



Explaining the negative returns to volatility claims: An equilibrium approach[☆]

Bjørn Eraker^{a,*}, Yue Wu^b

^a Wisconsin School of Business, Department of Finance, 975 University Avenue, Madison, WI 53706, USA

^b Moody's Analytics, 405 Howard Street #300, San Francisco, CA 94105, USA

ARTICLE INFO

Article history:

Received 14 December 2015

Revised 5 May 2016

Accepted 1 June 2016

Available online 10 May 2017

JEL classification:

G12

G13

C22

C58

Keywords:

Variance risk premium

VIX futures

VIX ETN

Dynamic equilibrium

Jump-diffusion

ABSTRACT

We study the returns to investing in VIX futures, VIX Exchange Traded Notes (ETNs), and variance swaps. We document substantial negative return premia for these assets. For example, the constant maturity portfolio of 1-month VIX futures loses about 30% per year over our sample period (2006–2013). We investigate if these findings are consistent with dynamic equilibrium. We derive a model based on present value computation that endogenizes stock prices, the VIX index, and its associated derivative contracts. The model explains the negative return premia as well as several other stylized features of the VIX futures, ETNs, and variance swap data.

© 2017 Published by Elsevier B.V.

1. Introduction

In 2004, the Chicago Board Options Exchange (CBOE) Futures Exchange introduced cash settled futures contracts on the CBOE VIX volatility index. While initially sparsely traded, the VIX futures market has become very liquid in recent years. In addition to the futures market itself, since 2009, more than a dozen VIX futures Exchange Traded Notes (ETNs) have been introduced, allowing retail investors to trade VIX futures through regular brokerage

accounts. The ETNs follow simple, pre-specified, dynamic trading programs, and in most cases offer constant maturity exposure to n -month futures positions.

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

The first objective of our paper is to provide descriptive statistics on the average returns to VIX futures positions and the associated ETNs. Szado (2009); Alexander and Korvilas (2012), and Whaley (2013) report negative annualized VIX futures returns. We collect futures data from

[☆] We are grateful for comments by the editor, Bill Schwert, and an anonymous referee. We also thank Grigory Vilkov, Mark Ready, Kjell Nyborg, and seminar participants at the European Winter Finance Summit and the University of Wisconsin for useful comments.

* Corresponding author.

E-mail addresses: beraker@bus.wisc.edu (B. Eraker), yue.wu@moodys.com (Y. Wu).

January 2006 to May 2013 and confirm these findings. For example, if someone invested in VIX futures in January 2006 and rolled the position at end-of-day futures prices reported by the CBOE, she would have lost more than 97% of the initial investment by the end of March 2013. This corresponds to an annualized return of about -30% . This number is staggering considering that during the first part of the sample period the investor would have more than doubled the initial investment through the peak of the 2008 financial crisis. Not surprisingly, the VIX ETNs perform as badly, if not worse, than the underlying futures. In fact, since the first two VIX ETNs were introduced on January 30, 2009, the VXX and VXZ, which offer exposure to short and medium term futures, respectively, have lost an average of 34 and 14 basis points per day (simple returns).

The second and major objective of our paper is to ask: are the negative average returns consistent with returns from a present value based equilibrium model? Specifically, we use the equilibrium model of Eraker and Wang (2015) to derive equilibrium VIX futures prices. This model is based on a dynamic present value framework where investors discount a distant cash flow using time-varying discount rates. We show that the model produces a sizable volatility risk premium. To understand where this premium is coming from, we detail the main ingredients of the model here.

The large negative return premium to volatility assets in our model is linked to the *volatility feedback effect*. The fact that volatility shocks and stock prices are strongly negatively correlated is well known, and many authors have suggested this is caused by a volatility feedback effect. For example, French et al. (1987) conclude "...we interpret this negative relation as evidence of a positive relation between risk premia and ex ante volatility." Campbell (1991) suggests decomposing realized returns into revisions in expected cash flows and revisions in expected discount rates, or expected rates of return. Any model based on present value computation that has time-varying expected rates of return will, according to Campbell's decomposition, have an endogenous negative correlation between shocks to expected returns and realized returns. Our model implies that expected returns are proportional to a time-varying variance factor. Shocks to this variance factor correlate negatively with returns and the magnitude of the correlation depends on the representative agent's risk aversion.

A primary objective in our analysis is to fully endogenize this negative correlation. To see why this is important, note first that prices of volatility derivatives, such as VIX futures, depend positively on spot volatility. Since volatility negatively correlates with stock prices, volatility claims are negative beta assets. Since volatility claims have negative market beta it is useful to consider a Capital Asset Pricing Model (CAPM) style equilibrium: in order to deliver a large negative premium for volatility assets they would need to be negatively correlated with the market portfolio. Since the prices of volatility derivatives are positive increasing functions of spot market volatility, a key component in generating a negative risk premium is that spot volatility itself is negatively correlated with the mar-

ket. Our model does this, and our baseline specification generates a volatility-return correlation of -0.61 at the estimated parameter values.

Both diffusive and jump shocks to cash flow volatility are priced in equilibrium and the market price of risk is a function of risk aversion and the "deep" parameters that govern the dynamics of volatility. Rather obviously, the market price of volatility risk depends on the parameters that govern the size of volatility shocks. Also, important to note, it is inversely related to the speed of volatility mean reversion. Intuitively, investors demand a higher risk compensation when shocks to volatility have a longer lasting effect. This is analogous to long run risk models.

The model generates an upward sloping equilibrium futures curve (contango) in steady-state. This means that, ceteris paribus, investors who purchase VIX futures pay more than the value of the spot VIX at expiration of the futures contract, on average. The equilibrium model produces a negative premium in all states of the world, whether or not the VIX is above or below its steady-state value. Even if the futures curve is in backwardation (downward sloping), the futures may imply a negative risk premium because the physical speed of mean reversion will be faster than the Q measure speed of mean reversion implicit in the futures prices. These pricing implications are entirely equilibrium outcomes. If the representative agent in the model is risk neutral, none of these pricing implications hold. In particular, there is no volatility risk premium, the steady-state futures curve is essentially flat, and the expected return on VIX futures is zero.

Our paper is connected to the extant literature in several ways. Our theoretical model is related to long run risk models (Bansal and Yaron, 2004) that deliver large volatility risk premia such as those of Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011). Other theoretical justifications for large volatility risk premia include the heterogeneous beliefs model of Buraschi et al. (2014). Bollerslev et al. (2009); Andersen and Todorov (2013) among others show that the volatility risk premium can predict stock market returns. Eraker (2012) shows that a large volatility risk premium is consistent with large negative equity options returns such as those found empirically in Bondarenko (2003); Bakshi and Kapadia (2003), and Eraker (2013), among others. Broadie et al. (2007) conclude that jump-risk premium, not volatility risk premium, is the primary driver of risk premia in option returns. Recently, Andersen and Todorov (2013) proposed a model with a self-exciting jump process but find that this "tail factor" has no incremental power in predicting equity return above the level of volatility itself. This empirical finding lends support to the specification of models in which jump-risks are not disentangled from the diffusive variance, as in our model.

In our model, jump and volatility risk premia are obtained endogenously and both are increasing in the level of risk aversion. Simplified, if agents are risk averse, they care about the volatility of future cash flows. Their aversion toward high volatility is similar across diffusive and jump driven increments to volatility. Yet, the equilibrium price process we use has characteristics that are similar to existing reduced-form, no-arbitrage models. Our

most general model has a two-factor volatility specification where jumps in the volatility endogenously lead to negative jumps in the equilibrium price. This is similar, but not identical to, the volatility co-jump models in Duffie et al. (2000); Eraker et al. (2003); Bandi and Reno (2012), and Andersen and Todorov (2013) where the correlations between volatility and prices are assumed.

While our paper, to our knowledge, is the first to attempt a structural explanation for the negative VIX futures return premium, several papers fit statistical models to futures prices and judge the resulting empirical model fit using root-mean-square error (RMSE) or other distance metrics based on the difference between the model and market prices. For example, Zhang and Zhu (2006) analyze the model fit based on Heston (1993), while (Lin, 2007) and (Zhu and Lian, 2012) analyze models in the more general class of Duffie et al. (2000). These papers generally conclude that models with more complicated volatility dynamics (i.e., jumps) produce better statistical fit. Egloff et al. (2010) find that the two-factor model of Bates (2000) outperforms Heston's model. Some studies from related markets include Song (2012) who studies returns on VIX options. He finds that both diffusive volatility-of-volatility and volatility jumps are important in capturing VIX option returns. Carr and Wu (2006) study a sample of returns to variance-swap contracts. Their sample, collected from 1990 to 2005, contains strikingly large positive (negative) returns to sellers (buyers) of variance swaps.

In our empirical examination we first confirm the large negative returns to VIX futures reported elsewhere. We verify that the negative returns to futures translate into correspondingly negative returns to VIX ETNs that invest in long positions. The returns are particularly bad for short maturity futures and VIX ETNs. Yet, we show that our equilibrium model is generating returns that are almost identical to those we observe in our sample. In our empirical implementation, we first estimate our return equilibrium model. This is done using Bayesian Markov chain Monte Carlo (MCMC) sampler, extending the method in Eraker et al. (2003) to a structural setting. The advantage of the structural model is that we estimate the risk aversion of the representative agent from returns data alone. We therefore recover the pricing kernel without the use of additional data from derivatives markets. This contrasts empirical studies of reduced-form, no-arbitrage models that simultaneously use derivatives and returns data to back out market risk prices, as in Pan (2002); Chernov and Ghysels (2000); Eraker (2004); Andersen and Todorov (2013), and Jackwerth and Vilkov (2014).

We show that the equilibrium model can explain the negative returns to VIX futures and VIX ETNs almost exactly. In particular, the model that includes volatility jumps fits all moments of maturities that are less than 4 months. For 4- and 5-month futures, the model underestimates the variability (standard deviation) of the futures returns. We argue that the model's inability to account for the return standard deviation of longer maturity contracts is consistent with similar model failures in the affine term structure literature in capturing low frequency movements. We demonstrate that a generalized model that allows for a second volatility factor (i.e., a "central tendency" factor)

can be calibrated such that all the moments of the VIX futures data are matched.

The rest of the paper is organized by the following. In the next section we present basic descriptive evidence on the returns to VIX futures contracts and VIX ETNs. Section 3 presents the equilibrium framework and structural parameter estimates. Section 4 presents our empirical evaluation of the model and compares it to data on VIX futures and VIX ETNs as well as variance swaps. Section 5 concludes.

2. VIX futures and ETN returns

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

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

$$T = w_t(T_1 - t) + (1 - w_t)(T_2 - t). \quad (2)$$

$r_p(t)$ is the return of the portfolio at day t . The portfolio involves two VIX future contracts: a front-month future with expiration date T_1 and price $F_t(T_1)$, and a next-month future with expiration date T_2 and price $F_t(T_2)$. The portfolio keeps a constant maturity T by rolling the front-month future to the next-month future on a daily basis. w_t is the weight in the front-month future after the rolling at the end of day t . When w_t goes to zero, the next-month future becomes the front-month future and the next rolling cycle starts. For example, the n -month constant maturity portfolio rolls the n th-month future to the $(n+1)$ th-month future.

We report the average daily arithmetic return,

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

the average daily log-return,

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

as well as the annualized geometric return

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

While the geometric return is known to be biased for the expected annual return (Blume, 1974), it represents a

Table 1

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

Maturity	R^1	R^2	R^3	Std	Skewness	Kurtosis
1 Month	−0.12	−0.20	−39.40	3.98	0.83	6.42
2 Month	−0.07	−0.11	−24.65	3.00	0.62	6.02
3 Month	−0.01	−0.04	−10.15	2.47	0.66	6.11
4 Month	−0.03	−0.05	−12.09	2.21	0.79	7.28
5 Month	−0.01	−0.03	−7.33	2.01	0.70	6.90

monotonic transformation of the total return over the sample period and we include it for this reason.

Table 1 presents measures of return and higher order return moments. Though introduced in 2004, low liquidity in the first 2 years leads us to start the sample in January 2006. As can be seen from the table, VIX futures averaged negative returns over the sample period. Largest were the losses for the short-term maturity futures, with 1-month contracts losing an average of 12 basis points per day which roughly annualizes to $252 \times -0.12\% = -30\%$. In terms of log-returns the 1-month futures averaged -20 basis points per day, or -50.4% annually. The annualized geometric return was -39.4%. The returns tend to increase with maturity, and the 5-month contract loses a comparably small amount, with “only” -7.33% per annum geometric average loss. Both the average loss for the short maturity contracts, as well as the comparably smaller loss for the long maturities are interesting features of the data. These features of the data are seen in Fig. 1 which plots the value and log-value of a dollar invested in rolled positions in VIX futures on January 3rd, 2006.

In order to get a first pass at whether or not the returns are consistent with an equilibrium story, we compute α 's using the standard Fama–French risk factors as well as a VIX-return factor. This factor is computed by sorting the Center for Research in Security Prices (CRSP) stocks on their exposure, β_{VIX} , with respect to changes in the VIX index. We follow the standard Fama–French approach and compute the VIX factor (VF) as the return to the largest minus the smallest β_{VIX} quintile portfolio.¹

The results are reported in Table 2. As can be seen, all specifications give rise to large negative α 's. The short maturity one- to 2-month α 's are statistically significantly negative while the longer maturities are not. We also report tests of the null hypothesis $\alpha_i = 0 \forall i = 1, \dots, 5$ which are rejected for all model specifications. Note that while VIX futures have large positive betas with respect to the VIX factor (bottom panel), the coefficients are close to zero for the specifications that include the CRSP value-weighted (MKT) factor. This is due to collinearity between VF and other factors. In particular, VF has a -93% correlation with MKT returns and 70% correlation with daily changes in the VIX index.

¹ At the beginning of each calendar year, all CRSP stocks are sorted into five portfolios based on their exposure to the changes in the VIX index for the past 2 years. The VIX factor is the equal-weighted return of the first quintile portfolio minus that of the fifth quintile portfolio.

To understand where the negative returns come from, we present the average values of the VIX spot and the various maturity futures over the sample period in Table 3. This table shows that on average, the futures curve is in contango and prices monotonically increase with maturity of the contract. This, mechanically, is the reason why long positions in VIX futures lose money on average. Consider, for example, an investor who buys a 1-month contract and holds it until it expires. Her average return would be $20.57/21.48 - 1 = -4.24\%$ per month, or -40.52% per year when compounded. This is close to the annually average compounding geometric return in Table 1. Similarly, the annualized average 1-month holding period returns for 2–7-month contracts are reported in the row labeled “Implied return” in Table 3. While not identical, the numbers are on the same order of magnitude as the actual returns reported in Table 1. This shows that in order to understand why the returns to VIX futures are so low, we must understand why the futures curve is on average severely upward sloping in the data. An equilibrium explanation for the negative returns, therefore, will need to generate a sharply upward sloping steady-state futures curve.

2.1. Returns to ETNs

The first Exchange Traded Notes (ETNs) linked to VIX futures were introduced in January 2009. Table 4 gives an overview of the characteristics of various VIX ETNs. The VXX and VXZ offer long exposures to 1- and 5-month futures, respectively. Large negative returns earned on VIX ETNs are often considered a result of their unfortunately timed inception. It is not true, however, that the decimation of these securities' values is solely a consequence of the directional move in the VIX over the sample period. This is evidenced by the fact that these securities also lose value during periods of no change in the underlying VIX index. For example, during the period March 1, 2010 to June 21, 2012, the VIX went from 19.26 to 20.08, a marginal positive change, but the VXX lost 82.74% of its value over this sample period while the VXZ lost 29.21%. Clearly, the directional move in the VIX was not the reason why the VXX lost almost 83% of its value over this period!

The performance of the VIX ETNs is closely tied, if not identical, to the performance of the synthetic portfolio that rolls futures positions at daily closing prices according to Eqs. (1) and (2). To see this, Fig. 2 shows the log value of our synthetic 1-month portfolio alongside VXX's log net asset value. As observed, the two are highly correlated

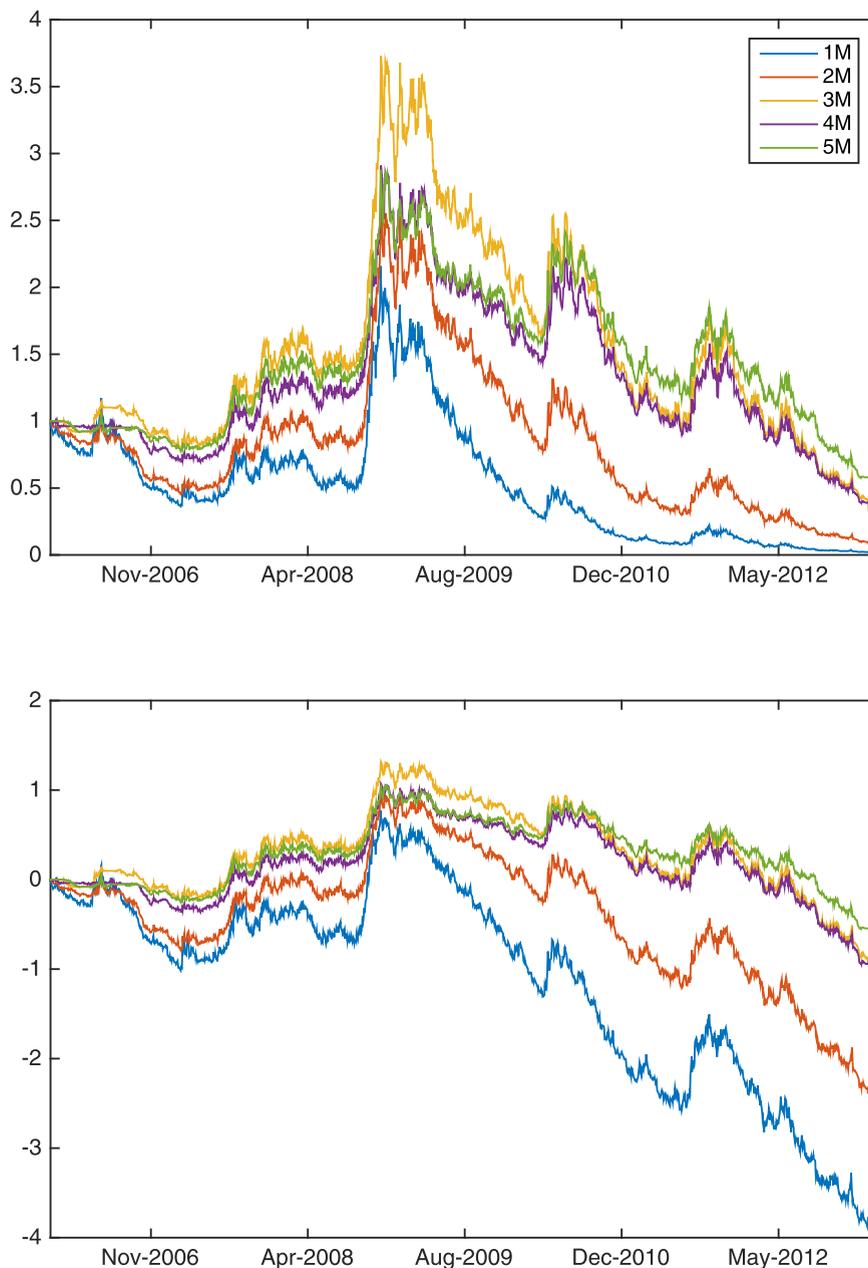


Fig. 1. Value of portfolios that roll the futures contracts at closing prices. Top: value of \$1 invested in Jan 2006. Bottom: log-value of \$1 invested in Jan 2006.

and essentially identical with the exception that the VXX depreciates at a slightly higher rate over the sample period. Specifically, the annualized geometric return to the VXX is -65.33% vs. -64.78% for the synthetic portfolio. The difference is only about 0.55% per year, of which 0.89% comprises its management fee. The remaining -34 basis points are presumably due to differences in execution between the VXX and our synthetic portfolio.

It is also interesting to study the relative performance of the three ETNs, VXX, TVIX, and XIV, which span the 1-month space with single long, double long, and single

short positions, respectively. Since the TVIX and XIV can be replicated by trading in the VXX, we compute the terminal values of the replicating portfolios and compare them to the respective terminal values of the XIV and TVIX. These results are reported in Table 5. We use two performance measures, $G1$ and $G2$. $G1$ denotes the total return difference (annualized) between actual trade price of TVIX and XIV and the value of synthetic securities. $G2$ denotes the corresponding differential based on the net asset value. The TVIX does about 161 basis points better than the synthetic security of VXX per annum based on the net asset

Table 2

Factor regressions.

Using the VIX futures logarithmic returns, this table reports the intercept and the betas of factor regressions including our VIX factor, VF, and the Fama–French factors MKT, SMB and HML. $t(\alpha)$ denotes a standard t -test of the null hypothesis $\alpha = 0$. ϑ_n is a test statistic for the null hypothesis, $\alpha = 0$ for all maturities (see Campbell et al., 1997), and $p(\vartheta)$ is the associated p -value. *, **, *** denote the estimate is statistically significant at 10%, 5%, and 1% level, correspondingly.

	α	$t(\alpha)$	VF	MKT	SMB	HML
Full model, $H_0 : \alpha = 0, \vartheta_n = 3.98, p(\vartheta_n) = 0.00$						
1 Month	-0.17***	-2.86	-0.38**	-2.35***	-0.70***	0.20**
2 Month	-0.08**	-1.90	-0.14	-1.68***	-0.33***	0.11*
3 Month	-0.02	-0.61	-0.07	-1.35***	-0.19***	0.06
4 Month	-0.03	-0.99	-0.16*	-1.28***	-0.16***	0.07
5 Month	-0.01	-0.43	-0.14*	-1.07***	-0.21***	0.04
VF, MKT, $H_0 : \alpha = 0, \vartheta_n = 3.99, p(\vartheta_n) = 0.00$						
1 Month	-0.17***	-2.89	0.15	-1.97***		
2 Month	-0.08**	-1.93	0.10	-1.51***		
3 Month	-0.02	-0.64	0.08	-1.25***		
4 Month	-0.03	-1.01	-0.04	-1.19***		
5 Month	-0.01	-0.45	0.03	-0.96***		
MKT, $H_0 : \alpha = 0, \vartheta_n = 3.99, p(\vartheta_n) = 0.00$						
1 Month	-0.17***	-2.89		-2.08***		
2 Month	-0.09**	-1.93		-1.58***		
3 Month	-0.02	-0.64		-1.30***		
4 Month	-0.03	-1.01		-1.16***		
5 Month	-0.01	-0.46		-0.98***		
VF, $H_0 : \alpha = 0, \vartheta_n = 3.85, p(\vartheta_n) = 0.00$						
1 Month	-0.17***	-2.69	2.52***			
2 Month	-0.09**	-1.80	1.92***			
3 Month	-0.02	-0.61	1.58***			
4 Month	-0.03	-0.95	1.39***			
5 Month	-0.02	-0.45	1.18***			

Table 3

Average VIX futures prices.

This table reports the summary statistics of the extrapolated constant maturity (1–7-month) VIX futures. “Implied return” is the annualized 1-month holding period return of the VIX future contract based on the average VIX future term structure. The sample data are at daily frequency from March 2004 to May 2013.

	Spot	Futures maturity						
		1	2	3	4	5	6	7
Mean	20.57	21.48	22.10	22.45	22.70	22.94	23.14	23.29
Median	17.61	19.51	20.94	21.80	22.31	22.75	23.19	23.44
Std. dev.	10.11	8.79	8.05	7.51	7.13	6.87	6.68	6.52
Implied return		-40.52%	-28.93%	-17.18%	-12.44%	-11.86%	-9.89%	-7.46%
	Eigenvalues of the correlation matrix							
	6.85	0.14	0.01	0.00	0.00	0.00	0.00	

Table 4

Investment objectives and descriptive statistics for selected VIX ETNs.

This table summarizes investment objectives and performance of some VIX Exchange Traded Notes (ETNs). The benchmarks for VIX ETNs with 1-month and 5-month horizons are VIX Short-Term Futures Index and VIX Mid-Term Futures Index, respectively. The leverage of VIX ETN reflects the return objective as a multiple (1x, 2x) or an inverse multiple (-1x) times the performance of its benchmark on a daily basis. \bar{r} and $\hat{\text{Std}}(r)$ are the average and standard deviation of daily arithmetic returns in percent, respectively.

Ticker	First date	Leverage	Horizon	\bar{r}	$\hat{\text{Std}}(r)$
VXX	29-Jan-2009	1x	1 Month	-0.34	3.94
VXZ	29-Jan-2009	1x	5 Month	-0.14	2.00
TVIX	30-Nov-2010	2x	1 Month	-0.67	8.06
TVIZ	30-Nov-2010	2x	5 Month	-0.38	4.20
VIIIX	30-Nov-2010	1x	1 Month	-0.30	4.27
VIIIZ	30-Nov-2010	1x	5 Month	-0.19	2.08
XIV	30-Nov-2010	-1x	1 Month	0.25	4.27
ZIV	30-Nov-2010	-1x	5 Month	0.15	1.99
VIXY	04-Jan-2011	1x	1 Month	-0.26	4.34
VIXM	04-Jan-2011	1x	5 Month	-0.17	2.12
IVOP	19-Sep-2011	-1x	1 Month	0.22	2.95
UVXY	04-Oct-2011	2x	1 Month	-1.13	8.57
SVXY	04-Oct-2011	-1x	1 Month	0.49	4.33

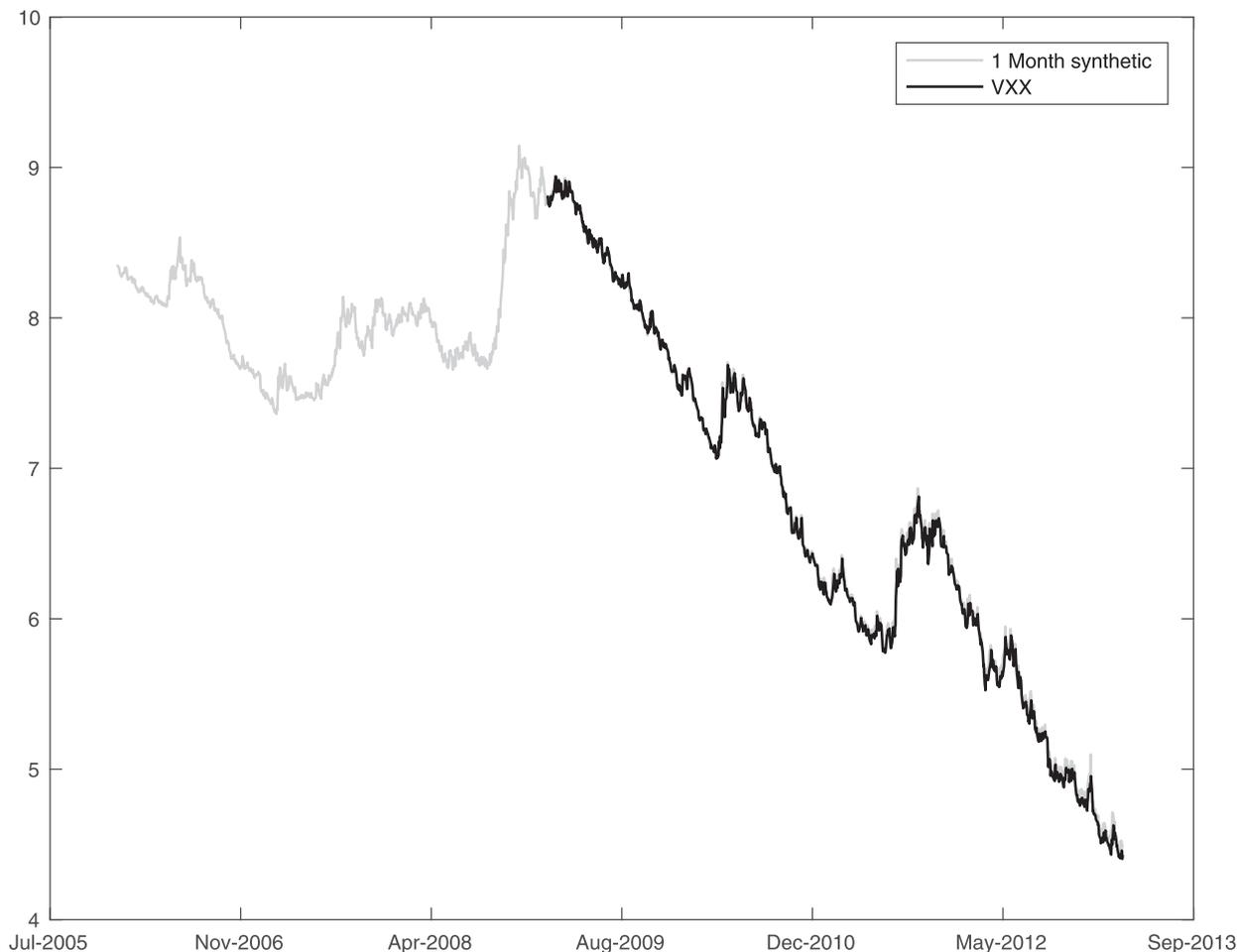


Fig. 2. Log-value of the VXX and a synthetic 1-month rolling VIX futures position.

Table 5

VIX ETN relative performance.

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

	G1	G2
TVIX	3.36	1.61
XIV	-1.80	-1.90

value. Based on the actual trade price, VXX gains 336 basis points on the synthetic counterpart. The XIV loses 180 and 190 basis points.

Table 5 reveals that the market value of the TVIX tends to deviate from its net asset value. Most VIX ETNs tend to track their net asset values relatively closely. The TVIX is an exception to this, and this stock has at times traded significantly above its net asset value. In particular, on March 21, 2012, the TVIX was traded 89% above its net asset value after the issuer, Credit Suisse, had announced it would stop share issuance. On March 22, Credit Suisse announced that it would issue more shares, resulting in a collapse in the spread between the price and the net asset value. The stock lost almost 60% of its value over the next 2 days. The stock has traded about 8% above fair value on average since then.

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

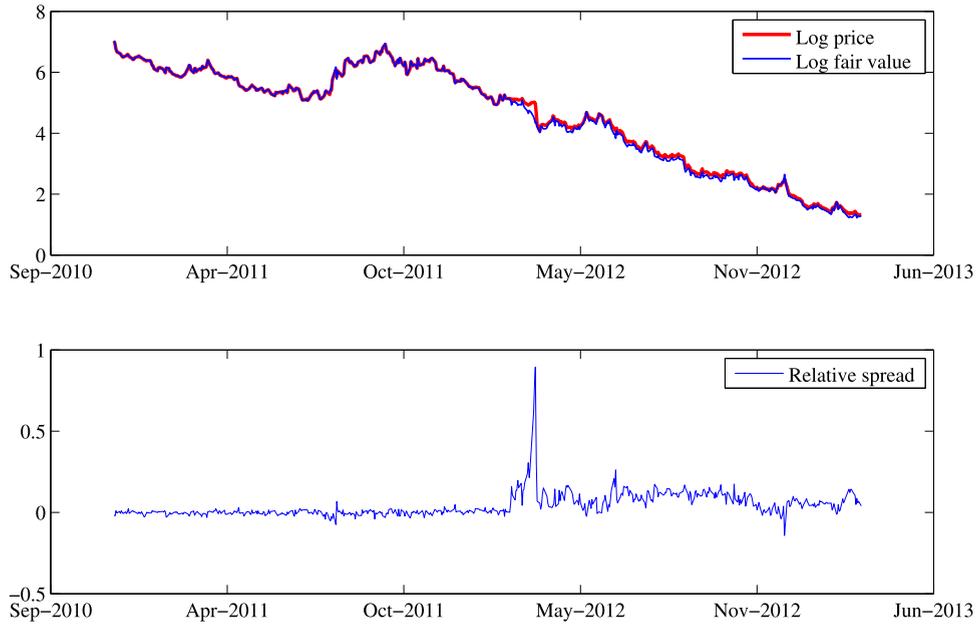


Fig. 3. Top: price of the TVIX and its Net Asset Value (NAV). Bottom: the relative difference between price and NAV.

sis regimes. This mechanism is essential in the equilibrium model.

3. An equilibrium model of VIX futures

In the following section we outline an equilibrium framework to understand and analyze the returns to VIX futures, their ETNs, and variance swaps. Our model aims to explain how VIX futures earn high negative expected returns. We take a first pass at this by estimating the (negative) beta of the futures returns and see whether the CAPM can explain the returns. With a beta of -2 and a market risk premium of 5% to 7.5% , we end up at a return premium of -10% to -15% if we assume a zero risk free rate, which is approximately right for the latter part of the sample period. The negative beta version gives us the right sign, but the CAPM still explains only half or less of the magnitude of the negative returns to 1-month VIX futures. In the next section we develop a model that endogenizes the negative volatility beta.

3.1. Model

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

of return can be equivalently represented as a function of the variance of R_m or \tilde{x}_1 .

We extend this simple equilibrium construction in the following way. First, let the terminal time period be denoted T such that the terminal cash flow is \tilde{x}_T . Second, imagine that claims to \tilde{x}_T can be traded in the capital market at, say, N discrete times prior to and including the terminal date T . P_t represents the equilibrium price of the claim to \tilde{x}_T . There are two different shocks to the price-cash flow shocks and discount rate shocks.

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

In order to incorporate a time-varying discount rate, or expected rate of return, we assume that the cash flow increments have persistent time-varying stochastic volatility. Previewing the results of our model, the expected rate of return, or discount rate for the terminal cash flow, is an increasing function of the volatility of the ϵ_i s. Thus, if the volatility of these cash flow shocks increases, the volatility of \tilde{x}_T is increased as well. If we think about today's price as the certainty equivalent of \tilde{x}_T , it is clear that by increasing its volatility, 1) we increase the volatility of the current price, and 2) the price responds negatively. Both effects are important: the former introduces stochastic, time-varying volatility into the stock price and the latter introduces an endogenous *leverage effect* caused by *volatility feedback*.

We now leave the discrete time setup above in favor of a continuous time economy. As before, agents trade claims to a terminal cash flow \tilde{x}_T . We assume that \tilde{x}_T is the ter-

minimal value of a stochastic process x which is exogenous, $\tilde{x}_T = x_T$. We call a claim to \tilde{x}_T “the stock market” and assume a unit net supply of stock and zero net supply of bonds.

The cash flow is assumed to be the terminal value of x_t which follows

$$\frac{dx_t}{x_t} = \mu dt + \sigma_t dB_t^x, \quad (3)$$

$$d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \sigma_v \sigma_t dB_t^v + \xi_t dN_t, \quad (4)$$

$$\xi_t \sim \text{Exp}(\mu_\xi), \quad (5)$$

$$\text{Corr}(dB_t^x, dB_t^v) = 0. \quad (6)$$

σ_t is the volatility of x_t and is driven by a diffusion, B_t^v , and a compound Poisson process, $\xi_t dN_t$, where the counting process, N_t , has Poisson arrivals with intensity l_0 . We assume that the parameters are chosen so that σ_t^2 is stationary and positive. Under this setup, the (relative) cash flow shocks dx_t/x_t are uncorrelated and have persistent stochastic volatility σ_t . Note that even when the planning horizon $T - t$ is large, σ_t impacts the variance of the terminal claim \tilde{x}_T . This follows because x_t is a random walk. Thus, any increase in the variance of its increments will yield a higher variance of the terminal value x_T . This is true even if T is large, and even if volatility persistence is low. With higher volatility persistence, this effect is magnified.

3.2. Equilibrium stock prices

Appendix A.3 derives the equilibrium price

$$P_t = \frac{\mathbb{E}_t\{u'(\tilde{x}_T)\tilde{x}_T\}}{\mathbb{E}_t\{u'(\tilde{x}_T)\}} e^{-(T-t)r}. \quad (7)$$

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

$$\mathbb{E}_t e^{u'X_{t+\tau}} = e^{\alpha(u,t,t+\tau) + \beta(u,t,t+\tau)'X_t} \quad (8)$$

where the functions $\alpha: \mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ and $\beta: \mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^N$ solve ordinary, first order differential equations.

By defining $X_t = \{\ln x_t, \sigma_t^2\}$ and setting $u_{1-\gamma} = (1 - \gamma, 0)$ and $u_{-\gamma} = (-\gamma, 0)$ in the numerator and denominator of (7), along with (8), we have the price

$$\begin{aligned} P_t &= \frac{\mathbb{E}_t\{e^{(1-\gamma)\ln \tilde{x}_T}\}}{\mathbb{E}_t\{e^{-\gamma \ln \tilde{x}_T}\}} e^{-r_f(T-t)} \\ &= \frac{e^{\alpha(u_{1-\gamma}, t, T) + \beta(u_{1-\gamma}, t, T)'X_t}}{e^{\alpha(u_{-\gamma}, t, T) + \beta(u_{-\gamma}, t, T)'X_t}} e^{-(T-t)r_f} \\ &= x_t e^{-r_f(T-t) + \lambda_0(t, T) + \lambda_\sigma(t, T)\sigma_t^2}, \end{aligned} \quad (9)$$

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

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

$$+ \lambda_\sigma(t, T) d\sigma_t^2. \quad (10)$$

Note that since the (log) return depends on $\lambda_\sigma d\sigma_t$, diffusive shocks $\sigma_v \sigma_t dB_t^v$ and jumps, $\xi_t dN_t$ have the same effect, λ_σ , on the stock price. This can also be accomplished in a reduced-form model, but the extant literature has focused mostly on models in which the jumps and diffusive shocks to volatility can have different impact (see Eraker et al., 2003 and Bandi and Reno, 2016 for examples).

3.2.1. Infinite horizon limit

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

$$d \ln P_t = r_f dt + \sigma_t dB_t^x + \lambda_\sigma d\sigma_t^2 + \lambda_0(\sigma_t^2) dt \quad (11)$$

where

$$\begin{aligned} \lambda_0(\sigma_t^2) &= -\kappa\theta\lambda_\sigma - \frac{1}{2}\sigma_t^2 - l_0(\varrho(\beta_2(u_{1-\gamma})) \\ &\quad - \varrho(\beta_2(u_{-\gamma}))) \end{aligned} \quad (12)$$

$$\beta_2(u_{1-\gamma}) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_v^2(\gamma^2 - \gamma)}}{\sigma_v^2} \quad (13)$$

$$\beta_2(u_{-\gamma}) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_v^2(\gamma^2 + \gamma)}}{\sigma_v^2} \quad (14)$$

$$\lambda_\sigma = \beta_2(u_{1-\gamma}) - \beta_2(u_{-\gamma}) < 0 \quad (15)$$

$$\varrho(h) = E(e^{h\xi_t}) = 1/(1 - \mu_\xi h). \quad (16)$$

A detailed derivation can be found in Appendix A.3.

The unconditional equity premium (in log return) in the economy defined as λ_0 is

$$\begin{aligned} \lambda_0 &= \mathbb{E}\lambda_0(\sigma_t^2) = -\kappa\theta\lambda_\sigma - \frac{1}{2}\left(\theta + \frac{l_0\mu_\xi}{\kappa}\right) \\ &\quad - l_0(\varrho(\beta_2(u_{1-\gamma})) - \varrho(\beta_2(u_{-\gamma}))). \end{aligned} \quad (17)$$

There are two immediate properties of λ_σ that are important. First, there is an upper limit on the amount of risk aversion, γ , that produces a well defined equilibrium (such that λ_σ is real). Second, when the equilibrium is well defined, λ_σ is negative for $\gamma > 0$ and exactly zero for $\gamma = 0$. The sign follows from the fact that $\gamma^2 + \gamma > \gamma^2 - \gamma$. The fact that λ_σ is negative is important as it generates the leverage effect in the model.

3.3. Risk neutral dynamics

Proposition 1. If equilibrium exists, the pricing kernel follows

$$\begin{aligned} \frac{dM_t}{M_t} &= -r_f dt - \gamma \sigma_t dB_t^x - \eta \sigma_v \sigma_t dB_t^v + (e^{-\eta\xi_t} - 1) dN_t \\ &\quad - l_0(\varrho(-\eta) - 1) dt. \end{aligned} \quad (18)$$

Under the equivalent measure Q , the variance process follows

$$d\sigma_t^2 = \kappa^Q(\theta^Q - \sigma_t^2)dt + \sigma_v \sigma_t dB_t^{v,Q} + \xi_t^Q dN_t^Q. \quad (19)$$

The risk neutral parameters are

$$\kappa^Q = \kappa + \eta\sigma_v^2 \tag{20}$$

$$\theta^Q = \frac{\theta\kappa}{\kappa^Q} \tag{21}$$

$$l_0^Q = l_0\varrho(-\eta) \tag{22}$$

$$\mu_\xi^Q = \mu_\xi\varrho(-\eta) \tag{23}$$

$$\eta = -\beta_2(u_{-\gamma}) < 0. \tag{24}$$

Under the risk neutral measure, the stock price follows

$$\frac{dP_t}{P_t} = r_f dt + \sigma_t dB_t^{x,Q} + \lambda_\sigma \sigma_v \sigma_t dB_t^{v,Q} + (e^{\lambda_\sigma \xi_t^Q} - 1) dN_t^Q - l_0^Q (\varrho^Q(\lambda_\sigma) - 1) dt. \tag{25}$$

A few notes on the functional form of the transformation of the probability measure are in order. First, it follows immediately that η is negative for $\gamma > 0$. The market price of the Brownian risk dB_t^v is $\eta\sigma_v^2\sigma_t^2$. As for the jump-risk, the following reward-to-risk ratio illustrates the equilibrium rewards for jump-risk

$$\eta_J \equiv \frac{\mathbb{E}\xi_t - \mathbb{E}^Q\xi_t}{Std(\xi_t)} = \frac{\mathbb{E}\xi_t}{Std(\xi_t)}(1 - \rho(-\eta)). \tag{26}$$

η_J is also negative since $\rho(-\eta) > 1$ for $\eta < 0$. It is somewhat misleading to coin this measure a market price of jump-risk. Jump-risks are characterized and priced not only according to their mean and standard deviation, but also their higher order moments. Nevertheless, as long as the representative agent is risk averse, the model features a negative variance risk premium from both diffusive and jump-risk. We explore this in some depth below in connection with variance futures (see Proposition 2). It is also clear from expressions (22) and (23) that jump arrivals are more frequent and jump sizes are larger under the risk neutral relative to the objective probability measure. Our model also contrasts reduced-form models in that there is a tightly specified relationship between the risk adjustments for jumps and diffusive risk premia.

The stock price follows

$$d \ln P_t = r_f dt + (\lambda_0(\sigma_t^2) - \gamma\sigma_t^2)dt + \lambda_\sigma d\sigma_t^2 + \sigma_t dB_t^Q \tag{27}$$

under the risk neutral measure where $d\sigma_t^2$ is given by Eq. (19).

Of essential interest is the τ period ahead conditional variance of the log-return under Q measure, which we can show to be a linear function of the spot variance σ_t^2

$$\text{Var}_t^Q(\ln P_{t+\tau}) = a(\tau) + b(\tau)\sigma_t^2. \tag{28}$$

Notice that VIX is defined as

$$\text{VIX}_t = \text{Std}_t^Q(\ln P_{t+21}) = \sqrt{a(21) + b(21)\sigma_t^2} \tag{29}$$

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

3.4. Dynamic equilibrium effects

The parameters λ_σ and η are important endogenous parameters. Fig. 4 shows both parameters as functions of risk aversion and speed of volatility mean reversion, κ . As seen, both decrease in γ . It is clear from Eq. (11) that λ_σ determines the endogenous impact of volatility shocks on stock prices. Since λ_σ increases (in absolute value) in γ , the correlation between volatility shocks and stock prices is negative (since λ_σ is negative) and increasing in absolute value as γ increases. We illustrate this through scatter plots of daily volatility changes and daily returns simulated from the model using $\gamma = 3$ and $\gamma = 8$ in Fig. 5. The correlations are -0.23 and -0.60 , respectively.

The equilibrium impact of shocks to volatility on stock prices, as captured through λ_σ , is an important distinguishing feature of our model relative to standard reduced-form models. For example, standard models such as Heston's (1993) model for pricing options specify an exogenous correlation between shocks to volatility and prices. In Heston's model this correlation is not tied to risk aversion, and the correlation can be large or small independent of volatility risk premium and equity premium. In our model, risk premia for all assets increase monotonically as a function of risk aversion, through λ_σ . The expressions (20)–(23) restrict volatility risk premia for jump and diffusive shocks to be functions of η . This contrasts the reduced-form literature where there is no explicit link between jump and diffusive risk premia.

Our model also leaves no room for volatility shocks that are not priced. Reduced-form models, such as the Stochastic-Volatility-with- Independent-Jumps (SVIJ) model estimated in Eraker et al. (2003), are not nested, since jumps in volatility can occur without price impact. Importantly, equilibrium implies not only that volatility shocks are priced, but that the relative impact of small (Brownian) and large (jumps) shocks to volatility load with a factor λ_σ on stock returns. Our model also rules out the conditional CAPM as that model does not feature priced volatility shocks, see Eraker and Wang (2015) for a discussion of how the conditional CAPM cannot be reconciled with dynamic present value computation.

Fig. 6 illustrates some of the key pricing implications through a comparison of the P and Q measures for the stock price. The difference between the P and Q densities increases as γ is increasing. The plot also shows that the two different values of γ give widely different Black–Scholes implied volatilities for the underlying stock which suggests that the model can generate a substantial volatility risk premium. This is of course key in explaining the negative returns to VIX futures.

3.5. Variance futures

The nonlinearity introduced by the square root in (29) necessitates numerical computation of futures prices. While we can do this through a single one-dimensional numerical integration (see Appendix A.4), for purposes of illustration, we discuss the shape of a hypothetical futures contract written on VIX-squared. Theoretical variance futures prices can be computed directly, and since the

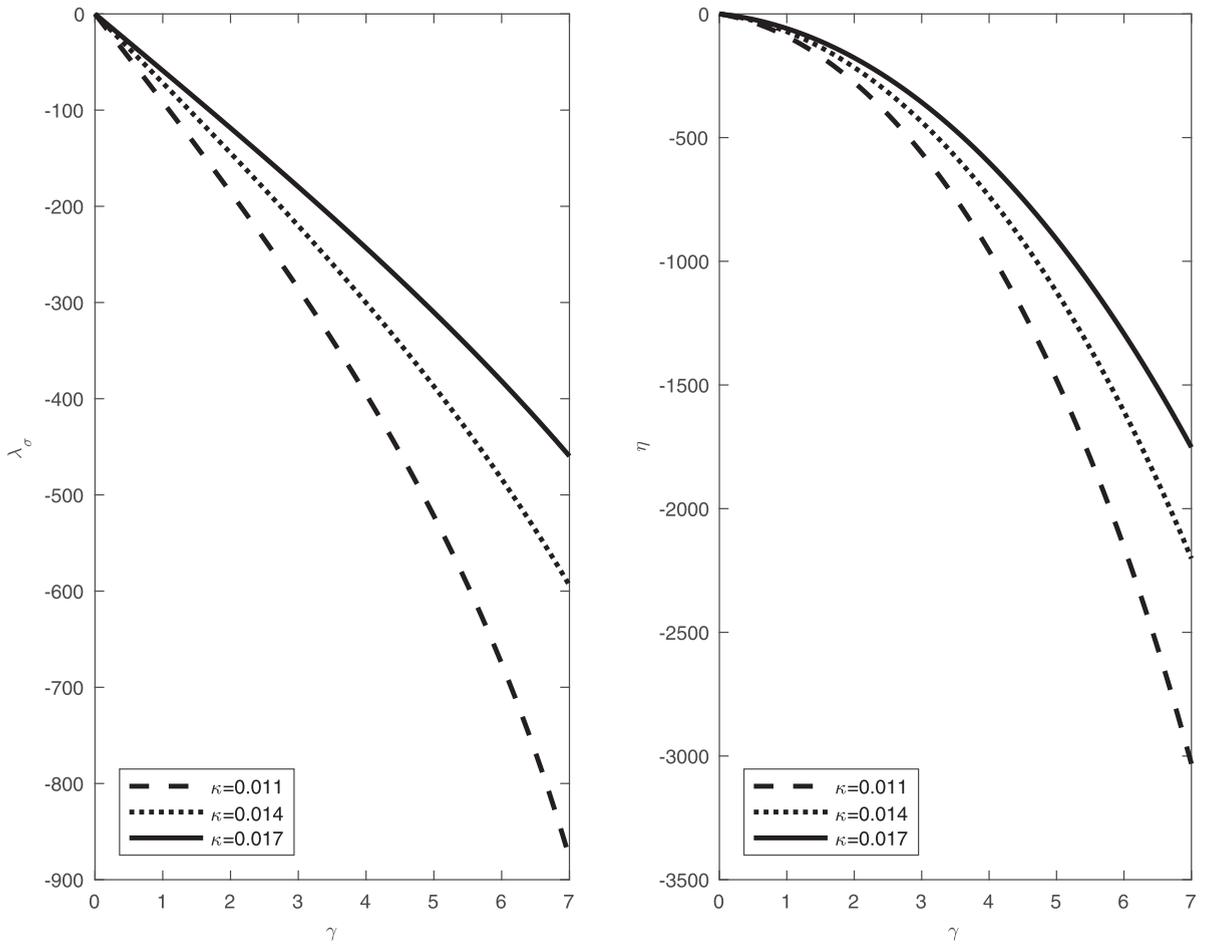


Fig. 4. The volatility feedback coefficient λ_σ and market price of volatility risk, η , for different values of risk aversion γ and volatility mean reverting speed κ .

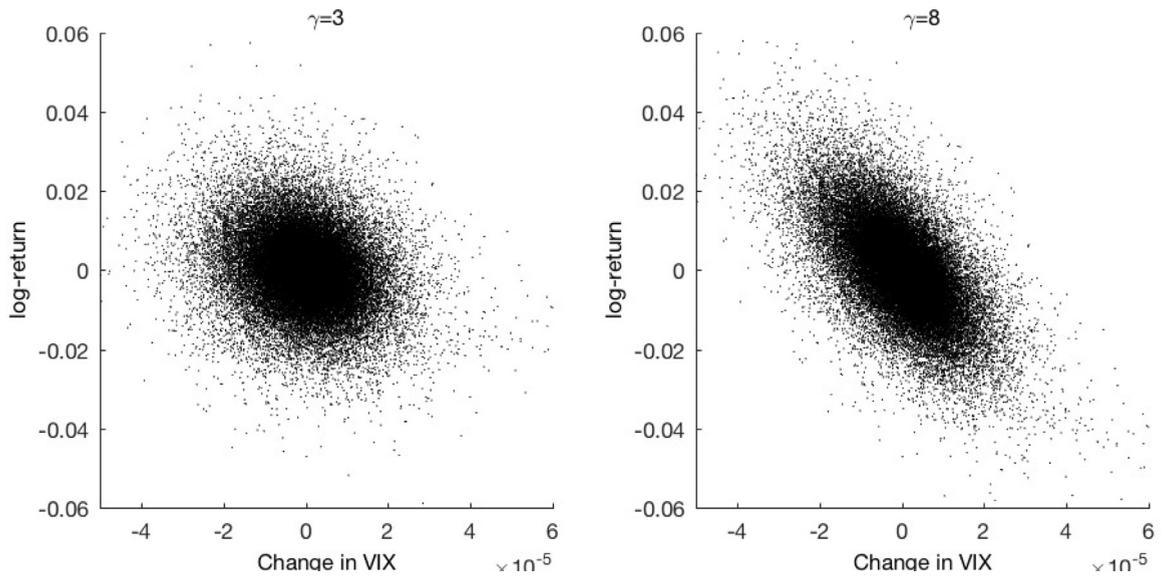


Fig. 5. Scatter plots of contemporaneous daily changes in volatility and stock returns. The data are simulated from the model using $\gamma = 3$ and 8, respectively.

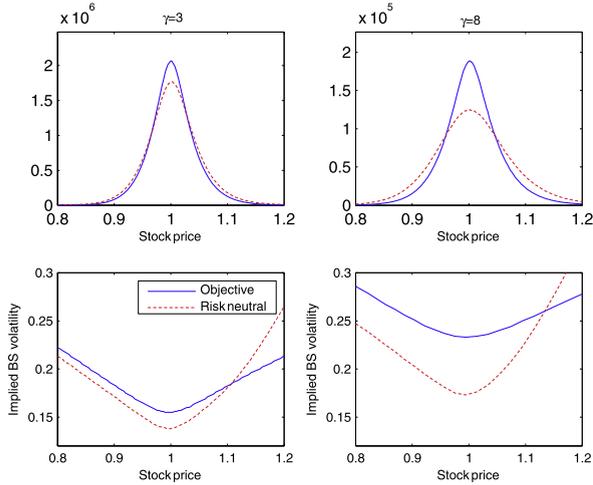


Fig. 6. Top: objective (P) and risk neutral (Q) 1-month conditional transition densities for the aggregate stock market for $\gamma = 3$ and 8 . Bottom: implied Black-Scholes volatility. The “objective measure” (solid) implied Black-Scholes (BS) volatility is computed assuming zero-volatility and jump-risk premia.

squared VIX is a linear function of σ_t^2 , it inherits the properties of the risk premia embedded into the differences between the objective P and risk neutral Q probability measures. These hypothetical variance futures are priced

$$\begin{aligned} F_t^{\text{VIX}^2}(\tau) &= \mathbb{E}_t^Q \{ \text{VIX}_{t+\tau}^2 \} = \mathbb{E}_t^Q \{ a(21) + b(21)\sigma_{t+\tau}^2 \} \\ &= a(21) + b(21)\mathbb{E}_t^Q \{ \sigma_{t+\tau}^2 \}. \end{aligned} \quad (30)$$

In steady-state, the futures curve is

$$F_t^{\text{VIX}^2}(\tau) = a(21) + b(21)\mathbb{E}^Q \{ \sigma_{t+\tau}^2 \mid \sigma_t^2 = \mathbb{E}(\sigma_t^2) \}. \quad (31)$$

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

Proposition 2. If $\gamma > 0$

1. The futures curve is upward sloping (contango) in steady-state.
2. Long positions in VIX-squared futures earn negative expected returns irrespective of the state of the economy.
3. VIX-squared futures have negative market betas.

It is worth commenting on some of these features. There is a popular notion among practitioners and some academics that if the futures curve is upward sloping, they will capture a negative roll by purchasing the futures. This is true only in relation to the physical drift rate of the underlying VIX or σ_t^2 processes. For example, if the underlying σ_t^2 process is in steady-state, ($\sigma_t^2 = \mathbb{E}(\sigma_t^2)$), the negative roll is indicative of the actual expected returns because the expected terminal value of the VIX-squared is its present value $\mathbb{E}_t(\text{VIX}_{t+\tau}^2) = \text{VIX}_t^2$. A positive slope of the futures curve reflects only the risk premium in steady-state. The larger the risk aversion in our model, the steeper the slope, and the higher (lower) the risk premium earned by short (long) positions. It is important to understand that the expected return to a long position in VIX futures is determined not by the shape of the futures curve per se, but rather by the shape of the curve relative to the expected

value of the VIX at expiration of the futures. Accordingly, in principle, if the futures curve were upward sloping but the objective measure drift was greater than that implied by the futures curve, the expected return to a long VIX futures position would have been positive. This, however, is never true in our model economy and investors earn a strictly negative premium.²

3.6. A two-factor model

Our basic one-factor benchmark model can be extended to a two-factor model by allowing θ to follow a square root, Cox–Ingersoll–Ross (CIR), stochastic process. We elaborate on the specification of this model in [Appendix A.1](#).

3.7. VIX futures

How do the actual VIX futures prices differ from the hypothetical VIX-squared futures prices we analyze in the preceding section? To answer this question, [Fig. 7](#) shows the various term structures under different assumptions about initial volatility states, σ_t . We compare two things: the actual, model-implied futures curve labeled Q (signifying expectation under Q) and the corresponding objective measure expectation, labeled P . The difference between the Q and the P expectations is due to the volatility risk premium. The right-hand graphs show the expected negative returns for the holding period, $E_t^P(\text{VIX}_{t+\tau})/E_t^Q(\text{VIX}_{t+\tau}) - 1$, to a long futures position.

Examining the plots to the left in [Fig. 7](#), we see the obvious relations between spot volatility and the shape of the futures curve: when spot volatility is high (low) the curve is in backwardation (contango). When in steady-state the futures curve is first convex and then concave. The convexity created at the short end of the curve is due to a Jensen’s inequality term created by the concavity of the square root function.³ The total expected holding period returns are negative for all maturities and for all initial values of spot volatility.

4. Empirical analysis

We estimate the model using structural likelihood-based estimation. The estimation approach is similar to that of [Eraker et al. \(2003\)](#) (hereby EJP) in that we draw the parameters of the model from the posterior distribution $\Theta_j \sim p(\Theta \mid \mathcal{Y}_T)$ using MCMC sampling. The latent conditional variances, σ_t^2 , the jump times, dN_t , and jump sizes, ξ_t , are drawn from the respective conditional posterior distributions in a manner similar to that of EJP. The main difference is that our structural model leads to non-standard posterior distributions. Since our structural model

² The variance risk premium is a function of the difference in drift rate for σ_t^2 under the two measures. It is easy to see that the Q minus P drift, $-\eta\sigma_t^2\sigma_t^2$, is a positive number regardless the level of σ_t^2 . $\kappa^Q < \kappa$ and $\theta^Q > \theta$ suggest that the physical mean reversion rate is faster than the risk neutral and the physical mean reversion level is lower, which gives a negative risk premium for VIX long futures.

³ The Jensen’s inequality term is $E_t(\text{VIX}_{t+\tau}) - \sqrt{E_t(\text{VIX}_{t+\tau}^2)} < 0$ since the square root function is concave.

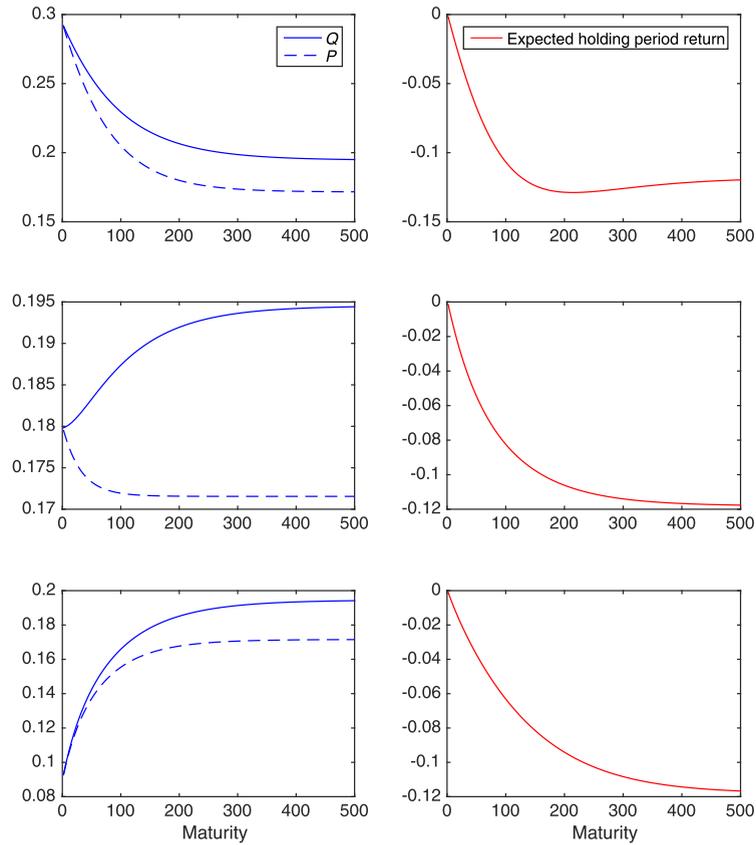


Fig. 7. Left: VIX futures curve and the objective measure expected payoff in the case of high, medium (steady-state), and low initial values of spot variance, σ_t^2 . Right: expected holding period return, $\mathbb{E}_t^P(VIX_{t+\tau})/\mathbb{E}_t^Q(VIX_{t+\tau}) - 1$, to a long futures position.

gives rise to complicated nonlinear relationships between the “deep” parameters and the affine model coefficients such as λ_σ , it is not possible to draw any of the parameters directly from conditional distributions. We therefore draw the entire parameter vector directly from the posterior using a Metropolis Hastings draw.

4.1. MCMC estimation

Bayesian inference for stochastic volatility models have become rather routine and a substantial literature is devoted to developing algorithms for sampling from the posterior distributions of these models. This literature was initiated by [Jacquier et al. \(1994\)](#) and has since seen numerous refinements see, for example, [Kim et al. \(1998\)](#) and references therein.

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

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

Our method for carrying out MCMC sampling follows the general recipe of [Eraker et al. \(2003\)](#), but differs in

some important ways as the conjugacy offered by the reduced-form models in that paper is lost in our structural model. For example, in EJP virtually all the model parameters have known conditional posterior distributions which follow from conditioning on the simulated values of volatilities, jump sizes, and times. This leads to conditionally Gaussian errors associated with the Brownian increments. For our model, the λ -parameters that determine the equilibrium unconditional return and volatility-feedback effect, are nonlinear functions of risk aversion and the statistical parameters that determine the dynamics of volatility. For this reason, the entire parameter simulation step will have to be done using Metropolis-Hastings.

4.2. Estimation results

[Table 6](#) reports structural estimates of the parameters in the model. All parameter estimates should be interpreted to be based on a unit of time being 1 day. This makes the time-series parameters directly comparable to estimates based on daily returns. Depending on whether the model includes jumps or not, κ is estimated to be 0.014 and 0.0091, respectively. These estimates imply daily autocorrelations of $\exp(-0.014) = 0.9861$ and $\exp(-0.0091) = 0.991$. These numbers are broadly consistent with estimates reported elsewhere see [Singleton \(2006\)](#) for a review.

Table 6

Model parameter estimates.

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

	SVJ	SV
$\theta \times 10000$	0.677 (0.058)	0.758 (0.069)
κ	0.014 (0.002)	0.0091 (0.001)
$\sigma_v \times 10000$	10.788 (0.357)	9.716 (0.307)
λ	0.002 (0.000)	
$\mu_v \times 10000$	0.390 (0.096)	
γ	7.945 (0.760)	5.898 (0.677)

The most interesting parameter in Table 6 is the risk aversion parameter, γ . We estimate this to be more than ten posterior standard deviations away from zero for the Stochastic-Volatility-with-Volatility-Jumps (SVJ) model. This contrasts conflicting evidence from the conditional CAPM literature where the risk aversion parameters are typically found to be statistically insignificant.

Our model generates an endogenous negative correlation between shocks to volatility and prices. Our estimates imply a correlation of -0.613 for the diffusive shocks and -1 for the jumps. The perfect negative correlation is an equilibrium outcome because σ_t^2 is a priced risk factor. As per the argument in Eraker and Wang (2015), any shock to the expected rate of return of the market will lead to a negative equilibrium price response unless that shock is perfectly offset by a simultaneous increase in future expected cash flow from holding the asset.⁴ Our estimates are broadly consistent with estimates reported in the reduced-form literature. In this literature, using time-series of returns only, the correlation is estimated to be -0.62 (Andersen et al., 2002), -0.48 (Eraker et al., 2003), and recently Ait-Sahalia et al. (2013) estimate the correlation to be -0.77 using high frequency based realized variance (RV) along with a bias-correction to account for noise in the RV estimator. Gualtieri and Sizova (2015) estimate the parameter to be -0.61 using MCMC. A second strand of the literature employs joint return/options data to recover estimates of ρ and find -0.57 (Bakshi et al., 1997), -0.53 (Pan, 2002), -0.58 (Eraker, 2004). Finally, Bandi and Reno (2016) estimate a non-parametric generalization of the model in Eraker et al. (2003) where the diffusive correlation is estimated to be -0.48 and the correlation between jumps in returns and spot variance is close to -1 . The sample correlation between changes in squared VIX

⁴ It is possible to set up an equilibrium model where the jumps to volatility are correlated with jumps in cash flows (our x_t process). If so, the simultaneous occurrence of a positive volatility shock and a positive shock to expected cash flow will partially offset, so as to increase the correlation to a number above -1 .

Table 7

VIX futures return moments: Data vs. model.

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

Maturity	R^1	R^2	R^3	Std.	Skewness	Kurtosis
Data						
1 Month	-0.12	-0.20	-39.40	3.92	0.56	5.95
2 Month	-0.07	-0.11	-24.65	2.97	0.41	5.85
3 Month	-0.01	-0.04	-10.15	2.45	0.49	5.88
4 Month	-0.03	-0.05	-12.09	2.18	0.61	6.86
5 Month	-0.01	-0.03	-7.33	2.00	0.54	6.67
SV						
1 Month	-0.09	-0.19	-36.10	4.50	-0.13	3.29
2 Month	-0.07	-0.14	-28.15	3.56	-0.10	3.17
3 Month	-0.06	-0.11	-22.57	2.92	-0.08	3.20
4 Month	-0.05	-0.08	-18.41	2.44	-0.07	3.27
5 Month	-0.05	-0.07	-15.19	2.06	-0.06	3.36
SVJ						
1 Month	-0.10	-0.19	-36.14	4.15	1.00	26.33
2 Month	-0.08	-0.13	-26.29	3.06	0.94	23.10
3 Month	-0.06	-0.09	-19.60	2.34	0.91	21.85
4 Month	-0.05	-0.07	-14.84	1.82	0.89	21.26
5 Month	-0.04	-0.05	-11.37	1.43	0.88	20.97
Two factor						
1 Month	-0.10	-0.22	-40.37	4.57	4.24	71.84
2 Month	-0.07	-0.14	-29.20	3.49	4.05	68.15
3 Month	-0.06	-0.10	-21.29	2.75	3.42	55.66
4 Month	-0.04	-0.07	-15.76	2.23	2.62	40.24
5 Month	-0.03	-0.05	-12.04	1.84	1.82	25.81

and S&P 500 returns is -0.72 using data from January 1990 to January 2016.

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

Table 7 reports the results. We include the first three estimators of the average returns from the data which we previously reported in Table 1 for convenience. Note that the higher order moments are of logarithmic daily returns. Although slightly higher, the returns produced by the SVJ model are close to what we see in data for 1-month maturity contracts. For longer maturities both models overstate the size of the negative return premium. Both models fit average returns surprisingly well. The models also do well in fitting the higher order moments, with the exception of the SVJ model's too high kurtosis. In general, the simulated model returns have moments in the ballpark of that seen in the real VIX futures data. We further test the null hypothesis of zero difference between the model and the data below.

To carry out these significance tests, we use a model-based bootstrap. We wish to avoid the use of test statistics based on asymptotic normality because the higher order

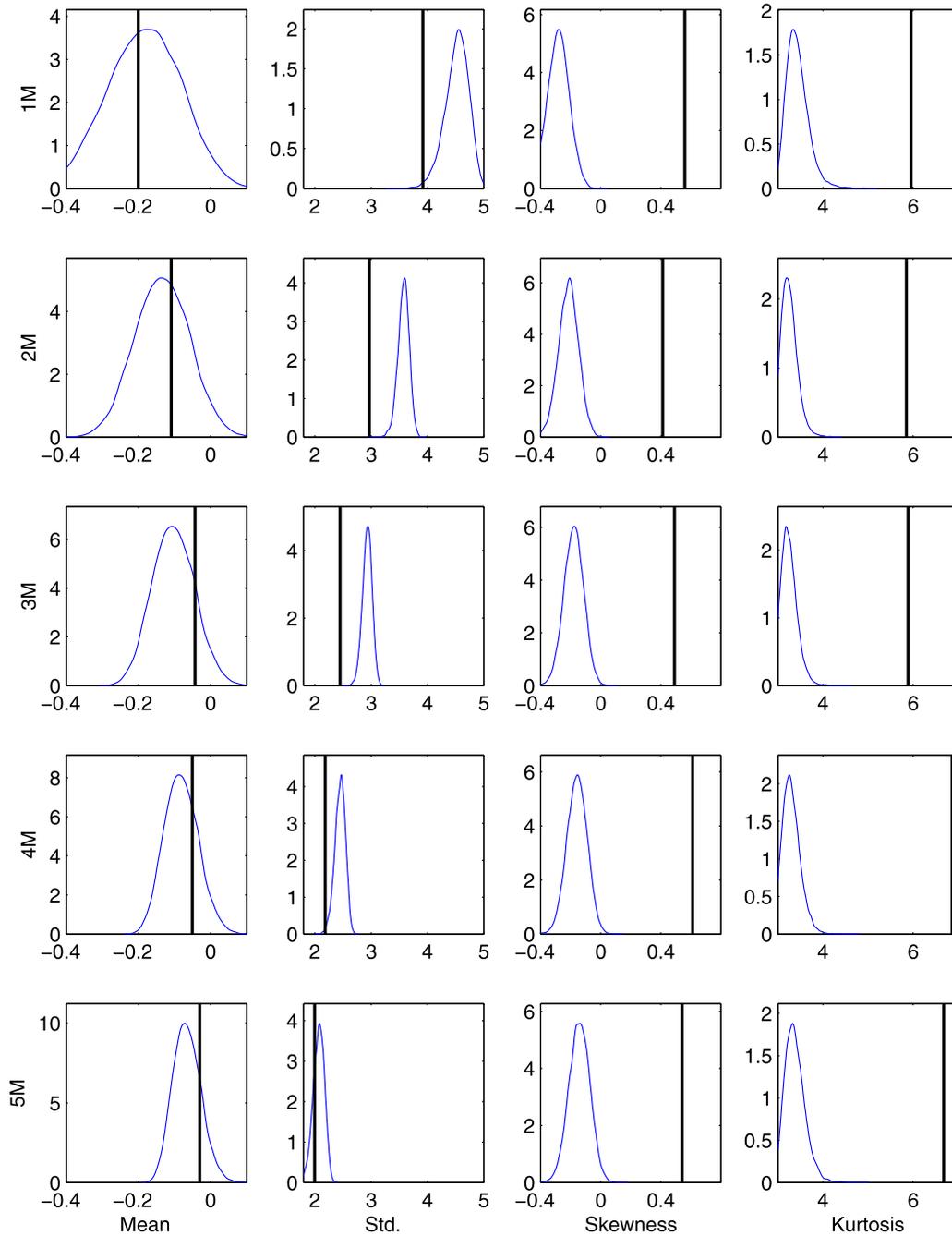


Fig. 8. Sampling distributions from the SV model vs. data: we simulate a panel of futures returns, compute the time-series moments and plot the cross section distributions of the moments under the SV model. The small-sample densities are estimated by kernel densities and compared to the sample moments in the data, shown in the vertical bars.

moments are non-normally distributed in small samples. The evidence is presented in Figs. 8 and 9, where kernel-smoothed densities of moments implied by the model and the corresponding data moments, are represented by vertical bars. The non-parametric density estimates should be interpreted as a model-based bootstrap of the sampling distribution for each respective moment.

In examining Figs. 8 and 9 we see that the average returns for the stochastic-volatility (SV) model are insignificantly different between the real data and the model-generated data. For the standard deviations and higher order moments, the results are very different. The SV model basically fails to fit any moment higher than the first. The standard deviations are all significantly different between

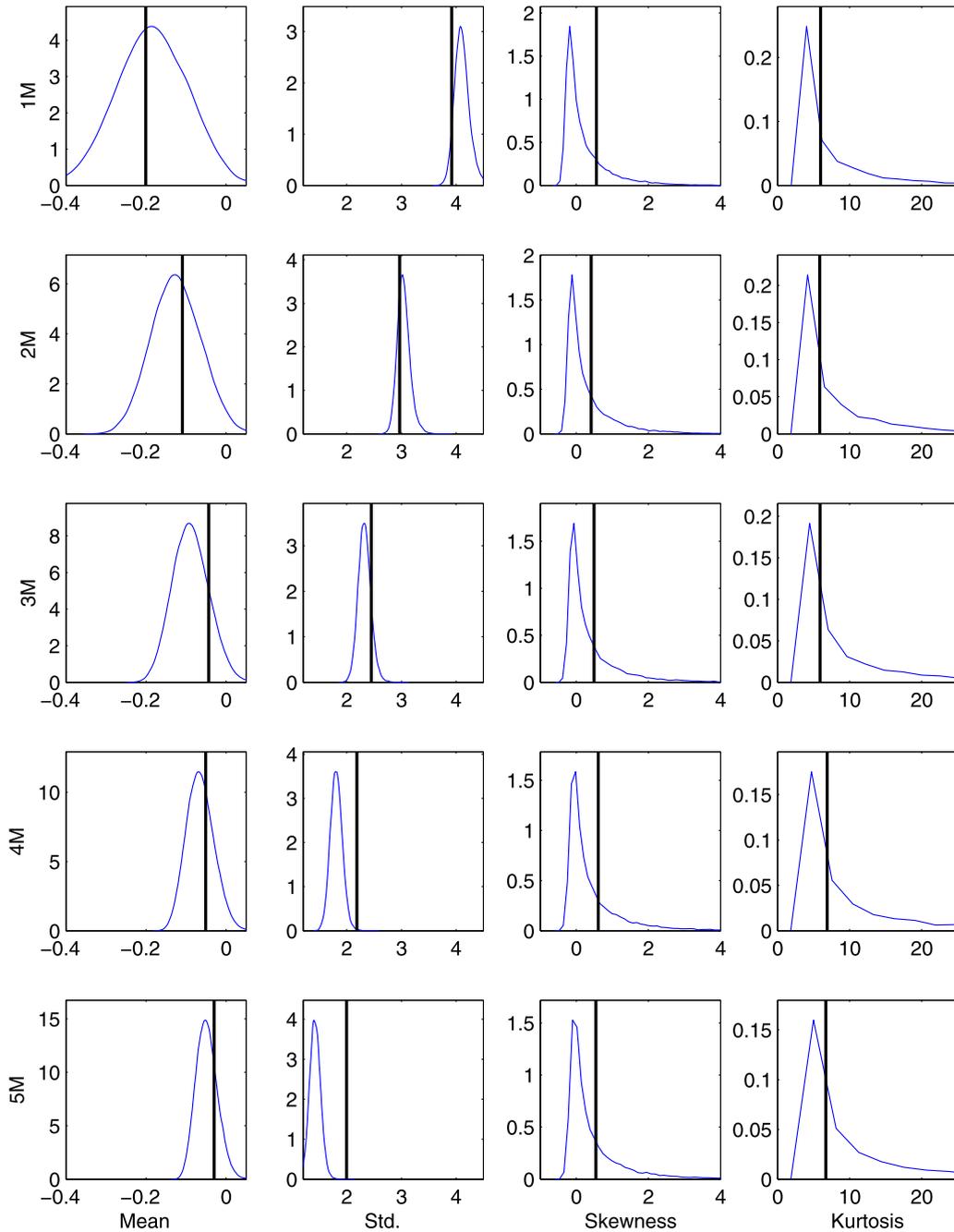


Fig. 9. Sampling distributions from the SVVJ model vs. data: we simulate a panel of futures returns, compute the time-series moments, and plot the cross section distributions of the moments under the SVVJ model. The small-sample densities are estimated by kernel smoothing and compared to the sample moments in the data, shown in the vertical bars.

the model and the data. The skewness and kurtosis are so far off that the vertical bars are outside of the empirical support of the sampling distributions under the model.

On the other hand, for the SVVJ model, the moments are surprisingly well matched. With the exception of the standard deviation for the 4- and 5-month maturity contracts, the data and the model are insignificantly different. From Table 7 we see that the SVVJ model generates sample kurtosis coefficients that range from 20 to 27 which

compares to coefficients that range from 5.95 to 6.86 in the data. This may seem like a large difference, however, Fig. 9 reveals that the sampling distribution under the model has most of its mass below what we find in the data. In fact, the medians of the small-sample distributions for the kurtosis coefficients under the SVVJ model range from 4.4 to 5. Thus, while it may seem from Tables 1 and 7 that the SVVJ model generates too high kurtosis, the differences are in fact almost non-existent.

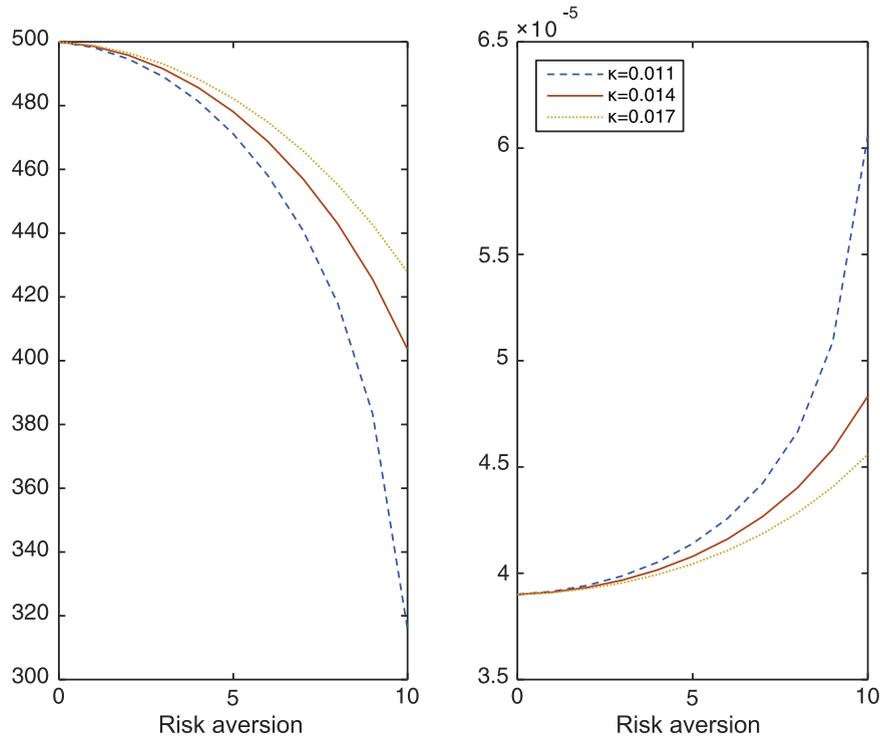


Fig. 10. Jump characterizations under the risk neutral measure. Left: the expected number of days for a jump to occur. Right: the expected size of a jump.

Next, we examine how the improvement in goodness of fit for the SVVJ model relates to jump-risk. The estimated parameters imply, on average, that it takes 500 days and 444 days for a jump to occur under physical and risk neutral measures, respectively. Once a jump does occur, the estimated average jump sizes are 3.9×10^{-5} and 4.4×10^{-5} , correspondingly. Fig. 10 shows that the jump intensity and size both increase with risk aversion and are inversely related to the mean reverting speed of the variance process.

The variance risk premium and the long run average returns for VIX futures and ETNs are fundamentally driven by the difference between average level of volatility under the physical measure and the risk neutral measure. Therefore, to quantify the relative contribution of the diffusion and the jump, we compute how much that difference is due to each component. Let $\bar{\theta}$ and $\bar{\theta}^Q$ be the unconditional means of the variance process under the respective measures. We know

$$\begin{aligned} \bar{\theta} &= \theta + \frac{l_0 \mu_\xi}{\kappa} \\ \bar{\theta}^Q &= \theta^Q + \frac{l_0^Q \mu_\xi^Q}{\kappa^Q} \\ &= \frac{\kappa \theta + l_0 \mu_\xi \rho^2 (-\eta)}{\kappa + \eta \sigma_v^2}. \end{aligned}$$

Let Λ be the difference between $\bar{\theta}$ and $\bar{\theta}^Q$, then

$$\Lambda = \frac{-\eta \sigma_v^2 \theta}{\kappa + \eta \sigma_v^2} + \frac{l_0 \mu_\xi (\kappa \rho^2 (-\eta) - \kappa - \eta \sigma_v^2)}{\kappa (\kappa + \eta \sigma_v^2)}.$$

The first term is due to the diffusion risk and the second term is due to the jump-risk. Using the estimated parameter values, we can compute the fraction of Λ due to each source of risk. As we can see from Fig. 11, the jump plays a more and more important role as risk aversion increases and mean reverting speed slows down.

4.3. Variance swaps

In this section we consider our models' abilities to explain the returns to variance swaps. A variance swap is a zero-cost contract that pays the difference between an agreed upon price, called the swap rate (VS) and physical, realized variance over some time period, say t, \dots, T . The realized variance leg of the swap is simply the sum of squared daily log-returns,

$$RV_{t,T} = \sum_{s=t+1}^T (\ln R_s)^2. \tag{32}$$

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

$$VS_{t,T} = \mathbb{E}_t^Q (RV_{t,T}). \tag{33}$$

Variance swaps trade in the over-the-counter (OTC) market. There is no public source for these OTC transactions. Fortunately, Eq. (33) suggests that any measure of Q variance is an estimator for the variance swap rate. A number of recent papers, including Dew-Becker, Giglio, Le, and Rodriguez (2015) (DGLR) and (Johnson, 2015) use S&P 500 option-implied variance. A variety of possible estimators of

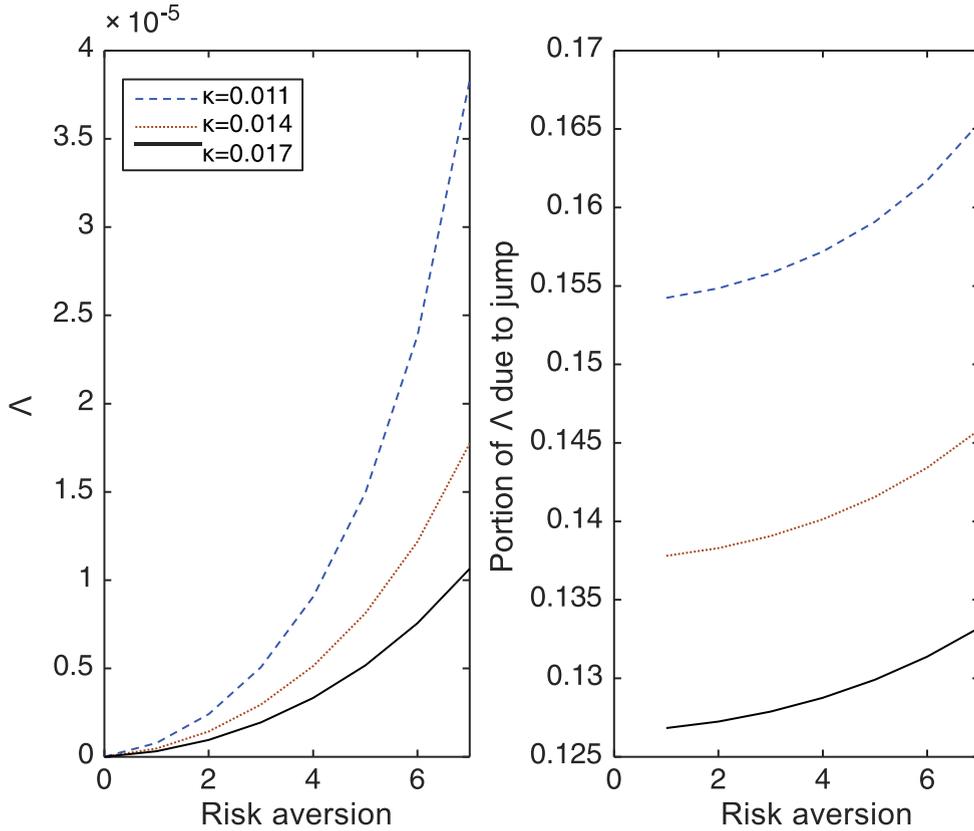


Fig. 11. Left: difference of average variances under risk neutral and physical measure denoted as Λ . Right: portion of Λ due to jump.

Q-variance, including the VIX method, can be applied to estimate the variance swap rate. We apply the VIX method with the exception that we ignore the squared drift term⁵ to make the computations consistent with the definition (33).

Given a variance swap rate, we compute 1-month holding period returns

$$R_{t,t+22} = \frac{RV_{t,t+22} + VS_{t+22,T}}{VS_{t,T}} - 1. \tag{34}$$

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

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

Table 8 presents various return estimates from the extant literature, along with our own estimates. We include

these five different data sources in order to convey that the returns are largely in the same ballpark. Three of the data sets include the financial crisis (DGLR, Johnson, and our own data). Those data sets produce returns that range from -17% to -26% for the 1-month maturity and close to zero returns for the longest horizon. The Bloomberg data yield negative returns across the entire maturity spectrum, but the Bloomberg sample period is particular as the data start in October of 2008, the height of the financial crisis. The data from Egloff et al. (2010) also produce negative returns across all maturities, but this sample stops before the financial crisis in 2007.

Table 9 presents the moments of the variance swap returns from our sample data alongside the model-implied returns. At first glance it would appear that our one-factor models do a poor job while the two-factor model matches the data. In particular, the one-factor models are incapable of matching the large negative return for the 1-month variance swaps. For example, the SVJ model generates an average return of -9.34% which compares to -21.08% in the data. At longer maturities the discrepancies are less significant. These findings echo the results in Dew-Becker, Giglio, Le, and Rodriguez (2015) who show that the Drechsler and Yaron (2011) model cannot match the large negative returns to 1-month claims while delivering too negative returns for longer maturity contracts.

To understand why our one-factor model fails to explain one market while almost perfectly explaining an-

⁵ The formula for squared VIX estimates $E_t^Q(RV_{t,t+22}) - E_t^Q(Ret_{t,t+22})^2$ so we ignore the last term.

Table 8

Returns to variance claims.

The table reports average monthly returns and standard deviations for variance claims, including variance swaps. We report average returns from Dew-Becker, Giglio, Le, and Rodriguez (2015) (DGLR), Johnson (2015), Bloomberg (BB), Egloff et al. (2010) (ELW), as well as our own estimates. All numbers are in percent.

		Period	Type	1M	2M	3M	6M	12M
DGLR	Mean	96–14	Option	–25.6	–5.6	0.8	0.5	1.8
	Std			69.0	47.8	34.0	19.7	17.4
Johnson	Mean	96–14	Option	–17.3	–8.7	–5.1	–3.6	–1.1
	Std			150	85.7	53.3	29.1	23.0
BB	Mean	08–15	Option	–33.2	–23.2	–19.1	–12.3	–6.9
	Std			63.3	52.7			
ELW	Mean	96–07	OTC	–	–21.3	–13.4	–7.2	–4.2
	Std			–	35.1	29.2	19.8	12.9
This paper	Mean	96–15	Option	–21.1	–9.7	–5.9	–1.4	0.4
	Std			71.8	63.1	50.9	38.5	28.1

Table 9

Variance swap returns: Data vs. model.

This table compares average 1-month holding period returns of variance swaps in data and model simulations. All numbers are in percent.

Maturity	Data	SV	SVJ	Two factor
1 Month	–21.08	–3.41	–9.34	–20.28
2 Month	–9.68	–4.34	–6.75	–10.08
3 Month	–5.90	–4.27	–5.53	–6.80
6 Month	–1.43	–3.57	–3.70	–3.59
12 Month	–0.39	–2.52	–2.18	–1.93

other, closely related market, we note two facts. First, pay-offs to variance swaps depend on physical, realized variance, whereas VIX futures do not. This means that if somehow the VIX itself is inflated by some constant across time relative to physical variance, this would show up in variance swap returns but not necessarily in VIX futures returns. Second, our variance swap rates are in fact not market rates, but rather swap rates that are computed from the midpoint of the bid and asks from S&P 500 options prices.

Next, we examine whether the use of midpoints between the bids and the asks for the underlying options affects our returns computations. Fig. 12 plots the average variance swap term structure in our sample data. The solid line represents the midpoint of the bids and the asks, which is what the CBOE uses to compute the VIX index, while the dashed and dotted lines represent the term structures computed from bids and asks. There is a dramatic difference between the rates computed from bids and asks. At the short end (1 month) the difference is 17% (in variance units), while for a 12-month contract it is about 10%.

Table 10 presents two pieces of evidence relating the bid-ask spreads in the underlying options to the variance swap returns. In Panel A we simply compute the variance swap rates from bids, midpoints, and asking prices. The results show that indeed, the returns are higher when computed from bids rather than asking prices. In Panel B we examine the returns to a “price taker” who will initiate a trade at either bids or asks, hold the position for 1 month, and then liquidate the position at the other side of the market. We refer to a buyer as a trader who is purchasing the variance swap at the asking price and liquidating

Table 10

Variance swap returns from bid and ask Data.

The table reports the returns to variance swaps computed from Option-Metrics using a modified version of the VIX formula. In Panel A we report returns to variance swaps based on bid, midpoints, and ask prices. In Panel B we report returns to price takers, defined as a trader who crosses the market to make trade, selling at the bid and buying at the ask price.

Panel A: Returns based on bids, midpoints, and asks					
	1M	2M	3M	6M	12M
Bids					
Mean	–14.27	–7.35	–4.87	–0.93	0.60
Std	78.48	64.76	51.70	39.65	28.33
Midpoints					
Mean	–21.08	–9.68	–5.90	–1.43	0.39
Std	71.80	63.15	50.92	38.52	28.12
Asks					
Mean	–26.77	–11.62	–6.75	–1.83	0.25
Std	66.41	62.08	50.48	37.76	28.17
Panel B: Returns to price takers					
Buyers					
Mean	–26.77	–19.31	–15.68	–10.24	–7.87
Std	66.41	55.79	45.64	35.00	26.25
Sellers					
Mean	14.27	–1.55	–5.31	–8.45	–9.60
Std	78.48	72.14	57.37	43.07	30.93

at the bid 1 month later, and vice versa. A seller realizes a positive return for a 1-month swap, and negative returns for any other maturity contract. This stems from the bid-ask being so large that by adding the spread to the cost of making the roundtrip trade, both long and short positions in long maturity contracts are unprofitable.

5. Concluding remarks

In this paper we show that the average returns earned on volatility and variance derivatives are very negative. We argue that the negative returns are consistent with equilibrium. Though the size of the negative return premium is not consistent with a traditional CAPM, which delivers a risk premium of –15% per year, we show that our dynamic equilibrium models are capable of explaining the –30% per year average return to a 1-month (rolled) futures position. The SVJ model, which includes volatility jumps, is our preferred one-factor model. We also present a two-

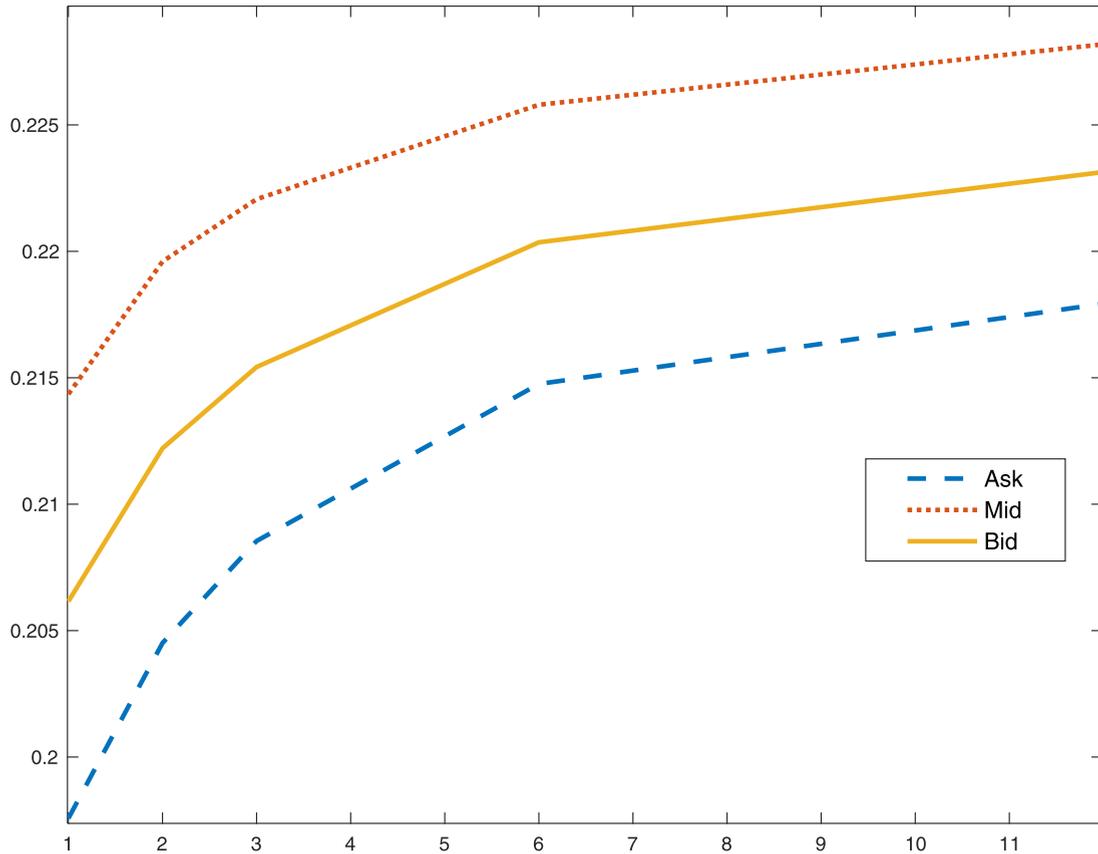


Fig. 12. Average term structure of variance swap rates computed from OptionMetrics using best bid, best ask, and the midpoint. All numbers are in annualized standard deviation units.

factor calibrated model that produces returns to these assets which are broadly consistent with what we observe in the data, including the returns to variance swaps computed from the midpoint of underlying bid and ask data for S&P 500 Index (SPX) options.

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

The overall size of the volatility risk premium in our model depends on risk aversion, volatility persistence, and the specification of volatility jumps. Large jumps lead to

discontinuities in stock prices reminiscent of financial crisis. Jumps imply a higher short term volatility risk premium. Expected returns, accordingly, are a steeper function of maturity under the SVVJ model than the SV model. The SV model comparably generates a higher volatility risk premium by assigning a slightly lower speed of volatility mean reversion. This leads to a somewhat higher negative return premium for long term futures contracts. It also leads to more volatility in futures returns at longer maturities (i.e., 5 months) in a manner that is consistent with market data.

Acknowledgments

We are grateful for comments by the editor, Bill Schwert, and an anonymous referee. We also thank Grigory Vilkov, Mark Ready, Kjell Nyborg, and seminar participants at the European Winter Finance Summit and the University of Wisconsin for useful comments.

Appendix A

A1. A two-factor volatility model

As demonstrated in Section 4.2, the SVVJ model captures all but one aspect of the observed futures returns—

it significantly underestimates the variability of longer term contracts. Since the SVVJ model is a single factor affine volatility model, it cannot generate long run memory behavior in conditional volatility, as empirically documented by [Bollerslev and Mikkelsen \(1996\)](#) among others. [Bates \(2000\)](#) and [Chernov et al. \(2003\)](#) propose two-factor nested conditional variance specification in a no-arbitrage, reduced-form model. We propose a similar model,

$$\frac{dx_t}{x_t} = \mu dt + \sigma_t dB_t^x \quad (35)$$

$$d\sigma_t^2 = \kappa(\theta_t - \sigma_t^2)dt + \sigma_v \sigma_t dB_t^v + \xi_t dN_t \quad (36)$$

$$d\theta_t = \kappa_\theta(\theta - \theta_t)dt + \sigma_\theta \sqrt{\theta_t} dB_t^\theta \quad (37)$$

$$\xi_t \sim \text{Exp}(\mu_\xi) \quad (38)$$

$$N_t \sim \text{Poisson}(l_0 t) \quad (39)$$

$$\text{Corr}(dB_t^i, dB_t^j) = 0, \quad i, j \in \{x, v, \theta\} \text{ and } i \neq j. \quad (40)$$

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

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

$$d \ln P_t = r_{f,t} dt + \lambda_0(\sigma_t^2, \theta_t) dt + \lambda_\sigma d\sigma_t^2 + \lambda_\theta d\theta_t + \sigma_t dB_t^x, \quad (41)$$

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

$$VIX_t^2 = \text{Var}_t^Q(\ln P_{t+21}) = a + b\sigma_t^2 + c\theta_t, \quad (42)$$

where we can solve for constants a , b , and c (see [Appendix A.4](#)). Since VIX_t^2 is a function of two processes with different autocorrelation functions, the autocorrelation for the VIX_t^2 itself is a mixture, and thus displays long-run memory-like behavior.

It's difficult to take our two-factor volatility model and estimate it using return data alone. The purpose of our exercise here is to demonstrate that the two-factor model is capable of matching the moments of the futures returns data. In [Fig. 13](#) we show the sampling distributions for the VIX futures data under the model, using calibrated parameters. The parameters are $\kappa = 0.0198$, $\sigma_v = 0.0019$, $\theta \times 10000 = 0.2167$, $\gamma = 5.4981$, $l_0 = 0.0018$, $\mu_v \times 10000 = 1.6933$, $\kappa_\theta = 0.0068$, and $\sigma_\theta \times 10000 = 5.22$. The data moments are all well inside the tails of the sampling distributions suggesting that our two-factor volatility specification provides a plausible description of the true data generating process. Note that the standard deviation of the longer maturity contracts is matched almost exactly.

A2. Solving the stochastic-differential-equation (SDE) system

Define the state variable X_t in this economy as $(\ln x_t, \sigma_t^2, \theta_t)$ and assume X_t follows an affine process equivalent to

$$dX_t = (K_0 + K_1 X_t) dt + \sigma(X_t) dB_t + \begin{pmatrix} 0 \\ \xi_t dN_t \\ 0 \end{pmatrix} \quad (43)$$

$$K_0 = (\mu \quad 0 \quad \kappa_\theta \theta)' \quad (44)$$

$$K_1 = \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ 0 & -\kappa & \kappa \\ 0 & 0 & -\kappa_\theta \end{pmatrix} \quad (45)$$

$$\sigma(X_t) \sigma(X_t)' = H_0 + \sum_{i=1}^3 H_i \cdot X_t(i) \quad (46)$$

$$H_0 = \mathbf{0}_{3 \times 1} \quad (47)$$

$$H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (48)$$

$$H_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{pmatrix} \quad (49)$$

Therefore, $\mathbb{E}_t e^{u' X_t} = e^{\alpha(u,t,T) + \beta'(u,t,T) X_t}$, where α and β solve a system of ordinary differential equations (see [Duffie et al., 2000](#)), given by

$$\frac{\partial \alpha}{\partial t} = -K_0' \beta - l_0(\varrho(\beta_2) - 1) \quad (50)$$

$$\frac{\partial \beta}{\partial t} = -K_1' \beta - \frac{1}{2} \beta' H \beta \quad (51)$$

$$\alpha(u, T, T) = 0 \quad (52)$$

$$\beta(u, T, T) = u \quad (53)$$

$$\varrho(\beta_2) = \mathbb{E} e^{\beta_2 \xi_t} = \frac{1}{1 - \beta_2 \mu_\xi}. \quad (54)$$

The analytical solutions for $\beta_1(u, t, T)$ and $\beta_2(u, t, T)$ can be obtained as

$$\beta_1(u, t, T) = u(1), \quad (55)$$

$$\beta_2(u, t, T) = a_2(u) + (a_2(u) - a_1(u)) \frac{C(t, u)}{1 - C(t, u)}, \quad (56)$$

while,

$$a_2(u) = \frac{\kappa - \sqrt{\kappa^2 - \sigma_v^2(u(1)^2 - u(1))}}{\sigma_v^2}, \quad (57)$$

$$a_1(u) = \frac{\kappa + \sqrt{\kappa^2 - \sigma_v^2(u(1)^2 - u(1))}}{\sigma_v^2}, \quad (58)$$

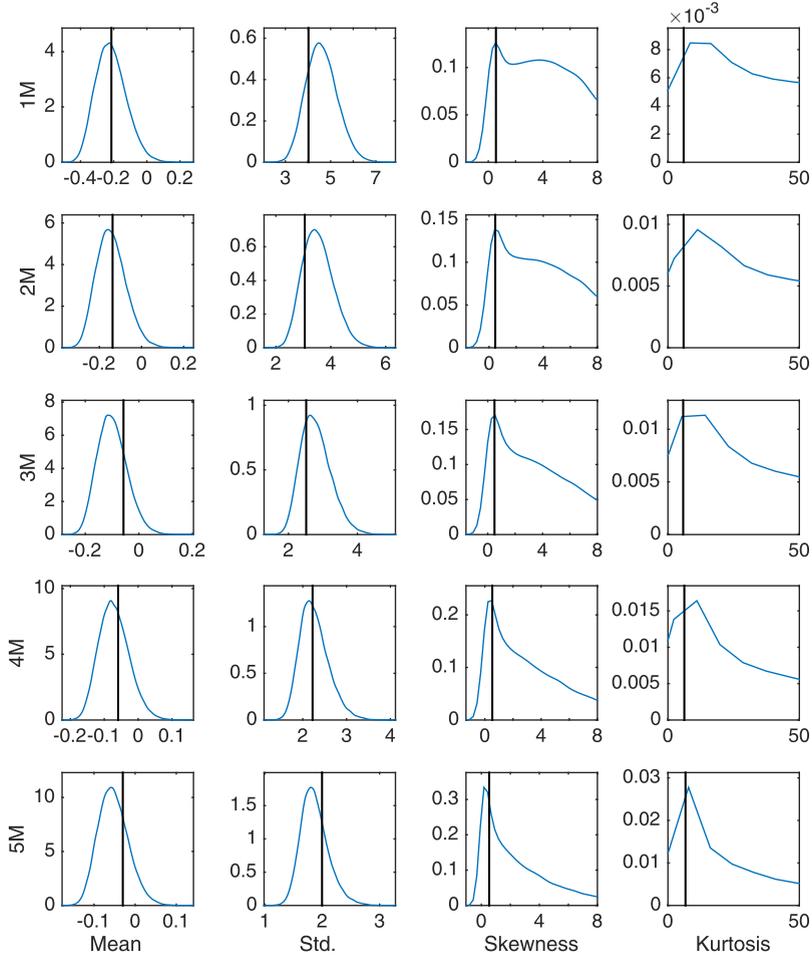


Fig. 13. Sampling distributions from the two-factor volatility model vs. data: We simulate a panel of futures returns, compute the time-series moments and plot the cross section distributions of the moments under the two-factor model. The small-sample densities are estimated by kernel densities and compared to the sample moments in the data, shown in the vertical bars.

$$c_1(u) = \frac{u(2) - a_2(u)}{u(2) - a_1(u)}, \quad (59)$$

$$c_2(u) = \frac{(a_2(u) - a_1(u))\sigma_v^2}{2}, \quad (60)$$

$$C(t, u) = c_1 e^{c_2(u)(T-t)}. \quad (61)$$

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

$$\beta_3(u) = \frac{\kappa_\theta - \sqrt{\kappa_\theta^2 - 2\kappa a_2(u)\sigma_\theta^2}}{\sigma_\theta^2}. \quad (62)$$

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

A3. Equilibrium stock returns

Let P_t denote the price of the risky asset at date t and s be the number of shares the representative agent holds.

The equilibrium asset price can be derived by solving the optimal portfolio problem

$$\max_s \mathbb{E}_t u(s\tilde{x}_T - (s-1)P_t e^{r_f(T-t)}). \quad (63)$$

From the first order condition and the fixed supply $s^* = 1$ of the risky asset we find that its price is

$$\begin{aligned} P_t &= \frac{\mathbb{E}_t u'(\tilde{x}_T)\tilde{x}_T e^{-r_f T}}{\mathbb{E}_t u'(\tilde{x}_T) e^{-r_f t}} = \frac{\mathbb{E}_t \tilde{x}_T^{1-\gamma} e^{-r_f T}}{\mathbb{E}_t \tilde{x}_T^{-\gamma} e^{-r_f t}} \\ &= \exp\{\alpha(u_{1-\gamma}, t, T) - \alpha(u_{-\gamma}, t, T) + (\beta'(u_{1-\gamma}, t, T) - \beta'(u_{-\gamma}, t, T))X_t - r_f(T-t)\} \end{aligned} \quad (64)$$

$$u_{1-\gamma} = (1-\gamma \quad 0 \quad 0)', \quad u_{-\gamma} = (-\gamma \quad 0 \quad 0)'. \quad (65)$$

If we take the limit $T \rightarrow \infty$, then β_s do not depend on t . Define

$$\lambda_\sigma = \beta_2(u_{1-\gamma}) - \beta_2(u_{-\gamma}), \quad (66)$$

$$\lambda_\theta = \beta_3(u_{1-\gamma}) - \beta_3(u_{-\gamma}). \quad (67)$$

Then

$$\ln P_t = \alpha(u_{1-\gamma}, t, T) - \alpha(u_{-\gamma}, t, T) - r_f(T - t) + \ln x_t + \lambda_\sigma \sigma_t^2 + \lambda_\theta \theta_t. \quad (68)$$

The dynamics of the stock price are

$$d \ln P_t = \left(\frac{\partial \alpha(u_{1-\gamma}, t, T)}{\partial t} - \frac{\partial \alpha(u_{-\gamma}, t, T)}{\partial t} \right) dt + r_f dt + d \ln x_t + \lambda_\sigma d\sigma_t^2 + \lambda_\theta d\theta_t \quad (69)$$

$$= r_f dt - (\kappa_\theta \theta \lambda_\theta + l_0 (\varrho(\beta_2(u_{1-\gamma})) - \varrho(\beta_2(u_{-\gamma})))) dt - \frac{1}{2} \sigma_t^2 dt + \sigma_t dB_t^x + \lambda_\sigma d\sigma_t^2 + \lambda_\theta d\theta_t. \quad (70)$$

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

$$d \ln P_t = r_f dt + \sigma_t dB_t^x + \lambda_\sigma \sigma_t dB_t^v + \lambda_\theta \sigma_\theta \sqrt{\theta_t} dB_t^\theta - \left(\frac{1}{2} + \kappa \lambda_\sigma \right) \sigma_t^2 dt + (\kappa \lambda_\sigma - \kappa_\theta \lambda_\theta) \theta_t dt - l_0 (\varrho(\beta_2(u_{1-\gamma})) - \varrho(\beta_2(u_{-\gamma}))) dt + \lambda_\sigma \xi_t dN_t \quad (71)$$

A4. VIX futures prices

By definition we have

$$VIX_t^2 = \text{Var}_t^Q(\ln P_{t+21}). \quad (72)$$

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

$$\begin{aligned} \Phi(u) &= \ln \mathbb{E}_t^Q e^{u \ln P_{t+21}} \\ &= \ln \mathbb{E}_t^Q e^{u \lambda_X' X_{t+21}} \\ &= \alpha(u \lambda_X, t, t+21) + \beta'(u \lambda_X, t, t+21) X_t \end{aligned} \quad (73)$$

$$\lambda_X := (1, \lambda_\sigma, \lambda_\theta)'. \quad (74)$$

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

$$\begin{aligned} b(\tau, \gamma) &= \frac{1}{\kappa_\gamma} + \frac{(1+2\gamma)^2 \sigma_v^2}{4(\kappa_\gamma)^3} \\ &\quad + \frac{e^{-\kappa_\gamma \tau} \sigma_v^2}{\kappa_\gamma^2} \lambda_\sigma (1 - e^{-\kappa_\gamma \tau}) (1 + 2\gamma + \lambda_\sigma \kappa_\gamma) \\ &\quad - \frac{e^{-\kappa_\gamma \tau} \sigma_v^2}{\kappa_\gamma^2} \left((1+2\gamma)^2 \left(\frac{\tau}{2} + \frac{e^{-\kappa_\gamma \tau}}{4\kappa_\gamma} \right) \right. \\ &\quad \left. + \kappa_\gamma (1+2\gamma) \tau \lambda_\sigma + \frac{\kappa_\gamma}{\sigma_v^2} \right) \end{aligned} \quad (75)$$

$$\kappa_\gamma = \sqrt{\kappa^2 - \sigma_v^2 (\gamma^2 + \gamma)} \quad (76)$$

The solution of b under P measure is a special case of letting $\gamma = 0$ and the VIX associated b is a special case of letting $\tau = 21$ in the above expression.

We adopt the analytical formula for VIX futures up to an integral by [Zhu and Lian \(2012\)](#), and price VIX futures by numerical integration:

$$\begin{aligned} F_t^{VIX}(t + \tau) &= \mathbb{E}_t^Q \sqrt{VIX_{t+\tau}^2} \\ &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \mathbb{E}_t^Q e^{-s VIX_{t+\tau}^2}}{s^{3/2}} ds \\ &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-as} \mathbb{E}_t^Q e^{u_1(s)' X_{t+\tau}}}{s^{3/2}} ds \\ &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-as} e^{\alpha(u_1(s), t, t+\tau) + \beta'(u_1(s), t, t+\tau) X_t}}{s^{3/2}} ds \\ &= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^\infty e^{-0.5s} (1 - e^{-aes} e^{\alpha(u_2(s), t, t+\tau) + \beta'(u_2(s), t, t+\tau) X_t}) ds, \end{aligned} \quad (77)$$

where

$$u_1(s) = (0, -bs, -cs)' \quad (78)$$

$$u_2(s) = (0, -be^s, -ce^s)'. \quad (79)$$

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

A5. MCMC details

Our MCMC estimator requires a Gibbs scheme where we cycle through draws from the conditional distributions for σ_t^2 , jump-times, N_t , jump sizes, ξ_t , and parameters, Θ . The respective conditional distributions are known to be proportional to the joint posterior, which is given by

$$\begin{aligned} p(\Sigma, \Xi, \mathcal{N}, \Theta | \mathcal{R}_T) &\propto \prod_{t=1}^T p(r_t, \sigma_t^2 | \xi_t, \sigma_{t-1}^2, \Theta) p(\xi_t | \Theta) p(\xi_t) l(\sigma_t) p(\Theta). \end{aligned} \quad (80)$$

We can now sample elements of Σ , Ξ , N , and Θ in a sequence of Metropolis draws. However, rather than sampling directly from this posterior we prefer to first transform the volatility processes through a log transform. Define $v_t = \ln \sigma_t^2$ and $\eta_t = \ln(\sigma_t^2 + \xi_t) - \ln \sigma_t^2$. The joint dynamics of r_t and v_t can be found through Ito's lemma. The Euler discretization is

$$r_t = \lambda_0 dt + \lambda_\sigma (\exp(v_t) - \exp(-v_{t-1})) + \exp(v_{t-1}) \epsilon_t^x \quad (81)$$

$$\begin{aligned} v_t &= v_{t-1} + \left(\kappa \theta_v - \frac{1}{2} \sigma_v^2 \right) \exp(-v_{t-1}) - \kappa - \eta_t dN_t \\ &\quad + \exp(-v_{t-1}) \sigma_v \epsilon_t^\sigma. \end{aligned} \quad (82)$$

A bivariate Gaussian specification for ϵ_t^x and ϵ_t^σ now implies the conditional density $p(r_t, v_t | v_{t-1}, \eta_t, dN_t)$. The

conditional posterior for η is

$$p(\eta_t | \Sigma, \mathcal{N}, \mathcal{R}_T, \Theta) \propto p(r_t, v_t | v_{t-1}, \eta_t, dN_t) p(\eta_t | \sigma_t^2) I_{\eta > 0}, \quad (83)$$

where

$$p(\eta_t | \sigma_t^2, \Theta) = \exp\left(-(\exp(\eta) - 1) \frac{\sigma_t^2}{\mu_v} + \eta_t\right) \frac{\sigma_t^2}{\mu_v}. \quad (84)$$

We draw η by proposing from a normal with mean $\mu_\eta^0 = (\mu_2 \omega_1^2 + \mu_1 \omega_2^2) / (\omega_1^2 + \omega_2^2)$, where $\mu_1 = \ln(\mu_v / \sigma_t^2)$, $\mu_2 = v_t - v_{t-1} - (\kappa \theta_v - \frac{1}{2} \sigma_v^2) e^{-v_{t-1}} - \kappa$, $\omega_2 = \sigma_v \exp(-\frac{1}{2} v_{t-1})$, and ω_1 is a constant. This proposal density approximates the target density through a normal approximation of $p(\eta_t | \cdot)$ in (83).

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

$$G(i) = \phi(v_t | v_{t-1}, \eta_t, dN_t = i, \Theta) p(\eta_t | \sigma_t^2, \Theta) p(dN_t = i), \quad (85)$$

where ϕ is a Gaussian density, then

$$p(dN_t = 1) = \frac{G(1)}{G(1) + G(0)} \quad (86)$$

is the binomial probability of $dN_t = 1$.

Finally, our algorithm requires simulating from

$$p(v_t | v_{t-1}, v_{t+1}, \eta_t, dN_t, \Theta) \propto p(r_t, v_t | v_{t-1}, \eta_t, dN_t) p(r_{t+1}, v_{t+1} | v_t, \eta_{t+1}, dN_{t+1}) \quad (87)$$

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

Fig. 14 shows the behavior of the MCMC sampler for 40,000 posterior draws of γ using different starting values. The figure suggests the ideal behavior of MCMC chains: as the sample size increases the respective posterior draws appear to settle in a subset of the parameter space, suggesting that they are drawn from the same stationary distribution.

A6. Proof of Proposition 1

Proof. From Appendix A.3, we know

$$\begin{aligned} P_t &= \frac{\mathbb{E}_t(X_T^{1-\gamma} e^{-r_f T})}{\mathbb{E}_t(X_T^{-\gamma} e^{-r_f T})} \\ &= \mathbb{E}_t \frac{\mathbb{E}_{t+\tau}(X_T^{1-\gamma} e^{-r_f T})}{\mathbb{E}_t(X_T^{-\gamma} e^{-r_f T})} \\ &= \mathbb{E}_t \left(\frac{\mathbb{E}_{t+\tau}(X_T^{-\gamma} e^{-r_f(t+\tau)})}{\mathbb{E}_t(X_T^{-\gamma} e^{-r_f t})} \frac{\mathbb{E}_{t+\tau}(X_T^{1-\gamma} e^{-r_f T})}{\mathbb{E}_{t+\tau}(X_T^{-\gamma} e^{-r_f(t+\tau)})} \right) \end{aligned}$$

$$\begin{aligned} &= \mathbb{E}_t \left(\frac{\mathbb{E}_{t+\tau}(X_T^{-\gamma} e^{-r_f(t+\tau)})}{\mathbb{E}_t(X_T^{-\gamma} e^{-r_f t})} P_{t+\tau} \right) \\ &= \mathbb{E}_t \left(\frac{M_{t+\tau}}{M_t} P_{t+\tau} \right). \end{aligned} \quad (88)$$

Therefore, the stochastic discount factor is

$$M_t = \mathbb{E}_t X_T^{-\gamma} e^{-r_f t} = e^{\alpha(u_{-\gamma, t, T}) + \beta'(u_{-\gamma, t, T}) X_t - r_f t}. \quad (89)$$

Following a similar approach for the dynamics of the stock price in Appendix A.3, we derive the dynamics of the pricing kernel as

$$\begin{aligned} d \ln M_t &= -r_f dt - \left(\frac{1}{2} \gamma^2 + \frac{1}{2} \sigma_v^2 \eta_2^2 \right) \sigma_t^2 dt \\ &\quad - \gamma \sigma_t dB_t^x - \eta \sigma_v \sigma_t dB_t^v \\ &\quad - l_0(\varrho(-\eta) - 1) dt - \eta \xi_t dN_t \end{aligned} \quad (90)$$

$$\eta := -\beta_2(u_{-\gamma}). \quad (91)$$

It follows that

$$\frac{dM_t^-}{M_t^-} = -r_f dt - l_0(\varrho(-\eta) - 1) dt - \gamma \sigma_t dB_t^x - \eta \sigma_v \sigma_t dB_t^v \quad (92)$$

and from the definition of $d \ln M_t^-$, we know $\ln M_t - \ln M_t^- = -\eta \xi_t dN_t$, therefore

$$\frac{M_t - M_t^-}{M_t^-} = e^{-\eta \xi_t dN_t} - 1 = (e^{-\eta \xi_t} - 1) dN_t. \quad (93)$$

Adding the above two equations together, we get

$$\begin{aligned} \frac{dM_t}{M_t} &= -r_f dt - l_0(\varrho(-\eta) - 1) dt + (e^{-\eta \xi_t} - 1) dN_t \\ &\quad - \gamma \sigma_t dB_t^x - \eta \sigma_v \sigma_t dB_t^v. \end{aligned} \quad (94)$$

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

$$dB_t^x = dB_t^{x, Q} - \gamma \sigma_t dt \quad (95)$$

$$dB_t^v = dB_t^{v, Q} - \eta \sigma_v \sigma_t dt \quad (96)$$

$$dX_t = (K_0^Q + K_1^Q X_t) dt + \sigma(X_t) dB_t^Q + \xi_t^Q dN_t^Q \quad (97)$$

$$K_0^Q = K_0 \quad (98)$$

$$K_1^Q = \begin{pmatrix} 0 & -\frac{1}{2} - \gamma \\ 0 & -\kappa - \eta \sigma_v^2 \end{pmatrix} \quad (99)$$

$$l_0^Q = l_0 \varrho(-\eta) \quad (100)$$

$$\mu_\xi^Q = \mu_\xi \varrho(-\eta). \quad (101)$$

The variance dynamics under the Q measure as in Eq. (19) can be easily derived from Eq. (97). Under the Q

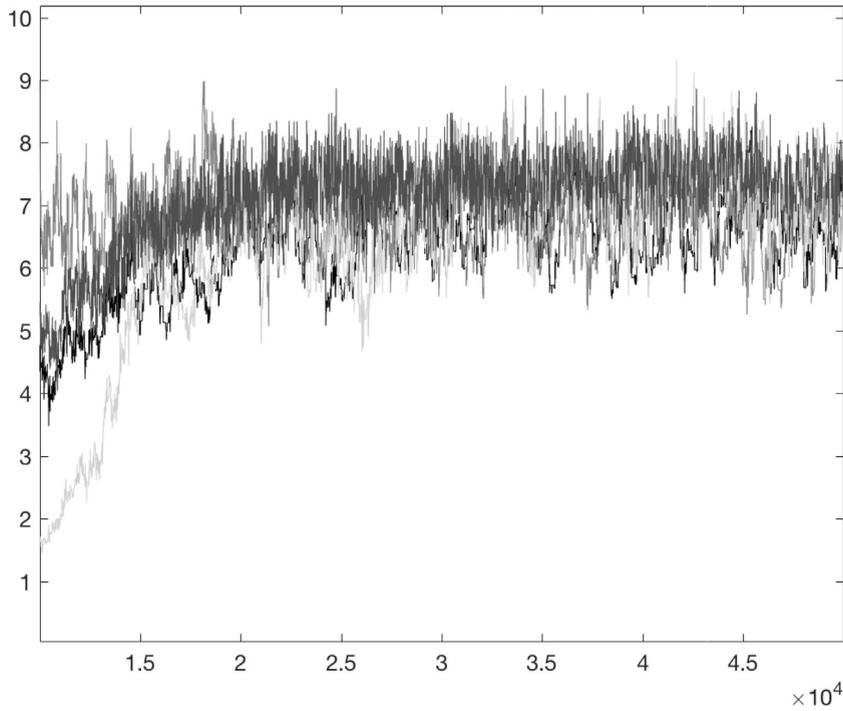


Fig. 14. Posterior MCMC draws of the risk aversion parameter γ using different starting values.

measure, state variables are still affine and thus the moment generating function of future state vector can be derived similarly as under the objective measure as in Appendix A.2.

By plugging in the expressions for the diffusions under Q in Eqs. (95) and (96) to Eq. (71), we recover the dynamics of the stock price under Q as

$$\begin{aligned}
 d \ln P_t &= r_f dt + \sigma_t dB_t^{x,Q} + \lambda_\sigma \sigma_v \sigma_t dB_t^{v,Q} \\
 &+ \left(-\frac{1}{2} - \frac{1}{2} \lambda_\sigma^2 \sigma_v^2 \right) \sigma_t^2 dt \\
 &- l_0 (\varrho(\beta_2(u_{1-\gamma})) - \varrho(\beta_2(u_{-\gamma}))) dt + \lambda_\sigma \xi_t^Q dN_t^Q.
 \end{aligned}
 \tag{102}$$

Realize

$$\begin{aligned}
 &l_0 (\varrho(\beta_2(u_{1-\gamma})) - \varrho(\beta_2(u_{-\gamma}))) \\
 &= l_0 \varrho(\beta_2(u_{-\gamma})) \left(\frac{\varrho(\beta_2(u_{1-\gamma}))}{\varrho(\beta_2(u_{-\gamma}))} - 1 \right) \\
 &= l_0^Q \left(\frac{1}{\frac{1 - \beta_2(u_{1-\gamma})\mu_\xi}{1 - \beta_2(u_{-\gamma})\mu_\xi}} - 1 \right) \\
 &= l_0^Q \left(\frac{1}{1 - \frac{\beta_2(u_{1-\gamma})\mu_\xi - \beta_2(u_{-\gamma})\mu_\xi}{1 - \beta_2(u_{-\gamma})\mu_\xi}} - 1 \right) \\
 &= l_0^Q \left(\frac{1}{1 - \frac{\lambda_\sigma \mu_\xi}{1 - \beta_2(u_{-\gamma})\mu_\xi}} - 1 \right)
 \end{aligned}$$

$$\begin{aligned}
 &= l_0^Q \left(\frac{1}{1 - \lambda_\sigma \mu_\xi^Q} - 1 \right) \\
 &= l_0^Q (\varrho^Q(\lambda_\sigma) - 1).
 \end{aligned}
 \tag{103}$$

By substituting back into Eq. (102) using a similar approach as we did to get the dynamics of $\frac{dM_t}{M_t}$ from $\ln M_t$, we get

$$\begin{aligned}
 \frac{dP_t}{P_t} &= r_f dt + \sigma_t dB_t^{x,Q} + \lambda_\sigma \sigma_v \sigma_t dB_t^{v,Q} + (e^{\lambda_\sigma \xi_t^Q} - 1) dN_t^Q \\
 &- l_0^Q (\varrho^Q(\lambda_\sigma) - 1) dt.
 \end{aligned}
 \tag{104}$$

The stock has expected return equal to risk free rate under measure Q . \square

A7. Proof of Proposition 2

Proof. For the first claim: Note that the unconditional mean of σ_t^2 is $\mathbb{E}(\sigma_t^2) = \theta + l_0 \mu_\xi / \kappa$. The slope of the futures curve is determined by the last conditional expectation term in (31), as both $a(21)$ and $b(21)$ are positive. This conditional expectation is

$$F(\tau) = \sigma_t^2 e^{-\kappa \tau} + \left(\theta^Q + \frac{\mu_\xi^Q l_0^Q}{\kappa^Q} \right) (1 - e^{-\kappa \tau}),
 \tag{105}$$

whose slope is given by

$$\frac{dF(\tau)}{d\tau} = -\kappa^Q \sigma_t^2 e^{-\kappa \tau} + \left(\theta^Q + \frac{\mu_\xi^Q l_0^Q}{\kappa^Q} \right) \kappa^Q e^{-\kappa \tau}.
 \tag{106}$$

Substituting in the steady-state $\sigma_t^2 = \theta + \frac{\mu_\xi l_0}{\kappa}$ we find that the sign of $\frac{dF(\tau)}{d\tau}$ is determined by

$$-\theta - \frac{\mu_\xi l_0}{\kappa} + \theta^Q + \frac{\mu_\xi^Q l_0^Q}{\kappa^Q} \geq 0 \tag{107}$$

where the equality holds if and only if $\gamma = 0$.

For the second claim, we need to prove

$$\mathbb{E}^Q(\sigma_{t+\tau}^2 | \sigma_t^2) > \mathbb{E}^P(\sigma_{t+\tau}^2 | \sigma_t^2) \tag{108}$$

for all $\sigma_t^2 \in \mathbb{R}^+$. Define $\bar{\theta}$ and $\bar{\theta}^Q$ to be the unconditional means of the process under the respective measures. The inequality in (108) can be written

$$\sigma_t^2 e^{-\kappa^Q \tau} + \bar{\theta}^Q (1 - e^{-\kappa^Q \tau}) - (\sigma_t^2 e^{-\kappa \tau} + \bar{\theta} (1 - e^{-\kappa \tau})) > 0. \tag{109}$$

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

$$\bar{\theta}^Q (1 - e^{-\kappa^Q \tau}) - \bar{\theta} (1 - e^{-\kappa \tau}) > 0 \tag{110}$$

which is equivalent to

$$\bar{\theta}^Q - \bar{\theta} > \bar{\theta}^Q e^{-\kappa^Q \tau} - \bar{\theta} e^{-\kappa \tau}. \tag{111}$$

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

$$f'(\tau) = -\kappa^Q \bar{\theta}^Q e^{-\kappa^Q \tau} + \kappa \bar{\theta} e^{-\kappa \tau}. \tag{112}$$

Recall

$$\begin{aligned} \bar{\theta}^Q &= \theta^Q + \frac{l_0^Q \mu_\xi^Q}{\kappa^Q} \\ &= \frac{\theta \kappa}{\kappa^Q} + \frac{l_0^Q \mu_\xi^Q}{\kappa^Q} \end{aligned} \tag{113}$$

$$\bar{\theta} = \theta + \frac{l_0 \mu_\xi}{\kappa}. \tag{114}$$

Substituting back to Eq. (112), we get

$$\begin{aligned} f'(\tau) &= -(\kappa \theta + l_0^Q \mu_\xi^Q) e^{-\kappa^Q \tau} + (\kappa \theta + l_0 \mu_\xi) e^{-\kappa \tau} \\ &= \kappa \theta (e^{-\kappa \tau} - e^{-\kappa^Q \tau}) + (l_0 \mu_\xi e^{-\kappa \tau} - l_0^Q \mu_\xi^Q e^{-\kappa^Q \tau}). \end{aligned} \tag{115}$$

Since $\kappa^Q < \kappa$, $l_0^Q > l_0$ and $\mu_\xi^Q > \mu_\xi$, it is easy to see that $f'(\tau) < 0$, therefore $f(0) > f(\tau)$ for all τ , which is exactly the inequality (111) we need to prove.

For the third claim, note that the price can be written

$$\begin{aligned} d \ln P_t &= -(\mu + \kappa \theta \lambda_\sigma + l_0 (\varrho(\beta_2(u_{1-\gamma})) \\ &\quad - \varrho(\beta_2(u_{-\gamma})))) dt + r_f dt + d \ln x_t + \lambda_\sigma d\sigma_t^2 \end{aligned} \tag{116}$$

while the VIX-squared futures is

$$\begin{aligned} dF_t^{\text{VIX}^2}(\tau) &= b(21) dF(\tau) \\ &= b(21) e^{-\kappa^Q \tau} d\sigma_t^2 + b(21) e^{-\kappa^Q \tau} \kappa^Q \sigma_t^2 dt. \end{aligned} \tag{117}$$

The instantaneous covariance of the VIX-squared with the market is therefore

$$E_t \left\{ dF_t^{\text{VIX}^2}(\tau) d \ln P_t \right\} = \lambda_\sigma b(21) e^{-\kappa^Q \tau} E_t \left((d\sigma_t^2)^2 \right) dt \tag{118}$$

whose sign is determined by the sign of λ_σ . \square

References

- Ait-Sahalia, Y., Fan, J., Li, Y., 2013. The leverage effect puzzle: disentangling sources of bias at high frequency. *J. Financial Econ.* 109, 224–249.
- Alexander, C., Korovilas, D., 2012. Volatility exchange-traded notes: curse or cure? Available at SSRN 2062854.
- Andersen, T., Todorov, V., 2013. The risk premia imbedded in index options. Working Paper Kellogg.
- Andersen, T.G., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time models for equity returns. *J. Finance* 57, 1239–1284.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *J. Finance* 52, 2003–2049.
- Bakshi, G., Kapadia, N., 2003. Delta hedged gains and the negative volatility risk premium. *Rev. Financial Stud.* 16, 527–566.
- Bandi, F., Reno, R., 2012. Time-varying leverage effects. *J. Econom.* 196, 94–113.
- Bandi, F., Reno, R., 2016. Price and volatility co-jumps. *J. Financial Econ.* 119, 107–146.
- Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. *J. Finance* 59, 1481–1509.
- Bates, D.S., 2000. Post-'87 crash fears in S&P 500 futures options. *J. Econom.* 94, 181–238.
- Blume, M.E., 1974. Unbiased estimators of long-run expected rates of return. *J. Am. Stat. Assoc.* 69 (347), 634–638. doi:10.1080/01621459.1974.10480180.
- Bollerslev, T., Mikkelsen, H.O., 1996. Modeling and pricing long memory in stock market volatility. *J. Econom.* 73, 151–184.
- Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. *Rev. Financial Stud.* 22 (11), 4463–4492.
- Bondarenko, O., 2003. Why are put options so expensive? working paper, University of Illinois.
- Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: evidence from futures options. *J. Finance* 62 (3), 1453–1490.
- Buraschi, A., Trojani, F., Vedolin, A., 2014. When uncertainty blows in the orchard: comovement and equilibrium volatility risk premia. *J. Finance* 69 (1), 101–137.
- Campbell, J., 1991. A variance decomposition for stock returns. *Econ. J.* 101, 157–179.
- Campbell, J.Y., Lo, A., Mackinlay, C.A., 1997. *The Econometrics of Financial Markets*.
- Carr, P., Wu, L., 2006. A tale of two indices. *J. Derivatives Spring*, 13–29.
- Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. *J. Econom.* 116, 225–257.
- Chernov, M., Ghysels, E., 2000. Towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *J. Financial Econ.* 56, 407–458.
- Dew-Becker, I., Giglio, S., Le, A., Rodriguez, M., 2015. The price of variance risk. Technical Report. National Bureau of Economic Research.
- Drechsler, I., Yaron, A., 2011. What's vol got to do with it. *Rev. Financial Stud.* 24, 1–45.
- Duffie, D., Pan, J., Singleton, K.J., 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.
- Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. *J. Financial Quant. Anal.* 45, 1279–1310.
- Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. *J. Bus. Econ. Stat.* 19-2, 177–191.
- Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *J. Finance* 59, 1367–1403.
- Eraker, B., 2012. The volatility premium. Working paper.
- Eraker, B., 2013. The performance of model based option trading strategies. *Rev. Derivatives Stud.* 16 (1), 1–23.
- Eraker, B., Johannes, M.J., Polson, N.G., 2003. The impact of jumps in returns and volatility. *J. Finance* 53, 1269–1300.
- Eraker, B., Shaliastovich, I., 2008. An equilibrium guide to designing affine pricing models. *Math. Finance* 18-4, 519–543.
- Eraker, B., Wang, W., 2015. Dynamic present values and the intertemporal CAPM. Working paper, University of Wisconsin.
- French, K.R., Schwert, W., Stambaugh, R.F., 1987. Expected stock returns and volatility. *J. Financial Econ.* 19 (1), 3–29.
- Gualtieri, J., Sizova, N., 2015. Extreme events in stock market fundamental factors. Working paper Rice.
- Heston, S., 1993. Closed-form solution of options with stochastic volatility with application to bond and currency options. *Rev. Financial Stud.* 6, 327–343.
- Jackwerth, J. C., Vilkov, G., 2014. Asymmetric volatility risk: evidence from option markets. Available at SSRN 2325380.
- Jacquier, E., Polson, N.G., Rossi, P., 1994. Bayesian analysis of stochastic volatility models. *J. Bus. Econ. Stat.* 12, 371–389.

- Johnson, T. L., 2015. Risk premia and the VIX term structure. Available at SSRN 2548050.
- Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with arch models. *Rev. Econ. Stud.* 65, 361–393.
- Lin, Y.-N., 2007. Pricing VIX futures: evidence from integrated physical and risk-neutral probability measures. *J. Futures Markets* 27 (12), 1175–1217.
- Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *J. Financial Econ.* 63, 3–50.
- Singleton, K.J., 2006. *Empirical Dynamic Asset Pricing—Model Specification and Econometric Assessment*. Princeton University Press.
- Song, Z., 2012. Expected VIX option returns. Working paper, Fed Reserve
- Szado, E., 2009. VIX Futures and Options - A Case Study of Portfolio Diversification During the 2008 Financial Crisis. *J. Alternative Invest* (12) 68–85.
- Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation (with discussion). *J. Am. Stat. Assoc.* 82, 528–550.
- Whaley, R.E., 2013. Trading volatility: at what cost. *J. Portfolio Manage.* 40 (1), 95–108.
- Zhang, J.E., Zhu, Y., 2006. VIX futures. *J. Futures Markets* 26, 521–531.
- Zhu, S.-P., Lian, G.-H., 2012. An analytical formula for VIX futures and its applications. *J. Futures Markets* 32 (2), 166–190.