John, and Amin Shams, 2018, Manipulation in the VIX?, *Review of Financial Studies* 4, 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.
Horenstein, Alex, Aurelio Vasquez, and Xiao Xiao, 2019, Common factors in equity option returns, Working paper, University of Miami.
Huang, Darien, Christian Schlag, Ivan Shaliastovich, and Julian Thimme, 2019, Volatility-of-volatility risk, *Journal of Financial and Quantitative Analysis* 54, 2423–2452.
Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611–1631.
Johnson, Travis L., 2017, Risk premia and the VIX term structure, *Journal of Financial and Quantitative Analysis* 52, 2461–2490.
Lin, Yueh-Neng, and Chien-Hung Chang, 2009, VIX option pricing, *Journal of Futures Markets* 29, 523–543.
Lorenz, Friedrich, Karl Schmedders, and Malte Schumacher, 2020, Nonlinear dynamics in conditional volatility, Working paper, University of Muenster.
Martin, Ian, 2011, Simple variance swaps, Technical report, National Bureau of Economic Research.
Mencía, Javier, and Enrique Sentana, 2013, Valuation of VIX derivatives, *Journal of Financial Economics* 108, 367–391.
Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.
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.
Pohl, Walter, Karl Schmedders, and Ole Wilms, 2018, Higher-Order effects in asset pricing models with long-run risks, *Journal of Finance* 73, 1061–1111.
Roll, Richard, and Stephen A. Ross, 1982, An empirical investigation of the arbitrage pricing theory, *Journal of Finance* 2, 347–350.
Santa-Clara, Pedro, and Shu Yan, 2010, Crashes, volatility and the equity premium: Evidence from S&P 500 options, *Review of Economics and Statistics* 92, 435–451.
Seo, Sang Byung, and Jessica A. Wachter, 2019, Option prices in a model with stochastic disaster risk, *Management Science* 65, 3449–3469.
Serban, Mihaela, John P. Lehoczky, and Duane J. Seppi, 2008, Cross-sectional stock option pricing and factor models of returns, EFA 2008 Athens Meetings Paper, AFA 2009 San Francisco Meetings Paper.
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.

Vasquez, Aurelio, 2017, Equity volatility term structures and the cross section of option returns, *Journal of Financial and Quantitative Analysis* 52, 2727–2754.
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?, *Journal of Finance* 68, 987–1035.
Whaley, Robert E., 2013, Trading volatility: At what cost, *Journal of Portfolio Management* 40, 95–108.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.