The Risk Premia Embedded in Index Options*

Torben G. Andersen† Nicola Fusari‡ Viktor Todorov§

April 15, 2013

Abstract

We infer the dynamic relation between market risks and risk premia from an exploration of the time series of index option surfaces. The crucial step is to estimate a general parametric asset pricing model for the risk-neutral equity market dynamics using a panel of options on the S&P 500 index, while remaining fully nonparametric about the actual evolution of market risks. We find that the risk-neutral jump intensity, which controls the pricing of left tail risk, cannot be spanned by the market volatility (and its components), so an additional factor is required to account for its dynamics. This tail factor has no incremental predictive power for future equity return volatility or jumps beyond what is captured by the current and past level of volatility. In contrast, the novel factor is critical in predicting the future market excess returns over horizons up to one year, and it explains a large fraction of the future variance risk premium. We contrast our findings with those implied by structural asset pricing models that seek to rationalize the predictive power of option data. Relative to those studies, our findings suggest a wider wedge between the dynamics of market risks and the corresponding risk premia with the latter typically displaying a far more persistent reaction following market crises.

Keywords: Option Pricing, Risk Premia, Jumps, Stochastic Volatility, Return Predictability, Risk Aversion, Extreme Events.

JEL classification: C51, C52, G12.

*Andersen gratefully acknowledges support from CREATES funded by the Danish National Research Foundation. Todorov's work was partially supported by NSF Grant SES-0957330. We would like to thank Snehal Banerjee, Anna Cieslak, Ravi Jaganathan, Sergio Rebelo, Ivan Shaliastovich as well as seminar participants at Kellogg School of Management, Northwestern University, Duke University, the "Montreal 2013 Econometrics Conference: Time Series and Financial Econometrics," the 40^{th} Annual Meeting of the Danish Econometric Society, Sandbjerg, Denmark, the European University Institute, Florence, 2013 Workshop on "Measuring and Modeling Financial Risk and High Frequency Data" and the 2013 "International Conference in Financial Econometrics" at Shandong University, Jinan, China, for helpful comments.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: t-andersen@northwestern.edu.

‡Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: n-fusari@northwestern.edu.

§Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.

1 Introduction

Equity markets are subject to pronounced time-variation in volatility as well as abrupt shifts, or jumps. Moreover, these risk features are related in intricate ways, inducing a complex equity return dynamics. Hence, the markets are incomplete and derivative securities, written on the equity index, are non-redundant assets. This partially rationalizes the rapid expansion in the trading volume for contracts that offer distinct exposures to volatility and jump risks. From an economic perspective, it suggests that derivatives data provide important information for gauging the time-varying risk and risk pricing for the underlying asset. Indeed, recent evidence, stemming from parametric models, e.g., Broadie et al. (2009), as well as nonparametric studies, e.g., Bollerslev and Todorov (2011b), finds that the pricing of jump risk, implied by option data, to account for a significant fraction of the equity risk premium.

Standard no-arbitrage and equilibrium-based asset pricing models imply a tight relationship between the dynamics of the options and the underlying asset. This arises from the assumptions concerning the pricing of risk in the no-arbitrage setting and the endogenous pricing kernels implied by the equilibrium models. A prominent example is the illustrative double-jump model of Duffie et al. (2000) in which the return volatility itself follows an affine jump diffusion. In this context, the entire option surface is governed by the evolution of market volatility, i.e., the dynamics of all options is driven by a single latent Markov (volatility) process.

Recent empirical evidence reveals, however, that the dynamics of the option surface is far more complex. For example, the term structure of the volatility index, VIX, shifts over time in a manner that is incompatible with the surface being driven by a single factor. Likewise, Bates (2000) documents that a two-factor stochastic volatility model for the risk-neutral market dynamics provides a significant improvement over a one-factor version. Moreover, Bollerslev and Todorov (2011b) find that even the short-term option dynamics cannot be captured adequately by a single factor as the risk-neutral tails display independent variation relative to market volatility, thus driving a wedge between the dynamics of option surface and underlying asset prices.

The objective of the current paper is to characterize the risk premia, implied by the large panel of S&P 500 index options, and its relation with the aggregate market risks in the economy. As discussed in Andersen et al. (2012), the option panel contains rich information both for the evolution of volatility and jump risks and their pricing. Consequently, we let the option data speak for themselves in determining the risk premium dynamics and discriminating among alternative hypothesis regarding the source of variation in risk as well as risk pricing.

The standard approach in the no-arbitrage framework is to specify and estimate the risk premia

starting from a parametric model for the evolution of the underlying asset, see, e.g., Singleton (2006), Chapter 15, for a review. Risk premia are then introduced through a pricing kernel that ensures the implied risk-neutral dynamics lies within the same parametric class as the model under the statistical measure, but with parameter shifts that reflect the risk compensation. This approach tends to tie the stock market and option surface dynamics closely together. In particular, as in existing equilibrium-based models, equity risk premia are typically linear in equity risk, i.e., volatility and its components. In contrast, we find the options to display risk price variation that is largely unrelated to, and effectively impossible to identify from, the underlying asset prices alone.

This motivates our "reverse" approach of starting with a parametric model for the risk-neutral dynamics and estimating it exclusively from option data along with no-arbitrage restrictions based on model-free volatility measures constructed from the underlying asset data. In this manner, we avoid letting a (possibly misspecified) parametric structure for the $\mathbb{P}$ dynamics impact the identification of option risk premia. We then explore the risk premia dynamics by combining the extracted state vector with high-frequency data on the equity index.

We document that even a very general two-factor stochastic volatility model with jumps both in price and volatility as well as a time-varying jump intensity produces systematic biases in the fit of the option surface dynamics. The problems are particularly acute following periods like the Asian crisis in 1997 or the great recession originating in the Fall of 2008. After these events, market volatility reverted back towards its pre-crisis level relatively quickly, while the out-of-the-money put options remained expensive, inducing a pronounced steepening of the short-maturity (Black-Scholes) implied volatility curve for lower strikes. This type of variation in the option surface is difficult to generate within standard models (both no-arbitrage or equilibrium-based) as the priced jump tail risk, which in turn determines the out-of-the-money put prices, in these models is determined solely by the (components of) market volatility. This feature implies, in particular, that standard models tend to generate risk premia that are too low in the aftermath of crises.

We follow Andersen et al. (2012) and alleviate the shortcomings of the two-factor stochastic volatility model by introducing a third factor that drives the risk-neutral jump intensity. This novel tail factor is not part of the volatility dynamics although it may be correlated with the level of volatility. We model it as purely jump driven, with one component jointly governed by the volatility jumps while another component is independent of the volatility process. This feature allows the jump intensity to escalate – through so-called self-excitation of the jumps – in periods of crises when price and volatility jumps are prevalent, thus "magnifying" the response of the risk-neutral jump intensity following major (negative) market shocks. Relative to Andersen et al. (2012), we estimate

the model by minimizing, not the squared, but the *relative* squared option pricing error across the full sample. This reduces the weight assigned to highly turbulent periods where the bid-ask spreads and pricing errors increase sharply. We find that the tail factor improves the characterization of the option surface dynamics very significantly. In particular, the new model no longer undervalues short-maturity out-of-the-money puts in the aftermaths of crises. Hence, our extended risk-neutral model provides a more suitable basis for studying the dynamic properties of market risk premia.

In turn, the presence of this independently evolving tail factor implies that part of the risk premium dynamics cannot be captured by the state variables driving the underlying asset price dynamics. This suggests that current risk premia have predictive power for the evolution of future risk premia over and above what is implied by the volatility factors. This is indeed what we find. The novel tail factor is significant in forecasting future excess market returns for horizons up to one year while the volatility factors are insignificant. Similarly, the new tail factor is important for predicting the future variance risk premium in conjunction with one of the volatility components. Taken together, our findings rationalize why the variance risk premium provides superior forecasts for future returns relative to volatility itself, as documented in Bollerslev et al. (2009). The key is the existence of the new factor driving the left jump tails of the risk-neutral distribution.

Importantly, while the new jump factor has predictive power for risk premia, it contains no incremental information regarding the future evolution of volatility and jump risks for the underlying asset relative to the traditional volatility factors. Hence, our findings indicate that option markets embody critical information about the market risk premia and its dynamics which is essentially unidentifiable from stock market data alone. Moreover, the option surface dynamics contains information that can improve the modeling and forecasting of future volatility and jump risks, but such applications necessitate an initial step of disentangling the components of the risk premia that are independent from the volatility process. Overall, our empirical results suggest that there is a wedge between the stochastic evolution of risks in the economy and their pricing, the latter typically having a far more persistent response to (negative) tail events than the former.

Our finding of a substantial wedge between the dynamics of the option and stock markets presents a challenge for traditional structural asset pricing models. Specifically, the standard exponentially-affine equilibrium models with a representative agent equipped with Epstein-Zin preferences imply that the ratio of the risk-neutral and statistical jump tails is constant. On the contrary, the new factor, extracted from the option data, drives the risk-neutral jump tail but has no discernable impact on the statistical jump tail. We conjecture that this wider gap between fundamentals and asset prices may be accounted for through an extension of the preferences via

some form for time-varying risk aversion and/or ambiguity aversion towards extreme downside risk.

The rest of the paper is organized as follows. Section 2 describes our data and estimation method. Section 3 presents the results from fitting the option panel using a general version of the standard two-factor stochastic volatility asset pricing model along with our new extension. In Section 4 we exploit the estimation results from Section 3 to study the risk premium dynamics and its implication for return and variance predictability. Section 5 concludes. In the Appendix we show that our specific extension of the standard model, focusing on the jump intensity in the left tail, is crucial for the striking improvement in fit. If we simply add another stochastic volatility factor to the traditional model, the performance cannot match what we achieve in Section 3.

2 Data and Estimation Methodology

We start out by briefly describing our option and underlying asset price data as well as the parametric estimation method we employ for fitting the option panel.

2.1 Data

We use European style S&P 500 equity-index (SPX) options traded at the CBOE. We exploit the closing bid and ask prices reported by OptionMetrics, applying standard filters and discarding all in-the-money options, options with time-to-maturity of less than 7 days, as well as options with zero bid prices. For all remaining options, we compute the mid bid-ask Black-Scholes implied volatility. The data spans January 1, 1996, till July 21, 2010. Following earlier empirical work, e.g., Bates (2000) and Broadie et al. (2009), we sample every Wednesday.[1] The sample includes 760 trading days, and the estimation is based on an average of 234 bid-ask quotes per day. The nonparametric estimate of volatility needed for the penalization of the objective function below is constructed from one-minute high-frequency data on the S&P 500 futures covering the time span of the options. The same data set is used to construct measures of volatility and jump risks for the predictive regressions in Section 4. Finally, we also employ the returns on the SPY ETF traded on the NYSE, which tracks the S&P 500 index portfolio, and the 3 month T-bill rate to proxy for the risk-free rate, when implementing these predictive regressions.

2.2 Estimation

To define our estimator, we first introduce some notation. We denote European-style out-of-the-money option prices for the asset X at time t by $O_{t,k,\tau}$. Assuming frictionless trading in the options

[1] Due to extreme violations of no-arbitrage-conditions, we replaced October 8, 2008, with October 6, 2008.

market, the option prices are given as,

$$O_{t,k,\tau} \;=\; \begin{cases} \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau}(r_s-\delta_s)\,ds}\;(X_{t+\tau}-K)^+\right], & \text{if} \quad K > F_{t,t+\tau}\;, \\ \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau}(r_s-\delta_s)\,ds}\;(K-X_{t+\tau})^+\right], & \text{if} \quad K \le F_{t,t+\tau}\;, \end{cases} \tag{2.1}$$

where τ is the tenor, or time-to-maturity, of the option, K is the strike price, $F_{t,t+\tau}$ is the futures price for the underlying asset at time t referring to date $t+\tau$, for $\tau > 0$, $k = \ln(K/F_{t,t+\tau})$ is the log-moneyness, and r_t and δ_t are, respectively, the instantaneous risk-free interest rate and instantaneous dividend yield. Finally, we denote the Black-Scholes implied volatility corresponding to the option price $O_{t,k,\tau}$ by $\kappa_{t,k,\tau}$. This merely represents an alternative, and convenient, pricing convention for the options, as the Black-Scholes implied volatility is a deterministic and strictly monotone transformation of the ratio $\frac{e^{r_{t,t+\tau}}O_{t,k,\tau}}{F_{t,t+\tau}}$.

As for virtually all models used in option pricing, we specify the risk-neutral dynamics of X_t (as well as that of r_t and δ_t) via a Markov state vector, $\mathbf{S}_t$. In this setting, the model-implied option price at time t (quoted in Black-Scholes implied volatility) is given by $\kappa(k,\tau,\mathbf{S}_t,\theta)$, where θ denotes the parameter vector of the model.

Allowing for observation error in the option prices, we let $\overline{\kappa}_{t,k,\tau}$ represent the average of the observed bid and ask (Black Scholes option implied volatility) quotes. We estimate the day-by-day realization of the state vector as well as the model parameter vector by the following penalized weighted least squares procedure,

$$\left(\{\widehat{\mathbf{S}}_t^n\}_{t=1,\ldots,T},\,\widehat{\theta}^n\right) = \underset{\{\mathbf{Z}_t\}_{t=1,\ldots,T},\,\theta\in\Theta}{\operatorname{argmin}} \sum_{t=1}^{T}\left\{\frac{1}{N_t}\sum_{j=1}^{N_t}\frac{\left(\overline{\kappa}_{t,k_j,\tau_j}-\kappa(k_j,\tau_j,\mathbf{Z}_t,\theta)\right)^2}{\widehat{V}_t^n}+\lambda_n\frac{\left(\widehat{V}_t^{(n,m_n)}-\xi(\mathbf{Z}_t)\right)^2}{\widehat{V}_t^n}\right\}, \tag{2.2}$$

for a deterministic sequence of nonnegative numbers $\{\lambda_n\}$ (decreasing asymptotically to zero) and $\xi(\cdot)$ being the model-based map from the state vector to the spot diffusive asset volatility (this map can depend in general on the parameter vector). $\widehat{V}_t^n \equiv \widehat{V}_t^{(n,n)}$ and $\widehat{V}_t^{(n,m_n)}$ are nonparametric estimators of diffusive variation constructed from the intraday record of the log-futures price of the underlying asset, f, as follows,

$$\widehat{V}_t^{(n,m_n)} \;=\; \frac{n}{m_n}\sum_{i=tn-m_n+1}^{tn}(\Delta_i^n f)^2\,1_{\{|\Delta_i^n f|\le\alpha n^{-\varpi}\}}, \qquad \Delta_i^n f = f_{\frac{i}{n}} - f_{\frac{i-1}{n}}, \tag{2.3}$$

where $\alpha > 0$, $\varpi \in (0,1/2)$, and m_n denotes a deterministic sequence with either $m_n/n \to 0$ or $m_n = n$. When $m_n = n$, $\widehat{V}_t^{(n,m_n)}$ is the truncated variation (see e.g., Mancini (2009)), which is a

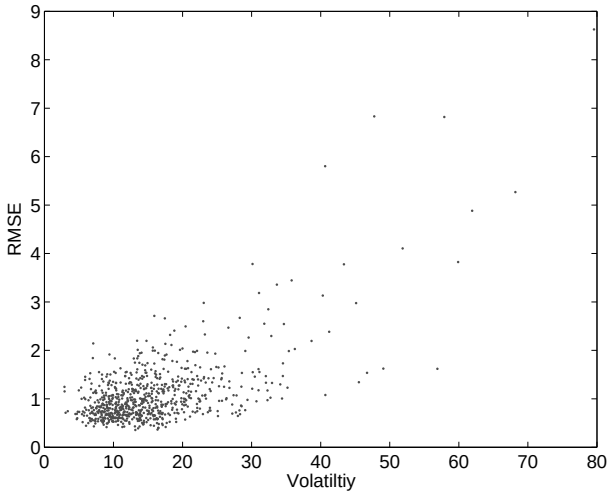


Figure 1: Option Pricing Root-Mean-Squared Errors as a Function of Volatility in the Three-Factor Model Estimated by Andersen et al. (2012).

consistent estimator of the integrated diffusive variance over the interval $[t-1,t]$. When $m_n/n \to 0$, $\widehat{V}_t^{(n,m_n)}$ is a consistent estimator of the spot variance at t and corresponds to the truncated variation computed over an (asymptotically shrinking) fraction of the day just prior to the option quote. In our implementation, we sample every minute over a 6.75 hours trading day, excluding the initial five minutes, resulting in $n = 400$. We employ $m_n = 300$ for $\widehat{V}_t^n$. Finally, the tuning parameters α and ϖ are determined in a data-driven fashion, as in Bollerslev and Todorov (2011b).

The estimator in (2.2) is based on minimizing the weighted mean squared error in fitting the panel of observed option implied volatilities, with a penalization term that reflects how much the model-implied spot volatility deviates from a model-free volatility estimate. The presence of $\widehat{V}_t^{(n,m_n)}$ in the objective function serves as a regularization device that helps identify the parameter vector by penalizing values that imply "unreasonable" volatility levels. Moreover, the standardization by $\widehat{V}_t^n$ allows us to weigh option observations on high and low volatility days differently to improve efficiency by accounting for the fact that the option error generally increases with market volatility. This is the point where our estimation procedure deviates from the original approach in Andersen et al. (2012). The rationale for our weighting scheme is apparent from Figure 1. It demonstrates that the option pricing errors are roughly linear in market volatility, thus implying a natural scaling of the errors by the estimated end-of-day spot volatility.

Andersen et al. (2012) develops the limiting distribution for the estimator based on an in-fill asymptotic scheme. The estimator is consistent for the parameter vector and the period-by-period realization of the (latent) state vector, while the distributional convergence to a mixed Gaussian

limit is stable, enabling feasible inference and formal diagnostic tests. These results apply in the presence of conditional heteroskedasticity of unknown form in the measurement errors for the option prices. We are not aware of alternative approaches that can handle systems of similar complexity and deliver feasible and reliable inference and specification tests. Beyond details on the formal estimation, inference and testing procedures associated with this approach, Andersen et al. (2012) also document encouraging finite sample performance through a large scale simulation study.

In practice, the optimization in equation (2.2) is most conveniently performed iteratively, exploiting two simple steps. For a given parameter vector, we first determine the optimal value of the state vector on each day of the sample. Second, for the concentrated objective function, we perform MCMC based estimation of the parameter vector with a chain of length $10,000$. For this new θ estimate, we then move back to step one and iterate the procedure.

3 Parametric Modeling of the Option Panel

We start our analysis by fitting the option panel using the data and estimation method described in the previous section. The objective is to capture the dynamics of the option surface succinctly through a parametric stochastic volatility model which may subsequently serve as a proper basis for the analysis of the equity and variance risk premiums. A pertinent issue is the number of latent state variables required to provide a suitable fit. The vast majority of empirical option pricing studies employs a single stochastic volatility factor, but the literature on estimating the return dynamics under the physical measure points to a minimum of two factors, which also is consistent with a number of recent option pricing studies. Consequently, we initially estimate a general two-factor stochastic volatility model that nests most parametric models in the empirical literature. Later, we extend this representation, guided by shortcomings in its fit to the option panel dynamics.

3.1 The Standard Dynamic Asset Pricing Model

The two-factor stochastic volatility model is given by the following specification for the risk-neutral dynamics for the price process X,

$$\begin{aligned}
\frac{dX_t}{X_{t-}} &= (r_t - \delta_t)\,dt + \sqrt{V_{1,t}}\,dW^{\mathbb{Q}}_{1,t} + \sqrt{V_{2,t}}\,dW^{\mathbb{Q}}_{2,t} + \int_{\mathbb{R}^2} (e^x - 1)\,\widetilde{\mu}^{\mathbb{Q}}(dt,dx,dy), \\
dV_{1,t} &= \kappa_1\,(\overline{v}_1 - V_{1,t})\,dt + \sigma_1\,\sqrt{V_{1,t}}\,dB^{\mathbb{Q}}_{1,t} + \int_{\mathbb{R}^2} y\mu(dt,dx,dy), \\
dV_{2,t} &= \kappa_2\,(\overline{v}_2 - V_{2,t})\,dt + \sigma_2\,\sqrt{V_{2,t}}\,dB^{\mathbb{Q}}_{2,t},
\end{aligned} \tag{3.1}$$

where $(W^{\mathbb{Q}}_{1,t}, W^{\mathbb{Q}}_{2,t}, B^{\mathbb{Q}}_{1,t}, B^{\mathbb{Q}}_{2,t})$ is a four-dimensional Brownian motion with $W^{\mathbb{Q}}_{1,t} \perp W^{\mathbb{Q}}_{2,t}$, $W^{\mathbb{Q}}_{1,t} \perp B^{\mathbb{Q}}_{2,t}$, and $W^{\mathbb{Q}}_{2,t} \perp B^{\mathbb{Q}}_{1,t}$, while $\text{corr}\left(W^{\mathbb{Q}}_{1,t}, B^{\mathbb{Q}}_{1,t}\right) = \rho_{d,1}$ and $\text{corr}\left(W^{\mathbb{Q}}_{2,t}, B^{\mathbb{Q}}_{2,t}\right) = \rho_{d,2}$; μ is an integer-valued

random measure on $\mathbb{R}_+ \times \mathbb{R}^2$ counting the jumps in the price, X, as well as the state vector, $(V_{1,t}, V_{2,t})$, with compensator (under the risk-neutral measure) $dt \otimes \nu_t^{\mathbb{Q}}(dx, dy)$ and $\widetilde{\mu}^{\mathbb{Q}}(dt, dx, dy) = \mu(dt, dx, dy) - dt\nu_t^{\mathbb{Q}}(dx, dy)$ is the associated martingale measure.

We adhere to the near universal specification of the risk-neutral compensator for the jump measure, proposed by Duffie et al. (2000), i.e.,

$$\nu_t^{\mathbb{Q}}(dx, dy) \;=\; \left\{ (c_0 + c_1 V_{1,t-} + c_2 V_{2,t-}) \; \frac{e^{-\frac{(x-\mu_x-\rho_j y)^2}{2\sigma_x^2}}}{\sqrt{2\pi}\sigma_x} \; \frac{e^{-y/\mu_v}}{\mu_v} \, 1_{\{y>0\}} \right\} dx \otimes dy.$$

The representation implies that the price and volatility processes jump simultaneously. The return jumps are normally distributed with mean μ_X and variance σ_X^2, while the volatility jumps are exponentially distributed with parameter μ_V and constrained to be positive. The correlation in the jump size of returns and volatility is governed by ρ_j. Finally, the jump intensity varies with the level of volatility, and is controlled by the coefficients c_0, c_1 and c_2.

By construction, the specification of the risk-neutral measure, $\mathbb{Q}$, determines how the various sources of randomness, or risk, in the return process is compensated, or priced. The above model has some notable features in this regard. First, volatility is governed by two separate factors, consistent with prior empirical evidence that volatility itself is not a Markov process. This implies, in particular, that two state variables span the dynamics of the entire option surface across time. Second, both the price and volatility trajectories display discontinuities. There is strong empirical evidence for both, including possible dependence between them. Third, the jump risk (under the risk-neutral measure) is time-varying. Strong time-variation of the jump risk, under both the statistical and risk-neutral measure, is documented by nonparametric means in Bollerslev and Todorov (2011a), respectively Bollerslev and Todorov (2011b).

Various constrained versions of the above model have been estimated in prior work. In particular, a version with constant jump intensity and excluding the $V_{2,t}$ factor was introduced by Duffie et al. (2000) and referred to as the double jump volatility model. Bates (2000) estimates a version of model (3.1) without volatility jumps, Pan (2002) considers a constrained one-factor volatility version without volatility jumps, Eraker (2004) and Broadie et al. (2009) similarly estimate a one-factor volatility version of model (3.1) but allow for volatility jumps.[2]

Table 1 provides the parameter estimates for the model, using the data and estimation method reviewed in Section 2. Several features of the risk-neutral measure are noteworthy. First, the

[2]Affine specifications for the asset price also arise naturally within equilibrium models, in which case the latent factors may be given a structural interpretation and components of the risk in the asset price process will be related to underlying fundamentals in specific ways; examples include Drechsler and Yaron (2011) and Drechsler (2013).

volatility factor, V_1, containing the jump component is the persistent one, while V_2 is a more transient volatility factor. Moreover, the first volatility factor has very strong negative correlation with the price, while the second one appears to have a weaker negative correlation with price, although this parameter is somewhat less precisely estimated. Second, the mean of the price jumps is around -16% with a standard deviation of approximately 6% and with an average of 0.6 jumps per year. The return jumps are the primary determinant of the relative price of short-maturity deep out-of-the-money put options, so the large negative jump size reflects the relative expensiveness of these options. At the same time, the estimates imply that positive return jumps are almost non-existent, which suggests – as we indeed confirm below – that the model will struggle with fitting observed out-of-the-money call option prices. Third, we note that, counter to usual intuition, the parameter estimates imply a positive dependence between price and volatility jumps. Of course, one should be cautious in interpreting the jump parameters if there is strong evidence of misspecification of this part of the model, as we subsequently argue is the case. Fourth, most of the jump intensity is driven by the time-varying component. This reflects the fact that short-maturity deep out-of-the-money options become relatively more expensive during periods of high volatility.

Table 1: Parameter Estimates for the two-factor model

Parameter	Estimates	Std.	Parameter	Estimates	Std.	Parameter	Estimates	Std.
ρ_1	-0.969	0.031	$\overline{v}_2$	0.008	0.000	c_2	6.363	0.297
$\overline{v}_1$	0.009	0.000	κ_2	17.976	0.782	μ_x	-0.193	0.003
κ_1	1.341	0.021	σ_2	0.074	0.054	σ_x	0.063	0.002
σ_1	0.138	0.006	c_0	0.006	0.004	μ_v	0.047	0.001
ρ_2	-0.742	0.525	c_1	25.557	0.575	ρ_j	0.638	0.023

Note: Parameter estimates of the two-factor model (3.1) for S&P 500 equity-index option data sampled every Wednesday over the period January 1996-July 2010.

We next explore some diagnostics that help identify situations in which the model struggles to fit the option surface. Andersen et al. (2012) develop Z-scores that reflect the fit to specific parts of the surface during a given trading period. They are obtained as the normalized statistics, $\mathcal{Z}_{\mathcal{K},t}/\sqrt{\widehat{\text{Avar}}\left(\mathcal{Z}_{\mathcal{K},t}\right)}$, based on the average pricing errors over the relevant part of the surface,

$$\mathcal{Z}_{\mathcal{K},t} = \sum_{j:k_j\in\mathcal{K}} \left(\overline{\kappa}_{t,k_j,\tau^*} - \kappa(k_j,\tau^*,\widehat{\mathbf{S}}_t^n,\widehat{\theta}^n)\right),$$

where $\mathcal{K}$ indicates a particular region of moneyness and τ^* represents a specified time to maturity. Andersen et al. (2012) show that these Z-scores are asymptotically standard normal under the null of correct model specification.

Figure 2 plots the Z-scores for each day in the sample corresponding to separate parts of the option surface. These regions span two maturity ranges – short-term and long-term options – as well as three categories of moneyness – out-of-the-money (OTM) puts, at-the money (ATM) options, and out-of-the-money calls – leading to a total of six regions for each trading day.[3] The two-factor model provides a reasonable fit to the average level of the short-term OTM put options, while there are fairly pronounced biases in the fit to the options in the other regions.[4] Generally, the OTM calls and the short-term ATM options are overpriced by the model, while the long-term OTM puts and ATM options are underpriced. The superior fit to the short-term OTM puts likely reflects the larger proportion of such options in our sample due to the active trading and quoting of these particular instruments. This implies that the corresponding option fit receives a larger weight in the criterion function in equation (2.2). Of course, overall, the estimation criterion should imply that the average (signed) pricing error across all the options in the sample is relatively small. This is, in fact, the case. For example, constructing the option portfolios – spanning all available strike prices – that mimic the 1-month and 12-month volatility index, VIX, produce sample averaged annualized volatilities of 20.01% and 20.87%, while the corresponding model-implied volatilities are 21.08% and 21.11%, respectively. In short, the average option price fit is not a particularly useful model diagnostic as model misspecification induces systematic, but largely offsetting, pricing errors in different regions of the volatility surface and over time.

Therefore, a critical dimension for assessing model adequacy is the extent of intertemporal pricing errors. The solid lines on Figure 2 indicate the two-sided 95% critical region. It is apparent that there are important failures in matching the dynamics of the option panel, particularly for OTM options. As before, the Z-scores for the short-maturity OTM puts are relatively low, reflecting their weight in the objective function. The model also fits the short-maturity ATM options reasonably well with a few notable exceptions. In particular, the model overvalues such options from early 1998 till the end of 1999 as well as from 2009 onwards, while it undervalues short-maturity ATM options during 2008. The short-maturity OTM calls are, on the other hand, almost invariably mispriced as

[3]The regions are defined as follows. Short maturity OTM put options: $\tau \leq 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \leq -1$; Short maturity ATM options: $\tau \leq 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \in (-1, 0.5]$; Short maturity OTM call options: $\tau \leq 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \geq 1$; Long maturity OTM put options: $\tau > 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \leq -1$; Long maturity ATM options: $\tau > 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \in (-1, 0.5]$; Long maturity OTM call options: $\tau > 60$ days and $\frac{k}{\sigma_t^{ATM}\sqrt{\tau}} \geq 1$.

[4]The time-averaged Z-scores are for short (maturity) OTM puts: -0.11; short ATM options: -0.61; short OTM calls: -0.99; long (maturity) OTM puts: 0.87; long ATM options: 0.38; long OTM calls: -1.63.

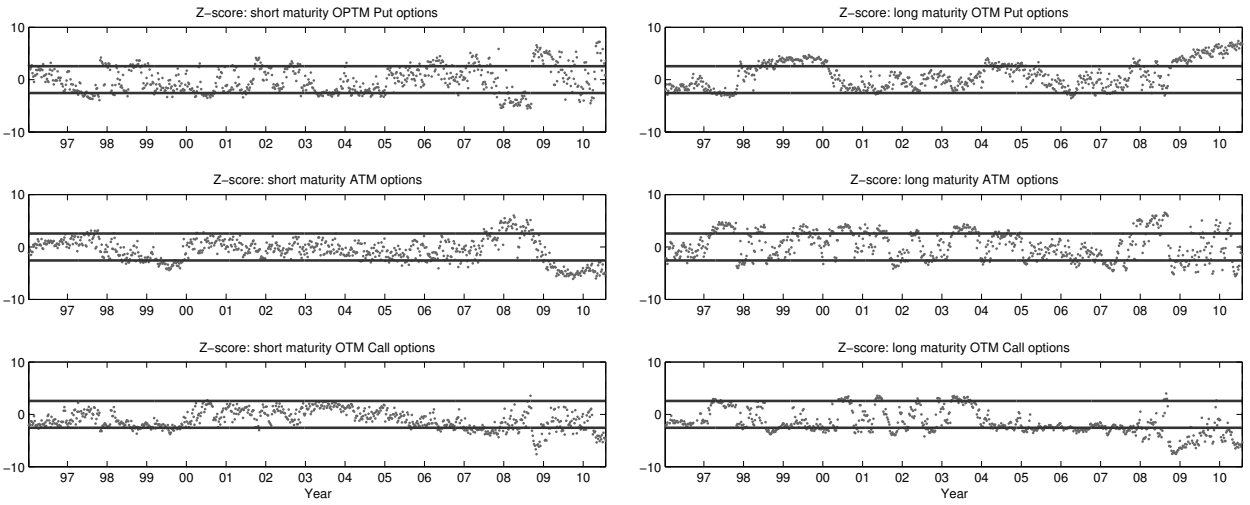


Figure 2: Two-factor model Z-scores, as defined in the main text.

periods with pronounced overvaluation alternates with periods of persistent undervaluation. Similarly, the long-maturity option prices are generally poorly fit in spite of the fact that we have included two separate volatility factors to provide added flexibility in accommodating shifts in the term structure of volatility. Specifically, there is a striking negative correlation in the pattern of over- and underpricing for the long-term OTM put and call options. When the long term puts are overpriced during 1997, the calls are underpriced, with the pattern reversing from 1998-2000. The same observations apply to 2004-2005 and from the onset of the financial crisis in 2008 through the end of the sample. Overall, Figure 2 reveals very persistent under- and overpricing of different regions of the option surface over time. This is formally confirmed by Panel A of Table 2 which reports the rejection rates for adequate fit to the different regions of the option surface. Invariably, the numbers of violations greatly exceed the corresponding nominal sizes.[5]

What is the origin of the mispricing in Figure 2? The persistent pattern of under- and overvaluation of different parts of the option surface suggest that it may stem from misspecification of the risk factors and, in particular, their time variation, i.e., the state vector dynamics is misrepresented. The implausible estimates for the parameters governing the jump process, as well as the prominent role played by jump risk for short-maturity options, suggest this feature is a natural starting point for a modified system. There are two potential sources of misspecification for the jump process, namely the size distribution and the intensity.

[5]Note that the tests across different days are not truly independent, even after conditioning on the filtration of the original probability space. This stems from the estimation error of the model parameter vector which relies on all the options in the sample. However, under correct model specification, the estimation error for the parameter vector is too small to induce strong ($\mathcal{F}$-conditional) asymptotic dependence between the Z-scores.

Focusing first on the jump size distribution, we note that the Gaussian assumption regarding the return jumps implies a negligible probability of observing positive price jumps. Moreover, the probability of small negative price jumps is very small relative to the probability of larger negative jumps, i.e., the jump intensity is not decreasing in the absolute value of the jump size. This feature is at odds with the nonparametric evidence in Bollerslev and Todorov (2011b) which finds the jump intensity decaying gradually with the (absolute) jump size.

Turning to the modeling of the jump intensity, we observe a recurrent pattern in Figure 2. During the turbulent periods of 1998-1999 and from 2009 to the end of the sample, the model manages to price short term OTM put options fairly accurately, but it comes at the cost of dramatically overvaluing the short term ATM options and undervaluing the longer term OTM put options. It is evident that the model is struggling because the short term OTM puts are expensive relative to the ATM options or, in other words, the volatility smirk is unusually steep. The model can only generate this type of gap between the two classes of options by enhancing the intensity of negative price jumps, but these are directly linked to the value of the state vector representing components of spot volatility. When the implied volatility rises to accommodate the valuation of the put options, the ATM options become overvalued. At the same time, it is clear that this elevation in volatility is insufficient to account for the higher valuation of the longer maturity OTM put options. These observations suggest the left tail of the jump distribution displays a persistent type of variation that cannot be spanned by a linear function of the volatility factors. A potential remedy is the introduction of an additional state variable that allows the left jump tail to display variation that is partially independent of the volatility factors. That is, the time-variation in the pricing of negative jump risk can evolve separately from the diffusive volatility, even if the two still may display some dependencies. Finally, we note that the model (3.1) ties the time-variation of positive and negative jumps to be identical, but there is no reason a priori to impose this restriction.

In summary, we find the steepening of the implied volatility curve and the flattening of the term structure of implied volatility for OTM put options to be critical features of the option surface dynamics during turbulent market conditions. Any such variation in the surface induces severe pricing violations for the model (3.1), because the jump intensity is an affine function of volatility state vector. In fact, using (model-free) expansions of the option price one may show that the short-maturity ATM Black-Scholes implied volatility provides a good approximation to the current (diffusive) volatility, see, e.g., Durrleman (2008), while short-maturity OTM put options are approximately proportional to the current level of the jump intensity, see, e.g., Carr and Wu (2003). Therefore, for a given level of volatility, the slope of the short-maturity implied volatility

curve in the model is essentially fixed. These observations motivate the extension we pursue below.

Table 2: Diagnostic Tests for S&P 500 Option Data

Test	**2F Model Nominal Size**		**3F Model Nominal Size**	
	1%	5%	1%	5%
Panel A: Fit to the Option Panel				
OTM, short-maturity puts	35.13%	52.89%	8.55%	19.74%
ATM, short-maturity	26.18%	36.32%	3.68%	11.58%
OTM, short-maturity calls	19.08%	36.84%	18.29%	31.58%
OTM, long-maturity puts	35.53%	48.95%	8.29%	21.71%
ATM, long-maturity	42.50%	58.42%	3.82%	11.58%
OTM, long-maturity calls	44.47%	66.32%	14.34%	32.37%
Panel B: Root-Mean Squared Error of IV Option Fit				
	2.55%		1.75%	

Note: Panel A reports rejection frequencies across the full sample for the option fit to specific portions of the option surface based on the Z-scores plotted on Figures 2 and 5. OTM puts, ATM options and OTM calls, and short- versus long-maturity options are defined in Figure 2. Panel B provides the root-mean-squared-error of the model-implied Black-Scholes volatilities relative to the market mid-quote implied Black-Scholes volatilities across all options used for estimation.

3.2 The Extended Model

The extended model stipulates the following risk-neutral equity index dynamics,

$$\begin{aligned}
\frac{dX_t}{X_{t-}} &= (r_t - \delta_t)\,dt + \sqrt{V_{1,t}}\,dW^{\mathbb{Q}}_{1,t} + \sqrt{V_{2,t}}\,dW^{\mathbb{Q}}_{2,t} + \int_{\mathbb{R}^2} (e^x - 1)\,\widetilde{\mu}^{\mathbb{Q}}(dt,dx,dy), \\
dV_{1,t} &= \kappa_1\,(\overline{v}_1 - V_{1,t})\,dt + \sigma_1\,\sqrt{V_{1,t}}\,dB^{\mathbb{Q}}_{1,t} + \mu_{v_1} \int_{\mathbb{R}^2} x^2\,1_{\{x<0\}}\,\mu(dt,dx,dy), \\
dV_{2,t} &= \kappa_2\,(\overline{v}_2 - V_{2,t})\,dt + \sigma_2\,\sqrt{V_{2,t}}\,dB^{\mathbb{Q}}_{2,t}, \\
dU_t &= -\kappa_3\,U_t\,dt + \mu_u \int_{\mathbb{R}^2} \left[(1-\rho_3)\,x^2\,1_{\{x<0\}} + \rho_3\,y^2\right]\mu(dt,dx,dy).
\end{aligned} \tag{3.2}$$

The jump measure μ has a compensator given by $dt \otimes \nu_t^{\mathbb{Q}}(dx, dy)$, where,

$$\nu_t^{\mathbb{Q}}(dx, dy) \;=\; \left\{ \left(c^- 1_{\{x<0\}} \lambda_- e^{-\lambda_- |x|} + c^+ 1_{\{x>0\}} \lambda_+ e^{-\lambda_+ x} \right) 1_{\{y=0\}} + c^- 1_{\{x=0,\; y<0\}} \lambda_- e^{-\lambda_- |y|} \right\} dx \otimes dy,$$

$$c^- \;=\; c_0^- + c_1^- V_{1,t-} + c_2^- V_{2,t-} + c_3^- U_{t-}, \qquad c^+ \;=\; c_0^+ + c_1^+ V_{1,t-} + c_2^+ V_{2,t-} + c_3^+ U_{t-}\,.$$

Before commenting on the modifications and extensions relative to model (3.1), we note some changes in the notation motivated by the new specification of the jump processes. The first term in the expression for the jump measure controls price-volatility co-jumps, exactly as in model (3.1), while the second term controls the independent jumps in the new factor, U. Furthermore, the jumps in U are either directly proportional to the volatility jumps (in V_1) or independent from them. Both jump components of U have the same marginal distribution - they only differ in their interaction with the other jump components of the model. To retain notational symmetry between the jump components of U, we define the second term of the jump measure, μ, over the squared negative realizations. In particular, this implies that all jumps in U are positive.

The distinct features of the extended model may be summarized as follows. First, the distribution of the price jumps is now exponential so we enforce the empirically realistic decay in the jump intensity as a function of the jump size. Moreover, we allow the left and right jump tail to have different tail decay parameters, i.e., we break the dependence between the behavior of the left and right jump tails. Second, the price and volatility co-jumps are now perfectly dependent, with the squared price jumps impacting the volatility dynamics. This is reminiscent of a GARCH specification for the volatility dynamics in discrete time. Given the proliferation of parameters in our extended model, this parsimonious formulation is desirable.[6] In addition, we allow only the negative price jumps to impact the volatility dynamics, as a more general dependence structure between volatility and price jumps cannot be well identified from the option panel. Third, the negative jumps now have an additional source of time variation, captured by the process U.[7] This feature of the extended model reflects the shortcomings of the standard model (3.1) documented in the previous section. In particular, U is driven, in part, by the negative price jumps which allow for crisis periods to elevate the jump intensity of the system. Importantly, U also contains an independent source of variation, and the degree of dependence between the jumps in U and X is captured by the parameter ρ_3. Thus, our specification encompasses scenarios of perfect dependence between the jump risks in V_1 and U (for $\rho_3 = 0$) and scenarios of V_1 and U being entirely

[6] Support for such modeling of the price-volatility jump dependence (albeit under $\mathbb{P}$) can be found in Todorov (2011).

[7] This feature can also be incorporated into the dynamics for the intensity of the positive price jumps, but since the relevant parameter is not significant, we end up imposing the constraint $c_3^+ = 0$ in our final specification.

independent state variables driving the dynamics of the option panel. In addition, we note that V_1 and U are connected not only through this common jump component, but also through the joint time-variation in the jump intensity. This latter feature generates "cross self-excitation" in which jumps in U excite the occurrence of jumps in V_1 and vice versa. Finally, the extended model (3.2) allows for separate time-variation in the intensity of positive and negative jumps via discrepancies in the parameters $c_1^{\pm}$ and $c_2^{\pm}$.[8]

The parameter estimates for model (3.2) are reported in Table 3. Although the basic features of the volatility factors, V_1 and V_2, are reminiscent of those obtained from model (3.1), there are several notable exceptions. In the current model, the volatility factor containing the jump component, i.e., V_1, has the lower degree of persistence. Also, both V_1 and V_2 now exhibit a very strong negative correlation with the diffusive price innovation. Not surprisingly, the jump parameters reveal even more substantial changes. First, the mean of the negative jumps is -4.7%, while the mean of the positive jumps is 2.1%. Furthermore, the left jump size tail is approximately twice as fat as the right one. In addition, the average jump frequency is 3.4 per year for the left tail and 2.6 per year for the positive tail. All these features represent a remarkable departure from model (3.1), which implies a more extreme negative jump size, occurring with a very low probability, and essentially no chance of positive price jumps. As before, the extended model implies that the jump intensity is driven almost exclusively by the time-varying component.[9] Notably, the c_1^- and c_1^+ as well as the c_2^- and c_2^+ coefficients are pairwise statistically distinct, thus indicating a different degree of time-variation in the left and right jump tail, consistent with the nonparametric evidence in Bollerslev and Todorov (2011b). Further, the state variable U plays a very strong role in the time variation of the left jump tail, accounting for about 80% of the mean jump intensity for the negative jumps. We also note that the state variable U is highly persistent, likely explaining why the volatility factor V_1 now is more transient. Finally, it is evident that the parameter ρ_3, determining the proportion of independent jump variation in U, cannot be estimated with precision.

To further clarify the distinct role played by the different factors in driving the dynamics of the option panel, Figure 3 depicts the shift in the short and long maturity implied volatilities across the strike prices due to a shock in each of the three state variables, perturbing them from their mean

[8] From an implementation standpoint, it is critical that the model (3.2) remains within the general affine class of Duffie et al. (2003). We refer to that paper for the ODE system of equations that must be satisfied by the coefficients of the associated conditional characteristic function. For the estimation, we impose the following conditions on the parameters to guarantee covariance stationarity of the three latent factors,

$$\kappa_1 > \frac{2c_1^-\mu_{v_1}}{\lambda_-^2}, \text{ and } \kappa_3 > \frac{2c_3^-\kappa_1\mu_u}{\kappa_1\lambda_-^2 - 2c_1^-\mu_{v_1}}, \text{ and } \kappa_2 < 0, \text{ and } \sigma_i^2 \leq 2\kappa_i\overline{v}_i, \quad i = 1, 2.$$

[9] In fact, the estimates of c_0^- were insignificant, and it was set to zero during the final estimation round.

Table 3: Parameter Estimates for the three factor model

Parameter	Estimate	Std.	Parameter	Estimate	Std.	Parameter	Estimate	Std.
ρ_1	-0.913	0.028	σ_2	0.110	0.006	c_2^-	0.913	3.730
$\overline{v}_1$	0.007	0.000	μ_u	1.756	0.647	c_2^+	14.269	4.986
κ_1	8.325	0.167	κ_3	0.522	0.080	c_3^-	19.836	5.460
σ_1	0.323	0.014	ρ_3	0.117	0.614	λ_-	21.157	0.240
ρ_2	-0.945	0.036	c_0^+	0.723	0.079	λ_+	48.365	2.053
$\overline{v}_2$	0.016	0.001	c_1^-	34.592	1.931	μ_{v_1}	11.602	0.262
κ_2	0.480	0.033	c_1^+	88.178	14.711			

Note: Parameter estimates of the three-factor model (3.2) for S&P 500 equity-index option data sampled every Wednesday over the period January 1996-July 2010.

sample values. To provide transparency regarding the critical modeling of the left tail dynamics, we calibrate the shocks so that the shift in V_1 and U have the identical effect on the intensity of negative price jumps. In practical terms, that implies that the shocks to V_1 and U have the same impact on the prices of short-maturity deep OTM put options.[10] Since V_2 has a minimal impact on the left jump tail, we calibrate the shock to this state variable to ensure that its effect on the ATM implied volatility falls within the range of the changes brought about by the shocks to V_1 and U. It is evident from Figure 3 that U, unlike V_1, has a negligible impact on the pricing of short term ATM options. Instead, it primarily influences the pricing of short term deep OTM put options. This is, of course, a reflection of the fact that U is not part of the diffusive volatility. In contrast, shocks to V_1 elevates the entire short-maturity implied volatility curve, with the effect being the largest for the ATM options. In other words, positive shocks to V_1 are ineffective in steepening the implied volatility curve. Turning towards longer maturity options, we again see the shock to V_1 having a smaller impact on the OTM puts than the ATM options, while a shock to U has a similar impact on the entire implied volatility function. In particular, the strong persistence of the U process ensures that it has a much stronger influence on the pricing of the long term options that V_1. Finally, we note that the influence of V_2 is mostly reserved for the ATM and OTM call options for both maturities. In sum, by highlighting the separate roles of the factors in driving the dynamics of the

[10] As the time-to-maturity of OTM put shrinks towards zero, the option price scaled by time-to-maturity multiplied by the current spot price converges to $\int_{\mathbb{R}}(e^k - e^x)^+ \nu_t^{\mathbb{Q}}(dx, dy)$, see e.g., Carr and Wu (2003).

option surface, Figure 3 underscores the critical function of U in accommodating time variation in both the steepness of the volatility smirk and the term structure of implied volatility.

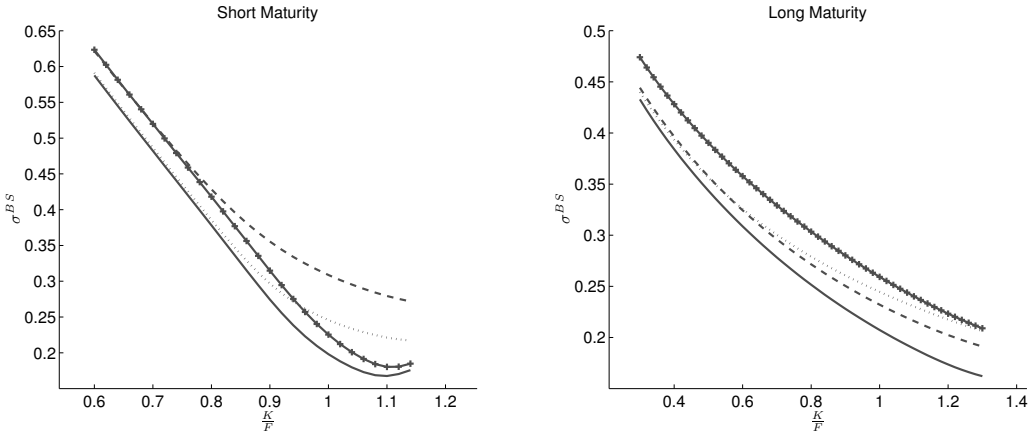


Figure 3: Sensitivity of option surface to changes in state variables in the three-factor model. Parameters are set equal to their estimates reported in Table 3. The solid line corresponds to the implied volatility for values of the three state variables equal to their sample means. All other lines correspond to changing one of the state variables from its sample mean value. The dashed line corresponds to changing V_1, the dotted line to changing V_2 and the x-line to changing U. The change in V_1 is three times its mean value; the change in U is such that the effect on the risk-neutral jump intensity of the negative jumps is the same as that due to changing V_1; the change in V_2 is twice its mean value. Short maturity is one-month and long maturity is one year.

Figure 4 provides a time series plot of the extracted diffusive stochastic variance, i.e., $V_{1,t}+V_{2,t}$, along with the third factor U_t. The two series exhibit similar low frequency variation throughout the sample, but with distinct differences in the medium and high-frequency dynamics. U reaches its maximum during the financial crisis, and then only, very gradually, reverts back to the pre-crisis level by the middle of 2010. In comparison, the volatility reaches its high in the Fall of 2008 but drops relatively quickly to normal levels by the middle of 2009. Subsequently, the volatility continues to drop well into 2010, when it then rises in response to the European sovereign debt crisis. A qualitatively identical pattern is observed in the aftermath of the second largest peak of U in the sample, corresponding to the Russian crises of October 1998. U remains elevated for more than a year, while the (diffusive) volatility only has a temporary spike. These episodes illustrate how U, following a crisis, often displays a much slower degree of mean reversion than the volatility process. The identical pattern may be found during the recent European debt crisis which had only a transient impact on our volatility estimate but a much larger and more persistent effect on U. Finally, the same effects are operative around the Asian crises in the Fall of 1997.

Figure 4 highlights how deep OTM put options following a crisis may become too expensive to be rationalized solely by the concurrent level of volatility. Importantly, however, not all episodes of elevated volatility fit this pattern. For example, the volatility spikes following the terrorist attack on 9/11, 2001, and the collapse of the tech bubble in 2002 did not presage any prolonged elevation of U. Thus, these episodes had a very limited impact on the pricing of tail risk, clearly separating the financial market impact of these events from those discussed above, and further highlighting the informational richness of the option panel.

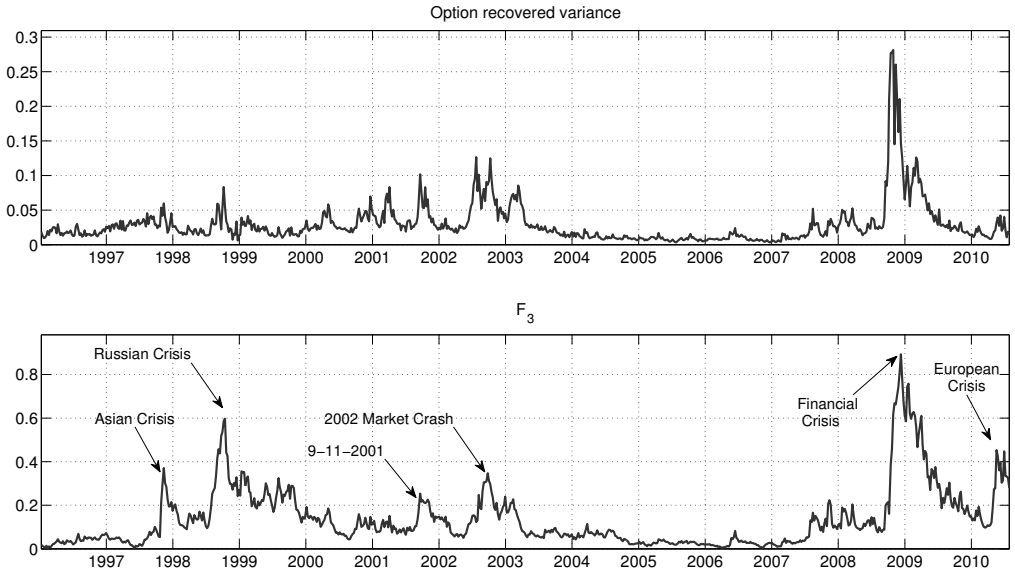


Figure 4: Volatility and Jump State Variable in the three-factor model.

We now consider model performance. First, in terms of average pricing, the portfolios of options mimicking the 1- and 12-month VIX index produce annualized implied volatilities of 20.01% and 20.87%, while the corresponding model-implied numbers are 21.08% and 20.97% respectively, thus generating a slightly better match than for model (3.1). Likewise, the unconditional fit to each of the six separate regions of the option surface defined earlier is improved, with the pricing for all ATM options as well as the short term OTM put options displaying essentially no bias.[11]

Turning to the dynamic performance, Figure 5 depicts the Z-scores for the extended model over the full sample for the six regions of the option surface. There is a visible improvement relative to

[11]The mean across the full sample for the Z-scores, plotted in Figure 5 below, are for the short (maturity) OTM puts: -0.07; short ATM options: -0.15; short OTM calls: -0.65; long (maturity) OTM puts: 0.53; long ATM options: -0.05; and long OTM calls: -1.11.

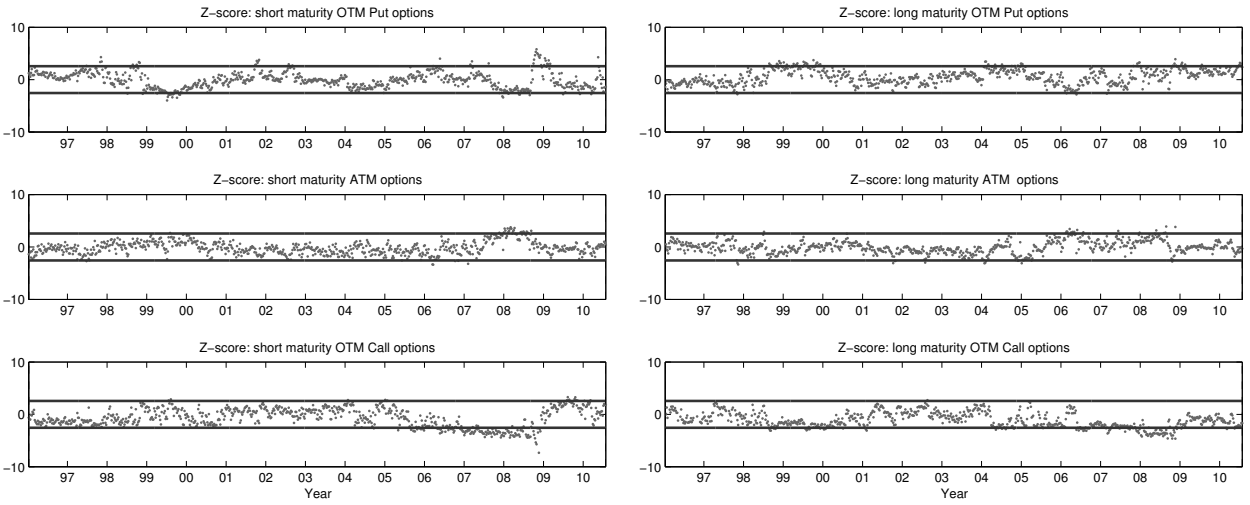


Figure 5: Three-factor model z-scores. The construction of the z-scores and definition of the regions of the option surface are exactly the same as in Figure 2.

the fit provided by model (3.1), including both turbulent and tranquil periods during the sample. In particular, following the Asian crises and early 2009, the short-maturity ATM options are no longer overvalued. Similarly, the fit to the long-term OTM puts improves to the point where the fairly mild undervaluation typically falls, point-wise, within the associated confidence band. However, during the 1997, 1998, and 2008 crises the short-maturity deep OTM puts continue to be (slightly) undervalued. Hence, there is still scope for additional improvement in the modeling of the jump tail risk. On the other hand, the mispricing is relatively short-lived, so it should have a limited impact on our subsequent analysis.[12]

Finally, the pricing of the OTM calls, both at short and long maturities, remains inadequate. For example, from the beginning of 2007 till the end of 2008 they are consistently overvalued. The same applies for the long-maturity OTM calls during 1999. This may be solved by introducing a separate factor driving the intensity of the positive jumps. Importantly, however, in the current model the connection between the mispricing of OTM puts and calls is significantly weakened, and it now performs significantly better on the critical OTM put side of the implied volatility curve.

The visual improvement in fit is confirmed by Panel A of Table 2. The rejection rates of the Z-score tests drop significantly compared to the two-factor model (3.1) for almost all regions of the surface. Nonetheless, consistent with our discussion, the fit is clearly not perfect, especially for

[12]In principle, the undervaluation can be rectified via a time-varying jump tail index, i.e., allowing the parameter λ_- to vary, see, e.g., the nonparametric evidence in Bollerslev and Todorov (2013). However, time variation in the tail index takes us outside the affine setting, implying that we lose the analytical tractability in pricing the option surface which is critical for our current implementation.

the OTM calls. In addition, Panel B of Table 2 indicates a substantial improvement in the overall RMSE for the option prices. We conclude that the model (3.2) provides a much improved fit to the option surface and that the more critical dynamic features are reasonably well approximated by the stipulated law of motion for the three state variables.

One may, of course, wonder whether our improvement stems from the mere addition of an extra state variable, or whether it is tied to our approach of dedicating the new state variable to capturing the jump tail dynamics. Thus, in Section 6, we present results from an alternative three-factor model in which the third factor is part of the diffusive volatility. This formulation is in line with the traditional way of building multi-factor volatility models. We find this alternative model to perform considerably worse in capturing the option dynamics, thus corroborating our hypothesis concerning the need to separate the risk-neutral jump intensity from the diffusive volatility.

4 Risk Premia Dynamics and Predictability

In this section we relate our findings based on the option panel to the underlying return data. In particular, we study the links between the state vector, or risk factors, extracted from the option panel with the volatility and jump risks inferred from the dynamics of the underlying stock market index. This sets the stage for a direct exploration of the associated equity and variance risk premiums and their relation with the option-implied factors. Given the superior performance, we rely on the three factor model (3.2) for tracking the dynamics of the option surface over time. We continue to avoid making strong assumptions regarding the dynamics of the market risks under the actual probability measure. Consequently, this part of our analysis is fully nonparametric, and we invoke only minimal stationarity conditions regarding the $\mathbb{P}$-dynamics.

4.1 Connecting the Information in the Option Panel and the Underlying Asset

To define risk premia, we must first develop some consistent notation concerning the pricing of each of the sources of risks in the model (3.2). We define

$$\begin{aligned} W_{1,t}^{\mathbb{P}} &= W_{1,t}^{\mathbb{Q}} - \int_0^t \lambda_s^{W_1}\, ds, \quad W_{2,t}^{\mathbb{P}} = W_{2,t}^{\mathbb{Q}} - \int_0^t \lambda_s^{W_2}\, ds, \\ B_{1,t}^{\mathbb{P}} &= B_{1,t}^{\mathbb{Q}} - \int_0^t \lambda_s^{B_1}\, ds, \quad B_{2,t}^{\mathbb{P}} = B_{2,t}^{\mathbb{Q}} - \int_0^t \lambda_s^{B_2}\, ds, \end{aligned} \tag{4.1}$$

where $W_{1,t}^{\mathbb{P}}$, $W_{2,t}^{\mathbb{P}}$, $B_{1,t}^{\mathbb{P}}$ and $B_{2,t}^{\mathbb{P}}$ are $\mathbb{P}$ Brownian motions and $\lambda_t^{W_1}$, $\lambda_t^{W_2}$, $\lambda_t^{B_1}$ and $\lambda_t^{B_2}$ denote the associated prices of risk. The compensator of the jump measure, μ, under the $\mathbb{P}$ measure is given by $dt \otimes \nu_t^{\mathbb{P}}(dx, dy)$, and the mapping $\nu_t^{\mathbb{P}}(dx, dy) \rightarrow \nu_t^{\mathbb{Q}}(dx, dy)$, defined for every jump size x and y and every point in time t, reflects the compensation for jump risk.

The dynamics of the stock price process under the physical probability measure $\mathbb{P}$ is then,

$$\frac{dX_t}{X_{t-}} \;=\; \alpha_t\,dt \;+\; \sqrt{V_{1,t}}\,dW^{\mathbb{P}}_{1,t} \;+\; \sqrt{V_{2,t}}\,dW^{\mathbb{P}}_{2,t} \;+\; \int_{\mathbb{R}^2}(e^x-1)\,\widetilde{\mu}^{\mathbb{P}}(dt,dx,dy), \tag{4.2}$$

where

$$\alpha_t-(r_t-\delta_t) \;=\; \lambda^{W_1}_t\sqrt{V_{1,t}} \;+\; \lambda^{W_2}_t\sqrt{V_{2,t}} \;+\; \int_{\mathbb{R}^2}(e^x-1)\,\nu^{\mathbb{P}}_t(dx,dy) \;-\; \int_{\mathbb{R}^2}(e^x-1)\,\nu^{\mathbb{Q}}_t(dx,dy), \tag{4.3}$$

is the instantaneous equity risk premium, reflecting compensation for diffusive and jump price risks.

Application of Itô formula yields the following representation for $\log(X_t)$ under $\mathbb{P}$,

$$d\log(X_t) \;=\; \left[\alpha_t-q^{\mathbb{P}}_t\right]dt \;+\; \sqrt{V_{1,t}}\,dW^{\mathbb{P}}_{1,t} \;+\; \sqrt{V_{2,t}}\,dW^{\mathbb{P}}_{2,t} \;+\; \int_{\mathbb{R}^2}x\,\widetilde{\mu}^{\mathbb{P}}(dt,dx,dy), \tag{4.4}$$

where $q^{\mathbb{P}}_t=\frac{1}{2}V_t+\int_{\mathbb{R}^2}(e^x-1-x)\,\nu^{\mathbb{P}}_t(dx,dy)$, and similarly under $\mathbb{Q}$ with α_t replaced by $r_t-\delta_t$ and all superscripts $\mathbb{P}$ replaced with $\mathbb{Q}$ in the expression above.

We may then define the (conditional) cum-dividend equity risk premium over the horizon τ,

$$\begin{aligned}\mathrm{ERP}^{\tau}_t \;\equiv\; & \frac{1}{\tau}\left[\mathbb{E}^{\mathbb{P}}_t\left(\log\left(\frac{X_{t+\tau}}{X_t}\right)+\int_t^{t+\tau}(\delta_s+q^{\mathbb{P}}_s)ds\right)-\mathbb{E}^{\mathbb{Q}}_t\left(\log\left(\frac{X_{t+\tau}}{X_t}\right)+\int_t^{t+\tau}(\delta_s+q^{\mathbb{Q}}_s)ds\right)\right]\\ & -\;\frac{1}{\tau}\left[\mathbb{E}^{\mathbb{P}}_t\left(\int_t^{t+\tau}r_s ds\right)-\mathbb{E}^{\mathbb{Q}}_t\left(\int_t^{t+\tau}r_s ds\right)\right]\\ =\; & \frac{1}{\tau}\,\mathbb{E}^{\mathbb{P}}_t\left(\int_t^{t+\tau}\left(\alpha_s-(r_s-\delta_s)\right)\,ds\right).\end{aligned} \tag{4.5}$$

Notice that since a long position in the market index involves a commitment of capital a part of the wedge in the $\mathbb{P}$ and $\mathbb{Q}$ expectations of the cum-dividend equity returns reflects compensation for the time-variation in the risk-free interest rate.[13]

We next define corresponding risk measures for the return variation. The most natural and popular one is the quadratic variation over $[t,t+\tau]$ which we denote $QV_{t,t+\tau}$. This captures the return variation over the given horizon and is given by,

$$\begin{aligned}QV_{t,t+\tau} \;&=\; QV^c_{t,t+\tau}+QV^j_{t,t+\tau},\\ QV^c_{t,t+\tau} \;&=\; \int_t^{t+\tau}(V_{1,s}+V_{2,s})\,ds,\\ QV^j_{t,t+\tau} \;&=\; \int_t^{t+\tau}\int_{\mathbb{R}^2}x^2\,\mu(ds,dx,dy),\end{aligned} \tag{4.6}$$

where we decompose the return variation into terms generated by the continuous and jump component of X, $QV^c_{t,t+\tau}$ and X, $QV^j_{t,t+\tau}$. Further, note that the (realized) quadratic variation is independent of the probability measure. The variance risk premium is defined as,

$$\mathrm{VRP}^{\tau}_t \;\equiv\; \frac{1}{\tau}\left[\,\mathbb{E}^{\mathbb{P}}_t\left(QV_{t,t+\tau}\right)-\mathbb{E}^{\mathbb{Q}}_t\left(QV_{t,t+\tau}\right)\right], \tag{4.7}$$

[13]This term is, of course, absent if we instead define the equity risk premium using futures on the market index.

which provides compensation for the variance risk associated with both the continuous and the jump component of X.

We are also interested in assessing directly the risks and risk premiums associated with jumps. In particular, we want to gauge the compensation for large price jumps and to allow for a separate risk premium for the negative versus positive jumps. We obtain direct measures of the jump risks by simply counting the number of "big" jumps over the relevant horizon,

$$LT^K_{t,t+\tau} \equiv \int_t^{t+\tau} \int_{\mathbb{R}^2} 1_{\{x \leq -K\}}\, \mu(ds, dx, dy), \qquad RT^K_{t,t+\tau} \equiv \int_t^{t+\tau} \int_{\mathbb{R}^2} 1_{\{x \geq K\}}\, \mu(ds, dx, dy), \tag{4.8}$$

where K is a prespecified threshold. We set $K = 0.5\%$ in the subsequent analysis.[14] From the properties of the compensator for a jump measure, we have,

$$\begin{aligned} LT^K_{t,t+\tau} &= \int_t^{t+\tau} \int_{\mathbb{R}^2} 1_{\{x \leq -K\}}\, \nu^{\mathbb{P}}_s(dx, dy)\, ds + \epsilon^L_{t,t+\tau}, \quad \mathbb{E}^{\mathbb{P}}_t(\epsilon^L_{t,t+\tau}) = 0, \\ RT^K_{t,t+\tau} &= \int_t^{t+\tau} \int_{\mathbb{R}^2} 1_{\{x \geq K\}}\, \nu^{\mathbb{P}}_s(dx, dy)\, ds + \epsilon^R_{t,t+\tau}, \quad \mathbb{E}^{\mathbb{P}}_t(\epsilon^R_{t,t+\tau}) = 0. \end{aligned} \tag{4.9}$$

Hence, up to martingale difference sequences, $LT^K_{t,t+\tau}$ and $RT^K_{t,t+\tau}$ measure the $\mathbb{P}$ jump intensity of "large" jumps.

The risk measures we have introduced, including, $QV^c_{t,t+\tau}$, $QV^j_{t,t+\tau}$, $\widehat{LT}^K_{t,t+\tau}$ and $\widehat{RT}^K_{t,t+\tau}$, are not directly observable. However, we can estimate them using high-frequency futures data. We introduce the following empirical measures,

$$\begin{cases} \widehat{LT}^K_{t,t+\tau} = \displaystyle\sum_{i:\frac{i}{n} \in (t,t+\tau]} 1_{\{\Delta^n_i f > (K \vee \alpha n^{-\varpi})\}}, \\ \widehat{RT}^K_{t,t+\tau} = \displaystyle\sum_{i:\frac{i}{n} \in (t,t+\tau]} 1_{\{\Delta^n_i f < -(K \vee \alpha n^{-\varpi})\}}, \end{cases} \tag{4.10}$$

$$\widehat{QV}_{t,t+\tau} = \sum_{i:\frac{i}{n} \in (t,t+\tau]} |\Delta^n_i f|^2, \quad \widehat{QV}^c_{t,t+\tau} = \sum_{i:\frac{i}{n} \in (t,t+\tau]} |\Delta^n_i f|^2 1_{\{|\Delta^n_i f| \leq \alpha n^{-\varpi}\}}, \tag{4.11}$$

where the tuning parameters α and ϖ are determined in the same way as for the estimation of $\widehat{V}^{(n,m_n)}_t$ in Section 2. Under fairly general conditions we can show, see, e.g., Jacod (2008),

$$\widehat{LT}^K_{t,t+\tau} \xrightarrow{\mathbb{P}} LT^K_{t,t+\tau}, \quad \widehat{RT}^K_{t,t+\tau} \xrightarrow{\mathbb{P}} RT^K_{t,t+\tau}, \quad \widehat{QV}_{t,t+\tau} \xrightarrow{\mathbb{P}} \widehat{QV}_{t,t+\tau}, \quad \widehat{QV}^c_{t,t+\tau} \xrightarrow{\mathbb{P}} \widehat{QV}^c_{t,t+\tau}. \tag{4.12}$$

[14] This (fixed) threshold K is large enough that we can separate returns exceeding this (absolute) level from diffusive volatility using 1-minute observations. Experiments with alternative cutoffs produced similar results.

Similarly, we estimate the equity and variance risk premiums using the relations,

$$\begin{aligned}
&\log\left(\frac{X_{t+\tau}}{X_t}\right) - \frac{1}{\tau}\int_t^{t+\tau}(r_s-\delta_s)\,ds = \mathrm{ERP}_t^\tau + \frac{1}{\tau}\,\mathbb{E}_t^{\mathbb{P}}\left(\int_t^{t+\tau} q_s^{\mathbb{P}} ds\right) + \epsilon_{t,t+\tau}^E, \quad \mathbb{E}_t^{\mathbb{P}}(\epsilon_{t,t+\tau}^E) = 0, \\
&\widehat{\mathrm{VRP}}_t^\tau = \frac{1}{\tau}\left[\widehat{QV}_{t,t+\tau} - \mathbb{E}_t^{\mathbb{Q}}\left(QV_{t,t+\tau}\right)\right] = \mathrm{VRP}_t^\tau + \epsilon_{t,t+\tau}^V, \quad \mathbb{E}_t^{\mathbb{P}}(\epsilon_{t,t+\tau}^V) = 0,
\end{aligned} \tag{4.13}$$

where $\mathbb{E}_t^{\mathbb{Q}}\left(QV_{t,t+\tau}\right)$ can be measured in model-free fashion via the VIX index computed by the CBOE from a portfolio of S&P 500 index options. Equation (4.13) shows that a martingale difference sequence separates $\widehat{\mathrm{VRP}}_t^\tau$ from VRP_t^τ, and, likewise, a martingale difference sequence separates the log excess cum-dividend returns on the underlying asset from the unobservable $\mathrm{ERP}_t^\tau + \frac{1}{\tau}\mathbb{E}_t^{\mathbb{P}}\left(\int_t^{t+\tau} q_s^{\mathbb{P}} ds\right)$. In principle, we can remove the term stemming from the convexity adjustment, i.e., $\frac{1}{\tau}\mathbb{E}_t^{\mathbb{P}}\left(\int_t^{t+\tau} q_s^{\mathbb{P}} ds\right)$, via a consistent estimator for $\int_t^{t+\tau} q_s^{\mathbb{P}} ds$ (again up to a martingale difference term) obtained from high-frequency data, e.g.,[15]

$$\int_t^{t+\tau} \widehat{q}_s^{\mathbb{P}}\, ds \;=\; \sum_{i:\frac{i}{n}\in(t,t+\tau]} \left[\frac{1}{2}\left|\Delta_i^n f\right|^2 1_{\{|\Delta_i^n f|\le \alpha n^{-\varpi}\}} \;+\; \left(e^{\Delta_i^n f} - 1 - \Delta_i^n f\right) 1_{\{|\Delta_i^n f|>\alpha n^{-\varpi}\}}\right].$$

In practice, this adjustment is minute and the results are virtually unchanged if we account for it.[16]

4.2 The Predictability of Equity and Variance Risk and Risk Premia

Equipped with empirically feasible estimators for the relevant risk measures and risk premia, we may now explore the relationship between the option-implied factors, $V_{1,t}$, $V_{2,t}$ and U_t, that drive the dynamics of the option surface and the various risk measures and risk premia associated with the underlying asset. We rely on alternative versions of the following predictive regression,

$$y_t = \alpha_0 + \alpha_1 V_{1,t} + \alpha_2 V_{2,t} + \alpha_3 U_t + \epsilon_t, \tag{4.14}$$

where the right hand side variable represents, in turn, jump and diffusive variance risk measures as well as risk premia, i.e., $y_t = \widehat{LT}_{t,t+\tau}^K$, $\widehat{RT}_{t,t+\tau}^K$, $\widehat{QV}_{t,t+\tau}^c$, $\log\left(\frac{X_{t+\tau}}{X_t}\right) - \frac{1}{\tau}\int_t^{t+\tau}(r_s-\delta_s)ds$ and $\widehat{\mathrm{VRP}}_t^\tau$. Given the relationships explicated in equations (4.10) and (4.13), it is evident that the regressions based on the alternative y_t variables will yield, asymptotically, the identical estimates as those based on the corresponding infeasible measures of interest, i.e., $\int_t^{t+\tau}\int_{\mathbb{R}} 1_{\{x\le -K\}}\nu_s^{\mathbb{P}}(dx,dy)$, $\int_t^{t+\tau}\int_{\mathbb{R}} 1_{\{x\ge K\}}\nu_s^{\mathbb{P}}(dx,dy)$, $QV_{t,t+\tau}^c$, ERP_t^τ and VRP_t^τ, respectively. Thus, the predictive regression in equation (4.14) speaks directly to the linkages between the option surface dynamics and the risks and risk premia associated with the equity-index market.

[15] For the overnight periods we just take one half of the squared return.

[16] We do not report those results to conserve space. They are available upon request.

In general, if the premia for the diffusive and jump risks are spanned by the three factors V_1, V_2 and U, then the expectation of y_t conditional on the time t information set, $\mathcal{F}_t$, in equation (4.14) should be functionally related to $V_{1,t}$, $V_{2,t}$ and U_t. Moreover, in the standard case, almost universally adopted in empirical option pricing applications, the measure change preserves the affine structure, so the conditional mean of y_t in equation (4.14) is a linear function of $V_{1,t}$, $V_{2,t}$ and U_t. Hence, the regression in (4.14) produces the optimal forecast for the volatility and jump risks using the information at time t. In this context, we note that, even conceptually, our extraction of the factors $V_{1,t}$, $V_{2,t}$ and U_t provides a considerably richer information set for forecasting the relevant volatility and jump realizations than what is available through reliance strictly on the history of the underlying asset returns. The latter may at best generate estimates of the path for $\{V_{1,s} + V_{2,s}\}_{s \leq t}$ and an associated measure of the return variation generated by jumps.[17]

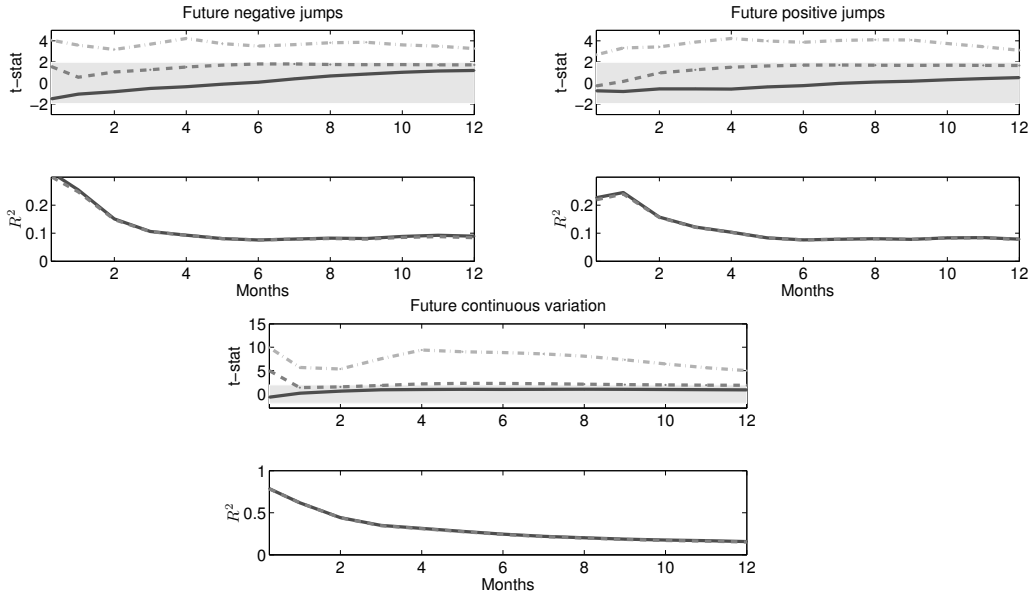


Figure 6: Predictive regressions for volatility and jump risks. The volatility and jump risk measures are defined in (4.10)-(4.11). For each regression, the top panels depict the t-statistics for the individual parameter estimates while the bottom panels indicate the regression R^2. The predictive variables are V_1 (dashed-dotted line), V_2 (dashed line) and $\widetilde{U}$ (solid line), where $\widetilde{U}$ is the residual from the linear projection of U on V_1 and V_2. Standard errors are constructed to account for the estimation error in $\widetilde{U}$. The dashed lines in the R^2 plots correspond to constrained regressions, including only V_1 and V_2.

[17]However, it is, of course, essential that we initially "correct" the option data for the relevant risk premia through the estimation of the parametric model (3.2).

The results from the predictive regressions are summarized on Figures 6 and 7. Given our relatively short sample, we compute the predictive regressions for horizons up to one year only. Our findings may be summarized as follows. First, Figure 6 shows that the state variables, extracted from the option panel, help predict the future evolution of risks. In particular, the plot pertaining to the count of positive and negative jumps demonstrates, rather convincingly, that the jump intensity varies under the statistical measure, $\mathbb{P}$, i.e., $\nu_t^{\mathbb{P}}(dx, dy)$ is truly a function of t.[18] At the same time, the three state variables V_1, V_2 and U have very different predictive power for the future volatility and jump risks of the underlying asset. Clearly, most of the predictability in Figure 6 stems from the volatility factors, V_1 and V_2. Once they are included in the regressions, U provides no incremental explanatory power for the future volatility and jump risk. This is evident both from the insignificant t-statistics corresponding to V_3 and the negligible drop in the R^2 value when excluding U from the regressions.[19] Hence, Figure 6 is consistent with a model for which the jump intensity, under $\mathbb{P}$, depends only on the two volatility factors V_1 and V_2, but not on U.[20]

Turning to the predictability of the equity and variance risk premia, displayed on Figure 7, we see that the three factors, extracted from the option panel, now take on very different roles. In particular, a significant part of the predictability of both the equity and the variance risk premia is due to the third factor U. For the variance risk premia, V_1 also contributes strongly to the predictive power at shorter horizons, while, importantly, for the equity risk premium U is the dominant explanatory factor for all horizons up through at least nine months. The similarity between the predictive regression results for the equity and variance risk premia in terms of the role of U is consistent with Bollerslev and Todorov (2011b) who find that the equity and risk premia have a common component stemming from the compensation for left tail jump risk.

Recall that U appears only in the risk-neutral intensity of the negative jumps in model (3.2). The common dependence of the equity and variance risk premia on U, coupled with the significant persistence of the latter, can rationalize the predictive power of the variance risk premium for future excess returns, documented in Bollerslev et al. (2009) and Drechsler and Yaron (2011). Our finding of a limited role for the two volatility factors in this context is consistent with earlier work on the risk-return tradeoff, see, e.g., French et al. (1987) and Glosten et al. (1993).[21] Finally, we

[18] For further nonparametric evidence of time variation in the $\mathbb{P}$-jump intensity, using only high-frequency futures data, see Bollerslev and Todorov (2011a).

[19] This is consistent with Andersen and Bondarenko (2011) who find that moderate-to-deep OTM put options have no predictive power for the future volatility of the underlying equity index returns once they control for the information provided by ATM and OTM call options.

[20] In general, since V_1 contains jumps with a time-varying intensity that loads on all the factors, V_1, V_2 and U, the conditional forecast of future volatility $V_{1,t} + V_{2,t}$ still depends on the current value of the third factor U.

[21] Predictive regressions for future excess returns involving only V_1 and V_2 produce insignificant t-statistics for both.

note that our results are not driven solely by the events surrounding the financial crisis. We obtain qualitatively identical results from subsamples that end in July 2007.[22]

4.3 Structural Implications of the Predictive Power of the Option Surface

Figures 6 and 7 demonstrate that the factor U, which drives a substantial part of the OTM short maturity put option dynamics, has no impact on the *actual* dynamics for the volatility and jump components of the underlying asset. In contrast, the factor has a critical effect on the *pricing* of volatility and jump risk. How can we rationalize these findings? To guide our intuition we compare the results in Figures 6 and 7 with those from the structural model of Drechsler and Yaron (2011) which links the option prices to fundamental macroeconomic risks, such as aggregate consumption and dividends.[23] In this model a representative agent with Epstein-Zin preferences faces consumption growth risk with time-varying conditional mean and volatility, both of which contain non-Gaussian shocks, i.e., jumps. Drechsler and Yaron (2011) show that the time-varying jump intensity – in turn governed by the volatility state – explains a significant part of the predictability in the variance risk premia of future excess returns. This structural model has three state variables which all are related to features of the growth rate of consumption, namely the conditional mean, the conditional variance, and the variation in the long run variance of consumption growth. Effectively, the last state variable serves as a second volatility factor.

On Figure 8 we plot the predictive regressions for future volatility and jump risks as well as equity and variance risk premia, using the concurrent level of the three state variables in the Drechsler and Yaron (2011) model as predictors. Not surprisingly, the long-run risk factor helps predict neither the future equity volatility and jump risks nor the equity and variance risk premia. In this structural setting, essentially all predictability stems from the two volatility factors. Moreover, the volatility factors provide the channel through which the past variance risk premia generate predictable movements in the equity risk premium. Thus, relative to our empirical finding, captured by Figures 6 and 7, the structural model ties the predictability of future volatility and jump risks closely together with the predictability of the equity and variance risk premia. Equivalently, the structural model implies a tight relationship between the dynamics of the option panel and the return dynamics of the underlying stock market. In contrast, our empirical results based on model (3.2) – which nests all standard two-factor affine volatility specifications – document a partial, and critical, decoupling between the factors driving the equity return dynamics and those governing the

[22]The documentation of these results are available upon request.

[23]This model builds on Eraker and Shaliastovich (2008) and generalizes many prior models in which consumption growth contains a small predictable persistent component, including the original work of Bansal and Yaron (2004).

pricing of risk, and thus the equity risk premium and variance risk premium.

The discrepancy between our findings and the implications of the Drechsler and Yaron (2011) model is not accidental. Any structural model of a representative agent with Epstein-Zin preferences and an affine specification for the fundamentals will tend to encounter a similar problem. This stems from restrictions on the jump measure change implied by the equilibrium pricing kernel. Formally, the part of the density $\frac{d\mathbb{Q}}{d\mathbb{P}}$ due to the change of the jump measure is, in general, given by,

$$\mathcal{E}\left(\int_0^t \int_{\mathbb{R}^n} (Y(\omega, s, \mathbf{x}) - 1)\widetilde{\mu}^{\mathbb{P}}(ds, d\mathbf{x})\right), \tag{4.15}$$

where the jump component of the state vector in the economy under the $\mathbb{P}$ probability measure is given as a (multivariate) integral of the form $\int_0^t \int_{\mathbb{R}^n} \mathbf{x}\, \widetilde{\mu}^{\mathbb{P}}(ds, d\mathbf{x})$; $Y(\omega, t, \mathbf{x}) = \frac{\nu_t^{\mathbb{Q}}(d\mathbf{x})}{\nu_t^{\mathbb{P}}(d\mathbf{x})}$ is the measure change for the jump intensity and $\mathcal{E}$ is the Doleans-Dade exponential.[24] In general $Y(\omega, t, \mathbf{x})$ can be stochastic. However, the restriction from the equilibrium model, with fundamentals having affine dynamics and a representative agent with Epstein-Zin preferences, is that the expression in equation (4.15) is exponentially affine in the state vector, see e.g., equation (2.22) of Eraker and Shaliastovich (2008). This, in particular, implies that in the equilibrium setting $Y(\omega, s, \mathbf{x})$ must be non-random and time-invariant, i.e., depend only on $\mathbf{x}$.[25] The result carries over to the price jumps (and their volatilities) as, in the equilibrium setting, they are exponentially affine in the state vector. In contrast, for our extended three-factor model, in which $\nu_t^{\mathbb{P}}(\cdot)$ does not involve U_{t-} (or loads only marginally on it), the ratio $\frac{\nu_t^{\mathbb{Q}}(dx,dy)}{\nu_t^{\mathbb{P}}(dx,dy)}$, driving the change of measure for the jumps (in price and volatility), is nonlinear in the state variables due to the fluctuations in our U factor.

It follows that, within the setting of Drechsler and Yaron (2011), the risk-neutral jump intensity is proportional to that under the actual probability measure, so the jump intensities are equivalent in terms of their time variation. Therefore, the jump risk premia are generated by changing the distribution of the jump size only. This implies that the time-variation of the jump intensity is "inherited" under an equivalent change of measure. Thus, this equilibrium based pricing kernel cannot "generate" new state variables in addition to those that drive the fundamental risks in the economy. In turn, this generates the tight link between the underlying asset dynamics and the option panel dynamics within the structural model of Drechsler and Yaron (2011).

There are several ways in which the link between the asset and option price dynamics may be relaxed within an equilibrium setting to potentially account for our empirical evidence. They

[24] For the expression in (4.15) and the definition of the Doleans-Dade exponential, see Jacod and Shiryaev (2003), Corollary III.5.22 and I.4.59, respectively.

[25] The restriction on $Y(\omega, s, \mathbf{x})$ is actually stronger. In the equilibrium model $Y(\omega, s, \mathbf{x})$ is an exponential function of jump size, see Theorem 1 of Eraker and Shaliastovich (2008). Hence, the jump measure change is implemented by exponential tilting, with the degree of tilting determined by the preference parameters of the representative agent.

all involve generalizing the preferences of the representative agent in some form.[26] One approach is to allow the representative agent's coefficient of risk aversion to vary over time. Du (2010) proposes a generalization of a habit formation model in which consumption growth is i.i.d. with rare jumps, building on the equilibrium models of Campbell and Cochrane (1999) and Barro (2006). In this model, and consistent with our findings, there is a wedge between the time-variation of the $\mathbb{P}$ and $\mathbb{Q}$ jump intensities: the former is constant while the latter is time-varying.[27] However, the time variation of the $\mathbb{Q}$ jump intensity is driven by the habit formation mechanism and is purely a function of the consumption surplus ratio. The latter, however, determines uniquely the diffusive volatility as well. This is at odds with our finding as the presence of U in the risk-neutral jump intensity drives a wedge between the time variation of the latter and that of the stochastic volatility. Thus, it remains an open question whether a model with external habit formation can decouple the option and asset price dynamics in a manner reminiscent of our empirical findings. It is also unclear whether the frequency and intensity of stock market jumps can be mapped into corresponding jumps in the consumption growth rate as implied by this model of habit formation.

An alternative way to relax the link between the option and asset price dynamics is by recognizing that agents do not directly observe the state vector and therefore need to filter the states from observables. The absence of perfect information plus aversion to ambiguity (particularly about extreme negative risks), like in Hansen and Sargent (2008), or lack of confidence in the estimate of the state vector, like in Bansal and Shaliastovich (2010), can make investors appear more risk averse than they would be in a perfect information setting. The key question is whether the ambiguity aversion or confidence risk variation can generate the required gap between the dynamics of the statistical and risk-neutral jump intensities.[28] For example, the tight link between the jump intensity under $\mathbb{P}$ and $\mathbb{Q}$ remains within the generalization of the Drechsler and Yaron (2011) model, developed by Drechsler (2013), in which agents are ambiguous about part of the dynamics, including the jumps in the conditional mean and variance of consumption growth. The representative agent's ambiguity drives an additional wedge between fundamental risks and asset prices and helps explain why variance risk premia have superior predictive power, relative to volatility itself, for future returns. Nonetheless, the linearity of the pricing kernel with respect to the state vector implies that no new state variable is "generated" going from $\mathbb{P}$ to $\mathbb{Q}$, thus ultimately rendering the model predictions incompatible with our empirical findings.

[26] Albuquerque et al. (2013) also argue for the need to generalize Epstein-Zin preferences to reduce the correlation between consumption and asset returns generated by an equilibrium model to better match its empirical counterpart.

[27] Our empirical findings, unlike the model of Du (2010), suggest strong time variation in the $\mathbb{P}$ jump intensity.

[28] Recall that, according to the estimates for (3.2), U does not covary perfectly with stochastic volatility.

5 Conclusion

We document that the standard exponentially-affine jump-diffusive specifications used in the empirical option pricing literature are incapable of fitting critical features of the option surface dynamics for the S&P 500 index, especially in scenarios involving significant shifts in the volatility smirk. We extend the risk-neutral volatility model to include a separate state variable which is crucial in capturing the time variation of priced downside tail risk. This new factor has no incremental explanatory power, beyond the traditional volatility factors, for the future evolution of volatility and jump risks. On the other hand, relative to the volatility components, the new factor provides critical, and superior, information for the time variation in the equity and variance risk premia.

Our findings demonstrate that the pricing in the option market is closely integrated with the underlying asset market. Moreover, the option panel embodies critical information regarding equity risk pricing that cannot be extracted directly from the underlying asset price dynamics. The wedge between the two probability measures arises primarily from the varying degree of compensation for downward tail jump risk. Our results suggest that time-varying risk aversion and/or ambiguity aversion, driven in part by the presence of large shocks, must be incorporated into structural asset pricing models if they are to explain the joint dynamics of the equity and option markets.

6 Appendix: An Alternative Three-Factor Model

In this appendix we estimate an alternative three-factor stochastic volatility model for the risk-neutral dynamics from the option panel and compare its performance to the three-factor model (3.2) used in the main text. This alternative model is more in line with the typical representation in the literature. It stipulates the following stock price evolution under the risk-neutral measure:

$$\begin{aligned}
\frac{dX_t}{X_{t^-}} &= (r_t - \delta_t)dt + \sqrt{V_{1,t}}dW^{\mathbb{Q}}_{1,t} + \sqrt{V_{2,t}}dW^{\mathbb{Q}}_{2,t} + \sqrt{V_{3,t}}dW^{\mathbb{Q}}_{3,t} + \int_{\mathbb{R}^2}(e^x - 1)\widetilde{\mu}^{\mathbb{Q}}(dt, dx, dy),\\
dV_{1,t} &= \kappa_1(\overline{v}_1 - V_{1,t})dt + \sigma_1\sqrt{V_{1,t}}dB_{1,t} + \mu_{V_1}\int_{\mathbb{R}^2} x^2 1_{\{x<0\}}\mu(dt, dx, dy),\\
dV_{2,t} &= \kappa_2(\overline{v}_2 - V_{2,t})dt + \sigma_2\sqrt{V_{2,t}}dB_{2,t},\\
dV_{3,t} &= -\kappa_3 V_{3,t}dt + \mu_{V_3}\int_{\mathbb{R}^2}((1-\rho_3)x^2 1_{\{x<0\}} + \rho_3 y^2)\mu(dt, dx, dy),
\end{aligned} \tag{6.1}$$

where the jump measure μ has a compensator (under $\mathbb{Q}$) given by $dt \otimes \nu^{\mathbb{Q}}_t(dx, dy)$,

$$\nu^{\mathbb{Q}}_t(dx, dy) = \left\{\left(c^- 1_{\{x<0\}}\lambda_- e^{-\lambda_-|x|} + c^+ 1_{\{x>0\}}\lambda_+ e^{-\lambda_+ x}\right)1_{\{y=0\}} + c^- 1_{\{x=0,\ y<0\}}\lambda_- e^{-\lambda_-|y|}\right\}dx \otimes dy,$$

$$c^- = c^-_0 + c^-_1 V_{1,t-} + c^-_2 V_{2,t-} + c^-_3 V_{3,t-}, \quad c^+ = c^+_0 + c^+_1 V_{1,t-} + c^+_2 V_{2,t-} + c^+_3 V_{3,t-}.$$

Comparing the models (6.1) and (3.2), we note that they differ only in the role of the third factor V_3 which now is a regular volatility component, while the corresponding factor U in model (3.2) affects only the intensity of negative jumps. That is, for model (6.1), V_3 contributes to the (diffusive) price volatility as well as the negative jump intensity. Thus, model (6.1) is much closer to the standard way of introducing time-variation in jump intensity, see e.g., Bates (2000).

The parameter estimates are reported in Table 4. In Figure 9, we plot the sensitivity of the option surface to changes in the state variables, fixing the parameters at their point estimates. The main difference from Figure 3 is the options' sensitivity to the third factor (V_1 and V_2 switch roles in terms of persistence, hence the more pronounced effect of V_2 on long term options in Figure 9). In model (6.1), V_3 has a non-trivial effect on short-maturity ATM options as well as OTM puts unlike its role in model (3.2) where U impacts only the latter. Hence, model (6.1) will have difficulty separating volatility from the risk-neutral jump intensity.

Table 4: Parameter Estimates for Model (6.1)

Parameter	Estimate	Std.	Parameter	Estimate	Std.	Parameter	Estimate	Std.
ρ_1	−0.8877	0.0252	σ_2	0.0929	0.0084	c_2^-	146.3945	1.3153
$\overline{v}_1$	0.0150	0.0006	μ_{V_3}	5.5825	0.3689	c_2^+	26.1797	36.5320
κ_1	4.0525	0.1462	κ_3	17.2955	0.6228	c_3^-	116.3945	4.6420
σ_1	0.3469	0.0110	ρ_3	0.4696	0.0284	λ_-	17.1616	0.1048
ρ_2	−0.9891	0.0582	c_0^+	6.3008	0.2052	λ_+	78.3691	0.4252
$\overline{v}_2$	0.0034	0.0003	c_1^-	0.3621	1.5596	$\mu_{v_1,-}$	3.2444	0.4862
κ_2	1.2871	0.0433	c_1^+	292.9071	49.2676			

Note: Parameter estimates of the three-factor model (6.1) for S&P 500 equity-index option data sampled every Wednesday over the period January 1996-July 2010.

This conjecture is confirmed by the z-scores for an adequate fit to separate segments of the option surface for model (6.1), plotted in Figure 10. As for the two-factor model (3.1), model (6.1) severely undervalues long-maturity OTM put options from the beginning of 1998 till the end of 1999 as well as from early 2009, and it overvalues ATM puts in the latter of the two periods. Moreover, the formal tests for fit to the separate regions of the option surface have substantially higher rejection rates for model (6.1) than for (3.2). Note that models (6.1) and (3.2) have the

identical number of parameters, so that non-nested model assessment reduces to a comparison of their success in fitting the option surface. The three-factor model (3.2) produces a RMSE 1.75% for the options in our sample, while it is 1.89% for the three-factor stochastic volatility model (6.1).

Table 5: Diagnostic Tests for S&P 500 Option Data for Model (6.1)

Test	**Nominal Size**	
	1%	5%
Panel A: Fit to the Option Panel		
OTM, short-maturity puts	20.79%	40.39%
ATM, short-maturity	12.63%	25.79%
OTM, short-maturity calls	35.13%	53.03%
OTM, long-maturity puts	29.34%	40.13%
ATM, long-maturity	44.74%	60.00%
OTM, long-maturity calls	44.74%	66.84%
Panel B: Root-Mean Squared Error of IV Option Fit		
		1.89%

Note: Tests for fitting option surface by model (6.1). Notation as in Table 2.

We conclude that the specification of the state variables has a critical impact on the success in matching the dynamics of the observed option surface. In particular, separating the dynamics of the jump intensity from that of the stochastic volatility, as done in model (3.2), improves the overall fit to the option panel substantially.

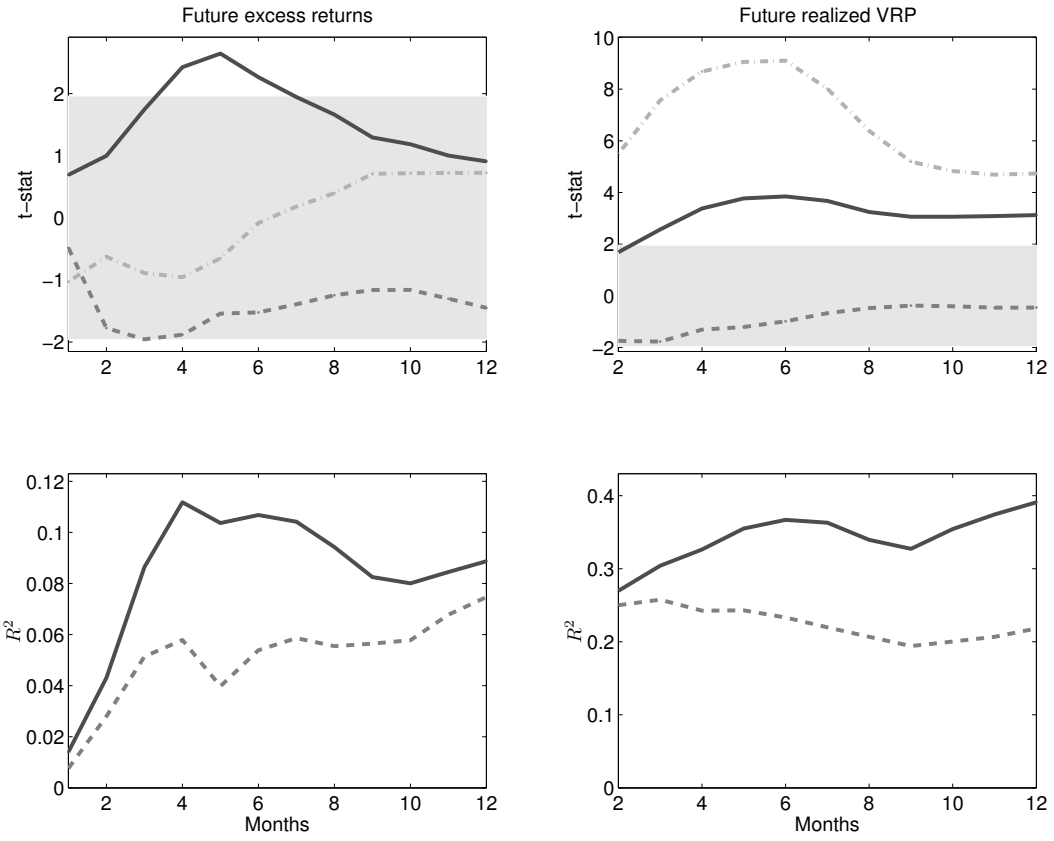


Figure 7: Predictive regressions for equity and variance risk premia. For each regression, the top panels depict the t-statistics for the individual parameter estimates while the bottom panels indicate the regression R^2. The predictive variables are V_1 (dashed-dotted line), V_2 (dashed line) and $\widetilde{U}$ (solid line), where $\widetilde{U}$ is the residual from the linear projection of U on V_1 and V_2. Standard errors are constructed to account for the estimation error in $\widetilde{U}$. The dashed lines in the R^2 plots correspond to constrained regressions, including only V_1 and V_2.

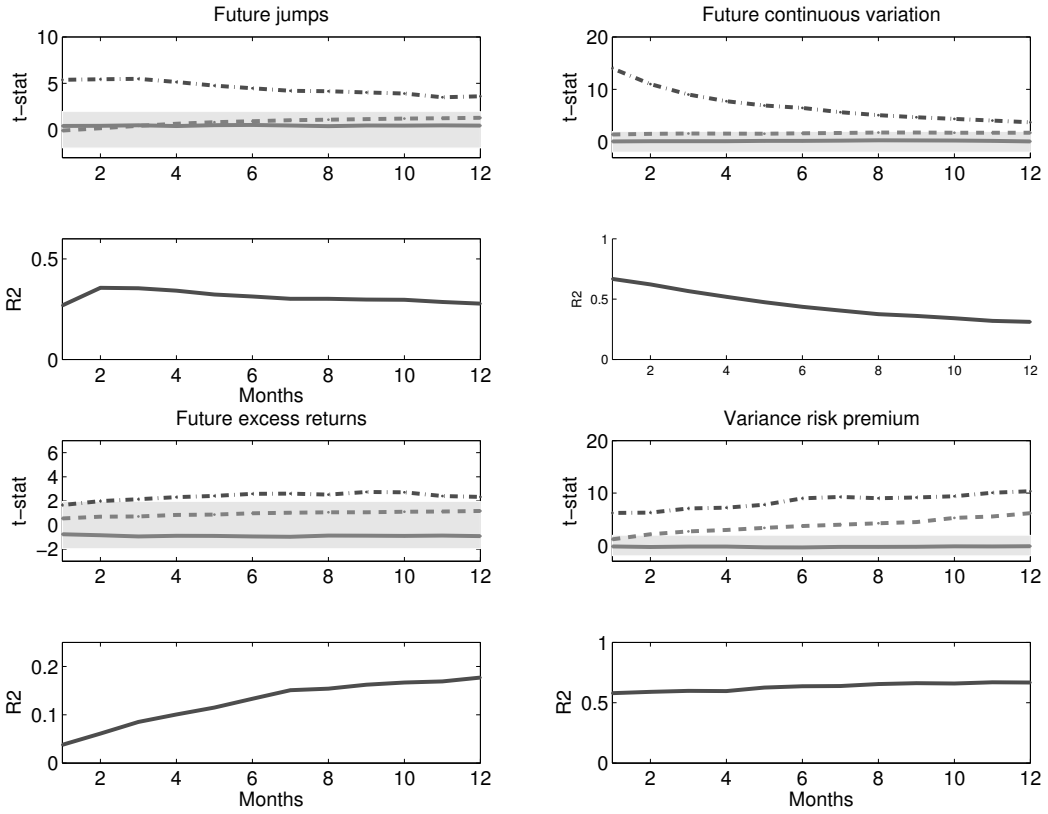


Figure 8: Predictive regressions implied by the Drechsler and Yaron (2011) structural model. The predictive variables are the conditional mean of consumption growth (solid line), stochastic volatility of consumption growth (dashed-dotted line) and central tendency of stochastic volatility (dashed line). The dashed lines in the R^2 plots correspond to constrained regressions including only the volatility state variables.

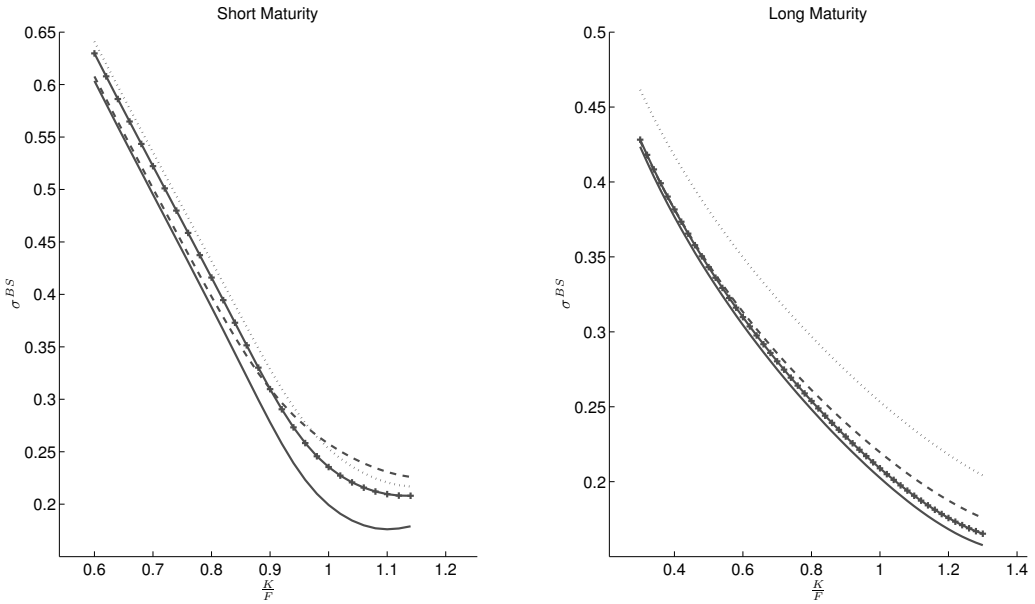


Figure 9: Sensitivity of option surface to changes in state variables in model (6.1). Parameters are set equal to their estimates reported in Table 4. The solid line corresponds to the implied volatility for values of the three state variables equal to their sample means. All other lines correspond to changing one of the state variables from its sample mean value. The dashed line corresponds to changing V_1, the dotted line to changing V_2 and the x-line to changing V_3. The change in V_1 is three times its mean value; the change in V_2 is its mean value; the change in V_3 is such that the effect on the risk-neutral jump intensity of the negative jumps is the same as that due to changing V_2. Short maturity is one-month and long maturity is one year.

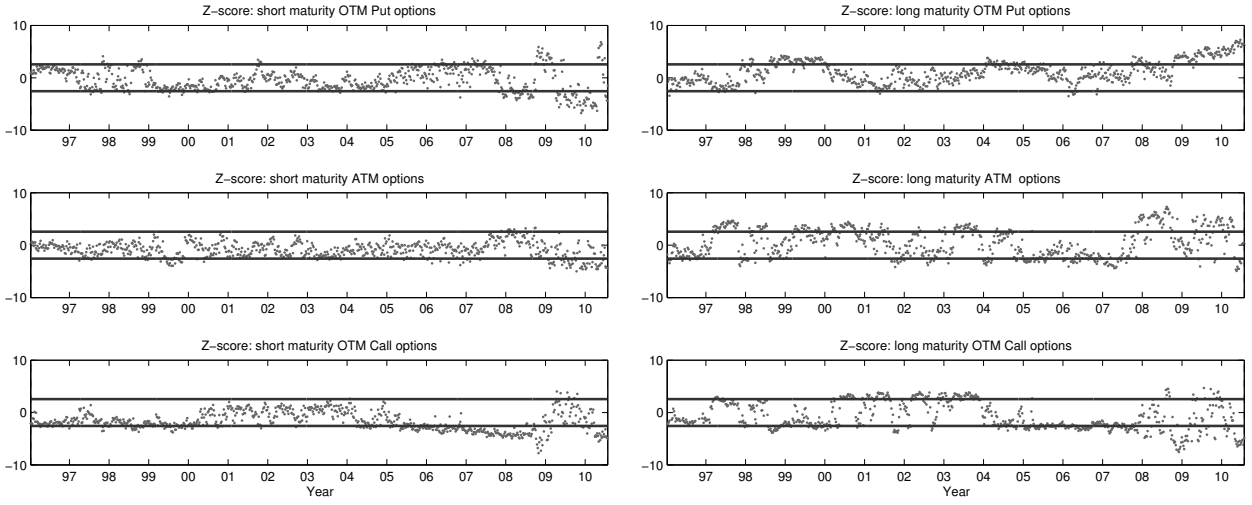


Figure 10: z-scores for model (6.1). The construction of the z-scores and definition of the regions of the option surface are exactly the same as in Figure 2.

References

Albuquerque, R., M. Eichenbaum, and S. Rebelo (2013). Valuation Risk and Asset Pricing. Working paper, Northwestern University.

Andersen, T. G. and O. Bondarenko (2011). Dissecting the Market Pricing of Volatility. Working paper, Northwestern University and University of Illinois at Chicago.

Andersen, T. G., N. Fusari, and V. Todorov (2012). Parametric Inference and Dynamic State Recovery from Option Panels. Working paper, Northwstern University.

Bansal, R. and I. Shaliastovich (2010). Confidence Risk and Asset Prices. *American Economic Review 100*, 537–541.

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

Barro, R. J. (2006). Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics 121*, 823–866.

Bates, D. S. (2000). Post-'87 Crash Fears in S&P 500 Future Options. *Journal of Econometrics 94*, 181–238.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies 22*, 4463–4492.

Bollerslev, T. and V. Todorov (2011a). Estimation of Jump Tails. *Econometrica 79*, 1727–1783.

Bollerslev, T. and V. Todorov (2011b). Tails, Fears and Risk Premia. *Journal of Finance 66*, 2165–2211.

Bollerslev, T. and V. Todorov (2013). Time Varying Jump Tails. Working paper, Duke University and Northwestern University.

Broadie, M., M. Chernov, and M. Johannes (2009). Specification and Risk Premiums: The Information in S&P 500 Futures Options. *Journal of Finance 62*, 1453–1490.

Campbell, J. and J. Cochrane (1999). By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy 107*, 205–251.

Carr, P. and L. Wu (2003). What Type of Process Underlies Options? A Simple Robust Test. *Journal of Finance 58*, 2581–2610.

Drechsler, I. (2013). Uncertainty, Time-Varying Fear, and Asset Prices. *Journal of Finance, forthcoming*.

Drechsler, I. and A. Yaron (2011). What's Vol Got to Do with It? *Review of Financial Studies 24*, 1–45.

Du, D. (2010). General Equilibrium Pricing of Options with Habit Formation and Event Risks. *Journal of Financial Economics 99*, 400–426.

Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine Processes and Applications in Finance. *Annals of Applied Probability 13(3)*, 984–1053.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*, 1343–1376.

Durrleman, V. (2008). Convergence of at-the-money Implied Volatilities to the Spot Volatility. *Journal of Applied Probability 45*, 542–550.

Eraker, B. (2004). Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *Journal of Finance 59*, 1367–1403.

Eraker, B. and I. Shaliastovich (2008). An Equilibrium Guide to Designing Affine Pricing Models. *Mathematical Finance 18*, 519–543.

French, K., W. Schwert, and R. Stambaugh (1987). Expected Stock Returns and Volatility. *Journal of Financial Economics 19*, 3–29.

Glosten, L., R. Jaganathan, and D. Runkle (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance 48*, 1779–1801.

Hansen, L. and T. Sargent (2008). *Robustness*. Princeton University Press.

Jacod, J. (2008). Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales. *Stochastic Processes and their Applications 118*, 517–559.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems For Stochastic Processes* (2nd ed.). Berlin: Springer-Verlag.

Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics 36*, 270–296.

Pan, J. (2002). The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study. *Journal of Financial Economics 63*, 3–50.

Singleton, K. (2006). *Empirical Dynamic Asset Pricing*. Princeton University Press.

Todorov, V. (2011). Econometric Analysis of Jump-Driven Stochastic Volatility Models. *Journal of Econometrics 160*, 12–21.