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

Bjørn Eraker
Aoxiang Yang *

June 28, 2021

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. A specific model aimed at capturing VIX options prices and other asset market data is shown to successfully replicate the salient features of consumption, dividends, and asset market data, including the first two moments of 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 which our model both captures.

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

1 Introduction

Since their introduction in 2006, options on the VIX index have become the second most traded contracts on the 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 towards understanding the factors that drive differences in the pricing of these two types of catastrophe insurance contracts.

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 a cutting-edge equilibrium model that reproduces salient features of VIX futures and options, SPX returns, SPX 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 aggregative 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,

[1] The Option Clearing Corporation cleared a total of 1.5 Trillion SPX trades vs. 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`

and then apply a novel Fourier-type payoff transform analysis to derive a semi-closed form (up to a single integral) formula for the value of VIX options.

While there are countless studies of equity options market data, relatively fewer papers study VIX options. Mencía and Sentana (2013) use a panel of VIX futures and options to fit a no-arbitrage based time-series model. 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 SPX and VIX options information to predict market returns (SPX), VIX futures returns, SPX and VIX options returns. Huang, Schlag, Shaliastovich, and Thimme (2019) derive a diffusion-based no-arbitrage model to explain negative delta-hedged VIX options returns. Both papers conclude that volatility of volatility risk is priced with a negative market risk price. 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 pricing of VIX options, SPX options, the equity premium, the variance risk premium, risk free rates while maintaining the discipline imposed by a consumption based, fully fledged 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 which tend to be negatively skewed, as with the SPX. The VIX returns distribution is much more heavily skewed to the right 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 ones. 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. Additionally, the shock to implied volatility of volatility (VVIX) is positively but imperfectly correlated with the level of VIX itself, suggesting that VIX option 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 only move independently of VIX 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, Schlag, Shaliastovich, and Thimme (2019) find that delta-hedged VIX options returns are statistically significantly negative on average. Their interpretation is that after controlling for directional volatility risk, volatility-of-volatility risk is priced. We compute average rates of return on VIX calls and find them to be 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.

[Figure 1 about here.]

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. There are some interesting features of the data. For SPX, option prices rose throughout March and dramatically on March 16 and 18, days in which the SPX index dropped 12% and 5.18%, respectively. A fortuitous investor who bought the farthest 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, the 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 better than SPX. The price of a 70 strike VIX call increased 400 fold from March 2 until 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 16th and 18th, VIX calls showed a much larger spike on the 18th. 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 and vice versa 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 there is one distinct common factor that drives about 80% of 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 through 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 through simulating data from our model and find that the model replicates these 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 variance risk premium 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-order moments of VIX option ex-post return distributions, including mean, variance, skewness, and kurtosis. It matches implied VIX volatility in several dimensions: the average ATM implied VIX volatility (i.e., VVIX) is almost matched identically; the average implied volatility surface over maturity and strike is similar to what we observe: it is concavely, vastly skewed to the right (as would be 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 of the implied volatility curve (as a function of the strike) from concave during low and average VIX periods to convex during market stress periods. We argue 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 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 risks, and risk-neutral state dynamics. We prove 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 non-zero risk premia.

The rest of the paper is organized as follows. Section 2 and 3 describes our sample of VIX options and presents reduced-form evidence, respectively. Section 4 presents our equilibrium model of VIX option pricing. Section 5 presents results from our model calibration exercise, and Section 6 summarizes our findings. The online appendix contains our general theory as well as model derivations and extensions.

2 Data

The sample was collected from the CBOE[2] and consists of data sampled at the one-minute interval over the period from January 2006 until 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 non-synchronous quotes that 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 pre-market on the VIX expiration day, typically the third or fourth Wednesday of the month. This contrasts the VIX itself, which is computed from midpoints. While in theory, VRO should differ little from the open value of the VIX on the settlement day, in practice, it may. Griffin

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

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 T maturity, and F_t a T maturity futures price. Keeping in mind that mean reversion will imply that F_t is below spot VIX when spot VIX is high (and vice versa), an ATM option ($K = F_t$) will have a strike that is below spot VIX when spot VIX is high, and above spot VIX when spot VIX is low. The fact that the 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.

3 Exploratory Data Analysis

3.1 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 and low VIX states in our sample. There are a number of features of the data that are worth commenting on.

[Figure 2 about here.]

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

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, seen in the red six-day maturity case in the top graph. In the low-VIX case, however, the implied volatility functions are uniformly forming 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 40-50 strike range were trading at implied volatilities below 100% in November of 2008. On the other hand, 83-day maturity 40-50 strike range options traded at above 100% implied volatilities in 2017. Our model successfully replicates all three main characteristics, as shown later.

[Table 1 about here.]

Table 1 left panel shows the average VIX implied volatility surface over strike and maturity, which should largely inherit the characteristics of implied volatility surface in the low-VIX case which occupies a vast majority of the time. As seen, the two predominant patterns for the low-VIX case 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 about here.]

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 thereby VIX ATM implied, is related to some persistent factor that is imperfectly correlated with volatility itself. In our structural model, therefore, we specify a structure in which aggregate consumption growth volatility, σ_t, is driven by exogenous shocks with two components. The first is a regular CIR-style diffusion term. Second, aggregate volatility is also discontinuous, with jumps arriving at a rate λ_t - an independent self-exciting diffusion process. In equilibrium, both VIX and vol-of-vol are non-linear functions of σ_t and λ_t. This modeling specification allows us to match not only the positive yet 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 1.

3.2 Option Returns

3.2.1 Descriptive Evidence

[Table 2 about here.]

[Table 3 about here.]

Much like traditional asset pricing research, recent developments in option research emphasize risk premia associated with factors-shocks. Coval and Shumway (2001) show that average returns to SPX options are statistically significantly negative. Their paper shows that even delta-neutral straddles that are immune to tail-risk experience large negative returns. Bondarenko (2003) reports Sharpe ratios of -0.38 and -3.93 for 4% and 6% OTM SPX puts, respectively, while Eraker (2012) 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 Rubenstein (1996), Bollerslev, Tauchen, and Zhou (2009) among others), which is interpreted as a volatility risk premium.

Table 2 presents summary statistics on returns to VIX options positions using data from January 2006 until June 2020 - a period covering both the 2008 financial crisis and the Covid-19 crash in March 2020. For comparison, Table 3 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 something 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 that VIX puts earn positive premiums, and they do: Table 2 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, Giglio, Le, and Rodriguez (2017) who document a sharply downward-sloping term structure of risk premia for variance claims.

[Figure 4 about here.]

Figure 1 suggests that the pre Covid-19 prices of VIX options were cheaper than those of SPX options. In the following 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 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 for portfolio managers is 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, Figure 4 presents portfolio values of investments that are constantly replenished with cash. Specifically, we assume that one invests a single dollar each month into a target security based on moneyness and maturity. The graphs show the cumulative P&L 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 time 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 one thousand 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 ETFs.

3.2.2 Principal Component Analysis

To further understand the connection between returns to SPX and VIX options we perform a 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) applies factor analysis for equity returns, Christoffersen, Fournier, and Jacobs (2017) and Alex Horenstein and Xiao (2019) analyze factor structures in equity options, and Johnson (2017) performs PCA on the VIX term structure.[3]

[Figure 5 about here.]

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 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 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.

3.3 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

[3]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. Differently, we apply PCA on different types of options - SPX and VIX options.

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 the index follows a geometric Brownian motion.[4] 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's (1993) 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 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 depend on (t, X_t) provided the data used to run the regression are sampled over a small time interval.

[Figure 6 about here.]

Figure 6 shows the results from regressions

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

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

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

a VIX call option index.[5] The regressions are run day by day using ten minute price changes from overlapping data sampled at the one-minute interval. The typical number of observations is 405 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 seen, during periods of market calm, the estimated slope coefficients are close to zero, and during periods of high volatility, the estimated slope coefficients are positive. This suggests 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 non-linear relationship between changes in OTM SPX puts and OTM VIX calls.

4 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 Appendix A. We first specify a specific economic environment and describe equilibrium VIX, and then perform a generalized Fourier payoff transform analysis to derive a pricing formula for VIX options as a single integral. Appendix C derives pricing formulas for VIX futures and SPX options also as single integrals.

4.1 The Model

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

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

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

where V_t represents the continuation value. The parameter β is the rate of time preference, γ is the relative risk aversion, and IES is implicitly set at one. Consumption, dividends, and in the end, asset

[5] 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.

prices and returns are influenced by a key variable, 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. The term $\xi_V dN_t$ is a jump term where N_t is a compounded Poisson process with instantaneous arrival intensity λ_t which itself follows a mean-reverting diffusion process, and ξ_V is a time-invariantly distributed random variable representing 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 $\xi_V > 0$, implying upward jumps in volatility are emphasized. We assume 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 a nutshell, we pursue the simplest framework that captures as many aspects of VIX derivatives data as possible.

As Cox, Ingersoll, 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 condition (Feller (1951)) is satisfied, implying a finite density at zero. The stationary distribution of λ_t is highly right-skewed, arising from the square root term multiplying the Brownian shock in (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. Therefore, the model implies that there are times when jumps to volatility can occur with high probability, but these times are themselves rare. For similar reasons, there is a σ_t term multiplying the Brownian shock in (9), helping both prevent σ_t^2 from falling below zero and correctly replicate the right-skewness in its distribution.

4.2 State-Price Density

Appendix C shows that the equilibrium value function of the representative agent is given by[6]

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

where

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

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

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

Assume 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 $\gamma > 1$. (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 agent. Both results are intuitive. Under recursive preferences the 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.

Appendix C shows the instantaneous risk-free rate is given by

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

[6]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 the data slightly. But this is a sacrifice for precise framework tractability. Importantly, for our purposes, the price-dividend ratio is not constant, as we will show.

where β represents the role of discounting, μ intertemporal smoothing, and $\gamma\sigma_t^2$ precautionary savings.[7] Appendix C also shows that the state-price density is given by

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

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

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

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

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.

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

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

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

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

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 Q measure. Equations (21) through (24) show that, for both σ_t^2 and λ_t, the mean reversion becomes slower and the long-run mean becomes higher under Q measure. Moreover, Appendix C shows that

[7] Note that r_t can become negative when σ_t^2 is sufficiently high. A standard arbitrage when 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).

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 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 that ξ_V is exponentially distributed, the jump size is adjusted upward in the sense that its mean is increased under Q. Specifically, let $\xi_V \sim \exp(\mu_\xi)$ under P, then $\xi_V^Q \sim \exp(\frac{\mu_\xi}{1-\mu_\xi b_2})$ under Q.

By now, we have drawn upon all 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.

4.3 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.[8] As a result, the state variable $\ln D_t$ remains redundant: dB_t^D 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 its apparent role to govern 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.

[8] 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 a Jensen's term which 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 dividend being log-linear in consumption: $D_t = C_t^\phi e^{\mu_D t} e^{\sigma_D B_t^D}$.

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

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

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,

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

Appendix C 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 equations (C.23) through (C.27) in Appendix C, and $G(\sigma_t^2, \lambda_t)$ is the price-dividend ratio function.

4.4 Equity Premium

We then discuss the equity premium. No-arbitrage implies the instantaneous equity premium conditional on no jumps occurring in our economy, as shown in Appendix C, is given by

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

with

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

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 CRRA risk premium which arises from the compensation for the diffusion risk in consumption, dB_t^C. The second component, $-b_2\frac{G_1}{G}\sigma_V^2\sigma_t^2$, captures the compensation for the diffusion risk in volatility, dB_t^V. Appendix C shows that $\beta_2(\tau)$ is negative for all τ as long as $1<\phi<2\gamma$ which we 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 the compensation for the diffusion risk in volatility jump intensity, dB_t^λ. Appendix C shows that given $\beta_2(\tau)$ is negative, $\beta_3(\tau)$ is also negative for all τ, which implies $G_2<0$ (i.e., the price-dividend ratio decreases in λ_t) and thus the third component also takes a positive value. In contrast, the last term captures the 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, at the times volatility jumps upward two things happen simultaneously: first, marginal utility jumps upward by a percentage equal to $e^{b_2\xi_V}$; second, the stock price jumps downward by a percentage equal to $1-\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}$. Therefore, investors demand a jump risk premium for holding equity.

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

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

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 return. 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.*

4.5 VIX

Having obtained equilibrium SPX, we turn to 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[9]

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

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[10]

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

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

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

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

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

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

[9] 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 $VIX_t^2 = -2\big(E_t^Q\big[\ln P_{t+1/12}\big] - \ln E_t^Q\big[P_{t+1/12}\big]\big)$ as shown e.g., in Martin (2011) Result 5. We show in Appendix D 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 one. We thank an anonymous referee on this point.

[10] Our such log-linearization is highly accurate for two reasons. First, as in Seo and Wachter (2018), 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%.

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.[11] 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 noteworthy. First, in the reduced-form literature, VIX typically takes an affine or exponential affine structure (Mencía and Sentana (2013); Park (2016)) which delivers convenience for VIX option pricing. But in our model, 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 itself a negative (positive) premium with no ambiguity, as in the data. Examples include VIX futures and call options (put options), as we will verify quantitatively.

[11] Appendix C 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.

4.6 VIX Options

Our key focus is a (European) VIX call option which 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 pre-specified strike K.[12] No-arbitrage implies that the price of a VIX call option is given by

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

where it follows from (36) that

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

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

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

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 Appendix C.

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

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

[12] Note that the standard convergence of futures price to the underlying price as the time to maturity approaches zero holds regardless of whether the underlying is tradable or not. 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 at t into $\sqrt{a_{1/12}+c_{1/12}\sigma^2_{t+\tau}+d_{1/12}\lambda_{t+\tau}}$ units of consumption goods at $t+\tau$ (for any τ). Any such intertemporal consumption transfer however is ruled out in the model. A τ maturity VIX futures at t is in essence a random consumption strip which pays off $\sqrt{a_{1/12}+c_{1/12}\sigma^2_{t+\tau}+d_{1/12}\lambda_{t+\tau}}$ 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.

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) < 0$.[13] As the payoff structure we are looking at is not common, we prove the existence of relevant Fourier transforms in Appendix C. 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 did to our best knowledge.

5 Quantitative Analysis

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

5.1 Calibration

Table 4 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 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 large premia on VIX derivatives as those seen in the data. But this results in an excessively high dividend growth through the stock market leverage parameter ϕ, to counter which we set the adjustment term $\mu_D = -2\%$, finally generating 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 be 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 upon is $1 - 3\%$. For example, Bansal and Yaron (2004) document a volatility of 2.93% while Wachter (2013) documents a volatility of 1.34%.

[13] In the numerical section, we have used both the Riemann rule and the quadrature rule to approximate the integrals. They generate the same result. We also have compared the price obtained via integral with that via risk-neutral Monte Carlo simulation. We found the difference is negligible as long as the VIX option is not "too far OTM," which applies to all our reported results.

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, which value constitutes a source of debate, is set to one for tractability. Plus, a number of studies conclude that the reasonable values for this parameter should be somehow close to one (e.g., Vissing-Jørgensen (2002); Hansen, Heaton, and Li (2008); 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 from equity market data, 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 an average size of jump in VIX, 11.4, computed from historical VIX data from CBOE during the period 1990-2020.[14] 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%, slightly higher than 2.93% documented in Bansal and Yaron (2004) and higher than 1.34% in Wachter (2013).

[Table 4 about here.]

To calibrate the other model parameters, notably risk-aversion (γ), the diffusion parameter of the volatility (σ_V), the mean reversion of the jump intensity (κ^{λ}), the diffusion parameter of the jump intensity (σ_{λ}), and the dividend growth idiosyncratic volatility (σ_D), we design a coarse Simulated Methods of Moments procedure. Specifically, we search the parameter space to overall best match the following six data moments: mean VIX (19.3); standard deviation of VIX (7.4); average holding-period return on one-month ATM VIX call option (-48%); average one-month ATM VIX option

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

Black-76 implied volatility (0.69); monthly autocorrelation of one-month ATM VIX option Black-76 implied volatility (0.53); contemporaneous correlation between VIX and one-month ATM VIX option Black-76 implied volatility (0.48). The values and data sources for these moments are summarized in Tables 2 and 5. We match a majority of these moments well.

We calibrate risk aversion at 14, slightly higher than that in Bansal and Yaron (2004) and 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 a 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. 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 which impacts VIX option price). Finally, we set $\sigma_\lambda = 2.6$ and $\sigma_D = 0.1$ in order to match 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 σ_D as the dividend idiosyncratic risk contained in σ_D is not priced in equilibrium and thus only contributes to the constant component of the VIX index, thereby decreasing the exposure of VIX (derivative) returns to σ_t^2 and λ_t.

5.2 Simulation Results

5.2.1 General Moments

Table 5 displays a list of moments from a simulation of the model at calibrated parameters, as well as their counterparts in U.S. data. The model is discretized using an Euler approximation and simulated at a monthly frequency ($dt = 1/12$) for 100,000 months. Simulating the model at higher frequencies produces negligible differences in the results. We then aggregate 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 vs. 2.93% in the data. Notably, we match the correlation between consumption and dividend growths closely: 0.69 in the model vs. 0.59 in the data. This outperforms leading asset pricing models, as we see, 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 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.[15] 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 p/d ratios vary substantially over time and display an annualized autocorrelation that exceeds anything we could expect to generate with our model. This is a natural consequence of the fact that our model structure is geared toward explaining derivatives data and calibrated to do so at a relatively high frequency. Price-dividend ratios display annualized persistence that way exceeds that seen in high-frequency derivatives-based variables such as VIX and VVIX. Adding additional state-variables, such as in models by Campbell and Cochrane (1999) (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 Appendix D, 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 Appendix D, 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.

[Table 5 about here.]

Importantly, for the purposes of our study, we match the mean and standard deviation of the VIX index almost exactly. The fact that mean VIX (19.41) is higher than equity return objective volatility (17.71) illustrates the model's ability to generate a large unconditional variance risk premium close to that in the data. The monthly autocorrelation of the simulated VIX index is 0.8, close to 0.84 in

[15] 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 1927-2020.

the data. Turning to the model's ability to match key moments of VIX options data, we see that the average implied volatility for one-month ATM VIX options, denoted $E(imp\_vol_t)$, is estimated at 71.84 in the model simulations, which compares to 68.8 in the data. The model produces a volatility of the simulated VIX implied volatility, denoted $\sigma(imp\_vol_t)$, of 12.64 vs. 14.3 in the data - a slight miss on the low side. The model also produces a monthly autocorrelation of one-month ATM VIX implied volatility, denoted $AC_1(imp\_vol_t)$, of 0.49 vs. 0.53 in the data.

Our model matches the observed positive but imperfect correlation between the implied VIX volatility and VIX at 0.32 vs. 0.48 in the data. It is useful to consider this in relation to a model where 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 will be 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, Schlag, Shaliastovich, and Thimme (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, Schlag, Shaliastovich, and Thimme (2019) propose a model where 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.[16]

5.2.2 VIX Futures Returns

[Table 6 about here.]

Table 6 compares average returns and return standard deviations for VIX futures prices computed from data (see Eraker and Wu (2017)) and our model. We report average daily arithmetic and logarithmic returns, and return standard deviations by re-simulating our model at a daily frequency

[16]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.

to facilitate comparison with the results in Eraker and Wu (2017). For one-month contracts, both log and arithmetic returns are ballpark the same for the model as in the data. At longer horizons, the model generates a too high (negative) risk premium. This is well known in the variance-risk literature. In fact, Dew-Becker, Giglio, Le, and Rodriguez (2017) report positive returns to long-maturity variance swaps, a finding that cannot be reconciled with a negative volatility risk premium. Our model also matches the observed daily return standard deviations of VIX futures almost exactly, although these moments were never targeted in our parameter calibration. The Variance Risk Premium is always positive in our model which follows from the fact that VRP is a positive linear function of two positive state-variables.

[Figure 7 about here.]

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 vs. 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.

5.2.3 SPX Option Implied Volatilities

Before proceeding to our key focus, VIX options, we will 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 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.[17]

[Figure 8 about here.]

[17] 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.

5.2.4 VIX Option Implied Volatilities

Table 1 right panel reports VIX implied volatilities from our model and is thereby comparable to Table 1 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 vs. 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 Table 1 left panel. For example, at a 30 strike the data averages 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, on the contrary, that VIX mean reversion in the model is slightly faster than that in the data. The current parameter choices reflect a balance struck between those tensions.

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 1 right panel, Figure 10 shows what happens when we condition upon a high and low initial VIX. In the low VIX case, where we have 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 saw 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 saw on November 12, 2008 (Figure 2 top).

To understand this contrast, note that there are two forces in shaping VIX option implied volatility curves. First, equilibrium VIX-squared is linear in σ_t^2 and λ_t, but mainly driven 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 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) shows 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 oppositely 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 sharply decrease 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 implied volatility curve, not only in low-volatility times (Figure 10) but also in average times (Figure 9).

On the other hand, 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 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 a left skewness, VIX now has a particularly fat tail at the left end, making the implied volatility curve convex there.

[Figure 9 about here.]

[Figure 10 about here.]

5.2.5 VIX Option Returns

Figure 11 reports average returns on holding VIX call and put options to maturities. All returns are normalized to a monthly frequency. Some striking patterns are as follows. First, the model generates a negative (positive) premium 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 are insurances against possible spikes in σ_t^2 and λ_t and thus a negative premium is generated. Second, the model implies that, ceteris paribus, shorter maturity VIX options always carry a greater premium than longer maturity ones, showing that the shorter the maturity is, the more excessively expensive (cheap) the call (put) is. This is in principle consistent with the downward sloping term structure of VIX option implied volatility shown in Table 1 right panel. Below we show that this is true in the VIX implied volatility data. Third, the premia for both call and put are decreasing with moneyness, showing that the more out-of-the-money the VIX call (put) is, the more pronounced its role as a bet on (against) volatility and volatility-of-volatility. This is in principle consistent with the upward sloping VIX call option implied volatility curve across strike shown in Table 1 right panel.

[Figure 11 about here.]

Table 7 further reports returns to VIX options in our model with greater details. ITM (OTM) here indicates 15% in-the-money (out-of-the-money). Given VIX futures price is most of the time close to 20, 15% corresponds to 3 points, so that this table is directly comparable with Table 2. A few comments on the similarities and dissimilarities between the data averages and the model are in order. First, the model generates negative returns to call options, and positive returns to put options. This is consistent with the data. The model generates negative short maturity (one-month) call returns ranging from -14% (ITM) to -36% (OTM) which compare to -33% (ITM) and -60% (OTM) in the data. We need to keep in mind here that over the course of our sample, VIX spiked dramatically on several occasions, including the financial crisis of 2008, the events of February 5th, 2018 discussed in Section 3.2, and the Covid-19 crisis, leading to large positive returns for OTM VIX calls. For the longer maturity (six-month) call average returns, our model matches the data well: -45% (ITM) to -58% (OTM) in the model which compare to -53% (ITM) and -61% (OTM) in the data. Another commonality between the model and the data is the term structure of (absolute) returns on VIX call is downward sloping. For example, for ATM calls, the monthly return is -0.24 (-0.48) for one-month vs. $-0.53/6 \approx -0.09$ ($-0.59/6 \approx -0.1$) for six-month in the model (data).

[Table 7 about here.]

The model generates positive average returns on ITM, ATM, and OTM short and long maturity put options, consistent with the data in Table 2. 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 shows 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 vs. 123% in the model for ITMs, 305% vs. 212% (ATM), and 498% vs. 318% (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 vs. 68% in the model for ITMs, 101% vs. 109% (ATM), and 221% vs. 247% (OTM). At longer maturities, our model slightly undershoots return standard deviations for calls as well as for 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 vs. 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.

5.3 VIX Options as Hedges For SPX Options

[Table 8 about here.]

In order 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 ($N = 50,000$) realizations of state-variables one day ahead, compute option prices and VIX futures prices. We ask which one 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.

The results are presented in Table 8. Starting from the bottom we see that low values of the state-variables, consistent with a VIX at 12.7, imply that ATM SPX options returns are driven almost entirely by cash flow risks (D_t). OTM SPX options depend less on cash flow D_t 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 which probability 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 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.

[Figure 12 about here.]

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 above 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 8). First, during normal times when VIX is low (λ_t and σ_t^2 low), variation in D_t dominates 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 the variation in VIX options, the correlation between SPX and VIX options increases with these variables.

5.4 Comparative Static Analysis

[Figure 13 about here.]

In order to gain some additional insights into the workings of our model, we report the results of some comparative statics. In doing so, we also emphasize the necessity of recursive preferences

($\gamma > 1/\psi$) for the model to generate non-zero VIX derivatives premia. The left subplot of Figure 13 illustrates the co-movements of seven important steady-state conditional model moments with risk aversion. As shown, the equity premium is sensitive to risk aversion universally: 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 equity premium reflects compensations for three sources of risks corresponding to the model's three state variables. The pattern of the equity premium's variation with risk aversion reflects the fact that market price of risk for consumption growth increases linearly with γ, whereas the market prices of risks for volatility and its jumping risk only increase slowly with γ at the beginning and increasingly faster afterwards.

The right subplot of Figure 13 shows the impact of risk aversion on the market risk prices 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 variations in her value function in the future, which are affected by risks in σ_t^2 and λ_t. Both state variables therefore enter the agent's pricing kernel and are priced in equilibrium. However, these two state variables are by nature higher-order. Specifically, σ_t^2 measures the spot variance of consumption growth and thus is a second-order moment in terms of its relation with consumption, while λ_t governs the arrival intensity of jump in σ_t^2 and has third or even higher-order effects on consumption. Accordingly, market risk prices associated with σ_t^2 and λ_t increase relatively slowly with risk aversion.

Turning back onto the left subplot of Figure 13, it remains to check how other 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, a higher risk-neutral persistence, and mean of 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 variance risk premium increases.

The left subplot 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 values 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.

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

6 Concluding Remarks

This paper studies the properties of VIX derivatives prices, including the 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 where persistent state dynamics generate risk premia that exceed those seen under time-additive preferences by separating risk-aversion from intertemporal elasticity of substitution, as in Bansal and Yaron (2004), Eraker and Shaliastovich (2008), Drechsler and Yaron (2011), Wachter (2013), and many others. Our theoretical formulation mirrors the general framework outlined in Eraker and Shaliastovich (2008), but has the advantage that it does not require any linearization approximations in deriving the pricing kernel.

We use this modeling framework to specify a model that features a time-varying consumption volatility and time-varying intensity of jumps in that volatility process. This is different from the consumption disaster literature, as for example 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 variance risk premium, the correlation between VIX and VVIX, 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, we replicate the general shape of VIX option implied volatility functions, including the positive skewness and downward sloping term structure.

In equity and variance swap options, it is well known that implied volatilities exhibit convexity (i.e., smile) over strikes. In our VIX option data, the smile is actually a concave frown for the most

part of our sample, and particularly so when VIX is low. When VIX is high, it surprisingly changes to a convex smile. Even more surprisingly, our model replicates this empirical phenomenon.

We show that VIX options variations are 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 onto 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.

References

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

Alex Horenstein, Aurelio Vasquez, and Xiao Xiao, 2019, Common Factors in Equity Option Returns, *Working paper, Univ. of Miami.*

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

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-hedged gains and the negative market volatility risk premium, *The 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, *The 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(6), 1389–1414.

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

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

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

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

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

Bollerslev, T., and V. 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, *The 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 Working Paper.

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

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

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

Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs, 2017, The Factor Structure in Equity Options, *The Review of Financial Studies* 31, 595–637.

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

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

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

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

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

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

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

Duffie, Darrel, 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, 2012, The Volatility Premium, *working paper*.

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

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

Eraker, Bjørn, Ivan Shaliastovich, and Wenyu Wang, 2016, Durable Goods, Inflation Risk and Equilibrium Term Structure, *Review of Financial Studies* 29(1), 193–231.

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

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

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

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

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

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

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

Johnson, Travis L, 2017, Risk premia and the vix term structure, *Journal of Financial and Quantitative Analysis* 52, 2461–2490.

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

Lin, Yueh-Neng, and Chien-Hung Chang, 2009, VIX Option Pricing, *Journal of Futures Markets* 29(6), 523–543.

Lorenz, Friedrich, Karl Schmedders, and Malte Schumacher, 2020, Nonlinear dynamics in conditional volatility, *Available at SSRN 3575458*.

Martin, Ian, 2011, Simple variance swaps, Working paper, National Bureau of Economic Research.

Martin, Ian, 2017, What is the Expected Return on the Market?, *The Quarterly Journal of Economics* 132, 367–433.

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.

Piazzesi, Monica, and Martin Schneider, 2006, Equilibrium Yield Curves, *NBER Macroeconomics Annual 2006* MIT Press, 389–442.

Pohl, Walter, Karl Schmedders, and Ole Wilms, 2018, Higher order effects in asset pricing models with long-run risks, *The Journal of Finance* 73, 1061–1111.

Roll, R., and S. 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(2), 435–451.

Schneider, Paul, and Fabio Trojani, 2019a, (Almost) model-free recovery, *Journal of Finance* 74, 323–370.

Schneider, Paul, and Fabio Trojani, 2019b, Divergence and the Price of Uncertainty, *Journal of Financial Econometrics* 17, 341–396.

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

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, in *EFA 2008 Athens Meetings Paper, AFA 2009 San Francisco Meetings Paper*.

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

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

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

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.

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

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

Table 1: 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 data results with a sample over the Jan 2006 - June 2020 period. The right panel reports results computed from simulating the benchmark VIX model over 100,000 months.

	Data				Model			
Maturity (months)	1	2	3	6	1	2	3	6
Strike								
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

Table 2: VIX Option Returns

The table reports sample statistics on buy-and-hold returns to option positions in VIX. Returns are defined as $\text{payoff}_T/p_0 - 1$ where T is the expiration and p_0 is the price (midpoint) of the option one month or six 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. An OTM (ITM) is defined as a call option with a strike that is 3 points higher (lower) than ATM. Confidence intervals (CI) for the expected returns are computed by bootstrapping the return distribution. The sample period is Jan 2006 - 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 3: SPX Put Option Returns

The table reports sample statistics on buy-and-hold returns to OTM S&P 500 options. Returns are defined as $\text{payoff}_T/p_0 - 1$ 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 Jan 2006 to June 2020. 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

Table 4: Parameters for the VIX Model

The table reports parameter values for the VIX model in Section 4. The processes for log consumption, consumption growth volatility, volatility jump arrival intensity, and log dividend are respectively given by

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

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

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

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

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

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

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

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

Table 5: Simulation: Selected Model Moments

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

Table 6: Simulation: VIX Futures Returns

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

Maturity	R^1	R^2	Std
	Model		
1 month	-0.10	-0.18	3.75
2 month	-0.09	-0.14	3.12
3 month	-0.08	-0.12	2.71
4 month	-0.07	-0.10	2.40
5 month	-0.06	-0.08	2.14
	Data		
1 month	-0.12	-0.20	3.98
2 month	-0.07	-0.11	3.00
3 month	-0.01	-0.04	2.47
4 month	-0.03	-0.05	2.21
5 month	-0.01	-0.03	2.01

Table 7: 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 to be 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

Table 8: 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 upon a high-VIX, a medium-VIX, and a 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

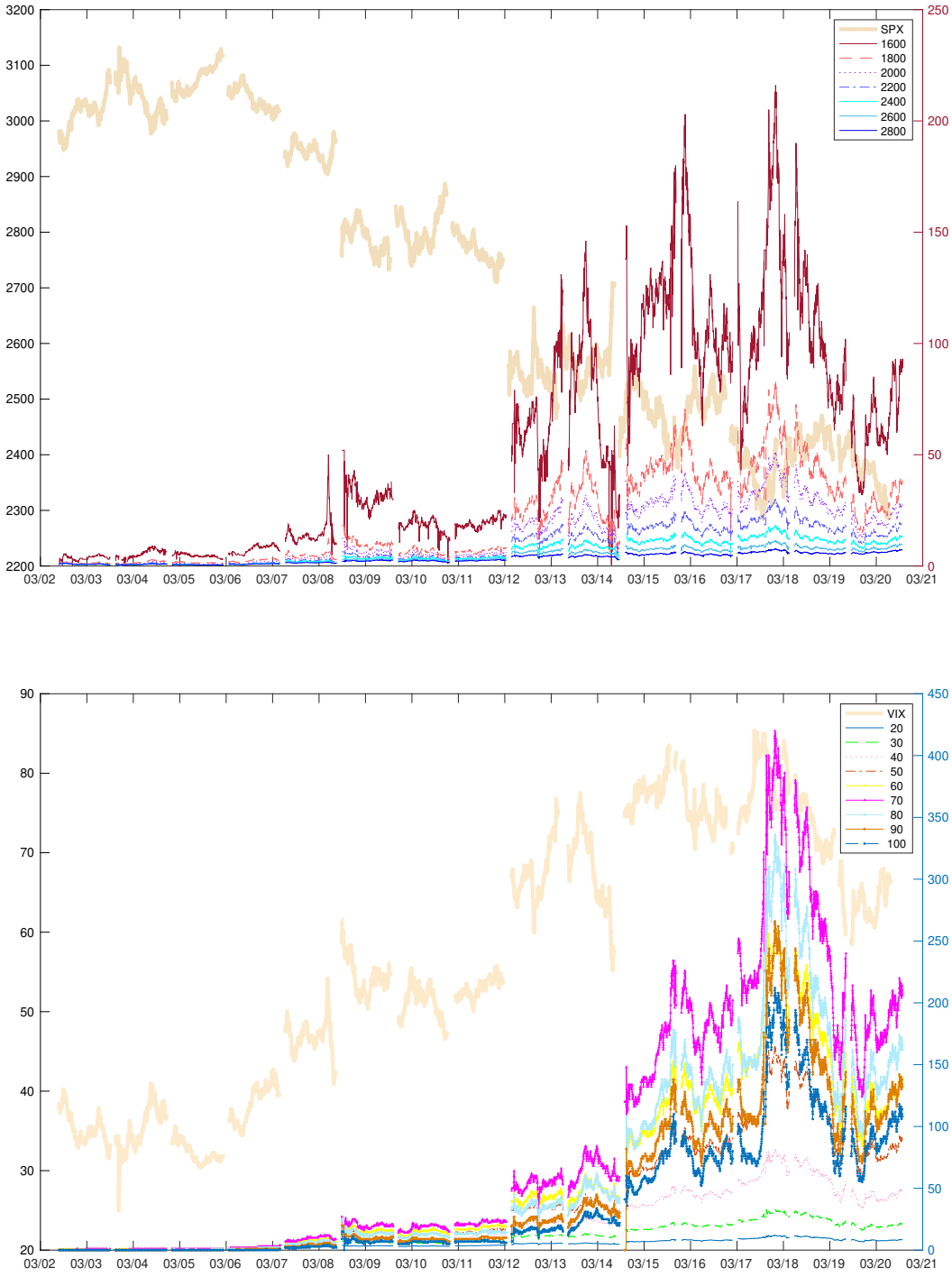


Figure 1: Returns to OTM SPX puts and VIX calls during the height of the Covid-19 crisis, March 2020.

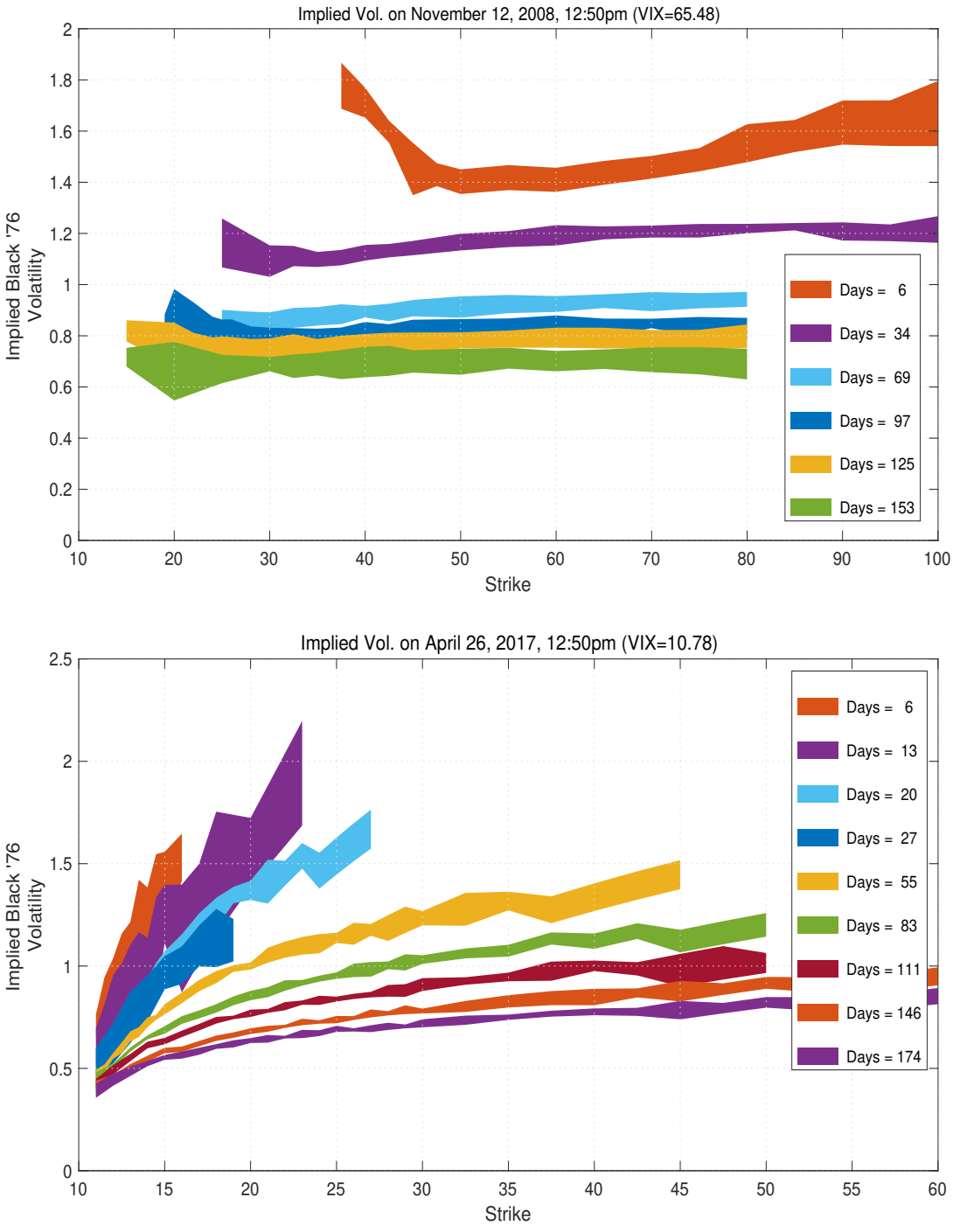


Figure 2: Implied VIX volatility on November 12, 2008 and April 26, 2017. The shaded areas represent the IV computed from bids and asks.

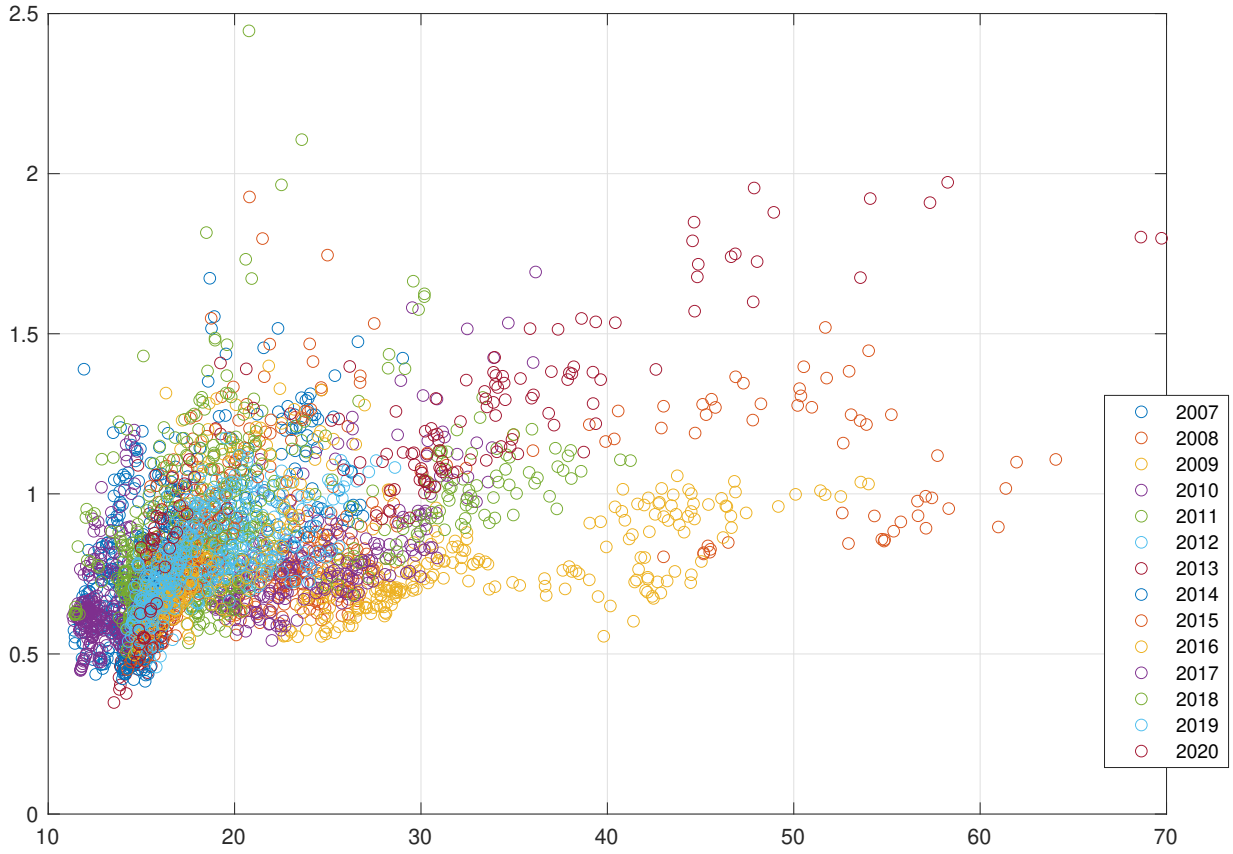


Figure 3: Scatter plot of one-month VIX futures prices vs. one-month ATM implied VIX volatility.

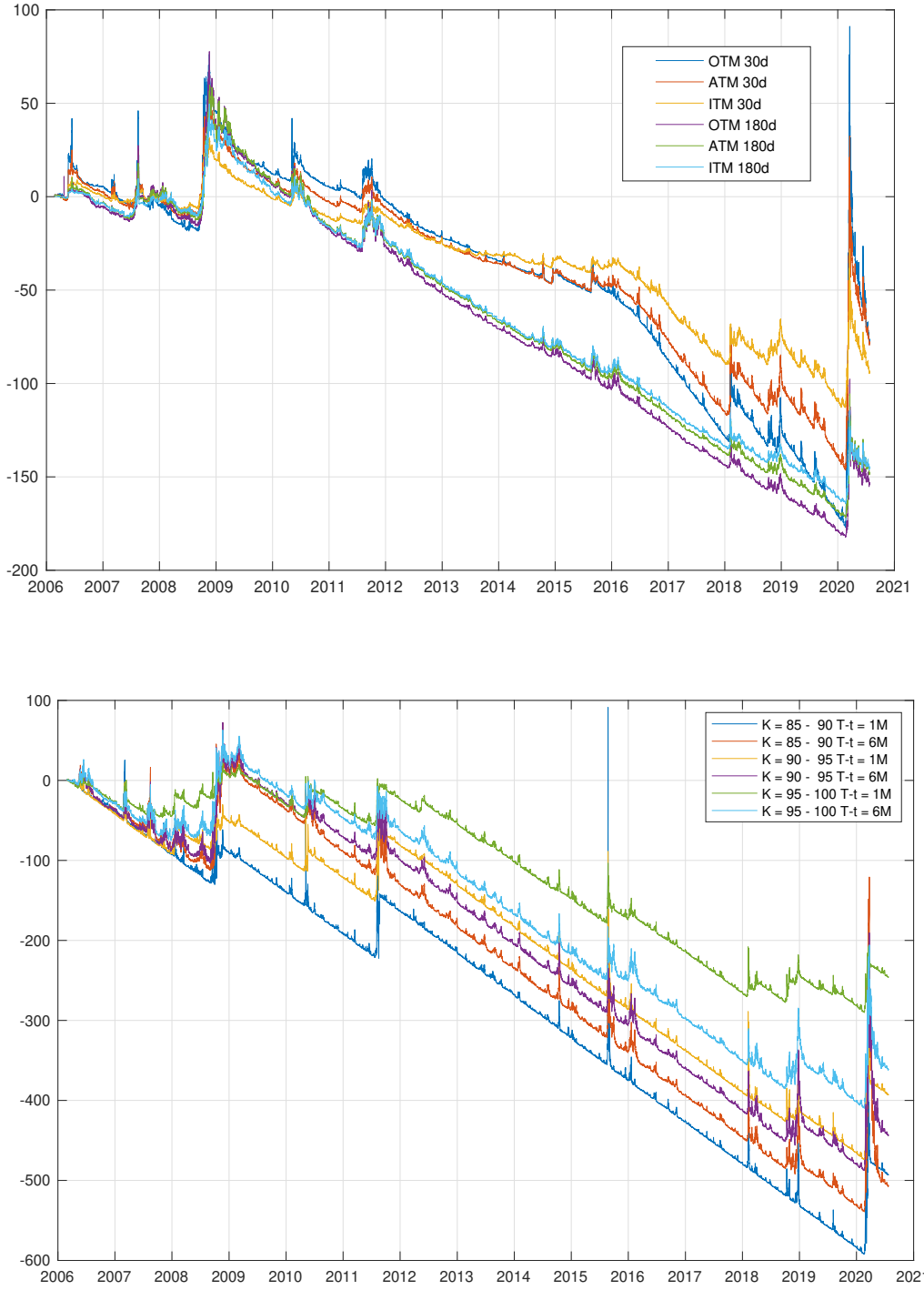


Figure 4: Marked-to-market value of 30 and 180 day VIX call options (top) and SPX put options (bottom).

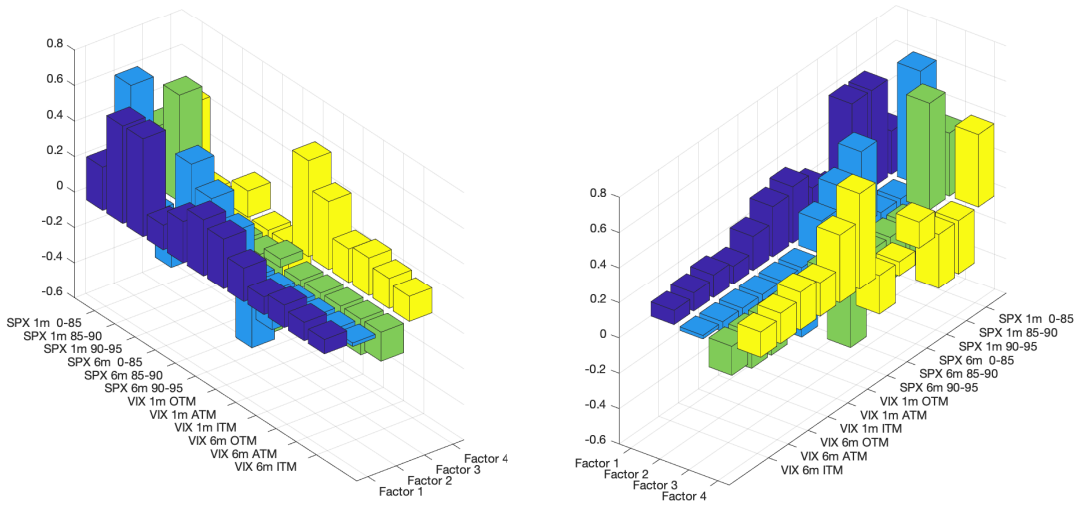


Figure 5: PCA factor loadings. The plots show the PCA factor loadings for the four first factors in the data as seen from two different angles.

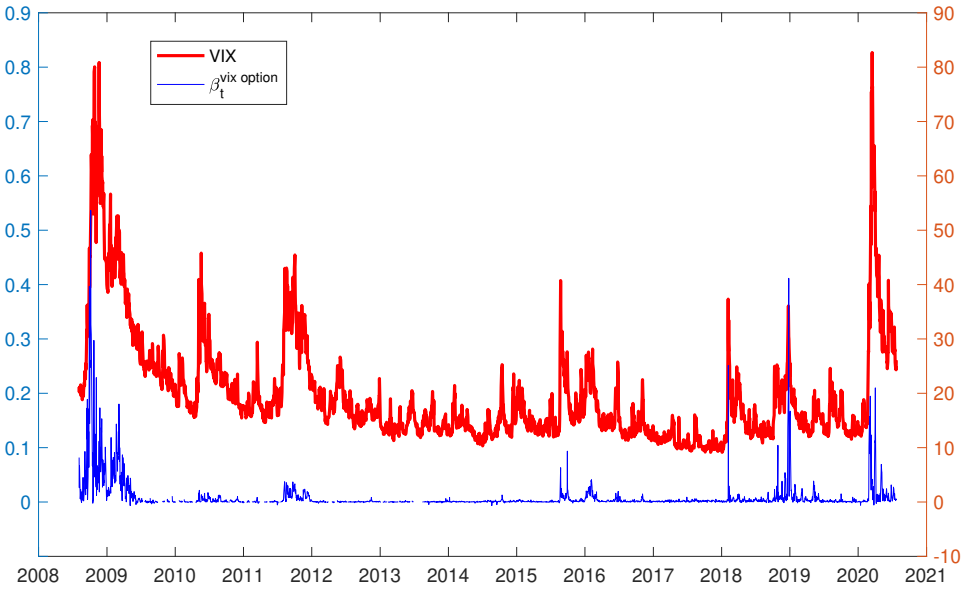


Figure 6: Hedge regressions. The figure shows $\beta_t^{\text{vix option}}$ computed through the regression, $\triangle P_t^i = \alpha_{i,t} + \beta_{i,t}^{\text{SPY}} \triangle SPY_t + \beta_{i,t}^{\text{VIX futures}} \triangle F_t + \beta_{i,t}^{\text{VIX option}} \triangle C_t + error_{i,t}$ where P_t is SPX put options, SPY_t is the S&P 500 index ETF, F_t is front month VIX futures and C_t is OTM VIX call options. The regression is run intraday using ten 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.

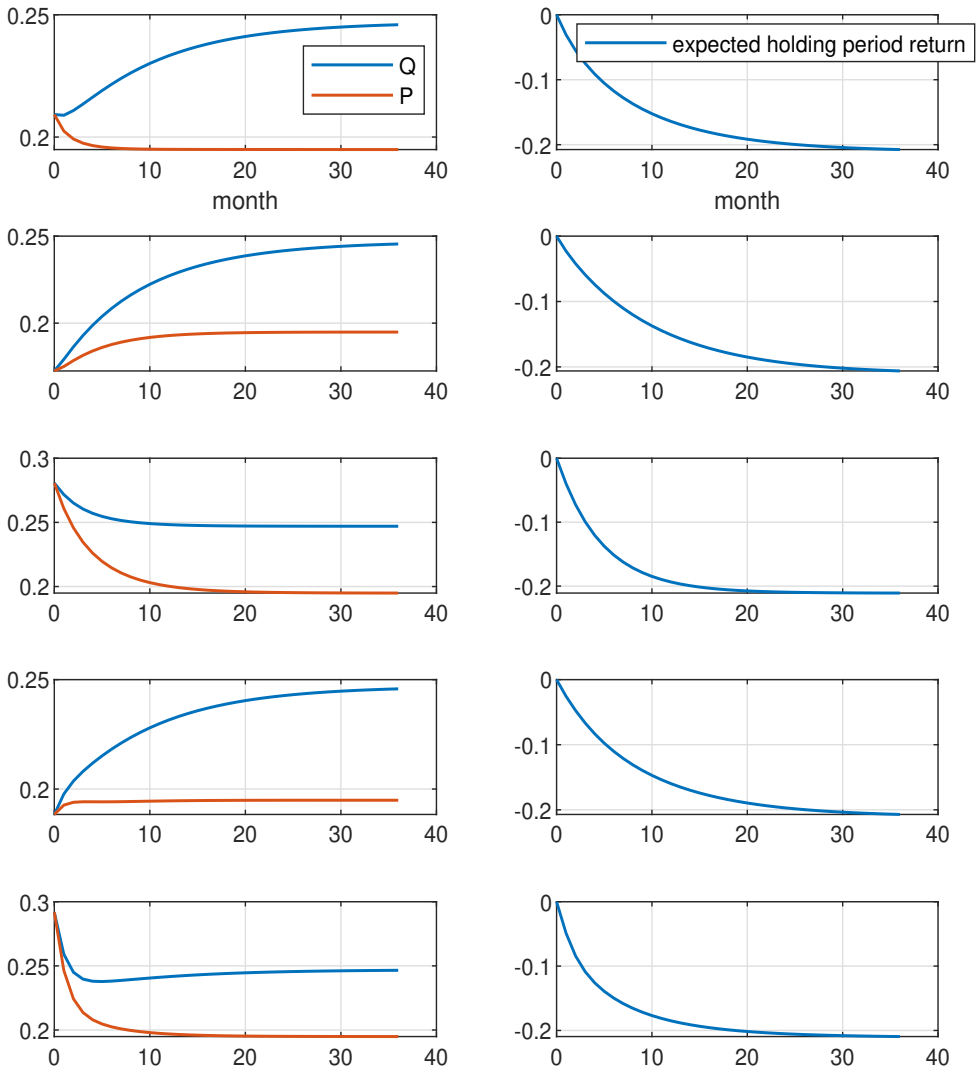


Figure 7: VIX futures curves and holding period returns. The figure illustrates conditional VIX futures term structures and conditional expected holding period returns on VIX futures in the model. Left: VIX futures curves (Q) and the objective-measure expected payoffs (P). Right: expected holding period 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 the following. First row: steady state σ_t^2 and λ_t; second row: low σ_t^2 and steady state λ_t; third row: high σ_t^2 and steady state λ_t; fourth row: steady state σ_t^2 and low λ_t; last row: steady state σ_t^2 and high λ_t.

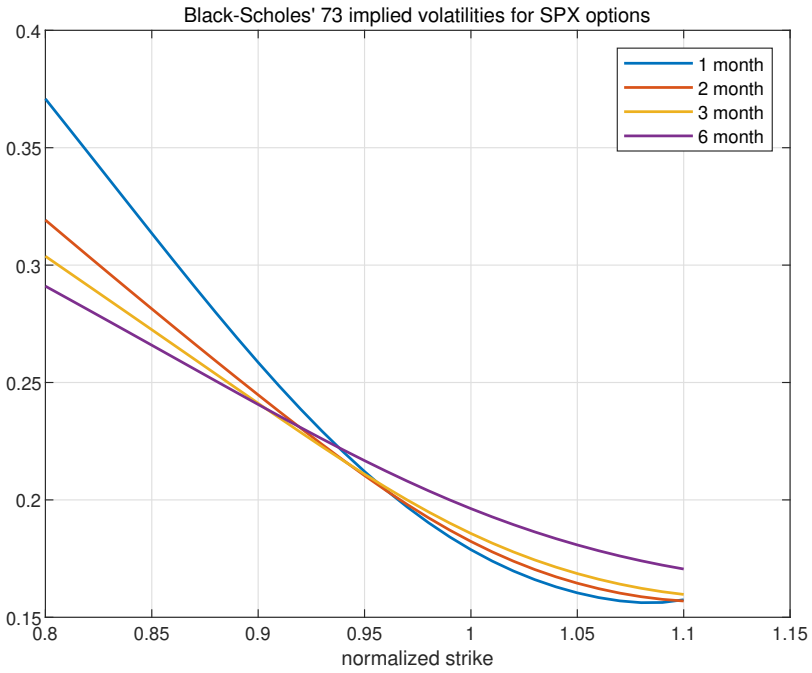


Figure 8: Black-Scholes implied volatility curves for SPX options. This figure plots (annualized) implied volatility curves computed from equating the Black-Scholes (1973) option pricing formula with the SPX option price in the model at steady state. The horizontal axis denotes strike normalized by SPX.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

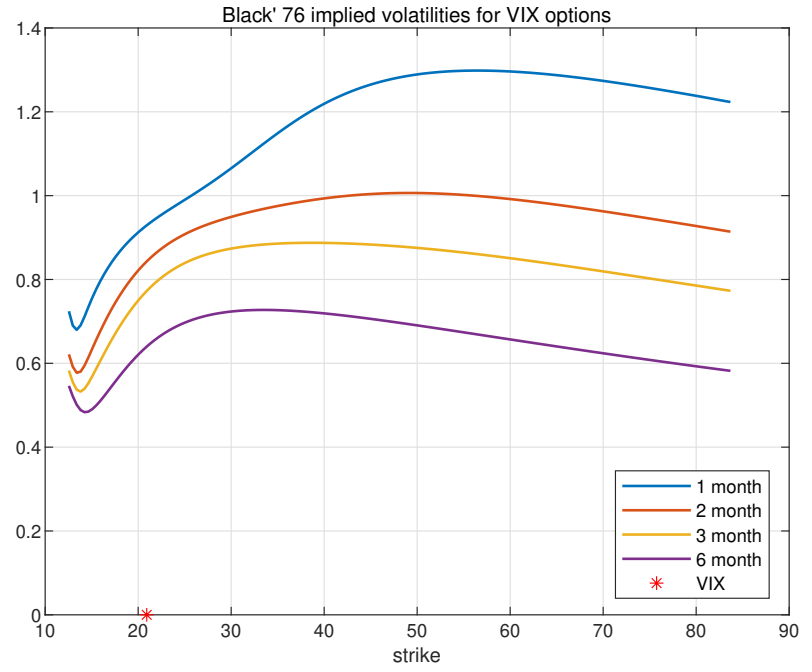


Figure 9: Black-76 implied volatility curves for VIX options. This figure plots (annualized) implied volatility curves computed from equating the Black (1976) futures option pricing formula with the VIX option price in the model at steady state. The horizontal axis denotes the absolute value of the strike.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

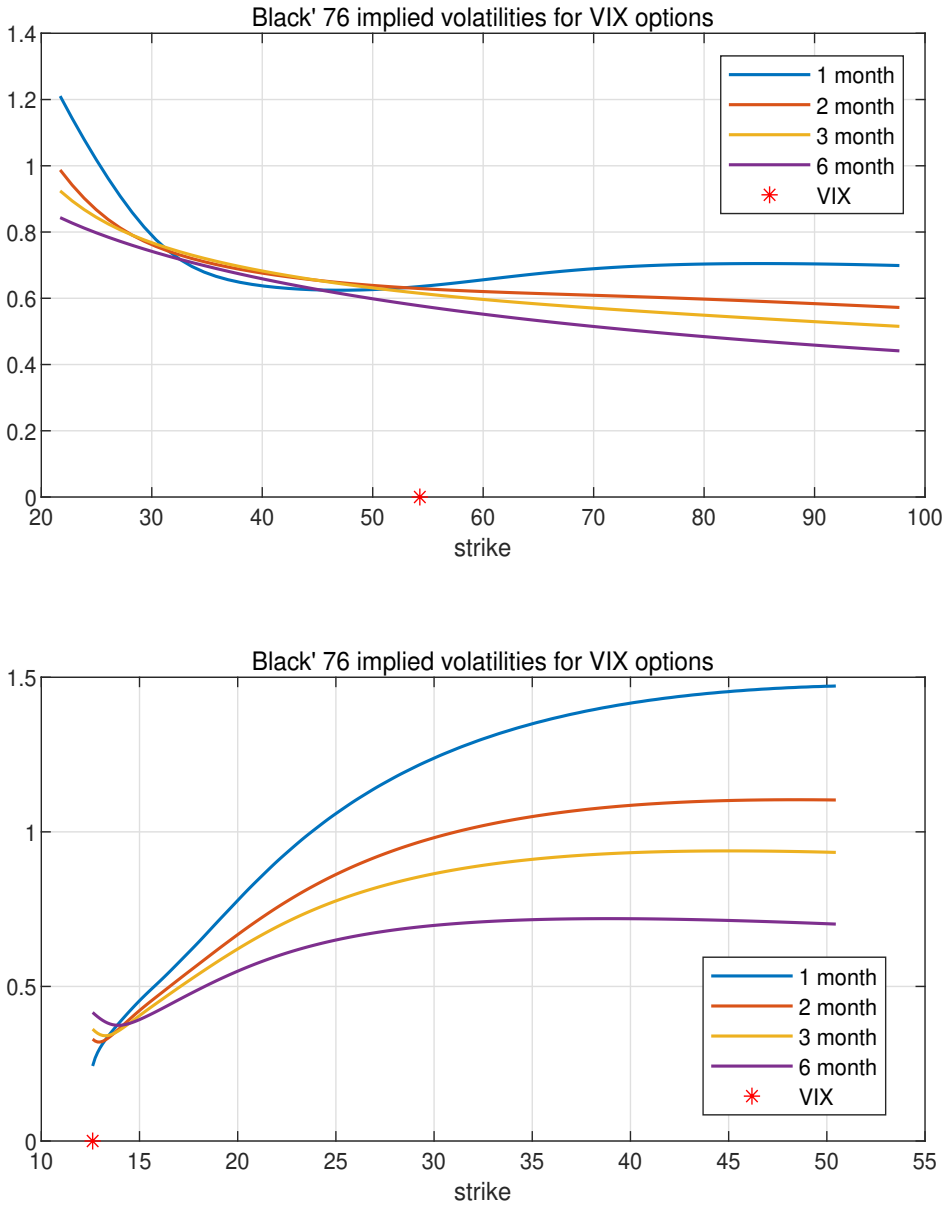
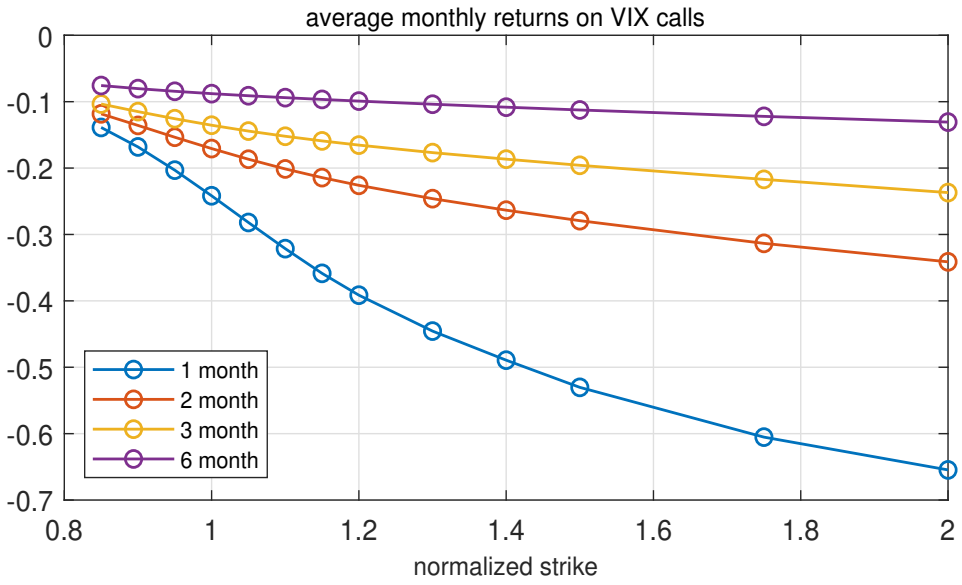


Figure 10: Black-76 implied volatility curves for VIX options: conditional analysis. The figure plots (annualized) implied volatility curves for VIX options in the model conditional on high and low initial VIX. In the upper case, we set both state variables very high: $\sigma_t^2 = 10\sigma_{ss}^2$ and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



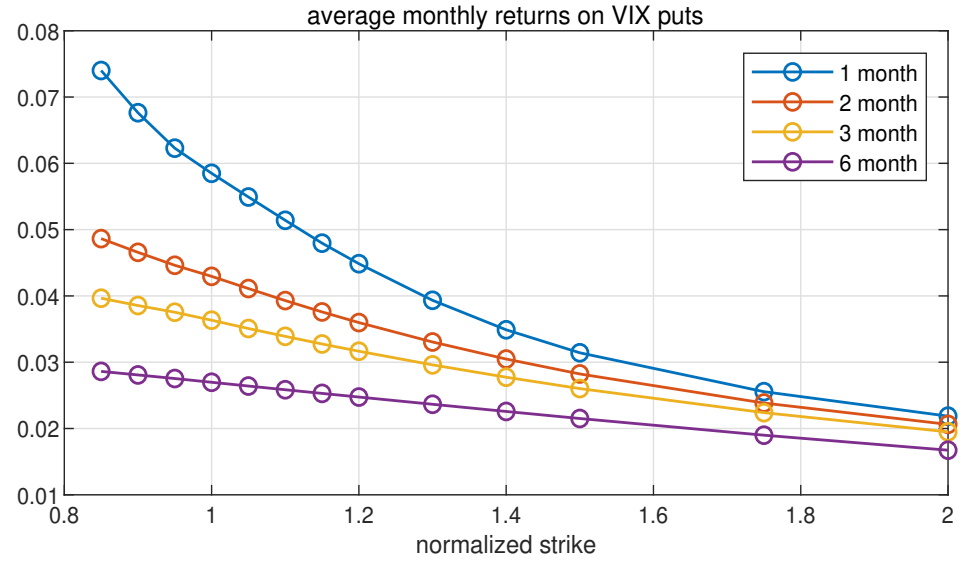


Figure 11: Average returns on VIX options. This upper (lower) figure plots average (monthly) returns on VIX calls (puts) in the model. In each case, we consider options with four maturities: 1, 2, 3, and 6 month, and the horizontal axis denotes the strike of relevant option normalized by its underlying asset price.

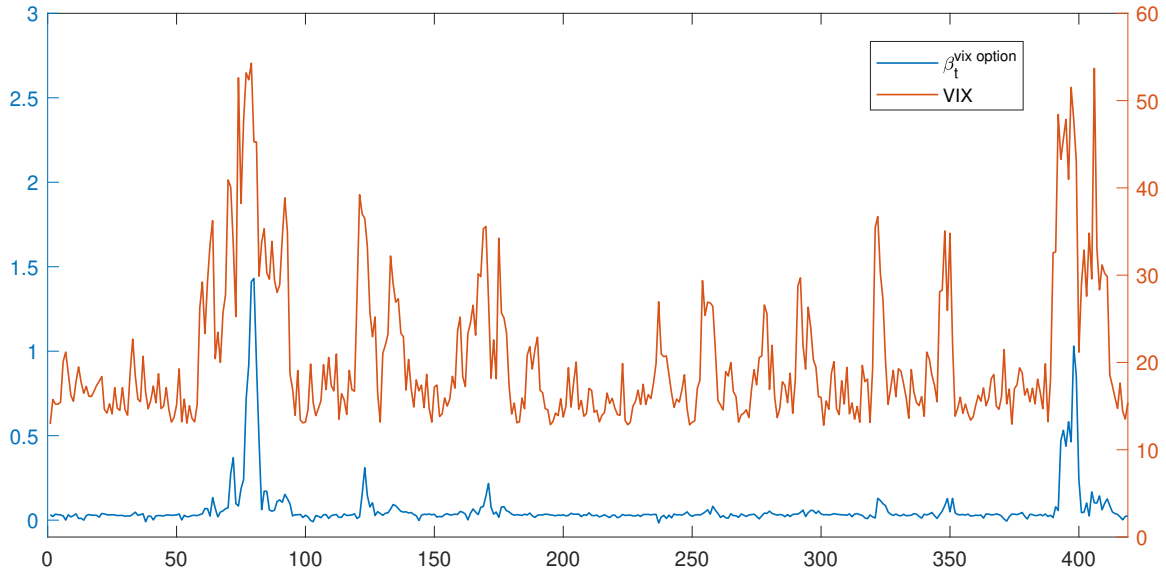


Figure 12: Hedge regressions. The figure shows $\beta_t^{\text{vix option}}$ computed through the regression under model-simulated data, $\Delta P_t = \alpha_t + \beta_t^{\text{SPX}} \Delta \ln SPX_t + \beta_t^{\text{vix futures}} \Delta F_t + \beta_t^{\text{vix option}} \Delta C_t + error_t$, where P_t is half-a-month 30% OTM SPX put normalized by SPX index (for stationarity in long-sample simulation), $\ln SPX_t$ is log market index, F_t is half-a-month VIX futures, and C_t is half-a-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.

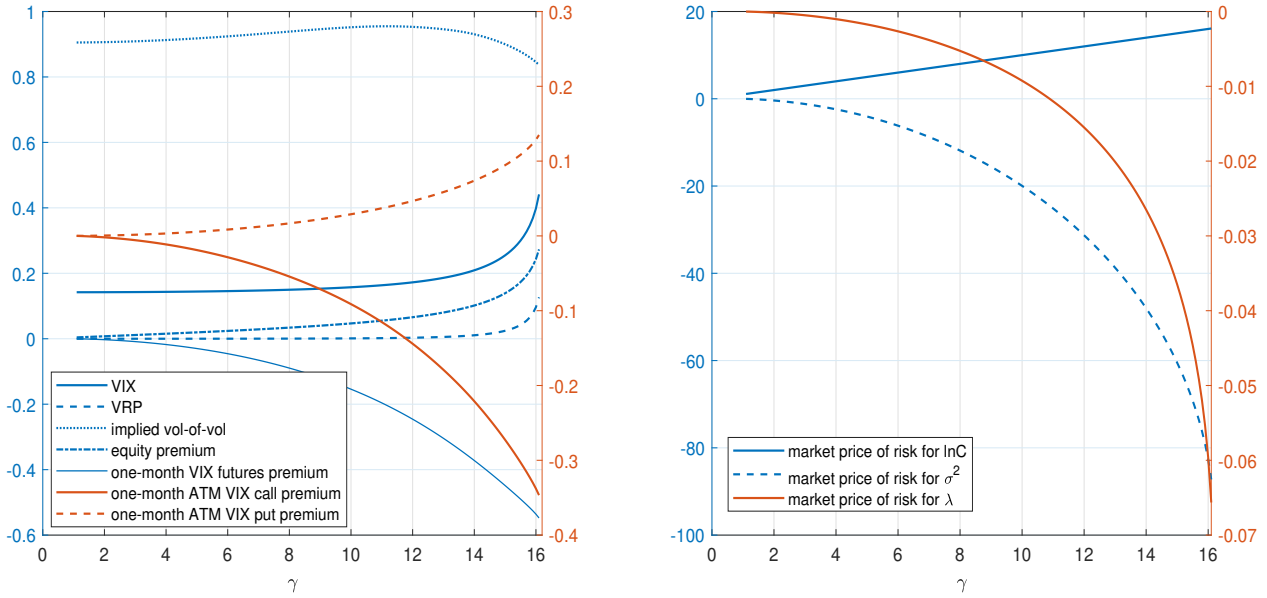


Figure 13: Comparative statics w.r.t. risk aversion. The figure illustrates steady-state conditional model moments (left) and market risk prices (right) as a function of risk aversion, γ. The left subplot reports VIX (VIX), variance risk premium (VRP), implied volatility of ATM VIX options (implied vol-of-vol), instantaneous equity premium (equity premium), one-month VIX futures premium, and one-month ATM VIX call and VIX put premiums. The right subplot reports the dependence of the market prices of risks for the model's three state variables, as represented by $(\gamma,-b_2,-b_3)$, on risk aversion, γ. The plot uses γ in the range from 1 to 16.1 - the upper limit for the existence of a model solution.

Internet Appendix

Appendix A. The General Model

This section presents a general equilibrium asset pricing model featuring recursive preferences for a representative agent in an economy with affine jump-diffusive states. Exact solutions are characterized for the value function, risk-free rate, state-price density, and induced risk-neutral measure. The model is set in real terms, such that interest rates are to be interpreted as real rates. The equity premium can be interpreted to be the same in real and nominal terms[18]

A.I Preferences

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

$$V_t = E_t \int_t^\infty f(C_s, V_s) ds \tag{A.1}$$

$$f(C, V) = \beta(1-\gamma)V(\ln C - \frac{1}{1-\gamma}\ln((1-\gamma)V)), \tag{A.2}$$

where V_t represents the continuation value. The parameter β is the rate of time preference, and γ is the relative risk aversion. The logarithm form in (A.2) indicates that we have set IES equal to one. It is well known that (A.2) is equivalent to the $\psi \to 1$ limit of the more general formulation

$$f(C, V) = \frac{\beta}{1-\frac{1}{\psi}} \frac{C^{1-\frac{1}{\psi}} - ((1-\gamma)V)^{\frac{1}{\theta}}}{((1-\gamma)V)^{\frac{1}{\theta}-1}}, \tag{A.3}$$

where $\theta = (1-\gamma)/(1-\frac{1}{\psi})$. The novel and appealing characteristic of the generalized preferences described in (A.3), sometimes named the Kreps-Porteus preferences, is that they break the tight link between intratemporal risk-aversion (γ) and intertemporal substitutability (ψ) and thus allow capturing the agent's preference for the timing of the resolution of uncertainty. For example, risk aversion is usually assumed larger than the reciprocal of the IES, $\gamma > 1/\psi$, in which case the

[18] Extensions of standard LRR models into nominal economies have been considered by Piazzesi and Schneider (2006), Eraker, Shaliastovich, and Wang (2016) among others. In these models equity and VIX derivative prices do not depend on inflation since inflation shocks are homoscedastic and therefore have constant risk-prices.

agent prefers early resolution of uncertainty. As we will show, this ensures that the compensations for various risks in our VIX model are of the right sign and quantitatively important. Note that although the preferences in (A.3) collapse into the familiar power utility with $\gamma = \frac{1}{\psi}$, in which case only risks to current consumption are priced, setting $\psi = 1$ and $\gamma > 1$ (i.e., (A.2)) still maintains the recursive preferences structure. From an empirical perspective, the relative size of the IES and one is also a source of debate. Several studies conclude that reasonable values for this parameter lie in a range close to one, or slightly lower than one (Vissing-Jørgensen (2002); Thimme (2017)), while the long-run risk literature (Bansal and Yaron (2004)) relies on an IES greater than one.

A.II State Variables

We follow Duffie, Pan, and Singleton (2000) and Eraker and Shaliastovich (2008) and assume that a set of n state variables, X_t, in the economy follow an affine diffusion-jump process. For an endowment-economy model to make sense, (log) consumption is taken to be part of the state-vector X_t. The other state variables are broadly defined. For example, one needs to include as an additional state variable expected consumption growth in a long-run risk model (Bansal and Yaron (2004)), the arrival intensity of the jumps to consumption in a time-varying disaster risk model (Wachter (2013)), and the price level or expected inflation in a nominal bond pricing model (Bansal and Shaliastovich (2013)). We emphasize that X_t includes all the state variables the agent needs to keep track of. Specifically, we fix the probability space $\{\Omega, \mathcal{F}, \mathcal{P}\}$ and the information filtration $\mathcal{F}_t$, and suppose that X_t is a Markov process in some state space $\mathcal{D} \subseteq \mathbb{R}^n$ with a stochastic differential equation representation

$$dX_t = \mu(X_t)dt + \Sigma(X_t)dB_t + \xi_t \cdot dN_t, \tag{A.4}$$

where B_t is an $\mathcal{F}_t$ adapted standard Brownian motion in $\mathbb{R}^n$. The term $\xi_t \cdot dN_t$ (element-by-element multiplication) captures n mutually conditionally independent jumps arriving with intensities respectively equal to the n elements of the vector $l(X_t)$ and jump sizes respectively equal to the n elements of the random vector ξ_t defined on $\mathcal{D}$.[19] Formally, each ith element of N_t is a Poisson process with time-varying intensity equal to the ith element of $l(X_t)$. We further assume that jump sizes ξ are

[19] Subsequently, a "$\cdot$" always represents an element-by-element multiplication; a "$\cdot/$" represents an element-by-element division. By assuming independent jumps, our model does not allow for exogenous co-jumps in state variables as seen in reduced-form models (Eraker, Johannes, and Polson (2003); Duffie, Pan, and Singleton (2000)). But note that because our model is an equilibrium one, jumps in state variables would endogenously translate to jumps in asset prices. For example, our specific VIX model endogenously implies co-jumps in SPX and volatility, consistent with reduced-form model specifications. Nevertheless, we emphasize that our general model remains tractable even with co-jumps in (A.4), in which case, to derive equilibrium, one needs to modify the HJB equation (A.14) and follow our derivation from there.

i.i.d. over time but not necessarily independent cross-sectionally. Their joint distribution is specified through the vector moment generating function $\varrho : \mathbb{C}^n \to \mathbb{C}^n$ (also called the "jump transform")

$$\varrho(u) = E[e^{u\xi}]. \tag{A.5}$$

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

$$\mu(X_t) = \mathcal{M} + \mathcal{K}X_t \tag{A.6}$$

$$\Sigma(X_t)\Sigma(X_t)' = h + \sum_i H_i X_{t,i} \tag{A.7}$$

$$l(X_t) = l + LX_t, \tag{A.8}$$

for $(\mathcal{M}, \mathcal{K}) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$, $(h, H) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$, $(l, L) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$. For X to be well defined, there are additional joint restrictions on the parameters of the model, which are addressed in Duffie and Kan (1996). To facilitate above matrix manipulations, note that $H = [H_1, H_2, ..., H_n]$ and $X_t = (X_{t,1}, X_{t,2}, ..., X_{t,n})'$. We assume an endowment economy and that the log consumption supply is always the first state variable of the economy. With a selection vector $\delta_c = (1, 0, 0, ..., 0)'$, this means

$$\ln C_t = \delta_c' X_t. \tag{A.9}$$

We make three additional assumptions. First, we rule out any intertemporal physical storage or transfer of consumption goods. Second, we assume the markets are not necessarily complete. In applying our theory, we take a pragmatic approach and assume an asset exists whenever we need to consider it. Examples include dividend claims, long-term riskless bonds, SPX options, and VIX derivatives, etc., which our general model keeps agnostic. However, because of the first assumption, market clearing implies the representative agent in equilibrium would strictly consume what she is endowed with each period C_t, implying that these assets' existences do not affect equilibrium consumption and thus the pricing kernel. Third, for the derivation of model equilibrium, we assume the consumption claim (a perpetual claim that exactly delivers the aggregate consumption as its dividend each period) and the instantaneous risk-free asset always exist (see Sections A.III, A.IV and A.V).[20]

[20] We note that all these assumptions are conventional in the endowment-based equilibrium asset pricing literature (e.g., Abel (1999); Campbell and Cochrane (1999); Campbell (2003); Bansal and Yaron (2004); Drechsler and Yaron (2011); Wachter (2013)).

A.III Dynamic Programming

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

$$\frac{S_t}{C_t} = A \tag{A.10}$$

for some constant A.[21] (A.9), (A.10) and a vector version of Ito's Lemma for jump-diffusion processes together imply

$$\frac{dS_t}{S_{t^-}} = \big(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c\big)dt + \delta_c'\Sigma(X_t)dB_t + \delta_c'\big[(e^{\xi} - \mathbf{1}) \cdot dN_t\big], \tag{A.11}$$

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

$$\frac{dS_t + C_{t^-}dt}{S_{t^-}} = \frac{dS_t}{S_{t^-}} + \frac{1}{A}dt. \tag{A.12}$$

Let r_t denote the instantaneous net risk-free rate. To solve for the value function, consider the Hamilton-Jacobi-Bellman equation for an investor who allocates wealth W_t between S_t and the risk-free asset. Let α_t be the fraction of wealth invested in the (risky) consumption claim S_t, and let C_t be the agent's consumption choice. Wealth then follows the process

$$\begin{aligned} dW_t = &\big[W_{t^-}\alpha_t(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{A} - r_t) + W_{t^-}r_t - C_{t^-}\big]dt \\ &+W_{t^-}\alpha_t\delta_c'\Sigma(X_t)dB_t + W_{t^-}\alpha_t\delta_c'\big[(e^{\xi} - \mathbf{1}) \cdot dN_t\big]. \end{aligned} \tag{A.13}$$

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

$$\begin{aligned} \sup_{\alpha_t, C_t} \Bigg\{ &J_W\big(W_t\alpha_t\big(\delta_c'\mu(X_t) + \frac{1}{2}\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c + \frac{1}{A} - r_t\big) + W_t r_t - C_t\big) + J_X'\mu(X_t) \\ &+\frac{1}{2}tr\Bigg(\begin{bmatrix} J_{XX} & J_{XW} \\ J_{XW}' & J_{WW} \end{bmatrix}\begin{bmatrix} \Sigma(X_t)\Sigma(X_t)' & W_t\alpha_t\Sigma(X_t)\Sigma(X_t)'\delta_c \\ W_t\alpha_t\delta_c'\Sigma(X_t)\Sigma(X_t)' & W_t^2\alpha_t^2\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c \end{bmatrix}\Bigg) \end{aligned}$$

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

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

$$+\sum_{i=2}^{n} E_{\xi_i}\big[J(W_t, X_{t,i}+\xi_i, X_{t,-i})\big]l_i(X_t) - J(W_t,X_t)\mathbf{1}'l(X_t) + f(C_t, J(W_t,X_t))\Bigg\} = 0, \qquad \text{(A.14)}$$

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

A.IV Value Function

In equilibrium, risk-free asset market clears: $\alpha_t = 1$, and the consumption claim market clears: $C_t = A^{-1}S_t = A^{-1}W_t$. Substituting these policy functions into (A.14) implies

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

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

$$+\sum_{i=2}^{n} E_{\xi_i}\big[J(W_t, X_{t,i}+\xi_i, X_{t,-i})\big]l_i(X_t) - J(W_t,X_t)\mathbf{1}'l(X_t) + f(A^{-1}W_t, J(W_t,X_t)) = 0. \qquad \text{(A.15)}$$

Conjecture that the solution to this equation, the equilibrium value function, takes the form

$$J(W,X) = \frac{W^{1-\gamma}}{1-\gamma}I(X). \qquad \text{(A.16)}$$

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

$$f(C,V) = \beta(1-\gamma)V(\ln C - \frac{1}{1-\gamma}\ln((1-\gamma)V)). \qquad \text{(A.17)}$$

Therefore,

$$f_C(C,V) = \beta(1-\gamma)\frac{V}{C}. \qquad \text{(A.18)}$$

The F.O.C. $f_C = J_W$, together with the derivative (A.18), the conjecture (A.16), and the condition that in equilibrium $J = V$, imply

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

Solving for A yields $A = \beta^{-1}$, i.e., $W_t/C_t = 1/\beta$. Having obtained the wealth-consumption ratio, we next solve for $I(X)$. It follows from (A.17) that

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

$$= \beta W^{1-\gamma}I(X)\Big(\ln\beta - \frac{\ln(I(X))}{1-\gamma}\Big). \tag{A.19}$$

Now substituting (A.16) and (A.19) into (A.15) yields

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

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

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

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

$$-\frac{1}{1-\gamma}I(X_t)\mathbf{1}'l(X_t) + \beta I(X_t)\Big(\ln\beta - \frac{\ln(I(X_t))}{1-\gamma}\Big) = 0, \tag{A.20}$$

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

$$I(X) = e^{a+b'X} \tag{A.21}$$

solves (A.20). Then

$$I_X(X) = bI(X) \tag{A.22}$$

$$I_{XX}(X) = bb'I(X). \tag{A.23}$$

Substituting (A.21) through (A.23) into (A.20) implies

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

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

$$-\frac{1}{1-\gamma}\mathbf{1}'l(X_t) + \beta\Big(\ln\beta - \frac{a+b'X_t}{1-\gamma}\Big) = 0. \tag{A.24}$$

Using equations (A.6) through (A.8) to rewrite $\mu(X_t)$, $\Sigma(X_t)\Sigma(X_t)'$ and $l(X_t)$, and making use of the property of the trace of matrices that $tr(AB) = tr(BA)$ whenever both AB and BA are defined, yield an equation linear in X_t

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

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

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

$$+\frac{1}{1-\gamma}\varrho(b)'\big(l+LX_t\big) - \frac{1}{1-\gamma}\mathbf{1}'\big(l+LX_t\big) + \beta\big(\ln\beta - \frac{a}{1-\gamma}\big) - \frac{\beta}{1-\gamma}b'X_t = 0. \tag{A.25}$$

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

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

$$+\Big(\varrho_1(1-\gamma+b_1) - \varrho_1(b_1)\Big)L'\delta_c + \frac{(1-\gamma)^2}{2}\delta_c'H\delta_c + (1-\gamma)\mathcal{K}'\delta_c - L'\mathbf{1} = 0, \tag{A.26}$$

which is a system of n equations for n unknowns $(b_1, b_2, ..., b_n)$. The system depends on the moment generating functions of the jump sizes, $\varrho(\cdot)$, and admits an explicit solution only in special cases. There are at least two cases in which (A.26) collapses into a quadratic equation system in b and can be easily solved with a relatively low dimension of n. First, if there are no jumps with state-dependent intensities, then $L = 0$. Second, if there is a jump only in consumption while the jump intensity does not depend on consumption itself, then one can also verify that (A.26) becomes a quadratic system. In many other cases, including our VIX model considered in Section 4, we solve these equations numerically.

While multiple solutions to (A.26) possibly exist, there are at least two ways to choose among them. First, Tauchen (2011) suggests choosing the solution which approaches a non-explosive limit as coefficients associated with X_t in $\Sigma(X_t)\Sigma(X_t)'$ approach zero (see equations (C.9) and (C.10) below). Second, Wachter (2013) suggests one can choose the solution that makes economic sense under a simple thought experiment. One such strategy can be to consider that jump size is identically equal

to zero while the jump intensity is not. Because essentially those jumps should have no economic consequence, the value function should somehow reduce to its counterpart under the standard diffusion model. Collecting constant terms in equation (A.25) results in the following characterization of a in terms of b

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

$$+\Big(\big(\varrho_1(1-\gamma+b_1) - \varrho_1(b_1)\big)\delta_c + \varrho(b) - \mathbf{1}\Big)'l + (1-\gamma)\beta\ln\beta\Big\}. \tag{A.27}$$

Equations (A.16), (A.21), (A.26), and (A.27) and the relation $W_t/C_t = 1/\beta$ together fully characterize the equilibrium value function. Because in most settings $\gamma > 1$, equations (A.16) and (A.21) imply that state variable X_i with a positive (negative) associated coefficient b_i would be negatively (positively) correlated with the value function, i.e., negatively priced. As shown in Section A.VI, the equilibrium prices of risks for X_t can be summarized by the following $n \times 1$ vector

$$\lambda = \gamma\delta_c - b. \tag{A.28}$$

Intuitively, except for log consumption, any state variable positively (negatively) correlated with the value function commands a positive (negative) market price of risk. Because log consumption affects the value function additionally through W, its market price of risk has an additional term γ. Now if $\gamma = 1 \equiv 1/\psi$, the Duffie-Epstein preferences collapse into the familiar CRRA preferences, and thus, as one can easily verify, (A.26) admits $b = (0, 0, ..., 0)'$ as a solution and from (A.28) $\lambda = (\gamma, 0, 0, ..., 0)$, which means only innovations to consumption are priced. Therefore, while consumption is the only priced factor in CRRA utility models, Duffie-Epstein preferences imply that all state variables are potentially priced.[22]

A.V Risk-Free Rate

Taking the derivative of (A.14) with respect to portfolio choice α_t and setting it to zero imply

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

[22] It is known that, for a state variable to be priced, it ultimately has to influence consumption in some systematic way. However, we emphasize that since our framework delivers exact solutions via a guess and verify method, one can always include a state variable that one suspects should be priced as part of X_t, say $X_{t,i}$, and solve equation (A.26) to check whether it is priced or not, that is, whether $b_i \neq 0$.

$$+E_{\xi_1}\big[\big(1+\alpha_t(e^{\xi_1}-1)\big)^{-\gamma}\big(e^{\xi_1}-1\big)e^{b_1\xi_1}\big]l_1(X_t)=0. \tag{A.29}$$

Evaluating the above equation under equilibrium condition $\alpha_t = 1$ and rearranging yield

$$\begin{aligned} r_t = \underbrace{\beta+\delta_c'\mu(X_t)+(\frac{1}{2}-\gamma)\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c+\delta_c'\big[\big(\varrho(b+1-\gamma)-\varrho(b-\gamma)\big)\cdot l(X_t)\big]}_{\text{CRRA Preferences}} \\ + \quad \underbrace{\delta_c'\Sigma(X_t)\Sigma(X_t)'b}_{\text{Duffie-Epstein Preferences}} \quad . \end{aligned} \tag{A.30}$$

The terms above the first bracket in (A.30) arise even if we assume CRRA preferences. β represents the role of discounting, $\delta_c'\mu(X_t)$ intertemporal smoothing, and $(\frac{1}{2}-\gamma)\delta_c'\Sigma(X_t)\Sigma(X_t)'\delta_c$ precautionary savings due to diffusion risks in consumption[23]. $\delta_c'\big[\big(\varrho(b+1-\gamma)-\varrho(b-\gamma)\big)\cdot l(X_t)\big]$ represents the representative agent's response to jump risks in consumption. Suppose the jump size for consumption is always negative, $\varrho_1'(\cdot)<0$, then $\varrho_1(b+1-\gamma)-\varrho_1(b-\gamma)$ is negative regardless of b. Intuitively, an increase in the probability of a downward jump in consumption, $l_1(X_t)$, increases the representative agent's desire to save and thus lowers the risk-free rate. The term above the second bracket in (A.30) represents the representative agent's saving motive response to risks in the economy that would only arise under Duffie-Epstein preferences. $\delta_c'\Sigma(X_t)\Sigma(X_t)$ captures the comovement between the diffusion in consumption and that in each state variable, while b determines the sign of the influence of each state variable on the marginal utility of consumption. Multiplication of them, if positive (negative), summarizes an additional aspect the representative agent likes (dislikes) about the diffusion risks in the economy. To better understand this point, think about volatility as the second state variable. Assume that the comovement between consumption and volatility diffusions is negative, i.e., $(\delta_c'\Sigma(X_t)\Sigma(X_t)')_2<0$, and that volatility positively (negatively) affects marginal utility (utility), i.e., $b_2>0$. Multiplying them would yield a negative push on the risk-free rate. Intuitively, because the times at which consumption is low are also times at which volatility and marginal utility are high, the representative agent dislikes and wants to avoid being impacted by this source of diffusion risk. She will thus have another precautionary saving motive, which pushes down the risk-free rate.

A.VI State-Price Density

Calculation of equilibrium prices and rates of returns is simplified considerably by using the state-price density and the induced risk-neutral measure, which reflect the equilibrium compensation the agent requires for bearing various risks in the economy. Unlike time-additive preferences, recursive

[23] Here, $\frac{1}{2}$ arises from applying Ito's Lemma and working with log consumption, which is quantitatively not important.

preferences imply that the state-price density depends explicitly on the value function. In particular, Duffie and Lions (1992) and Duffie and Skiadas (1994) show that the state-price density associated with the preferences in (A.1) and (A.2) is equal to

$$\pi_t = \exp\Big\{\int_0^t f_V(C_s, V_s)ds\Big\} f_C(C_t, V_t), \tag{A.31}$$

where f_C and f_V denote the derivatives of f with respect to the first and second argument, respectively.[24] To derive the state-price density explicitly, substitute $W_t = \beta^{-1}C_t$ and (A.21) into (A.16). Then taking (A.16) into (A.18) and taking the partial derivative of (A.19) with respect to V yield, respectively,

$$f_C(C_t, J(W(C_t), X_t)) = \beta^\gamma C_t^{-\gamma} e^{a+b'X_t} \tag{A.32}$$

$$f_V(C_t, J(W(C_t), X_t)) = \beta(1-\gamma)\ln\beta - \beta(a + b'X_t) - \beta. \tag{A.33}$$

Finally, substituting (A.32) and (A.33) into (A.31) and noting $\ln C_t = \delta_c' X_t$ yield

$$\pi_t = \beta^\gamma \exp\Big\{\eta t - \beta b' \int_0^t X_s ds + a + (b' - \gamma\delta_c')X_t\Big\}, \tag{A.34}$$

where $\eta = \beta(1-\gamma)\ln\beta - \beta a - \beta$. Applying Ito's Lemma on (A.34) then implies

$$\frac{d\pi_t}{\pi_{t^-}} = \mu_{\pi,t}dt + (b' - \gamma\delta_c')\Sigma(X_t)dB_t + \Big(e^{(b-\gamma\delta_c)\cdot\xi} - \mathbf{1}\Big)' dN_t, \tag{A.35}$$

where

$$\mu_{\pi,t} = \eta - \beta b'X_t + (b' - \gamma\delta_c')\mu(X_t) + \frac{1}{2}(b' - \gamma\delta_c')\Sigma(X_t)\Sigma(X_t)'(b - \gamma\delta_c). \tag{A.36}$$

Alternatively, given the form of the risk-free rate in (A.30), we can solve for $\mu_{\pi,t}$ following a familiar no-arbitrage argument. As no-arbitrage implies that $\pi_t e^{\int_0^t r_s ds}$ is a martingale, $E_t[d(\pi_t e^{\int_0^t r_s ds})]$ must be zero. Using Ito's Lemma, we can obtain

$$\mu_{\pi,t} = -r_t - \big(\varrho(b - \gamma\delta_c) - \mathbf{1}\big)' l(X_t). \tag{A.37}$$

[24] Section 6 of Duffie and Skiadas (1994) shows that three conditions are required to make the above statement. First, our aggregator function $f(C, V)$ (equation (A.2)) is continuous so it is progressively measurable. Second, as argued in Duffie and Epstein (1992), though our Kreps-Porteus aggregator $f(C, V)$ is a typical example that does not satisfy the Lipschitz condition, the existence of the stochastic differential utility V_t and the legitimacy of a pricing kernel as in (A.31) are proved by PDE methods in a Markov setting by Duffie and Lions (1992) under a growth condition in consumption. Third, the growth condition requires there existing a constant K such that $f(C, 0) \leq K(1 + |C|)$ for all possible C, which is trivially satisfied as from (A.2) $f(C, 0) = 0$ for all $C > 0$ (since we specify dynamics of $\ln C_t$, C_t is always positive) noting that $\lim_{V\to 0} V \ln V = 0$. Therefore, our general theory does not place any specific restrictions on state variable joint dynamics beyond its affine structure. We thank an anonymous referee for asking to check these conditions.

Making use of (A.27), (A.26) and (A.30), one can verify that (A.37) is equivalent to (A.36). We now define

$$\Lambda_t = \Sigma(X_t)'\lambda, \tag{A.38}$$

where, recall (A.28), $\lambda = \gamma\delta_c - b$ determines the market prices of risks in the different components of X_t such that if $\lambda_i = 0$ then innovations to $X_{t,i}$ are not priced. λ_i is related, through $\varrho_i(\cdot)$, to the price of jump risk with jump size ξ_i in the ith state variable. $\Lambda_{t,i}$ is literally the total price of the Brownian motion risks associated with $X_{t,i}$.

In Appendix B, we prove a convergence result: the equilibrium state-price density in (A.34) is an $IES \to 1$ limit of the more general state-price density approximately solved in Eraker and Shaliastovich (2008), thanks to the similarity between the state variable dynamics used in the two papers. Equivalently, the error in the log-linear approximation in that paper vanishes as the IES approaches one. We derive our model following different methods than Eraker and Shaliastovich (2008), but we will discuss the similarities in the IES=1 case. The benefit of focusing on this special case is that the equilibrium becomes tractable and not subject to any approximation errors, without losing any desirable features of recursive preferences.

A.VII Risk-Neutral Measure

With the state-price density at hand we can price assets with arbitrary state-dependent payoffs. It is sometimes more convenient, and even necessary, to work with state variable dynamics under the risk-neutral measure induced by the state-price density π_t. The following theorem is a generalization of Proposition 5 in Duffie, Pan, and Singleton (2000) (proof follows the same vein) to the settings of our model.

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

$$dX_t = (\mathcal{M}^Q + \mathcal{K}^Q X_t)dt + \Sigma(X_t)dB_t^Q + \xi_t^Q \cdot dN_t^Q, \tag{A.39}$$

where

$$\mathcal{M}^Q = \mathcal{M} - h\lambda \tag{A.40}$$

$$\mathcal{K}^Q = \mathcal{K} - H\lambda \tag{A.41}$$

$$dB_t^Q = dB_t + \Lambda_t dt \tag{A.42}$$

defines a Brownian motion under the risk-neutral measure.

The Q jump-arrival intensities are given by

$$l^Q(X_t) = l(X_t) \cdot \varrho(-\lambda). \tag{A.43}$$

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

$$\varrho^Q(u) = E^Q[e^{u\xi}] = \varrho(u - \lambda)./\varrho(-\lambda), \tag{A.44}$$

where element-by-element multiplication and division are respectively performed in (A.43) and (A.44).

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

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

$$\Lambda^{Jump} \equiv \big(E^P(\xi) - E^Q(\xi)\big)./Std^P(\xi) = \mathbf{1} - \varrho(-\lambda). \tag{A.45}$$

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

A.VIII Contribution of the General Model

The general model presented in this section 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 risks, and risk-neutral state dynamics. In contrast, Duffie and Epstein (1992), Duffie and Lions (1992), and Duffie and Skiadas (1994) define and characterize utility function on an exogenous consumption process without assuming affine state dynamics; thus, equilibrium risk price is not solved, that is, their terminal result is equation (A.31). Duffie, Pan, and Singleton (2000) do not take a stance on investors' preferences and impose an exogenous state-price density. In other words, market prices of risks (A.28), state-price density (A.34), and risk-neutral dynamics (Theorem A.1) are uncharacterized in DE92, DL92 and DS94, exogenous in Duffie, Pan, and Singleton (2000), and endogenous to investors' preferences and endowment dynamics in our model. Our theory has not only ready, wide asset pricing applications (e.g., our VIX model in Section 4) but also nests several existing applied models in the literature as special cases (e.g., Benzoni, Collin-Dufresne, and Goldstein (2011); Wachter (2013); Seo and Wachter (2018)).

Appendix B. Proof of Convergence of State-Price Density

Eraker and Shaliastovich (2008) solve a general equilibrium pricing model with an affine jump-diffusion structure for the state variable dynamics identical to those we use and a continuous-time limit of the discrete-time Epstein-Zin preferences. Because there is generally no closed-form solution to that model when the IES parameter is different from one, they also use a continuous-time limit of the log-linear approximation (see Campbell and Shiller (1988)) to maintain model tractability. We prove that the state-price density exactly solved in our paper is an $IES \to 1$ limit of the state-price density approximately solved in Eraker and Shaliastovich (2008).[25]

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

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

Because the affine structures for state variables are identical between both papers, the difference in the state-price densities can only arise from the difference in λ, which completely determines the

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

market prices of various risks in the economy for both papers. The λ in Eraker and Shaliastovich (2008) is given by

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

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

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

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

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

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

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

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

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

where the first equality follows from (B.2), the second from the fact that $\theta \equiv \frac{1-\gamma}{1-1/\psi} \to \infty$ as $\psi \to 1$, and the third and the fourth from (B.5). Note because $B \to 0$ and $\theta \to \infty$ as $\psi \to 1$, it is possible that λ approaches a well-defined limit, but when it does, from (B.6) the limit can only be

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

where the second equality follows from the definition of the vector χ in Eraker and Shaliastovich (2008). Therefore, it follows from comparing (A.28) with (B.7) that what is only left to show becomes

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

or

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

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

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

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

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

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

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

This is actually true since we exactly reproduce (A.26) after substituting $\chi = (1-\gamma)\delta_c + b$ into (B.11).

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

Appendix C. Solutions to the VIX Model

C.I Value Function and State-Price Density

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

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

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

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

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

and the jump intensities are summarized by

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

Define $b = (b_1, b_2, b_3)'$. Then (A.26) implies that b should solve

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

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

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

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

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

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

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

Once again we choose the negative root in (C.10). After solving out b we apply (A.27) to obtain

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

Then apply (A.30) to obtain the risk-free rate

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

where

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

Applying (A.35) then implies the state-price density obeys

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

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

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

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

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

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

C.II Equity Price

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

Riccati Equations for Discounted Characteristic Function

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

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

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

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

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

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

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

Together with boundary conditions, (C.21) implies

$$\beta_1(\tau) = \phi, \forall\tau. \tag{C.25}$$

Then as long as $1 < \phi < 2\gamma$ the solution to (C.22) takes the following closed form

$$\beta_2(\tau) = \frac{2(\phi-1)(\gamma - \frac{1}{2}\phi)\big(1 - e^{-\eta_\phi\tau}\big)}{(\eta_\phi + b_2\sigma_V^2 - \kappa^V)\big(1 - e^{-\eta_\phi\tau}\big) - 2\eta_\phi} \tag{C.26}$$

where

$$\eta_\phi = \sqrt{(b_2\sigma_V^2 - \kappa^V)^2 + 2(\phi-1)(\gamma - \frac{1}{2}\phi)\sigma_V^2}. \tag{C.27}$$

Note that since $1 < \phi < 2\gamma$ the term inside the square root is guaranteed to be positive. Moreover, $\eta_\phi > |b_2\sigma_V^2 - \kappa^V| \geq b_2\sigma_V^2 - \kappa^V$, implying that the denominator $(\eta_\phi + b_2\sigma_V^2 - \kappa^V)\big(1 - e^{-\eta_\phi\tau}\big) - 2\eta_\phi$ is strictly negative for all τ. The above arguments establish that $\beta_2(\tau) < 0$ for all τ. Noting the similarity between the forms of equations (C.22) and (C.23), it then follows that a similar argument as above can establish that $\beta_3(\tau) < 0$ for all τ, as long as $\varrho\big(\beta_2(\tau) + b_2\big) - \varrho(b_2) < -\epsilon$ for all τ for some $\epsilon > 0$. But the latter is actually satisfied since $\beta_2(\tau) < 0, \forall\tau$ and $\varrho(\cdot)$ is an increasing function. We have shown that $\beta_2(\tau) < 0, \beta_3(\tau) < 0$ for all τ. This is important because the sign of $\beta_2(\tau)$ and $\beta_3(\tau)$ respectively finally determines how the price-dividend ratio $G(\sigma_t^2, \lambda_t)$ responds to σ_t^2 and λ_t. The fact that $\beta_2(\tau) < 0, \beta_3(\tau) < 0$ for all τ implies that $G(\sigma_t^2, \lambda_t)$ is decreasing in both σ_t^2 and λ_t (recall (28)). This completes the proofs of propositions 1 and 2. Generally, (C.23) and (C.24) do not admit closed-form solutions, which we solve numerically.

C.III Equity Premium

We solve for the equity premium analytically. A familiar no-arbitrage condition on the equity market is that the discounted gains process $\pi_t P_t + \int_0^t \pi_s D_s ds$ is a P-martingale (for its derivation, see e.g.,

Appendix A.IV. of Wachter (2013)). It follows that the drift term in $E_t\big[\frac{d\left(\pi_t P_t+\int_0^t \pi_s D_s ds\right)}{\pi_{t^-}P_{t^-}}\big]$ is zero. Using Ito's Lemma, this implies

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

$$+\frac{1}{2}\frac{G_{11}}{G}\sigma_V^2\sigma_t^2+\frac{1}{2}\frac{G_{22}}{G}\sigma_\lambda^2\lambda_t+\frac{D_{t^-}}{P_{t^-}}+\lambda_t E\Big[e^{b_2\xi_V}\frac{G(\sigma_t^2+\xi_V,\lambda)}{G(\sigma_t^2,\lambda)}-1\Big]=0, \tag{C.28}$$

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

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

$$+\frac{1}{2}\frac{G_{11}}{G}\sigma_V^2\sigma_t^2+\frac{1}{2}\frac{G_{22}}{G}\sigma_\lambda^2\lambda_t+\frac{D_{t^-}}{P_{t^-}}+\lambda_t E\big[e^{b_2\xi_V}-1\big], \tag{C.29}$$

where we have used (A.37) to substitute r_t. Rearranging (C.28) properly and then substituting into (C.29) give rise to the expression for the equity premium conditional on no jumps occurring

$$\mu_{P,t}+\frac{D_{t^-}}{P_{t^-}}-r_t=\big[\phi\sigma_t,\frac{G_1}{G}\sigma_V\sigma_t,\frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\big]\Lambda_t+\lambda_t E\Big[e^{b_2\xi_V}(1-\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)})\Big]. \tag{C.30}$$

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

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

where

$$\sigma_{P,t}=\big[\phi\sigma_t,\frac{G_1}{G}\sigma_V\sigma_t,\frac{G_2}{G}\sigma_\lambda\sqrt{\lambda_t}\big]'. \tag{C.32}$$

C.IV Equity Price Dynamics under Q

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

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

$$+\sigma_{P,t}'dB_t+\sigma_D dB_t^D+\big[\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}-1\big]dN_t, \tag{C.33}$$

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

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

$$+\sigma'_{P,t}dB_t+\sigma_D dB_t^D+\ln\big[\frac{G(\sigma_t^2+\xi_V,\lambda_t)}{G(\sigma_t^2,\lambda_t)}\big]dN_t. \tag{C.34}$$

By plugging in the expressions for the diffusions under Q measure in equation (C.17) into (C.34) and replacing ξ_V and N_t with their counterparts under Q, we recover the dynamics of the log equity price under Q as

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

$$+\frac{1}{2}(\frac{G_{11}}{G}-\frac{G_1^2}{G^2})\sigma_V^2\sigma_t^2+\frac{1}{2}(\frac{G_{22}}{G}-\frac{G_2^2}{G^2})\sigma_\lambda^2\lambda_t\Big)dt+\sigma'_{P,t}dB_t^Q+\sigma_D dB_t^D+\ln\big[\frac{G(\sigma_t^2+\xi_V^Q,\lambda_t)}{G(\sigma_t^2,\lambda_t)}\big]dN_t^Q. \tag{C.35}$$

C.V VIX

By definition, $VIX^2(X_t)=\mathrm{Var}_t^Q\big[\ln P_{t+1/12}\big]=\mathrm{Var}_t^Q\big[\ln\tilde{P}_{t+\frac{1}{12}}\big]+1/12\sigma_D^2$, where we have separated the dividend idiosyncratic risk out and

$$\ln\tilde{P}_t\equiv\phi\ln C_t+g_1^*\sigma_t^2+g_2^*\lambda_t \tag{C.36}$$

denotes the portion of the log equity price that only involves systematic risk. The conditional cumulant generating function for $\ln\tilde{P}_{t+1/12}$ is given by

$$\Phi(u)=\ln E_t^Q e^{u\ln\tilde{P}_{t+1/12}} \tag{C.37}$$

$$=\ln E_t^Q e^{u\lambda_X' X_{t+1/12}} \tag{C.38}$$

$$=\alpha(u\lambda_X,t,t+1/12)+\beta'(u\lambda_X,t,t+1/12)X_t \tag{C.39}$$

where

$$\lambda_X\equiv(\phi,g_1^*,g_2^*)'. \tag{C.40}$$

Therefore, using the property of the cumulant generating function, we see that $\mathrm{Var}_t^Q\big[\ln\tilde{P}_{t+1/12}\big]=\tilde{a}_{1/12}+b_{1/12}\ln C_t+c_{1/12}\sigma_t^2+d_{1/12}\lambda_t$, where $\tilde{a}_{1/12},b_{1/12},c_{1/12}$ and $d_{1/12}$ are the second derivatives w.r.t. u of $\alpha(u\lambda_X,t,t+1/12)$, $\beta_1(u\lambda_X,t,t+1/12)$, $\beta_2(u\lambda_X,t,t+1/12)$ and $\beta_3(u\lambda_X,t,t+1/12)$ evaluated at

$u = 0$, respectively. Under appropriate technique conditions (see Duffie, Pan, and Singleton (2000)), let $\alpha(\tau), \beta(\tau)$ solve

Riccati Equations for Cumulant Generating Function

$$\dot{\beta}(\tau) = \mathcal{K}^{Q'}\beta(\tau) + \frac{1}{2}\beta(\tau)'H\beta(\tau) + L^{Q'}\big(\varrho^{Q}(\beta(\tau)) - 1\big) \tag{C.41}$$

$$\dot{\alpha}(\tau) = \mathcal{M}^{Q'}\beta(\tau) + \frac{1}{2}\beta(\tau)'h\beta(\tau) + l^{Q'}\big(\varrho^{Q}(\beta(\tau)) - 1\big) \tag{C.42}$$

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

$$\dot{\beta}_1(\tau) = 0 \tag{C.43}$$

$$\dot{\beta}_2(\tau) = \frac{1}{2}\sigma_V^2\beta_2^2(\tau) + (b_2\sigma_V^2 - \kappa^V)\beta_2(\tau) + \frac{1}{2}\beta_1^2(\tau) - (\frac{1}{2} + \gamma)\beta_1(\tau) \tag{C.44}$$

$$\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2\beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\big(\beta_2(\tau) + b_2\big) - \varrho(b_2) \tag{C.45}$$

$$\dot{\alpha}(\tau) = \mu\beta_1(\tau) + \kappa^V\theta^V\beta_2(\tau) + \kappa^\lambda\theta^\lambda\beta_3(\tau). \tag{C.46}$$

Together with boundary conditions, the solutions are

$$\beta_1(\tau) = u\phi, \forall\tau \tag{C.47}$$

$$\dot{\beta}_2(\tau) = \frac{1}{2}\sigma_V^2\beta_2^2(\tau) + (b_2\sigma_V^2 - \kappa^V)\beta_2(\tau) - \frac{1}{2}u\phi\big(2\gamma + 1 - u\phi\big) \tag{C.48}$$

$$\dot{\beta}_3(\tau) = \frac{1}{2}\sigma_\lambda^2\beta_3^2(\tau) + (b_3\sigma_\lambda^2 - \kappa^\lambda)\beta_3(\tau) + \varrho\big(\beta_2(\tau) + b_2\big) - \varrho(b_2) \tag{C.49}$$

$$\dot{\alpha}(\tau) = \mu\phi u + \kappa^V\theta^V\beta_2(\tau) + \kappa^\lambda\theta^\lambda\beta_3(\tau). \tag{C.50}$$

Here $\beta_1(\tau)$ has a closed-form solution, while $\beta_2(\tau), \beta_3(\tau), \alpha(\tau)$ do not.[26] It follows immediately from (C.47) that $b_{1/12} = 0$, i.e., VIX does not explicitly depend on current consumption. Finally, we can write $Var_t^Q\big[\ln \tilde{P}_{t+1/12}\big]$ as an affine function in σ_t^2 and λ_t

$$\mathrm{Var}_t^Q\big[\ln \tilde{P}_{t+\frac{1}{12}}\big] = \tilde{a}_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t, \tag{C.51}$$

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

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

$$VIX^2(X_t) = \frac{1}{12}\sigma_D^2 + \tilde{a}_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t. \tag{C.52}$$

For notational convenience, we compound the two constant terms in (C.52) and denote it as $a_{1/12}$. It follows that

$$VIX(X_t) = \sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}. \tag{C.53}$$

C.VI VIX Futures Pricing

The challenge in computing VIX futures price $F^{VIX}(X_t;\tau)$ is to properly deal with the square root in the expression of VIX. We adopt the numerical integration method in Appendix A.4. of Eraker and Wu (2017)

$$\begin{aligned}
F^{VIX}(X_t,\tau) &= E_t^Q\big[\sqrt{VIX_{t+\tau}^2}\big] \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1 - E_t^Q[e^{-sVIX_{t+\tau}^2}]}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1 - e^{-a_{1/12}s}E_t^Q[e^{-s(0,c_{1/12},d_{1/12})'X_{t+\tau}}]}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1 - e^{-a_{1/12}s}e^{\alpha(s,\tau)+\beta(s,\tau)'X_t}}{s^{3/2}}ds \\
&= \frac{1}{2\sqrt{\pi}}\int_{-\infty}^\infty e^{-s/2}\big(1 - e^{-a_{1/12}e^s}e^{\alpha(e^s,\tau)+\beta(e^s,\tau)'X_t}\big)ds.
\end{aligned} \tag{C.54}$$

The second equality follows from a mathematical result $\sqrt{x} = \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-e^{-sx}}{s^{3/2}}ds$ and Fubini's theorem to switch expectation and integral. The third equality follows from the expression of equilibrium VIX-squared (35). The fourth equality follows from the definition of the (undiscounted) characteristic function, where $\alpha(s,\tau)$ and $\beta(s,\tau)$ are the solutions (evaluated at τ) to the ODE system (C.43) through (C.46) with boundary conditions $\alpha(0) = 0; \beta(0) = (0, -c_{1/12}s, -d_{1/12}s)'$. The last equality follows from a change of variable to make the integrand bell shaped for easier numerical computation. To see the above mathematical result, consider a normal random variable with mean zero and standard deviation $s/\sqrt{2}$. Obviously, we have

$$1 = \frac{1}{\sqrt{2\pi}\frac{s}{\sqrt{2}}}\int_{-\infty}^{+\infty} exp\Big(-\frac{1}{2}\Big(\frac{t}{\frac{s}{\sqrt{2}}}\Big)^2\Big)dt,$$

or

$$s = \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-(\frac{t}{s})^2} dt,$$

or

$$\begin{aligned}
\sqrt{x} &= \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-\frac{t^2}{x}} dt \\
&= \frac{2}{\sqrt{\pi}} \int_0^{+\infty} \frac{x}{2} e^{-xs} s^{-\frac{1}{2}} ds \\
&= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} s^{-\frac{1}{2}} d(1 - e^{-xs}) \\
&= \frac{1}{\sqrt{\pi}} s^{-\frac{1}{2}} (1 - e^{-xs}) \Big|_{s=0}^{+\infty} - \frac{1}{\sqrt{\pi}} \int_0^{+\infty} (1 - e^{-xs})(-\frac{1}{2} s^{-\frac{3}{2}}) ds \\
&= \frac{1}{2\sqrt{\pi}} \int_0^{\infty} \frac{1 - e^{-xs}}{s^{3/2}} ds,
\end{aligned}$$

where the second line follows from a change of variable $t = x\sqrt{s}$, and the fourth an integral by parts.

C.VII SPX Option Pricing

Before pricing VIX options, we first price a (European) SPX option, which value in the current model is homogenous of degree one in SPX. To facilitate computation, we divide the standard no-arbitrage option pricing equation through by SPX for a normalization. Specifically, let $P^{SPX}(X_t, \tau, K)$ denote the normalized price of an SPX put option with maturity τ and normalized strike K. No-arbitrage then implies

$$P^{SPX}(X_t, \tau, K) = E_t^Q \Big[e^{-\int_t^{t+\tau} r_u du} \Big(K - P_{t+\tau}/P_t \Big)^+ \Big] \tag{C.55}$$

$$= E_t^Q \Big[e^{-\int_t^{t+\tau} r_u du} \Big(K - e^{\ln P_{t+\tau} - \ln P_t} \Big)^+ \Big]. \tag{C.56}$$

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

$$P^{SPX}(X_t, \tau, K) = E_t^Q \Big[e^{-\int_t^{t+\tau} r_u du} \Big(K - e^{\ln P_{t+\tau} - \ln P_t} \Big)^+ \Big] \tag{C.57}$$

$$= \frac{1}{2\pi} E_t^Q \Big[\int_{iz_i - \infty}^{iz_i + \infty} e^{-\int_t^{t+\tau} r_u du} e^{-iz(\ln P_{t+\tau} - \ln P_t)} \hat{\omega}(z) dz \Big], \tag{C.58}$$

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

$$\hat{\omega}(z) \equiv \int_{-\infty}^{+\infty} e^{izx}(K - e^x)^+ dx \tag{C.59}$$

$$= -\frac{K^{iz+1}}{z^2 - iz}, \tag{C.60}$$

for $z_i \equiv Im(z) < 0$ (we restrict the imaginary part of z to be negative because, as one can easily verify, the integral in (C.59) exists if and only if $Im(z) < 0$). Then, taking the expectation operator inside the integral in (C.58) yields

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

$$= -\frac{1}{2\pi} \int_{iz_i - \infty}^{iz_i + \infty} e^{iz(\ln \tilde{P}_t - \mu_D \tau) - \frac{1}{2}\sigma_D^2 z^2 \tau} \varrho_X^Q\big(- iz(\phi, g_1^*, g_2^*)', X_t, \tau\big) \frac{K^{iz+1}}{z^2 - iz} dz, \tag{C.62}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) < 0$. The second line follows from the definition of the complex-valued discounted characteristic function given in (27), equation (33), and the definition of $\ln \tilde{P}_t$ given in (C.36). For $z_i \equiv Im(z) > 1$, (C.62) is also the pricing formula for an SPX call option with maturity τ and normalized strike K, which follows from the well known fact that the generalized Fourier transform of the payoff function of SPX call option $(e^x - K)^+$ is also given by (C.60) but with $z_i \equiv Im(z) > 1$ (See Lewis (2001)). Using integral variable substitution $x = z - z_i i$, we obtain a numerically implementable pricing formula as

$$-\frac{1}{2\pi} \int_{-\infty}^{+\infty} Re\Big[e^{-(z_i - xi)(\ln \tilde{P}_t - \mu_D \tau) - \frac{1}{2}\sigma_D^2 (x + iz_i)^2 \tau} \varrho_X^Q\Big((z_i - xi)(\phi, g_1^*, g_2^*)', X_t, \tau\Big) \frac{K^{1 - z_i + xi}}{(x + z_i i)(x + (z_i - 1)i)} \Big] dx, \tag{C.63}$$

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

C.VIII VIX Call Option Pricing

Unlike an SPX option where normalization is convenient, it is more convenient to directly compute the VIX option price without normalization since there is no homogeneity property. No-arbitrage implies that the price of a VIX call option with maturity τ and strike K is given by

$$C^{VIX}(X_t, \tau, K) = E_t^Q \Big[e^{-\int_t^{t+\tau} r_u du} \Big(VIX_{t+\tau} - K \Big)^+ \Big] \tag{C.64}$$

$$= E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(\sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}} - K\Big)^+\Big]. \tag{C.65}$$

Using the Parseval identity, we obtain

$$C^{VIX}(X_t, \tau, K) = E_t^Q\Big[e^{-\int_t^{t+\tau} r_u du}\Big(\sqrt{a_{1/12} + c_{1/12}\sigma_{t+\tau}^2 + d_{1/12}\lambda_{t+\tau}} - K\Big)^+\Big] \tag{C.66}$$

$$= \frac{1}{2\pi} E_t^Q\Big[\int_{iz_i-\infty}^{iz_i+\infty} e^{-\int_t^{t+\tau} r_u du} e^{-iz(a_{1/12}+c_{1/12}\sigma_{t+\tau}^2+d_{1/12}\lambda_{t+\tau})}\hat{\omega}(z)dz\Big], \tag{C.67}$$

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

$$\begin{aligned}
\hat{\omega}(z) &\equiv \int_{-\infty}^{+\infty} e^{izx}(\sqrt{x}-K)^+ dx \\
&= \int_{K^2}^{+\infty} e^{izx}(\sqrt{x}-K)dx \\
&= \int_{K^2}^{+\infty} e^{izx}\sqrt{x}dx - K\int_{K^2}^{+\infty} e^{izx}dx \\
&= \frac{1}{iz}\int_{K^2}^{+\infty} \sqrt{x}de^{izx} - \frac{K}{iz}e^{izx}\Big|_{x=K^2}^{+\infty} \\
&= \frac{1}{iz}\sqrt{x}e^{izx}\Big|_{x=K^2}^{+\infty} - \frac{1}{iz}\int_{K^2}^{+\infty} e^{izx}d\sqrt{x} - \frac{K}{iz}e^{izx}\Big|_{x=K^2}^{+\infty},
\end{aligned}$$

where in the last line, we have used integral by parts. Functions in the first and third terms evaluated at $+\infty$ are defined (and equal to 0) if and only if $z_i \equiv Im(z) > 0$. In this case, the two terms cancel out. Therefore,

$$\begin{aligned}
\hat{\omega}(z) &= -\frac{1}{iz}\int_{K^2}^{+\infty} e^{izx}d\sqrt{x} \\
&= \frac{1}{(-iz)^{3/2}}\int_{\sqrt{-iz}K}^{+\infty} e^{-u^2}du \\
&= \frac{1}{(-iz)^{3/2}}\Big(\frac{\pi}{2} - \int_0^{\sqrt{-iz}K} e^{-u^2}du\Big) \\
&= \frac{\sqrt{\pi}Ercf\big(K\sqrt{-iz}\big)}{2(-iz)^{\frac{3}{2}}}
\end{aligned}$$

where the second line follows from a change of variable $\sqrt{-iz}\sqrt{x} = u$, the third from a standard result $\int_0^{+\infty} e^{-u^2}du = \sqrt{\pi}/2$, and the fourth from the definition of the complex-valued complementary error function, $Ercf(\cdot)$, with an expression given by

$$Ercf(z) = 1 - \frac{2}{\sqrt{\pi}}\int_0^z e^{-u^2}du, \tag{C.68}$$

for any complex number z. Then, taking the expectation operator inside the integral in (C.67) and using the definition of the discounted characteristic function under risk-neutral measure as defined in (27), we can rewrite the VIX call price as

$$C^{VIX}(X_t,\tau,K)=\frac{1}{4\sqrt{\pi}}\int_{iz_i-\infty}^{iz_i+\infty}e^{-iza_{1/12}}\varrho_X^Q\Big(-iz(0,c_{1/12},d_{1/12})',X_t,\tau\Big)\frac{Ercf\big(K\sqrt{-iz}\big)}{(-iz)^{\frac{3}{2}}}dz, \quad \text{(C.69)}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i\equiv Im(z)>0$. Using integral variable substitution $x=z-z_ii$, we obtain a numerically implementable pricing formula as

$$\frac{1}{4\sqrt{\pi}}\int_{-\infty}^{+\infty}Re\Big[e^{(z_i-xi)a_{1/12}}\varrho_X^Q\Big((z_i-xi)(0,c_{1/12},d_{1/12})',X_t,\tau\Big)\frac{Ercf(K\sqrt{z_i-xi})}{(z_i-xi)^{\frac{3}{2}}}\Big]dx, \quad \text{(C.70)}$$

where we only need to take the real part of the complex-valued integrand because the VIX call price is theoretically guaranteed to be real.

C.IX VIX Put Option Pricing

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

$$\begin{aligned}\hat{\omega}(z)&\equiv\int_{-\infty}^{+\infty}e^{izx}(K-\sqrt{x})^+dx\\&=\int_{-\infty}^{K^2}e^{izx}Kdx-\int_0^{K^2}e^{izx}\sqrt{x}dx\\&=\frac{K}{iz}e^{izx}\Big|_{x=-\infty}^{K^2}-\frac{1}{iz}\int_0^{K^2}\sqrt{x}de^{izx}\\&=\frac{K}{iz}e^{izx}\Big|_{x=-\infty}^{K^2}-\frac{1}{iz}\sqrt{x}e^{izx}\Big|_{x=0}^{K^2}+\frac{1}{iz}\int_0^{K^2}e^{izx}d\sqrt{x}.\end{aligned}$$

Note we have applied a trick in the second line. We should have set both integral lower limits to zero, but did not. The second line is equivalent to transforming a continuous payoff function equal to K for all $x<0$. If the first integral also had a lower limit of zero, then the payoff function we are transforming is discontinuous at zero, i.e., it is equal to zero for all $x<0$ and K for $x=0$. This little twist on payoff function has no economic consequence since VIX-squared in equilibrium is always positive, but makes the transform for VIX put more compact. In the last line, we have used integral

by parts. The function in the first term evaluated at $-\infty$ is defined (and equal to 0) if and only if $z_i \equiv Im(z) < 0$. In this case, the first and second terms cancel out. Therefore,

$$\begin{aligned}\hat{\omega}(z) &= \frac{1}{iz}\int_0^{K^2} e^{izx} d\sqrt{x} \\ &= -\frac{1}{(-iz)^{3/2}}\int_0^{\sqrt{-iz}K} e^{-u^2} du \\ &= -\frac{\sqrt{\pi}}{2}\frac{1}{(-iz)^{3/2}}\Big(1 - Ercf(K\sqrt{-iz})\Big),\end{aligned} \tag{C.71}$$

where the second line follows from a change of variable $\sqrt{-iz}\sqrt{x} = u$, and the third from the definition of the $Ercf(\cdot)$ function. To derive the VIX put pricing formula, we insert equation (C.71) into (C.67), take the expectation operator inside the integral, and use the definition of the discounted characteristic function under risk-neutral measure as defined in (27), which gives

$$P^{VIX}(X_t, \tau, K) = -\frac{1}{4\sqrt{\pi}}\int_{iz_i-\infty}^{iz_i+\infty} e^{-iza_{1/12}} \varrho_X^Q\Big(- iz(0, c_{1/12}, d_{1/12})', X_t, \tau\Big)\frac{1 - Ercf(K\sqrt{-iz})}{(-iz)^{3/2}} dz, \tag{C.72}$$

where the integration is performed on any a strip parallel to the real axis in the complex z plane for which $z_i \equiv Im(z) < 0$. Using integral variable substitution $x = z - z_i i$, we obtain a numerically implementable pricing formula as

$$-\frac{1}{4\sqrt{\pi}}\int_{-\infty}^{+\infty} Re\Big[e^{(z_i - xi)a_{1/12}} \varrho_X^Q\Big((z_i - xi)(0, c_{1/12}, d_{1/12})', X_t, \tau\Big)\frac{1 - Ercf(K\sqrt{z_i - xi})}{(z_i - xi)^{\frac{3}{2}}}\Big] dx, \tag{C.73}$$

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

C.X Black-Scholes (1973) Implied Volatility

We are interested in the implied volatilities for SPX put options. The underlying asset of an SPX put option is the SPX. Therefore, we should resort to the option pricing formula in Black and Scholes (1973). Consistent with our SPX put pricing, we consider a normalized version of Black and Scholes (1973). Specifically, consider an SPX put option with normalized strike K, and time to maturity τ. Let the corresponding continuously compounded bond yield be $r_t(\tau)$ and dividend yield be q_t. Under the assumptions of Black and Scholes (1973), the normalized price of the SPX put option should be given by

$$BSP(1, K, \tau, r_t(\tau), q_t, \sigma) = e^{-r_t(\tau)\tau} K \cdot N(-d_2) - e^{-q_t\tau} \cdot N(-d_1), \tag{C.74}$$

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

$$d_1 = \frac{\ln(1/K) + (r_t(\tau) - q_t + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} \tag{C.75}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{C.76}$$

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

$$\begin{aligned} P^{SPX}(X_t, \tau, K) &= BSP(1, K, \tau, r_t(\tau), q_t, \sigma_t^{imp}) \\ &= BSP(1, K, \tau, r(X_t, \tau), q(X_t), \sigma^{imp}(X_t, \tau, K)), \end{aligned} \tag{C.77}$$

where $P^{SPX}(X_t, \tau, K)$ is given by (C.62), $r(X_t, \tau) = -\ln\left(\varrho_X^Q(0, X_t, \tau)\right)/\tau$ by the definition of the discounted characteristic function, and $q(X_t) = \ln(1 + 1/G(X_t))$ with the price-dividend ratio $G(X_t)$ given by (28).

C.XI Black-76 Implied Volatility

We are also interested in the implied volatilities for VIX call options. The underlying asset of a VIX option is not the VIX index, but instead a VIX futures contract with the same maturity as the option. Therefore, we should resort to the futures option pricing formula in Black (1976). Specifically, consider a VIX call option with strike K, time to maturity τ, and underlying price $F_t^{VIX}(\tau)$. Let the corresponding continuously compounded bond yield be $r_t(\tau)$. Under the assumptions of Black (1976), the price of the VIX call option should be given by[27]

$$BC(F_t^{VIX}(\tau), K, \tau, r_t(\tau), \sigma) = e^{-r_t(\tau)\tau}\left[F_t^{VIX}(\tau) \cdot N(d_1) - K \cdot N(d_2)\right], \tag{C.78}$$

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

$$d_1 = \frac{\ln(F_t^{VIX}(\tau)/K) + \sigma^2\tau/2}{\sigma\sqrt{\tau}} \tag{C.79}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{C.80}$$

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

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

$$\begin{aligned} C^{VIX}(X_t, \tau, K) &= BC(F_t^{VIX}(\tau), K, \tau, r_t(\tau), \sigma_t^{imp}) \\ &= BC(F^{VIX}(X_t, \tau), K, \tau, r(X_t, \tau), \sigma^{imp}(X_t, \tau, K)), \end{aligned} \tag{C.81}$$

where $C^{VIX}(X_t, \tau, K)$ is given by (C.69), $r(X_t, \tau) = -\ln\left(\varrho_X^Q(0, X_t, \tau)\right)/\tau$ by the definition of the discounted characteristic function, and $F^{VIX}(X_t, \tau)$ is given by (C.54).

Appendix D. Additional Model Results

D.I Additional Model Moments

We report additional model moments. Table 5 in the body text shows that our model can replicate the first and second-order moments of consumption and dividends. Panel A of Table D.1 further compares higher-order moments of consumption and dividend growth rates in the model and the data. As shown, the introduction of jumps in endowment volatility does not affect odd moments of fundamentals. Both skewnesses are nearly zero simply due to conditional log-normality. Model-implied kurtosis of consumption growth seems greater than that in the data - the presence of jumps in volatility thickens the tail of consumption growth distribution. However, the difference between model-implied and data kurtosis is statistically insignificant. Model-implied kurtosis of dividend growth seems to undershoot its data counterpart - because of the presence of idiosyncratic Brownian risk in dividend growth, jumps in volatility do not thicken the tail of dividend growth distribution as much.

Panel B of Table D.1 compares moments of the conditional variance risk premium in the model and the data. To ensure robustness, we define the conditional variance risk premium in two ways: $VRP_t = Var_t^Q[\ln P_{t+1/12}] - Var_t^P[\ln P_{t+1/12}]$, and $VRP_t = E_t^Q[QV_{t,t+1/12}] - E_t^P[QV_{t,t+1/12}]$, where $QV_{t,t+1/12} = \int_t^{t+1/12}[d\ln P, d\ln P]_s$ is the total quadratic variation of log market return between t and $t + 1/12$. It turns out our model does very well on this dimension, although we never explicitly target the variance risk premium in our parameter calibration.

D.II Price-Dividend Ratio in a LRR-Augmented VIX Model

The shortcoming concerning the price-dividend ratio shown in Table 5 can be straightforwardly addressed by introducing long-run risks into our benchmark model without causing other moments to

suffer. Following Bansal and Yaron (2004), we assume the conditional expected consumption growth x_t is a time-varying, persistent process, and keep all other aspects of our benchmark model unchanged. Specifically, log consumption, expected consumption growth, consumption growth volatility, volatility jump arrival intensity, and log dividend are respectively given by

$$\begin{aligned} d\ln C_t &= (\mu + x_t - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t^C \\ dx_t &= -\kappa^x x_t dt + \sigma_x \sigma_t dB_t^x \\ d\sigma_t^2 &= \kappa^V(\theta^V - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t \\ d\lambda_t &= \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda \\ d\ln D_t &= \phi d\ln C_t + \mu_D dt + \sigma_D dB_t^D, \end{aligned} \tag{D.1}$$

where κ^x is a small number. In equilibrium, long-run productivity x_t heavily drives the price-dividend ratio, and thus a persistent x_t process translates into a persistent price-dividend ratio. In order to match the moments overall well, we recalibrate the parameters. For example, equity premium would be excessively high if we introduce long-run risks without recalibrating other parameters. Table D.2 reports our recalibrated parameters. Compared with benchmark calibration, we have let risk aversion be slightly lower and σ_t^2 be slightly less persistent. Table D.3 shows that introducing long-run risks brings the volatility and autocorrelation of the log price-dividend ratio substantially closer to the data, with none of the other moments severely compromised. It is worth noting that VIX still has a two-factor (σ_t^2, λ_t) representation even with long-run risks, since VIX in equilibrium reflects only the effects of the higher-order moments of the fundamental. Figures D.1 and D.2 confirm that the new set of parameters does not compromise our model's main results: steady-state and low and high-VIX conditional VIX option implied volatility curves.

D.III Return Predictability in the Benchmark Model

Table D.4 reports predictive regression results at both short and long horizons. Panel A shows that our model generates sufficient excess return predictability by the log price-dividend ratio at relatively short horizons, that is, less than two years. However, it undershoots this predictability at longer horizons. Unsurprisingly, this is related to the fact that the log price-dividend ratio does not carry sufficiently persistent information about the equity premium far into the future. Panel B shows that our model implies little consumption growth predictability by the log price-dividend ratio across all horizons, consistent with the data.

Finally, Panel C confirms our model's ability to generate sufficient excess return predictability by the variance risk premium at relatively short horizons. The model-implied slope coefficients are somehow higher than their data counterparts reported in Bollerslev, Tauchen, and Zhou (2009), but closer to those reported in Drechsler and Yaron (2011) which re-estimate the variance risk premium using $S\&P$ 500 futures data to replace cash index data. Our model-implied R^2s are on par with the data at short horizons, and slightly higher than the data at longer horizons.

D.IV Return Predictability in a Three-Factor VIX Model

The insufficient excess return predictability by the log price-dividend ratio at longer horizons (Table D.4 Panel A) can be partially fixed in an extended three-factor model to include a slow-moving component of the volatility process. Following Drechsler and Yaron (2011), we assume the long-run average of variance without jumps θ_t is a time-varying, persistent process, and keep all other aspects of our benchmark model unchanged. Specifically, log consumption, consumption growth transitory volatility, consumption growth long-run volatility, transitory volatility jump arrival intensity, and log dividend are respectively given by

$$\begin{aligned} d\ln C_t &= (\mu - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t^C \\ d\sigma_t^2 &= \kappa^V(\theta_t - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t \\ d\theta_t &= \kappa^\theta(\theta^V - \theta_t)dt + \sigma_\theta \sqrt{\theta_t} dB_t^\theta \\ d\lambda_t &= \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda \\ d\ln D_t &= \phi d\ln C_t + \mu_D dt + \sigma_D dB_t^D, \end{aligned} \tag{D.2}$$

where κ_θ is a small number. In equilibrium, both the conditional equity premium and the price-dividend ratio would be heavily driven by θ_t, investors' long-run volatility expectation. To the extent θ_t is persistent, the price-dividend ratio would possess long-term predictive power for excess market returns. In order to match the moments overall well, we recalibrate the parameters, which are reported in Table D.5. Table D.6 shows that the three-factor extension not only leaves all the other moments unharmed, but also (partially) helps with the volatility and autocorrelation of the log price-dividend ratio. Table D.7 Panel A shows that long-term return predictability by the log price-dividend ratio is fixed to almost the same extent as in Drechsler and Yaron (2011). Panel C further shows that high-frequency return predictability by the variance risk premium is nearly unaffected.

Importantly, this extension does not compromise the model's ability to explain VIX options data. Intuitively, although VIX now has a three-factor (θ_t, σ_t^2, λ_t) representation, it still loads in an important way on the two relatively fast-moving factors, σ_t^2 and λ_t. Properties of VIX derivatives in our benchmark model are thus maintained. Implied volatility curves in Figures D.3 and D.4 confirm this point. Considering all these aspects, we think the three-factor model has better explanatory power for low-frequency and high-frequency data overall.[28] In contrast, our benchmark model provides the most parsimonious framework to explain all the high-frequency VIX derivatives data. Finally, we emphasize that all the various model extensions considered in Appendix D fall within the description of our general model.

D.V Precise Definition of VIX

As shown in Martin (2011) and Schneider and Trojani (2019b), the precise definition of VIX-squared implied by its option-replicating portfolio is given by

$$VIX_t^2 = -2E_t^Q\left[\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]}\right] \tag{D.3}$$

$$= -2E_t^Q\left[\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]} - \left(\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]} - 1\right)\right] \tag{D.4}$$

$$= -2E_t^Q\left[\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]} - \left(\exp\left(\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]}\right) - 1\right)\right] \tag{D.5}$$

$$= 2E_t^Q\left[\frac{1}{2!}\left(\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]}\right)^2 + \frac{1}{3!}\left(\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]}\right)^3 + \frac{1}{4!}\left(\ln\frac{P_{t+1/12}}{E_t^Q\left[P_{t+1/12}\right]}\right)^4 + ...\right], \tag{D.6}$$

where the second line follows from subtracting a zero-value term, the third an identity, and the fourth a Taylor expansion of the exponential function around zero. Hence, VIX-squared involves the second as well as all higher-order risk-neutral cumulants for log market return. The expression we use in the body text (equation (32)) actually only picks up the second-order cumulant term.

Within our exponential affine framework, equation (D.3) can be computed under the help of the risk-neutral cumulant generating function for log SPX (first term) and the undiscounted characteristic function (second term). We find that $VIX_t = \sqrt{0.0159 + 12.504\sigma_t^2 + 0.0207\lambda_t}$ under the original definition, and $VIX_t = \sqrt{0.0154 + 12.299\sigma_t^2 + 0.0185\lambda_t}$ under the precise definition. The precise definition leads to three differences. First, VIX is smaller, because the possibility of volatility up-

[28] We thank an anonymous referee for suggesting the three-factor model specification.

jumps, through the "volatility feedback effect," implies log market return is negatively skewed, i.e., a negative third-order cumulant. Second, the loading on λ_t decreases the most, as conditional skewness is governed mainly by the jump arrival intensity λ_t. Third, the changes in moments of VIX and VIX option prices are negligibly small, as shown in Tables D.8 and D.9. That the third and higher-order cumulant terms are much smaller than the second-order term is not surprising since our model abstracts from jumps in consumption which can make log market return heavily negatively skewed (Wachter (2013)).

D.VI Schneider and Trojani (2019)'s NCC

Schneider and Trojani (2019a) show that physical conditional moments of market returns can be bounded below if certain NCC (negative covariance condition) is satisfied. Specifically, their Proposition 3 shows that the inequality

$$E_t^P[R_T^n] \geq \frac{E_t^Q[R_T^{p+n}]}{E_t^Q[R_T^p]} \tag{D.7}$$

holds if the following $NCC(p,n)$ is satisfied for $p \in [0,1]$ and $p+n>1$:

$$Cov_t[M_T R_T^p, R_T^n] \leq 0, \tag{D.8}$$

where R_T is gross market return over $[t,T]$, and M_T corresponding pricing kernel. Moreover, for each n, the lower bound increases with p such that the tightest lower bound always obtains when $p=1$. In particular, for $n=1$ and $p=1$, the NCC reduces to Martin (2017)'s NCC, and the inequality reduces to a lower bound for conditional equity premium:

$$E_t^P[R_T] - R_{f,t} \geq \frac{Var_t^Q[R_T]}{R_{f,t}}. \tag{D.9}$$

We derive the parameter space (γ, ϕ) such that an instantaneous ($T = t+dt$) version of Schneider and Trojani (2019a)'s NCC for all $n \geq 1$ and $p=1$ is satisfied in any state of our model, fixing all the other parameters at their calibrated values. We do so because, as we analyze shortly, γ and ϕ are two key parameters that affect the NCC. We want

$$Cov_t\left[\Big(\frac{d\pi_t}{\pi_{t^-}}+1\Big)\Big(\frac{dP_t+D_{t^-}dt}{P_{t^-}}+1\Big), \Big(\frac{dP_t+D_{t^-}dt}{P_{t^-}}+1\Big)^n\right] \leq 0 \tag{D.10}$$

$$\Leftrightarrow \quad Cov_t\left[\frac{d\pi_t}{\pi_{t^-}}\frac{dP_t}{P_{t^-}} + \frac{d\pi_t}{\pi_{t^-}} + \frac{dP_t}{P_{t^-}}, n\frac{dP_t}{P_{t^-}}\right] \leq 0 \tag{D.11}$$

$$\Leftrightarrow \quad Cov_t\Bigg[\Big(-\Lambda_t' dB_t + \big(e^{b_2\xi_V}-1\big)dN_t\Big)\Big(\sigma_{P,t}' dB_t + \big[e^{g_1^*\xi_V}-1\big]dN_t\Big) - \Lambda_t' dB_t + \big(e^{b_2\xi_V}-1\big)dN_t$$

$$+\sigma_{P,t}' dB_t + \big[e^{g_1^*\xi_V}-1\big]dN_t, \sigma_{P,t}' dB_t + \big[e^{g_1^*\xi_V}-1\big]dN_t\Bigg] \leq 0 \tag{D.12}$$

$$\Leftrightarrow \quad Cov_t\Bigg[(\sigma_{P,t}-\Lambda_t)' dB_t + \big[e^{(b_2+g_1^*)\xi_V}-1\big]dN_t, \sigma_{P,t}' dB_t + \big[e^{g_1^*\xi_V}-1\big]dN_t\Bigg] \leq 0 \tag{D.13}$$

$$\Leftrightarrow \quad \Big[(\phi-\gamma)\phi + (b_2+g_1^*)g_1^*\sigma_V^2\Big]\sigma_t^2 \tag{D.14}$$

$$+\Big[(b_3+g_2^*)g_2^*\sigma_\lambda^2 + \frac{1}{1-\mu_\xi(b_2+2g_1^*)} - \frac{1}{1-\mu_\xi(b_2+g_1^*)} - \frac{1}{1-\mu_\xi g_1^*} + 1\Big]\lambda_t \leq 0, \tag{D.15}$$

where the second line obtains because: first, we can discard the term $D_{t^-}/P_{t^-}dt$ (which eventually surely generates higher-order terms than dt); second, we can then discard all the second and higher-order terms of dP_{t^-}/P_{t^-} in the binomial expansion whose values are negligible compared with the first-order term when dt is small. This means for short horizons NCCs are essentially the same for all $n \geq 1, p = 1$. The third and fourth lines follow from equations (C.14), (C.33) and (33), the fifth from some algebra and removing higher-order terms than dt, and the sixth and seventh from some algebra, removing higher-order terms than dt, and equations (C.15) and (C.32). In order for the last inequality to hold regardless of σ_t^2 and λ_t, both premultiplying terms have to be non-positive. This leads to the parameter space shown in Figure D.5.

We find Schneider and Trojani (2019a)'s NCC holds as long as γ exceeds ϕ by a small margin. This is of course true under our current calibration ($\gamma = 14, \phi = 2.7$). Intuitively, NCC is a condition that requires the product of the pricing kernel and market return, $M_T R_T$, and market return, R_T, being negatively correlated. In the model, state shocks drive the movements in M_T and R_T. On the one hand, increasing γ has two effects: first, it increases (absolute) market prices of risks b_2 and b_3, the pricing kernel's exposure to state shocks; second, it increases market index return's (absolute) exposure to state shocks, $|g_1^*|$ and $|g_2^*|$. On the other, increasing ϕ only increases market index return's exposure to state shocks, $|g_1^*|$ and $|g_2^*|$, but does not affect the pricing kernel. Thus, when ϕ is high, the correlation tends to be dominated by the movement in R_T and tends to be positive. When γ is high, the correlation tends to be shaped by the negative correlation between M_T and R_T and tends to be negative. Example 4a of Martin (2017) derives the NCC in a dividend-based (implicitly $\phi = 1$) equilibrium asset pricing model with Epstein-Zin recursive preferences with $\psi \geq 1$: $\gamma \geq 1$. So our NCC in an otherwise identical but more general consumption-based model, $\gamma > \phi$, is consistent with Martin (2017)'s result. Martin (2017) and Schneider and Trojani (2019a) suggest that in structural models the NCC is sensitive to preference parameters. We show that, in addition, the stock market leverage parameter is also important.

Table D.1: Additional Model Moments

The table reports additional model moments. Panel A reports higher-order moments of log consumption growth and log dividend growth, both in the benchmark model and the data. We use real consumption and dividend data, which both are from 1930 to 2020. Panel B reports moments of conditional variance risk premium, both in the benchmark model and the data. We define the conditional variance risk premium in two ways. The first set of numbers are based on the definition: $VRP_t = Var_t^Q[\ln P_{t+1/12}] - Var_t^P[\ln P_{t+1/12}]$. The second set of numbers are based on the definition: $VRP_t = E_t^Q[QV_{t,t+1/12}] - E_t^P[QV_{t,t+1/12}]$, where $QV_{t,t+1/12} = \int_t^{t+1/12}[d\ln P, d\ln P]_s$ is the total quadratic variation of log market return between t and $t+1/12$. Variance risk premium data moments are from Drechsler and Yaron (2011), Table 3 last column. The variance risk premium is monthly in basis points. The model-implied moments are the 5%, 50%, and 95% quantile values from 1,000 simulations with the same length as the data sample.

	Model			U.S. Data	Data Source
	5%	50%	95%		
Panel A: higher-order fundamental					
skewness(Δc_t)	-2.19	-0.10	1.80	-1.08	BEA Table
kurtosis(Δc_t)	4.66	7.55	17.31	4.84	BEA Table
skewness(Δd_t)	-1.00	-0.02	0.84	-0.88	CRSP
kurtosis(Δd_t)	2.72	3.96	8.24	8.67	CRSP
Panel B: variance risk premium					
mean	9.04	10.68	13.35	11.27	DY2011
std. dev.	4.64	6.73	10.83	7.61	DY2011
median	7.69	8.75	10.35	8.92	DY2011
min	2.88	3.07	3.36	3.27	DY2011
skewness	1.14	1.80	3.03	2.39	DY2011
kurtosis	3.96	6.99	15.69	12.03	DY2011
$AC(1)$	0.39	0.61	0.80	0.65	DY2011
mean	6.28	7.00	7.83	11.27	DY2011
std. dev.	2.81	3.44	4.33	7.61	DY2011
median	5.51	6.12	6.92	8.92	DY2011
min	2.28	2.40	2.60	3.27	DY2011
skewness	0.89	1.34	2.07	2.39	DY2011
kurtosis	3.34	5.15	9.66	12.03	DY2011
$AC(1)$	0.29	0.46	0.62	0.65	DY2011

Table D.2: Parameters for the LRR-Augmented VIX Model

The table reports parameter values for the extended VIX model with long-run risks. The processes for log consumption, expected consumption growth, consumption growth volatility, volatility jump arrival intensity, and log dividend are respectively given by

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

$$dx_t = -\kappa^x x_t dt + \sigma_x \sigma_t dB_t^x$$
$$d\sigma_t^2 = \kappa^V(\theta^V - \sigma_t^2)dt + \sigma_V \sigma_t dB_t^V + \xi_V dN_t$$
$$d\lambda_t = \kappa^\lambda(\theta^\lambda - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_t^\lambda$$
$$d\ln D_t = \phi d\ln C_t + \mu_D dt + \sigma_D dB_t^D,$$

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

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

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

In equilibrium, VIX has a two-factor representation: $VIX_t = \sqrt{a_{1/12} + c_{1/12}\sigma_t^2 + d_{1/12}\lambda_t}$.

Parameters values are interpreted in annual terms.	
Rate of time preference β	0.01
Relative risk aversion γ	11
Average growth in consumption μ	0.025
Mean reversion of expected consumption growth κ^x	0.16
Diffusion scale parameter of expected consumption growth σ_x	0.3
Mean reversion of volatility κ^V	4
Average consumption growth variance without jumps θ^V	0.0005
Diffusion scale parameter of consumption growth volatility σ_V	0.14
Average volatility jump size μ_ξ	0.003
Mean reversion of jump arrival intensity κ^λ	10
Average jump arrival intensity θ^λ	0.8
Diffusion scale parameter of jump arrival intensity σ_λ	3
Stock market leverage ϕ	2.5
Adjustment in dividend growth drift μ_D	0.00
Idiosyncratic risk in dividend growth σ_D	0.11

Table D.3: Selected Moments in the LRR-Augmented VIX Model

The table reports a list of moments in the extended VIX model with long-run risks and their counterparts in the U.S. data. The model is simulated at a monthly frequency (dt=1/12) for 100,000 months 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. $r^{C^{VIX}(\tau=1/12)}$ ($r^{P^{VIX}(\tau=1/12)}$) denotes net holding-period return on one-month ATM VIX call (put) option.

	Model	U.S. Data	Data Source
$E[\Delta c]$	2.39	1.80	BY2004
$\sigma(\Delta c)$	3.13	2.93	BY2004
$AC_1(\Delta c)$	0.44	0.49	BY2004
$E[\Delta d]$	5.99	4.61	CRSP
$\sigma(\Delta d)$	11.90	11.49	BY2004
$AC_1(\Delta d)$	0.33	0.21	BY2004
$corr(\Delta c, \Delta d)$	0.67	0.59	DY2011
$E[\exp(pd)]$	19.05	26.56	BY2004
$\sigma(pd)$	15.73	29.00	BY2004
$AC_1(pd)$	0.65	0.81	BY2004
$E[r_t^e - r_t^f]$	9.36	8.33	Ken French
$\sigma(r_t^e)$	17.51	18.31	CRSP
$E[r_t^f]$	1.10	0.86	BY2004
$\sigma(r_t^f)$	1.92	0.97	BY2004
$E[VIX_t]$	18.70	19.28	CBOE
$\sigma(VIX_t)$	6.00	7.42	CBOE
$AC_1(VIX_t)$	0.72	0.84	CBOE
$E[imp\_vol_t]$	70.25	68.80	CBOE
$\sigma(imp\_vol_t)$	12.57	14.30	CBOE
$AC_1(imp\_vol_t)$	0.48	0.53	CBOE
$corr(VIX_t, imp\_vol_t)$	0.50	0.48	CBOE
$E[r^{C^{VIX}(\tau=1/12)}]$	-0.23	-0.48	CBOE
$E[r^{P^{VIX}(\tau=1/12)}]$	0.09	0.21	CBOE

Table D.4: Predictive Regressions in the Benchmark VIX Model

This table reports R^2 and slope coefficients from regressing cumulative excess market returns onto lagged log price-dividend ratio and lagged variance risk premium and regressing consumption growth onto lagged log price-dividend ratio, using data from simulating the benchmark model over 100,000 months. The time-t conditional variance risk premium is calculated as: $VRP_t = Var_t^Q[\ln P_{t+1/12}] - Var_t^P[\ln P_{t+1/12}]$. Data regression results are also reported. Data in Panels A and B are from Wachter (2013). Data in Panel C are from Bollerslev, Tauchen, and Zhou (2009); numbers in parentheses in Panel C are from Drechsler and Yaron (2011). In Panel C's regressions, all numbers follow Drechsler and Yaron (2011)'s data measures: returns are annualized in percentage, and variance risk premiums are monthly in basis points.

Panel A: $\sum_{h=1}^{H}(r_{m,t+h} - r_{f,t+h}) = \alpha + \beta_1(p_t - d_t) + \varepsilon_{t+H}$

Horizon	1Y	2Y	4Y	6Y	8Y	10Y
			Model			
β_1	-1.07	-1.19	-1.24	-1.27	-1.27	-1.29
R^2	0.11	0.07	0.04	0.03	0.02	0.02
			Data			
β_1	-0.13	-0.23	-0.33	-0.48	-0.64	-0.86
R^2	0.09	0.17	0.23	0.30	0.38	0.43

Panel B: $\sum_{h=1}^{H}\Delta c_{t+h} = \alpha + \beta_1(p_t - d_t) + \varepsilon_{t+H}$

Horizon	1Y	2Y	4Y	6Y	8Y	10Y
			Model			
β_1	0.01	0.00	-0.01	-0.02	-0.03	-0.02
R^2	0.00	0.00	0.00	0.00	0.00	0.00
			Data			
β_1	-0.00	-0.01	-0.01	-0.01	-0.02	-0.01
R^2	0.00	0.01	0.02	0.02	0.03	0.02

Panel C: $\frac{1}{H}\sum_{h=1}^{H}(r_{m,t+h} - r_{f,t+h}) = \alpha + \beta_1 VRP_t + \varepsilon_{t+H}$

Horizon	1M	2M	3M	6M	9M	12M
			Model			
β_1	1.52	1.37	1.26	0.97	0.76	0.61
R^2	0.03	0.06	0.07	0.09	0.09	0.08
			Data			
β_1	0.47	0.70	0.56	0.36	0.20	0.14
β_1	(0.76)	(1.26)	(0.86)			
R^2	0.01	0.07	0.07	0.05	0.02	0.01
R^2	(0.02)	(0.04)	(0.06)			

Table D.5: Parameters for the Three-Factor VIX Model

The table reports parameter values for the extended three-factor VIX model. The processes for log consumption, consumption growth transitory volatility, consumption growth long-run volatility, transitory volatility jump arrival intensity, and log dividend are respectively given by

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

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

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

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

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

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

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

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

In equilibrium, VIX has a three-factor representation: $VIX_t = \sqrt{a_{1/12} + b_{1/12}\sigma_t^2 + c_{1/12}\theta_t + d_{1/12}\lambda_t}$.

Parameters values are interpreted in annual terms.	
Rate of time preference β	0.01
Relative risk aversion γ	17
Average growth in consumption μ	0.03
Mean reversion of transitory volatility κ^V	3.4
Diffusion scale parameter of transitory volatility σ_V	0.17
Average transitory volatility jump size μ_ξ	0.004
Mean reversion of long-run volatility κ^θ	0.6
Average long-run volatility θ^V	0.0004
Diffusion scale parameter of jump arrival intensity σ_θ	0.03
Mean reversion of jump arrival intensity κ^λ	10
Average jump arrival intensity θ^λ	0.5
Diffusion scale parameter of jump arrival intensity σ_λ	3
Stock market leverage ϕ	3
Adjustment in dividend growth drift μ_D	-0.02
Idiosyncratic risk in dividend growth σ_D	0.1

Table D.6: Selected Moments in the Three-Factor VIX Model

The table reports a list of moments in the three-factor VIX model and their counterparts in the U.S. data. The model is simulated at a monthly frequency (dt=1/12) for 100,000 months 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. $r^{C^{VIX}(\tau=1/12)}$ ($r^{P^{VIX}(\tau=1/12)}$) denotes net holding-period return on one-month ATM VIX call (put) option.

	Model	U.S. Data	Data Source
$E[\Delta c]$	2.91	1.80	BY2004
$\sigma(\Delta c)$	2.63	2.93	BY2004
$AC_1(\Delta c)$	0.25	0.49	BY2004
$E[\Delta d]$	6.73	4.61	CRSP
$\sigma(\Delta d)$	11.44	11.49	BY2004
$AC_1(\Delta d)$	0.25	0.21	BY2004
$corr(\Delta c, \Delta d)$	0.68	0.59	DY2011
$E[\exp(pd)]$	23.44	26.56	BY2004
$\sigma(pd)$	11.68	29.00	BY2004
$AC_1(pd)$	0.27	0.81	BY2004
$E[r_t^e - r_t^f]$	8.97	8.33	Ken French
$\sigma(r_t^e)$	18.32	18.31	CRSP
$E[r_t^f]$	0.84	0.86	BY2004
$\sigma(r_t^f)$	2.34	0.97	BY2004
$E[VIX_t]$	18.92	19.28	CBOE
$\sigma(VIX_t)$	7.48	7.42	CBOE
$AC_1(VIX_t)$	0.80	0.84	CBOE
$E[imp\_vol_t]$	67.81	68.80	CBOE
$\sigma(imp\_vol_t)$	17.30	14.30	CBOE
$AC_1(imp\_vol_t)$	0.55	0.53	CBOE
$corr(VIX_t, imp\_vol_t)$	0.39	0.48	CBOE
$E[r^{C^{VIX}(\tau=1/12)}]$	-0.21	-0.48	CBOE
$E[r^{P^{VIX}(\tau=1/12)}]$	0.08	0.21	CBOE

Table D.7: Predictive Regressions in the Three-Factor VIX Model

This table reports R^2 and slope coefficients from regressing cumulative excess market returns onto lagged log price-dividend ratio and lagged variance risk premium and regressing consumption growth onto lagged log price-dividend ratio, using data from simulating the three-factor model over 100,000 months. The time-t conditional variance risk premium is calculated as: $VRP_t = Var_t^Q[\ln P_{t+1/12}] - Var_t^P[\ln P_{t+1/12}]$. Data regression results are also reported. Data in Panels A and B are from Wachter (2013). Data in Panel C are from Bollerslev, Tauchen, and Zhou (2009); numbers in parentheses in Panel C are from Drechsler and Yaron (2011). In Panel C's regressions, all numbers follow Drechsler and Yaron (2011)'s data measures: returns are annualized in percentage, and variance risk premiums are monthly in basis points.

Panel A: $\sum_{h=1}^{H}(r_{m,t+h} - r_{f,t+h}) = \alpha + \beta_1(p_t - d_t) + \varepsilon_{t+H}$

Horizon	1Y	2Y	4Y	6Y	8Y	10Y
			Model			
β_1	-0.63	-0.90	-1.18	-1.27	-1.32	-1.38
R^2	0.11	0.12	0.11	0.09	0.07	0.06
			Data			
β_1	-0.13	-0.23	-0.33	-0.48	-0.64	-0.86
R^2	0.09	0.17	0.23	0.30	0.38	0.43

Panel B: $\sum_{h=1}^{H}\Delta c_{t+h} = \alpha + \beta_1(p_t - d_t) + \varepsilon_{t+H}$

Horizon	1Y	2Y	4Y	6Y	8Y	10Y
			Model			
β_1	0.00	0.00	-0.00	-0.01	-0.02	-0.03
R^2	0.00	0.00	0.00	0.00	0.00	0.00
			Data			
β_1	-0.00	-0.01	-0.01	-0.01	-0.02	-0.01
R^2	0.00	0.01	0.02	0.02	0.03	0.02

Panel C: $\frac{1}{H}\sum_{h=1}^{H}(r_{m,t+h} - r_{f,t+h}) = \alpha + \beta_1 VRP_t + \varepsilon_{t+H}$

Horizon	1M	2M	3M	6M	9M	12M
			Model			
β_1	1.49	1.33	1.19	0.89	0.64	0.55
R^2	0.03	0.05	0.06	0.07	0.06	0.06
			Data			
β_1	0.47	0.70	0.56	0.36	0.20	0.14
β_1	(0.76)	(1.26)	(0.86)			
R^2	0.01	0.07	0.07	0.05	0.02	0.01
R^2	(0.02)	(0.04)	(0.06)			

Table D.8: Precise Definition of VIX: I

The table reports a list of moments of VIX and VIX options in the benchmark model and their counterparts in the U.S. data. Column "Precise VIX" contains moments with VIX calculated using the precise definition given in equation (D.3). Column "Original VIX" contains moments with VIX calculated using the original definition given in equation (32). imp_vol_t denotes Black-76 implied volatility for one-month ATM VIX option. $r^{C^{VIX}(\tau=1/12)}$ ($r^{P^{VIX}(\tau=1/12)}$) denotes net holding-period return on one-month ATM VIX call (put) option, and so on.

	Precise VIX	Original VIX	U.S. Data	Data Source
$E[VIX_t]$	19.04	19.41	19.28	CBOE
$\sigma(VIX_t)$	7.53	7.51	7.42	CBOE
$AC_1(VIX_t)$	0.80	0.80	0.84	CBOE
$E[imp\_vol_t]$	71.81	71.84	68.80	CBOE
$\sigma(imp\_vol_t)$	13.43	12.64	14.30	CBOE
$AC_1(imp\_vol_t)$	0.50	0.49	0.53	CBOE
$corr(VIX_t, imp\_vol_t)$	0.34	0.32	0.48	CBOE
$E[r^{C^{VIX}(\tau=1/12)}]$	-0.25	-0.24	-0.48	CBOE
$E[r^{C^{VIX}(\tau=6/12)}]$	-0.54	-0.53	-0.59	CBOE
$E[r^{P^{VIX}(\tau=1/12)}]$	0.06	0.06	0.21	CBOE
$E[r^{P^{VIX}(\tau=6/12)}]$	0.16	0.16	-0.02	CBOE

Table D.9: Precise Definition of VIX: II

The table reports the average (annualized) Black-76 implied volatilities for VIX options by maturity and strike in the benchmark model, with VIX calculated using the precise definition given in equation (D.3). Comparison with Table 1 right panel in the body text shows that the difference in the entire implied volatility surface is quantitatively very small.

	Maturity (months)			
	1	2	3	6
Strike				
12	0.66	0.59	0.56	0.55
14	0.63	0.54	0.51	0.47
16	0.72	0.63	0.59	0.51
18	0.81	0.72	0.67	0.58
20	0.91	0.80	0.74	0.63
22	1.00	0.87	0.79	0.67
24	1.08	0.92	0.83	0.69
26	1.14	0.96	0.87	0.71
28	1.20	0.99	0.89	0.72
30	1.24	1.02	0.91	0.73
32	1.27	1.04	0.92	0.74
34	1.30	1.05	0.93	0.74
36	1.33	1.06	0.93	0.74
38	1.35	1.07	0.93	0.74
40	1.36	1.07	0.93	0.73

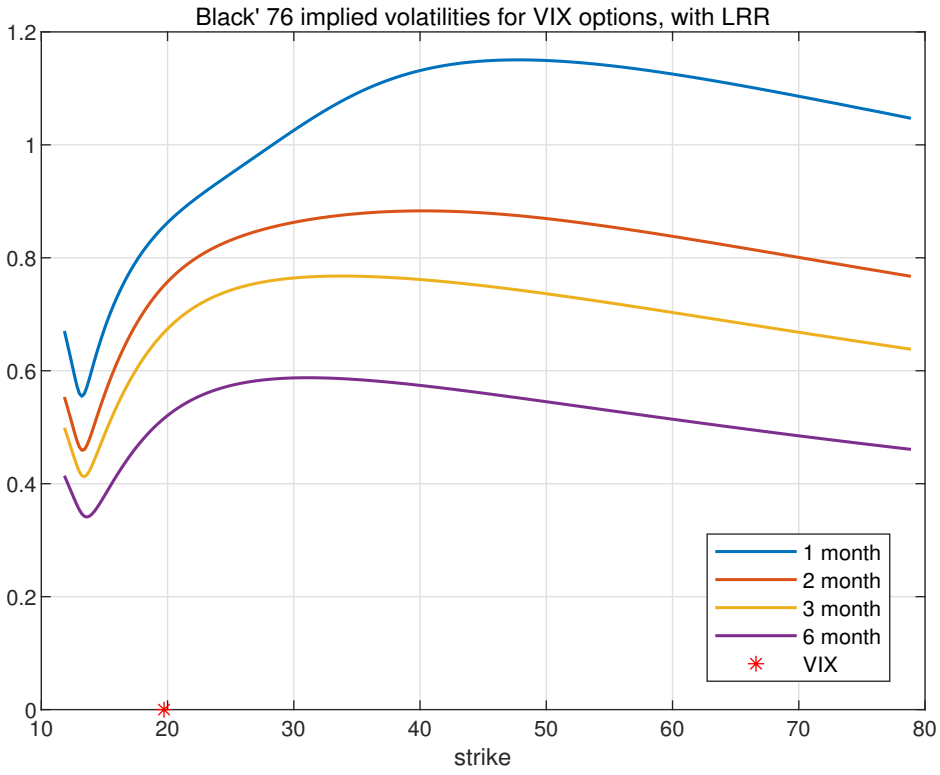


Figure D.1: Black-76 implied volatility curves for VIX options in the LRR-augmented VIX model. This figure plots (annualized) implied volatility curves for VIX options in the LRR-augmented VIX model at steady state. The horizontal axis denotes the absolute value of the strike. Implied volatilities are computed for VIX options with four maturities: 1, 2, 3, and 6 month.

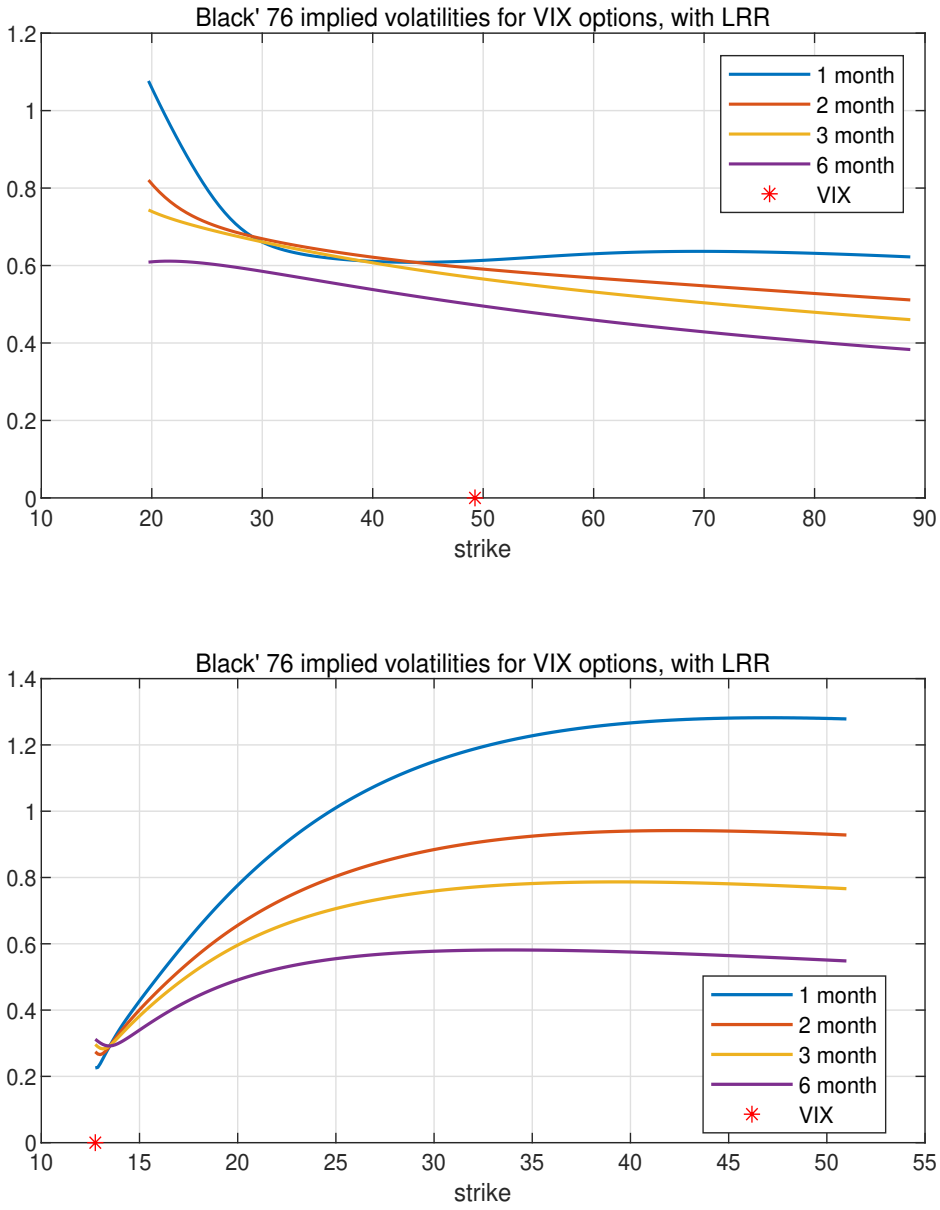


Figure D.2: Black-76 implied volatility curves for VIX options in the LRR-augmented VIX model: conditional analysis. The figure plots (annualized) implied volatility curves for VIX options in the LRR-augmented VIX model, conditional on high and low initial VIX. In the upper case, we set both state variables very high: $\sigma_t^2 = 10\sigma_{ss}^2$ and $\lambda_t = 10\lambda_{ss}$, implying a very high VIX, 49.3. In the lower case, we set both state variables at minimum values: $\sigma_t^2 = \lambda_t = 0$, implying a small value of VIX, 12.8.

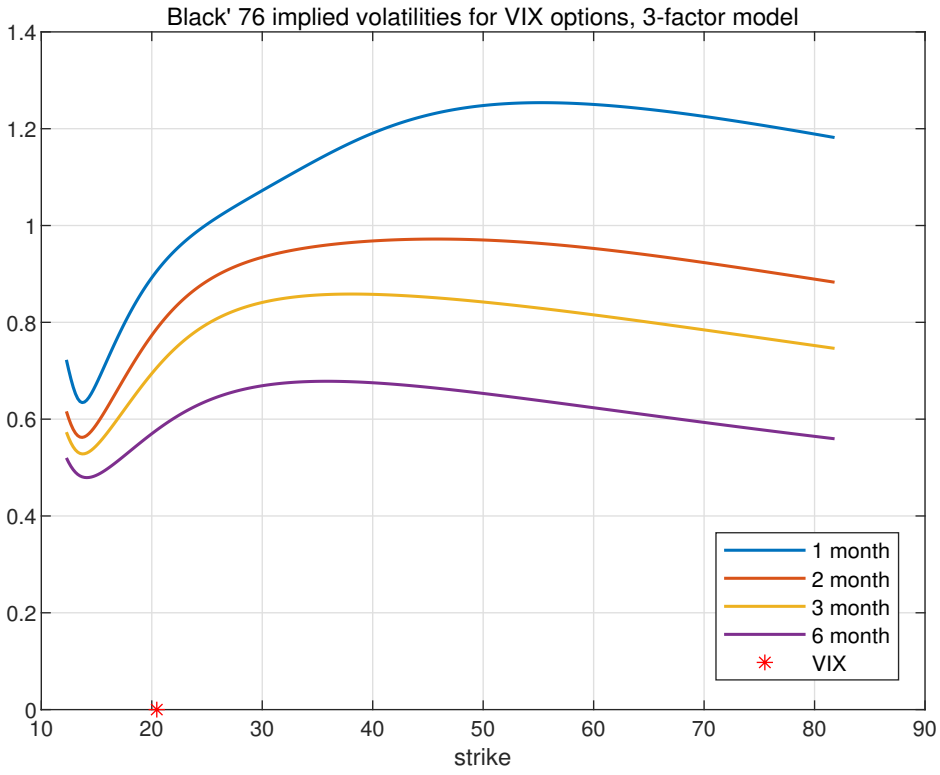


Figure D.3: Black-76 implied volatility curves for VIX options in the three-factor model. This figure plots (annualized) implied volatility curves for VIX options in the three-factor VIX model at steady state. The horizontal axis denotes the absolute value of the strike. Implied volatilities are computed for VIX options with four maturities: 1, 2, 3, and 6 month.

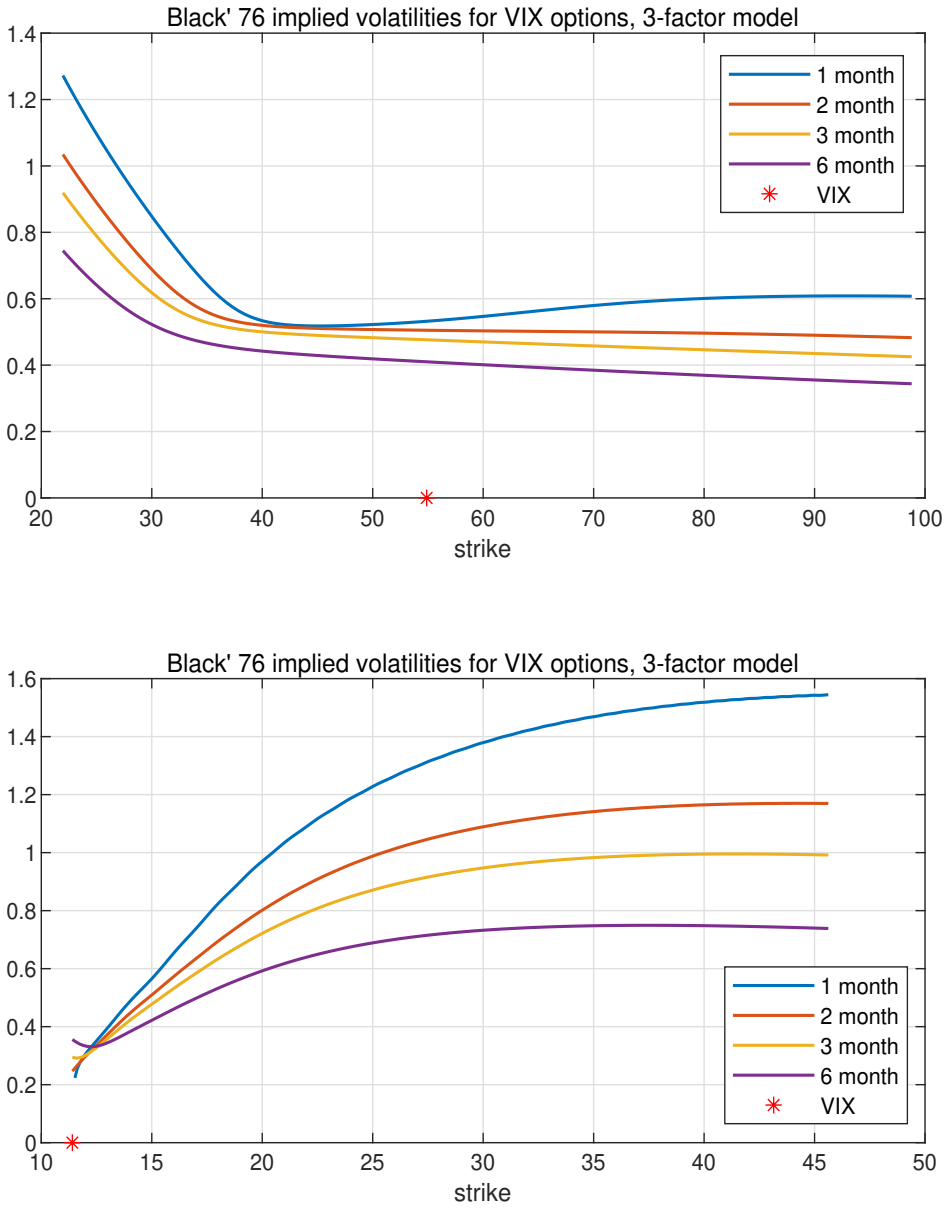


Figure D.4: Black-76 implied volatility curves for VIX options in the three-factor model: conditional analysis. The figure plots (annualized) implied volatility curves for VIX options in three-factor VIX model, conditional on high and low initial VIX. In the upper case, we set all state variables very high: $\sigma_t^2 = 10\sigma_{ss}^2$ $\theta_t = 10\theta_{ss}$, and $\lambda_t = 10\lambda_{ss}$, implying a very high VIX, 54.9. In the lower case, we set all state variables at minimum values: $\sigma_t^2 = \theta_t = \lambda_t = 0$, implying a small value of VIX, 11.4.

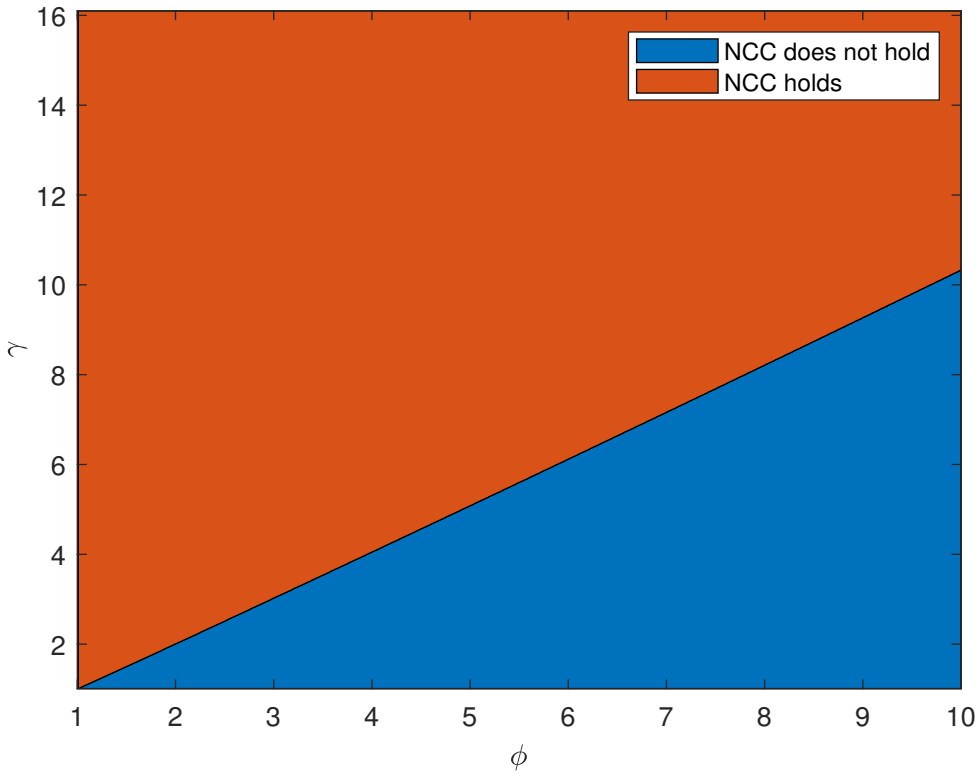


Figure D.5: Parameter space that satisfies Schneider and Trojani (2019)'s NCC. This figure plots the parameter space (γ, ϕ) such that Schneider and Trojani (2019)'s NCC $(n \geq 1, p = 1)$ holds and does not hold in the model. All the other parameters are fixed at values in the baseline calibration.