The Transient and The Persistent Variance Risk Premium

Abstract

This paper examines theoretically and empirically a two-factor stochastic volatility model. We adopt an affine two-factor stochastic volatility model, where aggregate market volatility is decomposed into two independent factors; a persistent factor and a transient factor. We introduce a pricing kernel that links the physical and risk neutral distributions, where investor's equity risk preference is distinguished from her variance risk preference. Using simultaneous data from the S&P 500 index and options markets, we find a consistent set of parameters that characterizes the index dynamics under physical and risk-neutral distributions. We show that the proposed decomposition of variance factors can be characterized by a different persistence and different sensitivity of the variance factors to the volatility shocks. We obtain negative prices for both variance factors, implying that investors are willing to pay for insurance against increases in volatility risk, even if those increases have little persistence. We also obtain negative correlations between shocks to the market returns and each volatility factor, where correlation is less significant in transient factor and therefore has a less significant effect on the index skewness. Our empirical results indicate that unlike stochastic volatility model, join restrictions do not lead to the poor performance of two-factor SV model, measured by Vega-weighted root mean squared errors.

JEL Classi ication: G10; G12; G13
Keywords: variance risk premium; two-factor stochastic volatility; GARCH; joint estimations;

1 Introduction

The dynamics of index return volatility and their role in pricing options have had a long history following the classic early works by Wiggins (1987) and Heston (1993), that recognized the volatility's stochastic nature and managed to derive closed form expressions for the resulting European options. Related early contributions were also by Duan (1995), Duan et al. (1999), and Heston and Nandi (2000) under GARCH return dynamics, with option prices derived either by closed form expressions or numerical methods. More recent studies, however, have pointed out that a single factor stochastic volatility (SV) or GARCH is not sufficient to simultaneously fit the persistence of volatility and the volatility of volatility. Two volatility factors, one with persistent dynamics and one with transient dynamics, are needed to model return volatility dynamics in both the underlying (P) and the risk neutral (Q) measures for the key S&P 500 index and its options.[1] Methodologically, these studies either fit numerically the two volatility factor return dynamics to both return and option data, or limit themselves only to one of the two returns, most often the option-implied ones.

This paper examines index option pricing under two SV factors, in an integrated theoretical and empirical framework, by reconciling the two markets where the underlying and the options are traded through a pricing kernel that contains the index return and the two volatility factors. Aggregate market volatility is decomposed into a more persistent volatility component, which has nearly a unit root, and a transitory volatility component, which has a more rapid time decay. We adopt an affine two-factor SV process for the underlying index returns and introduce an admissible pricing kernel to find the risk-neutral returns dynamic and to price European options.[2] We also introduce an associated component volatility model (bivariate GARCH model) and derive the corresponding pricing kernel linking the P- and Q-distributions under these dynamics. Although our study is not the first one to examine multifactor SV and GARCH models, it is to our knowledge the only one to present consistent P- and Q-parameter estimates both theoretically and empirically. Our paper has the same relationship to the cited bivariate SV and GARCH option pricing models as the Christoffersen et al. (2013) study had to the earlier Heston (1993) and Heston and Nandi (2000) models.

In our empirical work, we apply our theoretically derived two-factor SV and GARCH models to the S&P 500 index estimating the joint dynamics of returns and variances under the P and Q measures.[3] First, we derive two vectors of daily spot variances using the Particle

[1] See, for instance, Bollerslev and Zhou (2002), Alizadeh et al. (2002), and Chernov et al. (2003) for the P-returns and Bates (2000), Christoffersen et al. (2008), and Christoffersen et al. (2009) for the option-based Q-distribution.

[2] Note that the extracted risk-neutral dynamics are not restricted to the introduced admissible pricing kernel, where investor's variance risk preference is distinguished from her equity risk preference. In other words, we can obtain the risk-neutral dynamics without completely characterizing the equilibrium in economy. To do so, we specify a class of Radon-Nikodym derivatives and derive restrictions that ensure the existence of equivalent martingale measure, which makes the discounted stock price process a martingale.

[3] Joint estimation appropriately weights returns and option data and simultaneously address the model's ability to fit the time-series of returns and cross-section of option prices. The importance of joint estimation of the structural parameters of the underlying returns and volatility dynamics has been addressed in Bates

Filter (PF) method[4] and extend the conventional filtration procedure of similar studies by a novel procedure for the separation of the two variance components' paths. We then use a likelihood-based loss function that combines both underlying and option-based data to obtain a consistent set of structural parameters for the two-factor models. Our estimations are based on a 15-year data set that contains daily observations and the entire cross section of option prices, unlike earlier studies that limited themselves to a very short time series and weekly or monthly options in order to minimize the computational burden.[5] To the best of our knowledge, this is the first study that estimates consistent P- and Q-parameters from underlying index return and option data and reports variance risk premiums for a persistent and a transient component.

We find that one of the volatility factors is highly persistent (persistent component) while the immediate impact of volatility shocks on the other volatility factor is larger but short-lived (transient component). We also find the same level of persistence in the transient and persistent variance components when we only use option data, which is consistent with previous studies in option market. The unconditional transient and persistent variances are consistent with the average filtered spot transient and persistent variance components. Consistent with our intuition, we observe that the transient volatility component is much more volatile than the persistent volatility component. The same result holds when we use only option data.

We also find negative prices for both variance components, $\lambda_1 = -1.0798$ and $\lambda_2 = -1.0355$, implying that investors are willing to pay for an insurance against an increase in volatility risk, even if that increase has little persistence. To the best of our knowledge, this finding is novel for the option market, since none of the previous two-factor SV studies reports the price of the variance risk factors. It is, however, consistent with the findings in Adrian and Rosenberg (2008), who find negative and significant prices for both short-run and long-run volatility components in stock return data.[6]

We obtain a negative correlation between shocks to the market returns and each variance component, implying that both components are important in capturing the so-called leverage effect. Nonetheless, the point estimate of the transient correlation parameter ($\rho_2 = -0.2173$) is smaller in absolute value that that of the persistent one ($\rho_1 = -0.6918$), implying a weaker impact on the volatility smirk, on the return skewness and kurtosis, and on the the price of out-of-money put options. We observe the same pattern between correlation parameters when we estimate the model only with option data.

In the remainder of this section we complete the literature review. Extensive empirical evidence supports the presence of two volatility components in the dynamics of the market

(1996), Chernov and Ghysels (2000), Pan (2002), Eraker (2004), and Broadie et al. (2007) among others.

[4] For the application of PF in estimating the model parameters see Gordon et al. (1993), Johannes et al. (2009), Johannes and Polson (2009), Christoffersen et al. (2010), and Boloorforoosh (2014).

[5] See, for instance, Pan (2002) and Eraker (2004).

[6] Note that Adrian and Rosenberg (2008) introduce a discrete-time model where short-run and long-run volatility components are distinguished by construction whereas in our models we do not impose any restrictions on the variance dynamics other than variance shocks are independent.

returns. For, the P-distribution the relative performance of the two-factor SV structure compared to its one-factor counterpart in the dynamics of the exchange rate and equity returns has been examined in Bollerslev and Zhou (2002), Alizadeh et al. (2002), and Chernov et al. (2003).[7] These studies document that one-factor models are incapable of simultaneously fitting the persistence of volatility and the volatility of volatility. For instance, Chernov et al. (2003) suggest that the addition of a second volatility factor breaks the link between tail thickness and volatility persistence, leading to a significant improvement in capturing the return dynamics in affine models. They also find that when the second volatility factor is allowed to have its own correlation with returns, the correlation parameters can take on both positive and negative values, contrary to the findings in single factor volatility models, where the correlation parameter is always negative.

Similar considerations also hold for the Q-distribution. Earlier studies in the option markets such as Bakshi et al. (1997), Bates (2000), Jones (2003), and Egloff et al. (2010) have noted the problems with single factor SV models in the modeling of the volatility smirk.[8] Other empirical studies such as Derman (1999) note that the shape of the volatility smirk can be either flat or steep at a given volatility level, but stochastic volatility models cannot accommodate both at the same time for a given parametrization.[9] This problem in one factor SV models is more serious when estimating the model parameters using multiple cross-sections of options data, since the correlation between stock returns and variance is constant across all cross-sections of option contracts regardless of the level and shape of the volatility. Multiple SV models, on the other hand, can better capture the time-varying nature of the smirk as the correlation between stock returns and total volatility is stochastic.[10] Moreover, the conditional skewness and kurtosis are more flexible for given levels of conditional variance.

Inconsistencies in the joint estimation of the SV model are illustrated by Broadie et al. (2007), who note the failure of SV model to reconcile the P- and Q-estimates of certain structural parameters of the SV model (correlation coefficient and volatility of volatility) and conclude that the SV model is basically misspecified. They also show that the joint restrictions on the returns and volatility dynamics under the P and Q measures lead to the poor performance of the SV model, which cannot generate sufficient amounts of conditional skewness and kurtosis. Christoffersen et al. (2008) introduced a two-component GARCH model, which can generate more flexible skewness and volatility of volatility dynamics in capturing the dynamics of the S&P 500 index returns and in pricing European S&P 500 call options.

[7] There is also evidence that multifactor volatility model is needed to capture the term structure of the interest rates. See Dai and Singleton (2000, 2002) among others.

[8] Egloff et al. (2010, Page 1289) investigate the volatility term structure effect by incorporating mean reversion in variance dynamics. They show that the upward sloping autocorrelation term structure of variance swap rate quotes points to the existence of multiple variance risk factors and is evidence for non-zero market prices for variance risk factors.

[9] See Derman (1999).

[10] Christoffersen et al. (2009, Equation 15) show that the correlation between returns and total volatility in a two-factor SV model is stochastic. Such models, therefore, have more flexibility to fit the term structure of the volatility and to control the level and the slope of volatility smirk in cross-sections of option prices. See, for instance, Egloff et al. (2010) and Mencía and Sentana (2013).

Nonetheless, the absence of an explicit pricing kernel linking the P- and Q-distributions in that study necessitated either the use of an arbitrary price of volatility risk or the estimation of the risk neutral parameters by relying on the Q-distribution only. Similarly, Christoffersen et al. (2009) use only the Q-distribution to further explore multiple variance factors and find that it can generate stochastic correlation between total instantaneous volatility and stock returns. Our own empirical analysis confirms the advantages of the two-factor SV and GARCH models by using our theoretically integrated P and Q dynamics and finds that the joint restrictions do not lead to the poor performance of the two-factor SV model.

This paper proceeds as follows. Section 2 presents the theoretical model for pricing index options under SV and GARCH. Section 3 contains the description of the data set. In section 4, we discuss the methodology for estimation of the structural parameters that characterize the dynamics of index return and variance components under both P- and Q- distributions. Section 5 presents the estimation results. Section 6 investigates the performance of the model and reports in-sample goodness-of-fit statistics. Section 7 examines the stability of the model and measures the out-of-sample performance of the model. Section 8 concludes. The appendix provides the proofs of the most important theoretical results.

2 Model Setup

We start by a multiple-factor stochastic volatility dynamics that governs the market index returns under the P-distributions and then introduce a pricing kernel that links the P-dynamics to their risk-neutral counterparts by imposing appropriate martingale's restrictions on pricing kernel. We complete the index model by deriving a closed-form pricing equation for index options. We then introduce a GARCH model under physical distribution which is similar to our multiple-factor stochastic volatility model with two independent volatility dynamics. The risk neutral GARCH dynamics is also defined using a discrete-time analog of our continuous-time pricing kernel.

2.1 The Multifactor Stochastic Volatility Model

We assume the following two-factor stochastic volatility process governing the dynamics of the market index returns and variance under the physical distributions.

$$\begin{aligned} dS_t/S_t &= (r + \mu_1 v_{1,t} + \mu_2 v_{2,t})dt + \sqrt{v_{1,t}}dz_{1,t} + \sqrt{v_{2,t}}dz_{2,t} \\ dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}dw_{1,t} \\ dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}dw_{2,t} \end{aligned} \tag{1}$$

where, as in Christoffersen et al. (2009) we assume the stochastic structure (2).

$$
\begin{aligned}
&dw_{1,t} \cdot dz_{1,t} = \rho_1 dt, \ -1 \leq \rho_1 \leq +1 \\
&dw_{2,t} \cdot dz_{2,t} = \rho_2 dt, \ -1 \leq \rho_2 \leq +1 \\
&dw_{1,t} \cdot dw_{2,t} = 0 \\
&\rho_1^2 + \rho_2^2 \leq +1
\end{aligned} \tag{2}
$$

As in Heston (1993), θ_1 and θ_2 are unconditional average variance components, κ_1 and κ_2 capture the speed of mean reversion in each variance components, and σ_1 and σ_2 measure the volatility of variance components. The market equity risk premiums are denoted by $\mu_1 v_{1,t}$ and $\mu_2 v_{2,t}$. Following Bollerslev and Zhou (2006) we expect that μ_1 and μ_2 measure the persistent and transient "continuous-time" volatility feedback effects or risk-return trade-offs. The instantaneous correlation between shocks to the market returns and shocks to the persistent variance component is measured by ρ_1 and the instantaneous correlations between market returns and the transient variance component is given by ρ_2. As in Bollerslev and Zhou (2006), we expect that ρ_1 and ρ_2 account for persistent and transient "continuous-time" leverage (asymmetry) effect.

Note that (2) implies that the total return variance and the correlation between return and total variance are as follows.

$$
\begin{aligned}
&\text{Var}_t[dS_t/S_t] = v_{1,t}dt + v_{2,t}dt = v_t dt \\
&\text{Corr}_t[dS_t/S_t, dV_t] = \frac{\rho_1\sigma_1 v_{1,t} + \rho_2\sigma_2 v_{2,t}}{\sqrt{\sigma_1^2 v_{1,t} + \sigma_2^2 v_{2,t}}\sqrt{v_{1,t} + v_{2,t}}} dt
\end{aligned} \tag{3}
$$

We may then prove the following result.

Proposition 1. *The market index has the following dynamics under the risk-neutral measure:*

$$
\begin{aligned}
dS_t/S_t &= r dt + \sqrt{v_{1,t}} d\tilde{z}_{1,t} + \sqrt{v_{2,t}} d\tilde{z}_{2,t} \ , \\
dv_{1,t} &= \tilde{\kappa}_1(\tilde{\theta}_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}} d\tilde{w}_{1,t} \ , \\
dv_{2,t} &= \tilde{\kappa}_2(\tilde{\theta}_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}} d\tilde{w}_{2,t} \ ,
\end{aligned} \tag{4}
$$

where, $\tilde{\kappa}_1 = \kappa_1 + \lambda_1$, $\tilde{\kappa}_2 = \kappa_2 + \lambda_2$, $\tilde{\theta}_1 = \frac{k_1\theta_1}{k_1+\lambda_1}$, $\tilde{\theta}_2 = \frac{k_2\theta_2}{k_2+\lambda_2}$. *The market prices of risk factors are*

$$
\begin{aligned}
\psi_{1,t} &= \frac{\sigma_1\mu_1 - \rho_1\lambda_1}{\sigma_1(1-\rho_1^2)}\sqrt{v_{1,t}} \ , \ \psi_{2,t} = \frac{\sigma_2\mu_2 - \rho_2\lambda_2}{\sigma_2(1-\rho_2^2)}\sqrt{v_{2,t}} \ , \\
\psi_{3,t} &= \frac{\lambda_1 - \rho_1\sigma_1\mu_1}{\sigma_1(1-\rho_1^2)}\sqrt{v_{1,t}} \ , \ \psi_{4,t} = \frac{\lambda_2 - \rho_2\sigma_2\mu_2}{\sigma_2(1-\rho_2^2)}\sqrt{v_{2,t}} \ .
\end{aligned} \tag{5}
$$

One admissible pricing kernel that links the physical dynamics in (1) to the risk-neutral dynamics in (4) takes the following exponential affine form.

$$\frac{M_t}{M_0} = \big(\frac{S_t}{S_0}\big)^{\phi} \exp\Big[\delta t + \eta_1 \int_0^t v_{1,s} ds + \eta_2 \int_0^t v_{2,s} ds + \zeta_1(v_{1,t} - v_{1,0}) + \zeta_2(v_{2,t} - v_{2,0})\Big] \tag{6}$$

As in Christoffersen et al. (2013), $\{\delta, \eta_1, \eta_2\}$ governs the time-preferences, while $\{\phi, \zeta_1, \zeta_2\}$ governs the respected risk aversion to the index and variance risk factors, all of which are defined in the appendix.

Proof. See Appendix A. □

We note that the introduced nonlinear log pricing kernel in (6) is one way of "completing the market" and linking P- to Q- dynamics, where ζ_1, ζ_2 capture the nonlinearity of the log pricing kernel.[11] Transforming the physical dynamics in (1) into the risk neutral dynamics in (4) can also be done by assuming the following standard stochastic discount factor and without explicit assumptions about the investor's variance preferences. The proof of such a transformation can be found in Appendix B.

$$\frac{dM_t}{M_t} = -rdt - \psi_t' dW_t \ , \tag{7}$$

where $\psi_t \equiv [\psi_{1,t}, \psi_{2,t}, \psi_{3,t}, \psi_{4,t}]$ is the vector of market price of risk factors and $W_t \equiv [z_{1,t}, z_{2,t}, w_{1,t}, w_{2,t}]$ is the vector of innovations in return and variance.

To embed the options market data into the estimation of structural parameters, we determine a closed-from expression for the price of the European call options, with strike price K and time to maturity τ, by inverting the conditional characteristic function of the log spot index prices, $x_t = \ln(S_t)$.

$$C_t(S_t, K, v_{1,t}, v_{2,t}, \tau) = S_t P_1 - K e^{-r\tau} P_2 \ , \tag{8}$$

where,

$$\begin{aligned} P_1 &= \frac{1}{2} + \frac{1}{\pi}\frac{1}{S_t e^{r\tau}} \int_0^\infty \Re\Big[\frac{e^{-i\phi \ln K}\tilde{f}(v_{1,t}, v_{2,t}, \tau, \phi - i)}{i\phi}\Big] d\phi \ , \\ P_2 &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re\Big[\frac{e^{-i\phi \ln K}\tilde{f}(v_{1,t}, v_{2,t}, \tau, \phi)}{i\phi}\Big] d\phi \ , \end{aligned} \tag{9}$$

and where the risk-neutral conditional characteristic function of the natural logarithm of the index price at expiration, $x_{t+\tau}$, is

[11] Note also that ζ_1, ζ_2 affect a wedge between physical and risk neutral structural parameters of volatility dynamics.

$$\tilde{f}(v_{1,t}, v_{2,t}, \tau, \phi) \equiv \mathrm{E}_t^Q \left[\exp(i\phi x_{t+\tau}) \mid x_t\right] . \tag{10}$$

Since the two-factor SV model in (4) is an affine process, following Heston (1993), the conditional risk-neutral characteristic function in (10) has the following affine exponential form.[12]

$$\tilde{f}(v_{1,t}, v_{2,t}, \tau, \phi) = \exp\left[i\phi x_t + i\phi r\tau + A_1(\tau,\phi) + A_2(\tau,\phi) + B_1(\tau,\phi)v_{1,t} + B_2(\tau,\phi)v_{2,t}\right] , \tag{11}$$

where[13] for every $j = \{1, 2\}$

$$\begin{aligned}
A_j(\tau,\phi) &= \frac{\tilde{\kappa}_j\tilde{\theta}_j}{\sigma_j^2}\Big[(\tilde{\kappa}_j - \rho_j\sigma_j i\phi - d_j)\tau - 2\ln\big[\frac{1 - c_j e^{-d_j\tau}}{1 - c_j}\big]\Big] \\
B_j &= \frac{\tilde{\kappa}_j - \rho_j\sigma_j i\phi - d_j}{\sigma_j^2}\big[\frac{1 - e^{-d_j\tau}}{1 - c_j e^{-d_j\tau}}\big] \\
c_j &= \frac{\tilde{\kappa}_j - \rho_j\sigma_j i\phi - d_j}{\tilde{\kappa}_j - \rho_j\sigma_j i\phi + d_j} \\
d_j &= \sqrt{(\tilde{\kappa}_j - \rho_j\sigma_j i\phi)^2 + \sigma_j^2\phi(\phi + i)} .
\end{aligned} \tag{12}$$

2.2 The Component Volatility Model (Bivariate GARCH)

Since the seminal papers of Engle (1982) and Bollerslev (1986) several ARCH-type models have been proposed where the main difference is in parametrization of the conditional variance and asymmetry effect. Extensive empirical evidence examines the importance of conditional heteroskedasticity and variance mean reversion in modeling index returns and index options.

Note that in ARCH-type models volatility is considered as a deterministic process, whereas in case of SV models volatility has a fully stochastic nature.

Engle and Lee (1999) introduce a component extension to the simple GARCH(1,1) model where the unconditional mean of the conditional variance process is time-varying and provide empirical evidence that the component model provides a very good fit to return data. Christoffersen et al. (2008) consider an affine version of component volatility model of Engle

[12] Note that the conditional risk-neutral characteristic function of the natural logarithm of return, $x_{t+\tau} - x_t = \ln(S_{t+\tau}/S_t)$, can be defined with the same expression as (11) but without the first component, $i\phi x_t$.

[13] Following Duffie et al. (2000), the coefficients A_1, A_2, B_1, and B_2 are the solutions of a system of Riccati equations subject to appropriate boundary conditions. For the ease of computation, we modify these solutions based on the little Heston trap formulation of Albrecher et al. (2006).

and Lee (1999) by generalizing the affine Gaussian GARCH(1,1) Heston and Nandi (2000) as follows.

$$
\begin{aligned}
R_t &\equiv \ln(\frac{S_t}{S_{t-1}}) = r + (\mu - \frac{1}{2})h_t + \sqrt{h_t}z_t \\
h_t &= q_t + \beta_h(h_{t-1} - q_{t-1}) + \alpha_h\big((z_{t-1} - \gamma_h\sqrt{h_{t-1}})^2 - (1 + \gamma_h^2 q_{t-1})\big) \\
q_t &= w_q + \beta_q q_{t-1} + \alpha_q\big((z_{t-1} - 1)^2 - 2\gamma_q\sqrt{h_{t-1}}z_{t-1})^2\big)\ ,
\end{aligned}
\tag{13}
$$

where h_t is referred to as the total conditional variance, q_t as the long-run component of conditional variance, and therefore $h_t - q_t$ as the short-run component conditional variance with zero unconditional mean. This volatility component model is relatively simple since both of the volatility components, h_t and q_t, are characterized by nonlinear functions of a single innovation z_{t-1}. A richer model of return volatility includes multiple innovations.[14]

We introduce a component volatility model (bivariate GARCH model) which is similar to our two-factor stochastic volatility model in the sense that volatility components are independent. We extend the Heston and Nandi (2000) affine Gaussian GARCH(1,1) model that yields a closed-form option valuation formula similar to our SV model. Note that several studies investigate the limits of GARCH models as the time intervals become small and find that for a given GARCH model, there could be a several continuous-time limits and several GARCH models could converge to a continuous-time stochastic volatility model.[15] A discrete time analog of our SV model under the physical measure can be defined as follows.

$$
\begin{aligned}
R_t &\equiv \ln(\frac{S_t}{S_{t-1}}) = r + (\mu_1 - \frac{1}{2})h_{1,t} + (\mu_2 - \frac{1}{2})h_{2,t} + \varepsilon_{1,t} + \varepsilon_{2,t} \\
h_{1,t} &= w_1 + \beta_1 h_{1,t-1} + \alpha_1(z_{1,t-1} - \gamma_1\sqrt{h_{1,t-1}})^2 \\
h_{2,t} &= w_2 + \beta_2 h_{2,t-1} + \alpha_2(z_{2,t-1} - \gamma_2\sqrt{h_{2,t-1}})^2
\end{aligned}
\tag{14}
$$

where r is the daily continuously compounded interest rate, $\varepsilon_{1,t} = \sqrt{h_{1,t}}z_{1,t}$, $\varepsilon_{2,t} = \sqrt{h_{2,t}}z_{2,t}$, and $z_{1,t}$ and $z_{2,t}$ are standard normal distributions. $h_{1,t}+h_{2,t}$ is the conditional variance of the log return in period t. The autoregressive parameters β_1 and β_2 determine the persistence of the each variance component and the innovation parameters α_1 and α_2 determine the variance of variance and thus kurtosis in each variance component. γ_1 and γ_2 capture the so-called leverage effect, asymmetry in the response of each volatility component to positive versus negative return shocks. Note that in our specification, the conditional mean return is

$$
\mathrm{E}_{t-1}[S_t/S_{t-1}] = \mathrm{E}_{t-1}[\exp(R_t)] = \exp(r + \mu_1 h_{1,t} + \mu_2 h_{2,t})\ . \tag{15}
$$

[14] See for instance Feunou and Tédongap (2012), Christoffersen et al. (2010), and Khrapov and Renault (2016).

[15] See Corradi (2000).

The expected future variance is a linear function of current variance and long-run average (unconditional) variance.

$$\begin{aligned}\mathrm{E}_{t-1}[h_{t+1}] &= \mathrm{E}_{t-1}[h_{1,t+1} + h_{2,t+1}] \\ &= (\beta_1 + \alpha_1\gamma_1^2)h_{1,t} + (1 - \beta_1 - \alpha_1\gamma_1^2)\,\mathrm{E}[h_{1,t}] \\ &+ (\beta_2 + \alpha_2\gamma_2^2)h_{2,t} + (1 - \beta_2 - \alpha_2\gamma_2^2)\,\mathrm{E}[h_{2,t}]\end{aligned} \tag{16}$$

where $\mathrm{E}[h_{1,t}] \equiv \sigma_1^2 = (w_1 + \alpha_1)/(1 - \beta_1 - \alpha_1\gamma_1^2)$ and $\mathrm{E}[h_{2,t}] \equiv \sigma_2^2 = (w_2 + \alpha_2)/(1 - \beta_2 - \alpha_2\gamma_2^2)$ are long-run average (unconditional) component variance. We refer to $(\beta_1 + \alpha_1\gamma_1^2)$ and $(\beta_2 + \alpha_2\gamma_2^2)$ as the persistence of the variance component. A high level of persistence (close to one) implies that shocks that push variance away from its long-run average will persist for a long time. The conditional variance of h_{t+1} is also linear in past variance.

$$\mathrm{Var}_{t-1}[h_{t+1}] = \mathrm{Var}_{t-1}[h_{1,t+1} + h_{2,t+1}] = 2\alpha_1^2 + 4\alpha_1^2\gamma_1^2 h_{1,t} + 2\alpha_2^2 + 4\alpha_2^2\gamma_2^2 h_{2,t} \tag{17}$$

The conditional covariance between stock returns and variance is

$$\mathrm{Cov}_{t-1}(R_t, h_{t+1}) = \mathrm{Cov}_{t-1}(R_t, h_{1,t+1} + h_{2,t+1}) = -2\alpha_1\gamma_1 h_{1,t} - 2\alpha_2\gamma_2 h_{2,t}\ . \tag{18}$$

We transform the physical stock price process (14) to the corresponding risk neutral process using a discrete-time analog of the continuous-time pricing kernel (6).

$$\frac{M_t}{M_0} = \big(\frac{S_t}{S_0}\big)^{\phi} \exp\Big[\delta t + \eta_1 \sum_{s=1}^{t} h_{1,s} + \eta_2 \sum_{s=1}^{t} h_{2,s} + \zeta_1(h_{1,t+1} - h_{1,1}) + \zeta_2(h_{2,t+1} - h_{2,1})\Big] \tag{19}$$

where parameters $\{\delta,\ \eta_1,\ \eta_2\}$ govern the time-preference, and parameters $\{\phi,\ \zeta_1,\ \zeta_2\}$ govern the respected risk aversion to equity risk and to variance risk factors. Note that ζ_1 and ζ_2 capture the non-linearity of the log pricing kernel.

Proposition 2. *Given the physical GARCH process (14) and the pricing kernel (19), the risk neutral innovations may be characterized by the following transformations.*

$$\begin{aligned}z_{1,t}^* &= \sqrt{1 - 2\alpha_1\zeta_1}\big(z_{1,t} + (\mu_1 + \frac{\alpha_1\zeta_1}{1 - 2\alpha_1\zeta_1})\sqrt{h_{1,t}}\big) \\ z_{2,t}^* &= \sqrt{1 - 2\alpha_2\zeta_2}\big(z_{2,t} + (\mu_2 + \frac{\alpha_2\zeta_2}{1 - 2\alpha_2\zeta_2})\sqrt{h_{2,t}}\big)\end{aligned} \tag{20}$$

Hence, the corresponding risk-neutral GARCH process may be characterized as follows

$$
\begin{aligned}
R_t \equiv \ln(\frac{S_t}{S_{t-1}}) &= r - \frac{1}{2}h^*_{1,t} - \frac{1}{2}h^*_{2,t} + \sqrt{h^*_{1,t}}z^*_{1,t} + \sqrt{h^*_{2,t}}z^*_{2,t} \\
h^*_{1,t} &= w^*_1 + \beta_1 h^*_{1,t-1} + \alpha^*_1(z^*_{1,t-1} - \gamma^*_1\sqrt{h^*_{1,t-1}})^2 \\
h^*_{2,t} &= w^*_2 + \beta_2 h^*_{2,t-1} + \alpha^*_2(z^*_{2,t-1} - \gamma^*_2\sqrt{h^*_{2,t-1}})^2
\end{aligned} \tag{21}
$$

where conditional variance under physical and risk-neutral distributions are linked as

$$
h^*_{1,t} = \frac{h_{1,t}}{1 - 2\alpha_1\zeta_1} \ , \ h^*_{2,t} = \frac{h_{2,t}}{1 - 2\alpha_2\zeta_2} \tag{22}
$$

and for every $j = \{1, 2\}$ the parameters mapping may be given by

$$
\begin{aligned}
\alpha^*_j &= \frac{\alpha_j}{(1 - 2\alpha_j\zeta_j)^2} \\
w^*_j &= \frac{w_j}{1 - 2\alpha_j\zeta_j} \\
\gamma^*_j &= (\mu_j - \frac{1}{2} + \gamma_j)(1 - 2\alpha_j\zeta_j) + \frac{1}{2}
\end{aligned} \tag{23}
$$

Proof. The proof of this proposition is very similar to its continuous-time counterpart. We show that the GARCH model under physical measure (14) is linked to the GARCH model under risk-neutral measure (21) with the proposed pricing kernel (19) by specifying a set of sufficient conditions (20), (22), and (23). We first impose Euler equation for the risk-free asset and subsequently impose Euler equation for the underlying asset to find this parameters mapping. See Appendix C. □

Note that linking P- to Q- dynamics can also be done through a log-linear pricing kernel. But, log-linear pricing kernel within GARCH models does not incorporate directly the effect of variance premium on risk neutralization. However, variance dependent pricing kernel allows to directly incorporate the effect of variance premium as $-2\alpha\zeta$ in risk neutralization. A negative variance premium yields higher level of risk-neutral variances compared to the physical variances as $h^*_{1,t}$ exceeds $h_{1,t}$ and $h^*_{2,t}$ exceeds $h_{2,t}$. Negative variance premium also yields higher level of risk neutral innovation parameters α^*_1 and α^*_2 and hence increases the risk neutral variance persistence, $(\beta_1 + \alpha^*_1{\gamma^*_1}^2)$ and $(\beta_2 + \alpha^*_2{\gamma^*_2}^2)$.

3 Data

We obtain daily prices of S&P 500 index options from the OptionMetrics volatility surface data set, which is based on the midpoint of bid-ask quotes. Our sample of S&P 500 index options is from January 4, 1996 through December 29, 2011. We follow the data cleaning routine commonly used in the empirical option pricing literature: we remove options with implied volatility less than 5% and greater than 150%; we also follow the filtering rules in Bakshi et al. (1997) to remove options that violate various no-arbitrage conditions. We focus on out-of-the-money (OTM) options with maturity up to and including one-year and with 10% moneyness (spot price over strike price).[16,17] Our option-based optimization function minimizes the squared deviations between model and market option prices and therefore may put greater weight on expensive in-the-money (ITM) and long-maturity options.[18] Moreover, ITM S&P 500 call options are less liquid than OTM call options. To prevent such biases in our optimization, we discard all ITM options and use OTM S&P 500 put options and convert them into ITM call options. After cleaning, we have 345,710 S&P 500 index option quotes together with daily underlying returns. This is the dataset that we use to filter daily spot variances and to estimate a set of structural parameters.

Table (1) presents the descriptive statistics of the call option contracts in our sample sorted by moneyness (stock price over strike price) and day-to-maturity (DTM). Note that we focus on OTM option contracts, which means S/K is below 1 for OTM call contracts. After cleaning, we have 208,098 out-of-the-money call option contracts with an average day-to-maturity of 143 days, an average price of $35.59, an average implied volatility of 20.64%, and an average delta of 0.37. Table (2) reports the descriptive statistics of the put option contracts in our sample sorted by moneyness and day-to-maturity. After cleaning, we use 137,612 out-of-the-money (S/K is above 1) put option contracts with an average day-to-maturity of 136 days, an average price of $32.11, an average implied volatility of 24.34%, and an average delta of -0.29. Note that Panel C in Tables (1) and (2) reflect the well-known volatility smirk in index options, as implied volatility is larger for OTM put options (Table (2), Panel C) compared to the OTM call options (Table (1), Panel C).

[Table (1) about here]

[Table (2) about here]

The data for daily index level, index return, and the dividend yields are from CRSP. In our analysis we first adjust daily index level with dividend yields and then compute the option prices using the dividends adjusted returns. Risk-free interest rates for all maturities

[16] This range of moneyness implies that we keep OTM call options with moneyness less than 1.1 and OTM put options with moneyness greater than 0.9.

[17] As discussed in previous section, multiple-factor SV models could better capture the slope and the level of smirk compare to single-factor SV models. Therefore, unlike similar analysis, we undertake a more extensive calibration exercise by incorporating the information content of options on longer maturity horizons and wider moneyness ranges. For instance, Ait-Sahalia and Kimmel (2007, Section 7) only include short-maturity at-the-money S&P 500 Index Options; Eraker (2004) use 3,270 call options contracts recorded over 1,006 trading days; Jones (2003) models are estimated using a sample of 3537 S&P 100 index options from January 1986 to June 2000.

[18] See Huang and Wu (2004).

are estimated by linear interpolation between the closest zero-coupon rates using the Zero Coupon Yield Curve data from OptionMetrics.

4 Estimation Methodology

To estimate the parameters of two-factor stochastic volatility model of the index we follow the literature on the estimation of stochastic volatility models, where the main challenge is the estimation of unobserved latent volatilities. There are several approaches to estimate stochastic volatility model. Our own approach combines the information from underlying index and option markets to impose consistency between structural parameters under P and Q distributions, known as joint estimation. Therefore, we use a likelihood function that contains a return-based component and an option-based component, as in Santa-Clara and Yan (2010) and Christoffersen et al. (2013).[19] Here we do a joint-estimation by filtering the two vectors of daily spot variances, $\{v_{1,t}, v_{2,t}\}$, and simultaneously estimating a set of structural parameters of the dynamics of index returns and variances, including the market price of each variance component, $\Theta \equiv \{\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2, \rho_1, \rho_2, \lambda_1, \lambda_2\}$. Note that joint estimation allow us to have reliable prices of variance risk factors, as we can get a consistent set of structural parameters between the P and Q distributions.

Since the market variances are unobserved state variables, we first extract daily instantaneous persistent and transient variance components using the Particle Filter (PF) method. This optimal filtering methodology provides a tool for learning about unobserved shocks and states from discretely observed prices generated by continuous-time models.[20] Although we generally follow the conventional filtration procedure in the literature, we provide a novel approach to the challenge of filtering the two separate variance paths. Our proposed solution is not trivial and to the best of our knowledge is novel and constitutes a methodological contribution to the option pricing literature.

4.1 The Return Based Likelihood Function

To define the return-based likelihood function and filter spot variances, we start by discretizing the returns dynamics (1). Applying Ito's lemma to equation (1), gives the dynamics of logarithm of stock prices as follows.

[19] Consistency can also be imposed through moment-based and simulation-based methods; see Ait-Sahalia and Kimmel (2007), Eraker (2004), Jones (2003), Chernov and Ghysels (2000), and Pan (2002). Other approaches use only option-based data to estimate only the Q distribution; Bakshi et al. (1997), Bates (2000), Huang and Wu (2004), and Christoffersen et al. (2009).

[20] For the application of PF in estimating the model parameters see Gordon et al. (1993), Johannes et al. (2009), Johannes and Polson (2009), Christoffersen et al. (2010), and Boloorforoosh (2014).

$$\begin{aligned}
d\ln(S_t) &= (\mu - \frac{1}{2}(v_{1,t}+v_{2,t}))dt + \sqrt{v_{1,t}}dz_{1,t} + \sqrt{v_{2,t}}dz_{2,t} \ , \\
dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}dw_{1,t} \ , \\
dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}dw_{2,t} \ ,
\end{aligned} \tag{24}$$

where, $\mu \equiv r + \mu_1 v_{1,t} + \mu_2 v_{2,t}$. We discretize (24) using the Euler scheme.[21] Equation (25) models the relation between observed index prices and unobserved variances at time $t+\Delta t$ conditional on the time t variances.

$$\begin{aligned}
\ln(S_{t+\Delta t}) - \ln(S_t) &= (\mu - \frac{1}{2}(v_{1,t}+v_{2,t}))\Delta t + \sqrt{v_{1,t}\Delta t}\ z_{1,t+\Delta t} + \sqrt{v_{2,t}\Delta t}\ z_{2,t+\Delta t} \ , \\
v_{1,t+\Delta t} &= v_{1,t} + \kappa_1(\theta_1 - v_{1,t})\Delta t + \sigma_1\sqrt{v_{1,t}\Delta t}\ w_{1,t+\Delta t} \ , \\
v_{2,t+\Delta t} &= v_{2,t} + \kappa_2(\theta_2 - v_{2,t})\Delta t + \sigma_2\sqrt{v_{2,t}\Delta t}\ w_{2,t+\Delta t} \ .
\end{aligned} \tag{25}$$

Brownian shocks $z_{1,t+\Delta t}$, $z_{2,t+\Delta t}$, $w_{1,t+\Delta t}$, and $w_{2,t+\Delta t}$ are normal random variables with mean zero and variance one. From the first equation in (25) we use the observed daily index log-prices ($\ln(S_t)$, $\ln(S_{t+\Delta t})$) to first filter the daily return's shocks ($z_{1,t+\Delta t}$, $z_{2,t+\Delta t}$) and then, using the filtered shocks in returns and the last two equation in (25), we filter daily spot variances ($v_{1,t+\Delta t}$, $v_{2,t+\Delta t}$). Note that we filter filter the summation of return shocks $z_{1,t+\Delta t} + z_{2,t+\Delta t}$ as we cannot separate the daily observed shocks into two components, $z_{1,t+\Delta t}$ and $z_{2,t+\Delta t}$. Therefore, we rewrite the underlying dynamics as (26), given that the return shocks are uncorrelated and then discretize the dynamics.

$$\begin{aligned}
d\ln(S_t) &= (\mu - \frac{1}{2}(v_{1,t}+v_{2,t}))dt + \sqrt{v_{1,t}+v_{2,t}}dz_t \ , \\
dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}dw_{1,t} \ , \\
dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}dw_{2,t} \ ,
\end{aligned} \tag{26}$$

with the correlation structure:

$$\begin{aligned}
dw_{1,t}\cdot dz_t &= \rho_1 dt,\ -1 \le \rho_1 \le +1 \ , \\
dw_{2,t}\cdot dz_t &= \rho_2 dt,\ -1 \le \rho_2 \le +1 \ , \\
dw_{1,t}\cdot dw_{2,t} &= 0 \ .
\end{aligned} \tag{27}$$

We decompose the variance shocks into orthogonal components as in (28) and then discretize the return dynamics (26) using the Euler scheme and shock's decomposition (28).[22]

[21] According to Eraker (2004) and Li et al. (2008) the discretization bias of the Euler scheme is negligible for daily data.

[22] Note that the quadratic variations of the transformed using the proposed shocks decomposition (28) should remain the same as $\sqrt{dt}$.

$$
\begin{aligned}
dw_{1,t} &= \rho_1 dz_t + \sqrt{1-\rho_1^2}\; dB_{1,t} \\
dw_{2,t} &= \rho_2 dz_t - \frac{\rho_1\rho_2}{\sqrt{1-\rho_1^2}}\; dB_{1,t} + \sqrt{\frac{1-\rho_1^2-\rho_2^2}{1-\rho_1^2}}\; dB_{2,t} \\
\langle dB_{1,t}\ , dB_{2,t}\rangle &= 0
\end{aligned}
\tag{28}
$$

$$
\begin{aligned}
\ln(S_{t+\Delta t}) - \ln(S_t) &= (\mu - \frac{1}{2}(v_{1,t}+v_{2,t}))\Delta t + \sqrt{(v_{1,t}+v_{2,t})\Delta t}\; z_{t+\Delta t}\ , \\
v_{1,t+\Delta t} &= v_{1,t} + \kappa_1(\theta_1 - v_{1,t})\Delta t + \sigma_1\sqrt{v_{1,t}\Delta t}\; w_{1,t+\Delta t}\ , \\
v_{2,t+\Delta t} &= v_{1,t} + \kappa_2(\theta_2 - v_{2,t})\Delta t + \sigma_2\sqrt{v_{2,t}\Delta t}\; w_{2,t+\Delta t}\ ,
\end{aligned}
\tag{29}
$$

where, $z_{t+\Delta t}$, $w_{1,t+\Delta t}$, and $w_{2,t+\Delta t}$ are all $N(0,1)$. Now, using daily index log-returns, we proceed to filter the spot variances from the discretized model in (29) given the correlation structure in (28).

We follow Pitt (2002)[23] and adopt a particular implementation of the PF, which is referred to as the sampling-importance-resampling (SIR) PF. This implementation of PF method allow us to approximate the true density of the persistent variance component ($v_{1,t}$) and the transient variance component ($v_{2,t}$) using two sets of particles that are updated recursively through equations (29). In other words, we recursively simulate next period particles of each variance component until we have the empirical distributions of each variance factor over the entire sample. That is, given N particles of $\{v^j_{1,t}\}^N_{j=1}$, N particles of $\{v^j_{2,t}\}^N_{j=1}$, simulated return shocks, and $w_{1,t+\Delta t}$ and $w_{2,t+\Delta t}$ we generate the next period particles, N particles $\{v^j_{1,t+\Delta t}\}^N_{j=1}$ and another N particles $\{v^j_{2,t+\Delta t}\}^N_{j=1}$ at any time $t+\Delta t$.

We start by simulating return's shocks $z^j_{t+\Delta t}$ given the initial value of structural parameters Θ_0 and current variance particles $\{v^j_{1,t}, v^j_{2,t}\}$, on every day t and for every particle $j = 1, 2, ..., N$, according to (30). Then using (31) we simulate volatility shocks $w^j_{1,t+\Delta t}$ and $w^j_{1,t+\Delta t}$. Note that $\epsilon^j_{1,t+\Delta t}$ and $\epsilon^j_{2,t+\Delta t}$ are independent standard normal random variables.

$$
z^j_{t+\Delta t} = \Big[\ln(S_{t+\Delta t}/S_t) - (\mu - \frac{1}{2}(v^j_{1,t}+v^j_{2,t}))\Delta t\Big] / \sqrt{(v^j_{1,t}+v^j_{2,t})\Delta t} \tag{30}
$$

$$
\begin{aligned}
w^j_{1,t+\Delta t} &= \rho_1 z^j_{t+\Delta t} + \sqrt{1-\rho_1^2}\; \epsilon^j_{1,t+\Delta t} \\
w^j_{2,t+\Delta t} &= \rho_2 z^j_{t+\Delta t} - \frac{\rho_1\rho_2}{\sqrt{1-\rho_1^2}}\; \epsilon^j_{1,t+\Delta t} + \sqrt{\frac{1-\rho_1^2-\rho_2^2}{1-\rho_1^2}}\epsilon^j_{2,t+\Delta t}
\end{aligned}
\tag{31}
$$

[23] See Pitt (2002), Christoffersen et al. (2010), and Boloorforoosh (2014) for a detailed description of the PF algorithm.

Then, given the simulated return's shocks $\{z^j_{t+\Delta t}\}_{j=1}^N$ and simulated shocks to the persistent and transient variance components $\{w^j_{1,t+\Delta t}\}_{j=1}^N$ and $\{w^j_{2,t+\Delta t}\}_{j=1}^N$, we simulate next period variance particles $\{\tilde{v}^j_{1,t+\Delta t}\}$ and $\{\tilde{v}^j_{2,t+\Delta t}\}$, for every day t according to (32).

$$\begin{aligned}\tilde{v}^j_{1,t+\Delta t} &= v^j_{1,t} + \kappa_1(\theta_1 - v_{1,t})\Delta t + \sigma_1\sqrt{v_{1,t}\Delta t}\, w_{1,t+\Delta t} \\ \tilde{v}^j_{2,t+\Delta t} &= v^j_{2,t} + \kappa_2(\theta_2 - v_{2,t})\Delta t + \sigma_2\sqrt{v_{2,t}\Delta t}\, w_{2,t+\Delta t}\end{aligned} \tag{32}$$

This is the "Sampling Step," at the end of which we generate N possible daily values for the persistent variance component $v_{1,t+\Delta t}$ and another N possible daily values for the transient variance component $v_{2,t+\Delta t}$ over the entire sample. In the next step, "Importance Step," we evaluate importance of the sampled daily particles by assigning appropriate weights $\tilde{W}^j_{t+\Delta t}$ to the simulated daily particles using a multivariate normal distribution. Intuitively, these weights, $\tilde{W}^j_{t+\Delta t}$, are likelihood that the next day return at $t + 2\Delta t$ is generated by this set of particles. Then, the probability of each daily particle can be defined by normalizing the weights within each day according to (35). Note that these weights are the basis of our likelihood function under the P distribution.

$$(r_{t+2\Delta t}|\{\tilde{v}_{1,t+\Delta t}\ ,\tilde{v}_{2,t+\Delta t}\}) \sim N\big[(\mu - \frac{1}{2}(\tilde{v}_{1,t+\Delta t} + \tilde{v}_{2,t+\Delta t}))\Delta t\ ,(\tilde{v}_{1,t+\Delta t} + \tilde{v}_{2,t+\Delta t})\Delta t\big] \tag{33}$$

$$\tilde{W}^j_{t+\Delta t} = \frac{1}{\sqrt{2\pi(\tilde{v}^j_{1,t+\Delta t} + \tilde{v}^j_{2,t+\Delta t})\Delta t}} \cdot \exp\Big(-\frac{1}{2}\frac{\big(\ln(\frac{S_{t+2\Delta t}}{S_{t+\Delta t}}) - (\mu - \frac{1}{2}(\tilde{v}^j_{1,t+\Delta t} + \tilde{v}^j_{2,t+\Delta t}))\Delta t\big)^2}{(\tilde{v}^j_{1,t+\Delta t} + \tilde{v}^j_{2,t+\Delta t})\Delta t}\Big) \tag{34}$$

$$\breve{W}^j_{t+\Delta t} = \frac{\tilde{W}^j_{t+\Delta t}}{\sum_{j=1}^N \tilde{W}^j_{t+\Delta t}} \tag{35}$$

Note that combining independent shocks $z_{1,t}$ and $z_{2,t}$ in (26) imposes a restriction on the weights of daily variance particles. Therefore, the importance probability is assigned to the summation of return's shocks. However, estimation results show that the path of filtered spot persistent variance component and transient variance component in our two-factor SV model are not sensitive to this assumption. We investigate the sensitivity of our result to this weighting assumption by estimating daily spot variances using the two-step iterative approach, following Huang and Wu (2004). We do not observe significant difference between filtered spot variances in two-step iterative approach and those filtered with particle filter method.

In the last step, "Resampling Step," we find the empirical distribution of smoothly resampled daily particles. Following the Pitt (2002) algorithm, we draw smoothed daily particles by assigning uniform distributions to the raw daily particles for persistent and transient variance components. As in the sampling step, we start from the beginning of the sample period and recursively simulate the next period daily particles using the smoothly resampled daily particles. The procedure continues until we have the empirical distributions of the persistent and transient variance components over the entire sample.

Given the appropriate weights (35), we define the return-based likelihood function as follows.

$$LLR \propto \sum_{t=1}^{T} \ln \left(\frac{1}{N} \sum_{j=1}^{N} \breve{W}_t^j(\Theta) \right) \tag{36}$$

Our implementation uses the maximum likelihood importance sampling (MLIS) methodology to maximize LLR criterion. Note that return-based likelihood function (36) is a function of the structural parameters of the market model under P measure, $\Theta \equiv \{\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_1, \rho_1, \rho_2\}$. Note also that the filtered daily spot persistent variance component $v_{1,t}^P$ and transient variance component $v_{2,t}^P$ can be defined as the average of the smoothly resampled particles.

$$\hat{v}_{1,t}^P = \frac{1}{N} \sum_{j=1}^{N} v_{1,t}^j \ , \qquad \hat{v}_{2,t}^P = \frac{1}{N} \sum_{j=1}^{N} v_{2,t}^j \tag{37}$$

4.2 The Option Based Likelihood Function

In order to fully specify the market dynamics under the Q measure, we need to estimate a set of structural parameters for the market model under Q measure $\tilde{\Theta} \equiv \{\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_1, \rho_1, \rho_2, \lambda_1, \lambda_2\}$, a vector of daily spot persistent variance component $\hat{v}_{1,t}^Q$, and a vector of daily spot transient variance component $\hat{v}_{2,t}^Q$. Unobserved daily spot persistent and transient variance components under the Q measure can be filtered using the PF method. We follow the same procedure as described in (30)-(35) for the market variances under P measure while using structural parameters under Q measure, $\{\tilde{\kappa}_1, \tilde{\kappa}_2, \tilde{\theta}_1, \tilde{\theta}_2, \sigma_1, \sigma_1, \rho_1, \rho_2\}$. Note that $\tilde{\kappa}^i = \kappa^i + \lambda^i$ and $\tilde{\theta}^i = \frac{\kappa^i \theta^i}{\kappa^i + \lambda^i}$ for $i = \{1, 2\}$ according to the Proposition (1). We may obtain daily spot persistent and transient variance components under Q measure as the average of the smoothly resampled daily particles for each component of market variance.

$$\hat{v}_{1,t}^Q = \frac{1}{N} \sum_{j=1,Q}^{N} v_{1,t}^j \ , \qquad \hat{v}_{2,t}^Q = \frac{1}{N} \sum_{j=1,Q}^{N} v_{2,t}^j \tag{38}$$

Define the option-based likelihood function using a Vega-weighted loss function for the index options, where Vega is the Black-Scholes sensitivity of the option price with respect to volatility.[24] The Vega- weighted option pricing errors serves as an approximation to the implied volatility root mean squared errors,[25] which is a very popular loss function. This Vega-weighted loss function does not require a numerical inversion of the Black and Scholes (1973) model price and thus is helpful in large scale optimization problems such as ours.

Define normalized option pricing errors as follows.

$$\eta_n = (C_n^O - C_n^M(\tilde{\Theta}, \hat{v}_1^Q, \hat{v}_2^Q, S_t, K, \tau))/Vega_n \ , \qquad n = 1, \ldots, M \tag{39}$$

where C_n^O is the observed daily option prices and $C_n^M(\tilde{\Theta}, \hat{v}_1^Q, \hat{v}_2^Q, S_t, K, \tau)$ is the model price of index option n, according to pricing equation (8), given the filtered spot persistent and transient variance component and structural parameters under Q measure. M is the total number of index option contracts and $Vega_n$ is the Black and Scholes (1973) option Vega for the option n. Then we may obtain the option-based likelihood as follows.[26]

$$LLO \propto -\frac{1}{2}\Big(M \ln(2\pi) + \sum_{n=1}^{M}\big(\ln(s^2) + \eta_n^2/s^2\big)\Big) \ , \tag{40}$$

Combining the returns-based likelihood function (36) and the options-based likelihood function (40), we have the total likelihood function. Our implementation uses the nonlinear least squares importance sampling (NLSIS) estimation mythology to solve the following optimization and to estimate the structural parameters of the market model $\hat{\Theta}$ and $\hat{\tilde{\Theta}}$ and daily spot persistent and transient variance components.

$$\max_{\Theta, \tilde{\Theta}}\big(LLR + LLO\big) \ . \tag{41}$$

It is important to note that our optimization algorithm is iterative. Each iteration starts with an initial set of structural parameters, which then will be used to filter daily spot volatilities using the information content of index returns. Then, given spot volatilities and observed option prices, next set of optimal parameters can be reached by minimizing the option pricing errors over the entire sample. The procedure iterates until an optimal set of structural parameters is reached and thereby we obtain final vectors of transient and variance spot variance components.

[24] Note that while several loss functions have been used in option pricing literature, option theory does not suggest a specific loss function as pricing equations do not contain an error term. Therefore, the appropriate loss functions are defined according to econometric considerations as well as convenience.

[25] See for example Carr and Wu (2007) and Christoffersen et al. (2009).

[26] Note that we replace s^2 by its sample analog $\hat{s}^2 = \frac{1}{M}\sum_{n=1}^{N} \eta_n^2$.

5 Parameter Estimation Results

This section reports the filtered daily spot variance components together with the structural parameter estimates for the two-factor SV model. As described in the Data Section, we use a long time-series of daily S&P 500 index returns and the entire cross-section of S&P 500 option prices over the period from January 4, 1996 to December 29, 2011. Given the slow mean-reversion in the dynamic of market volatility, it is important to let the data set span a long time series. This is in particular important in our analysis as we decompose the overall market volatility into two independent components and would like to characterize the dynamics of transient and persistent variance components.

In what follows we set the market risk premium μ equal to the sample average daily index returns. We use 10% OTM index options and then put-call-parity to convert OTM puts into ITM calls. Table (3) reports structural parameter estimates (under P measure) that characterize the dynamics of index returns and its persistent and transient variance components. Panel A provides result of the joint estimation; a consistent set of parameters under P and Q measures. Therefore, the speeds of mean reversion and the unconditional mean of the persistent and transient variance components under Q-measure are linked to their P-measure equivalents through the market prices of the volatility risk factors ($\tilde{\kappa}_1 = \kappa_1 + \lambda_1$, $\tilde{\kappa}_2 = \kappa_2 + \lambda_2$, $\tilde{\theta}_1 = \frac{k_1\theta_1}{k_1+\lambda_1}$, $\tilde{\theta}_2 = \frac{k_2\theta_2}{k_2+\lambda_2}$).[27] To provide a basis for further comparison and to examine the goodness of fit of the two-factor SV model under the joint-estimation, we also estimate structural parameters using only option data. This result is provided in Panel C.

[Table (3) about here]

As discussed, the purpose of two-factor stochastic volatility model is to capture independent movements in the underlying returns and option prices over time. Consistent with previous studies in both discrete time GARCH models and continuous time stochastic volatility models, we find that one of the volatility factors is highly persistent and the other one is highly mean-reverting. In joint-estimation, we find that the first variance component is slowly mean-reverting with $\kappa_1 = 1.4271$ under physical measure while the rate of mean reversion in the second variance component is much higher with $\kappa_2 = 3.5874$ under the physical measure.[28] The point estimate of mean reversion parameters from option-based estimation is similar to those from joint estimation. Using options data only, we find that $\tilde{\kappa}_1 = 0.2267$ and $\tilde{\kappa}_2 = 2.9137$, which is consistent with the speed of mean reversion from joint estimation where under Q-measure $\tilde{\kappa}_1 = 0.3473$ and $\tilde{\kappa}_2 = 2.5520$.

To gain a better intuition about persistent and transient variance components we define the half-life ($T_{1/2}$) of a variance component as the number of weeks that it takes for autocorrelation of each variance component to decay to half of its weekly autocorrelation level.

[27] See Proposition (1).

[28] These value correspond to a daily variance persistence of $1 - 1.4271/365 = 0.9961$ for the first component and $1 - 3.5874/365 = 0.9901$ for the second component.

Half-life can be computed as $T_{1/2} = \ln(\phi/2)/\ln(\phi)$ where $\Delta t = 7/365$ and $\phi = \exp(-\kappa\Delta t)$, denoting weekly autocorrelation of time-series each variance component. The risk neutral point estimate of mean reversion speed in transient variance component implies a half-life around 15 weeks while it is 105 weeks in the persistent variance component, almost 7 times larger than its transient counterpart. These values confirm that first variance component is highly persistent while the second one is highly auto-correlated and thus the immediate impact of variance shocks on this component is larger but short-lived.

We observe that the unconditional persistent variance under P-measure is $\theta_1 = 0.0026$, which is much less than the unconditional transient variance $\theta_2 = 0.0171$. The unconditional risk neutral persistent and transient variance components are $\tilde{\theta}_1 = 0.0106$ and $\tilde{\theta}_2 = 0.0240$ which correspond to 10.30% and 15.49% volatility per year. Note that the unconditional variance of both components are consistent with the average filtered daily spot persistent variance and daily spot transient variance over the entire sample.

Consistent with our intuition, we observe a wide spread between the volatility of variance in the persistent and transient variance components. As a result of joint estimation we find that $\sigma_1 = 0.0855$ and $\sigma_2 = 0.3496$. This result is consistent with the option-based estimation where we find that transient variance component is much more volatile with $\sigma_2 = 0.5678$ compared to the persistent variance component with $\sigma_1 = 0.0958$. Higher level of volatility of variance in option-based estimation compared to the joint estimation is consistent with previous studies[29]

We find negative prices for both variance components where $\lambda_1 = -1.0798$ and $\lambda_2 = -1.0355$. These negative prices imply that investors are willing to pay for an insurance against an increase in volatility risk, even if that increase has little persistence. To the best of our knowledge none of the previous studies of two-factor stochastic volatility models in option market reports the prices of the variance risk factors as they either focused on the options market data or the underlying index returns data. Our negative prices for both variance components is consistent with asset pricing studies where the short-run and the long-run volatility components are priced cross-sectional asset pricing factors. Adrian and Rosenberg (2008) use a large cross-section of individual stocks over a very long period and find that prices of both short-run and long-run variance components are negative and highly significant. Therefore, our joint estimation results confirm that there is a consensus of opinions about the price of transient and persistent variance components among option traders and equity traders.

Our joint estimation results show that correlation between shocks to the index returns and shocks to the persistent variance component is $\rho_1 = -0.6918$. The correlation between shocks to the index returns and shocks to the transient variance component is $\rho_2 = -0.2173$. ρ_1 and ρ_2 captures asymmetry in the response of persistent and transient variance components to positive versus negative return shocks and can be considered as the persistent and transient

[29] For instance, Bates (2000) reports that option-based estimates of volatility of variance is larger than the one obtained from time-series-based estimates.

continuous time leverage (asymmetry) effect. The leverage effect induces negative skewness in index returns and thus yields a volatility smirk. Our results show that that leverage effect is more significant in the persistent variance component compared to the transient variance component. Therefore, persistent variance component has more significant effect on the dynamic of index skewness. Using the data from option market only, we find that $\rho_1 = -0.91$ and $\rho_1 = -0.49$. The higher absolute level of option implied correlation coefficients compared to those of joint estimation is partly related to the well documented fact that risk neutral distribution is more negatively skewed.

Our persistent and transient correlation coefficients are almost consistent with those of previous studies in option market. The average correlation coefficients in Christoffersen et al. (2009, Table 3) are $\rho_1 = -0.96$ for their first variance component and $\rho_2 = -0.83$ for their second variance component.[30] Bates (2000) also reports the structural parameter estimates of a two-factor SV model using 1988-1993 S&P 500 futures option prices. He obtains one set of structural parameters over the entire sample where $\rho_1 = -0.78$ and $\rho_2 = -0.38$. To provide a basis for comparison, we also estimate structural parameters using options data only over the same sample period and find $\rho_1 = -0.91$ and $\rho_2 = -0.49$. There are potential explanations for differences between the reported estimates of the correlation coefficients in these studies, not in the least, the very different data set and the very different time span. Despite differences in the magnitude of the coefficients, the point estimates for the correlation coefficients are negative for both persistent and transient variance components across all these studies. Further, the transient variance component has lower (in absolute value) level of correlation compared to the persistent variance components in all these studies.

To provide some empirical evidence on the difference between persistent and transient variance components over time, we plot the paths of filtered variance components. Figure (1) plots filtered time series of risk-neutral spot variance components of S&P 500 index based on our two-factor stochastic volatility model. Panel A shows time series of persistent variance component and Panel B shows time series of transient variance component. The blue plots are based on the Particle Filter method using data from both S&P 500 index and option markets (joint estimation) and the red plots are filtered spot variances using only S&P 500 options data.

[Figure (1) about here]

Naturally, the overall patterns of persistent and transient variance components filtered from joint estimation are consistent with those filtered from options data only. However, option implied variance components are more volatile in the sense that when variance increases, it tends to do more sharply compared to the one filtered based on joint estimation and thus exhibit more spikes. In particular, this pattern in more pronounced in the transient variance component (Panel B). The observed sharper spikes in option-based filtered variance

[30] Christoffersen et al. (2009) use data on European S&P 500 call option quotes over the period 1990-2004. Note that they estimate a separate set of structural parameters for every year in their sample.

in the two-factor SV model is consistent with previous studies of one-factor SV model. The smoother variance paths in joint-estimation is partly due to smooth resampling procedure in SIR PF method and partly due to imposed consistency between parameter estimates under P and Q measures.

To provide more intuition about the total risk neutral variance in our two-factor SV model, Figure (2) combines persistent and transient variance components and plots time series of total spot variance versus model-free option-implied VIX volatility index. As we expect, the time series of option implied total spot variance is closely related to the VIX volatility index. Further, the time series of total spot variance from joint estimation follow the same pattern as the VIX volatility index. However, due to joint restrictions, the total spot variance from joint estimation do not exhibits volatility spikes as large as those observed in the VIX volatility index.

[Figure (2) about here]

6 Model Performance and In-Sample Fit

We measure the goodness of fit using the following Vega-weighted root mean squared option pricing errors (Vega RMSE) as it is consistent with the loss function that we used in the the optimization routine.

$$\text{Vega RMSE} \equiv \sqrt{\frac{1}{N}\sum_{n,t}^{M}\Big(\frac{C^O_{n,t} - C^M_{n,t}(\hat{\hat{\Theta}}, \hat{v}^Q_{1,t}, \hat{v}^Q_{2,t})}{Vega_{n,t}}\Big)^2}\,, \tag{42}$$

where, $C^O_{n,t}$ is the observed price of index option n on day t, $C^M_{n,t}$ is the model price for the same index option on the same day, and $Vega_{n,t}$ is the Black-Scholes option Vega for the same option contract on the same day. To provide a reference for comparison, we also report the implied volatility root mean squared error (IVRMSE).

$$\text{IVRMSE} \equiv \sqrt{\frac{1}{N}\sum_{n,t}^{M}\big(IV^O_{n,t} - IV(C^M_{n,t}(\hat{\hat{\Theta}}, \hat{v}^Q_{1,t}, \hat{v}^Q_{2,t}))\big)^2}\,, \tag{43}$$

where, $IV^O_{n,t}$ is the Black-Scholes implied volatility of observed option n on day t and $IV(C^M_{n,t}(\hat{\hat{\Theta}}, \hat{v}^Q_{1,t}, \hat{v}^Q_{2,t}))$ is the Black-Scholes implied volatility of the model option price for the same index option on the same day.

Table (4) reports in-sample goodness-of-fit for the two-factor stochastic volatility model over the entire sample, 1996 through 2011 for various maturities. Panels A and B report in-sample fit for calls and puts separately. The right panel reports model fit based on the joint estimation while the left panel gives reports option-based fit. We find that the overall Vega-weighted RMSE of joint estimation and option-based estimation are 2.56% and 0.98% respectively. Note that the overall IVRMSE are 2.59% and 0.99% respectively, which means that Vega-weighted RMSE could be used as an approximation of IVRMSE. Overall, our two-factor SV model provides a better fit to call option contracts compared to put option contracts, which is consistent with the findings in one-factor stochastic volatility model.

Note that joint estimation imposes a consistency between physical and risk neutral parameters which are otherwise not identical. Such a restriction is not required in option-based estimation which could partly explain the better in-sample fit of option-based estimation compared to joint estimation. However, the reported RMSEs confirms that unlike stochastic volatility model, joint restrictions on return and variance dynamics under P and Q measures does not lead to the poor performance of the two-factor SV model.

Broadie et al. (2007) refer to the inconsistency between the option-based estimates of certain structural parameters in SV model and the parameter estimates from underlying time-series of returns and indicate that the SV model is basically misspecified. In particular, they state that the point estimates of the correlation coefficient and volatility of volatility are incompatible under the P and Q measures. They also show that the joint restrictions on the returns and volatility dynamics under the P and Q measures lead to the poor performance of the stochastic volatility model, measured by high level of RMSE. Using S&P 500 returns and futures options data over the period of 1987 through 2003, they find IVRMSE of 1.1% for the option-based estimation and 8.73% while imposing time-series consistency.

They note that this poor performance of SV model indicates the inability of the SV models to generate sufficient amounts of conditional skewness and kurtosis. This drawback in standard SV models is mainly attributed to the fact that the estimated conditional higher moments are highly correlated with the estimated conditional variance. By contrast, in-sample fit of our two-factor SV model is significantly improved relative to the Heston SV model. Further, the spread between Vega-weighted RMSE of joint estimation and option-based estimation is reduced significantly in the two-factor SV model versus the Heston SV model. The better performance of two-factor SV model is due to the fact that it can generate stochastic correlation between volatility and stock returns. This feature enables the two-factor SV model to better capture the conditional skewness and kurtosis.[31]

7 Model Stability and Out-of-Sample Performance

In order to examine the stability of the two-factor SV model of index and its out-of-sample performance, we divide the dataset into two subsample periods. The first subsample is from January 1996 through December 2003 and contains 169,800 daily option contracts. The

[31] Previous studies show that using the option data only two factor SV model improves on the benchmark SV model both in-sample and out-of-sample, see Christoffersen et al. (2009, Section 3.1).

second one is from January 2004 to December 2011 which contains 175,910 daily option contracts. Using both daily returns and option data we filter spot daily persistent variance path and transient variance path and repeat the joint estimation routine within each subsample. Table (5) reports the parameter estimates within each subsample (Panels A and B). For the sake of comparison, Panels C and D also report the parameter estimates from option-based estimation. The main results of the subsample tests are as follows.

First, we find that PF is a reliable filtering technique even within shorter sample period of 8 years. We observe that the time series of total spot daily variances under risk neutral measure is largely consistent with the time series of the VIX option implied volatility index within each subsample period.

Second, the parameter estimates within each subsample period is largely inline with those obtained from whole-sample estimates. Moreover, within each subsample period, the joint estimation results are also consistent with option-based parameter estimates. We find that point estimate for the transient mean reversion parameter is higher in the second subsample period while the opposite is true for the persistent mean reversion speed. Overall, the level and order of parameter estimates are almost consistent within both subsample periods and also across both estimation methods (joint estimation and option-based estimation).[32]

Third, the correlation coefficients between transient and persistent variance shocks and return shocks within subsample periods remain consistent with the ones estimated over the entire sample period and those reported in previous studies[33] in the sense that the magnitude of persistent correlation coefficient is higher than its transient counterpart. Further, the transient and persistent remain negative with the same order within two subsample periods, confirming our previous findings that investors are willing to pay to avoid transient and highly mean reverting volatility shocks.

Fourth, we evaluate our model fit within both subsample periods and report Vega RMSEs and IVRMSEs separately for calls and puts and for different maturities. Entries in Table (6) and Table (7) are inline with model fit over the entire sample period, reported in Table (4). Our joint estimation result show a better in-sample fit over the second subsample period as Vega RMSEs and IVRMSEs are reduced.

Last, in order to measure the out-of-sample performance of the two-factor SV model in capturing the behaviour of S&P 500 index options, we use the parameter estimates form the first subsample (1996-2003). Given the parameter estimates from the first subsample period, we use Particle Filter methods to filter risk neutral spot daily persistent and transient variance components over the second subsample period and then compute the IVRMSEs and

[32] Christoffersen et al. (2009, Table 3) report annual risk neutral parameter estimates for the two-factor SV model over the period 1990 through 2004 using data from S&P 500 index option data. Our option-based subsample parameter estimates are mostly consistent with their average annual result except for the volatility of volatility parameter. Apart from differences in the size of sample, this difference in point estimates may partly be explained by the fact that the annual parameter estimates in Christoffersen et al. (2009) does not satisfy the Feller condition. Feller (1951) shows that a square root process is strictly positive if $2\kappa\theta > \sigma^2$.

[33] See Section 6.

Vega RMSE over the second subsample (2004-2011). Table (8) reports the summary statistics of the out-of-sample performance for different maturities and for calls and puts separately. Comparing out-of-sample entries in (8) with those of in-sample in (7) over the same period supports the stable performance of the two-factor SV model either in joint-estimation or in option-based estimation.

8 Concluding Remarks

In this paper we investigate a two-factor stochastic volatility model where the aggregate market volatility is decomposed into a persistent and a transient volatility component. We extend the pricing kernel in Christoffersen et al. (2013), where investor's equity preference is distinguished from her variance preference, and introduce an admissible pricing kernel that links the proposed market dynamics under P and Q measures. We also discuss alternative pricing kernel for risk neutralization without separating equity and variance preferences. As the proposed two-factor specification is affine, we obtain a closed-from pricing expression for European call options. We use a long time-series of daily S&P 500 index returns and the entire cross-section of S&P 500 option prices over the same time span. We filter time series of persistent and transient spot variance components and simultaneously estimate a set of structural parameters that characterizes the dynamics of index return and variance components.

In empirical analysis, we show that the proposed decomposition of volatility can be characterized by different sensitivity of the variance components to the volatility shocks and different persistence in variance components. Consistent with the previous studies in both discrete time GARCH models and continuous time stochastic volatility models, we find that one of the volatility component is highly persistent and the other one is highly mean-reverting, where immediate impact of volatility shocks on the transient volatility component is bigger but short-lived. We obtain negative risk premium for both variance components, implying that investors are willing to pay for insurance against increases in volatility risk, even if such increases have little persistence. The negative risk premiums of both variance components are consistent with the findings in equity market where Adrian and Rosenberg (2008) find that short-run and long-run variance components are priced factors with negative risk premium. We also obtain negative correlations between shocks to the index returns and shocks to the transient and persistent variance components. In particular, we observe that the persistent correlation coefficient has more significant effect on the dynamics of index skewness.

Our model provides good fit to observed option prices both in- and out-of-sample, measured by Vega-weighted root mean squared option pricing errors and implied volatility root mean squared errors. More to the point, we find that unlike stochastic volatility model, joint restrictions on return and variance dynamics under P and Q measures does not lead to the poor performance of our two-factor SV model.

Appendix

A Proof of Proposition 1

We impose the condition that the product of the price of any traded asset and the pricing kernel under physical measure is a martingale. We also impose the condition that the discounted price of any traded asset under risk neutral measure is also a martingale. We show that the two-factor stochastic volatility process under physical measure in (1) are linked to its risk-neutral counterpart in (4) by the unique arbitrage free pricing kernel introduced in (6) and deduce restrictions on the time-preference parameters, $\{\delta, \eta_1, \eta_2\}$, risk-aversion (equity aversion) parameter, ϕ, and variance preference parameters (variance aversion), $\{\zeta_1, \zeta_2\}$. We close this proof by showing how physical Wiener processes $\{z_{1,t}, z_{2,t}, w_{1,t}, w_{2,t}\}$ are linked to risk neutral Wiener processes $\{\tilde{z}_{1,t}, \tilde{z}_{2,t}, \tilde{w}_{1,t}, \tilde{w}_{2,t}\}$ by equity premium $\{\mu_1, \mu_2\}$ and variance premium $\{\lambda_1, \lambda_2\}$ parameters.

Consider that index return under physical and risk-neutral measures follows the dynamics (A.1) and (A.2).

$$
\begin{aligned}
dS_t/S_t &= (r + \mu_1 v_{1,t} + \mu_2 v_{2,t})dt + \sqrt{v_{1,t}}dz_{1,t} + \sqrt{v_{2,t}}dz_{2,t} \\
dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}(\rho_1 dz_{1,t} + \sqrt{1-\rho_1^2}dB_{1,t}) \\
dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}(\rho_2 dz_{2,t} + \sqrt{1-\rho_2^2}dB_{2,t})
\end{aligned}
\tag{A.1}
$$

$$
\begin{aligned}
dS_t/S_t &= rdt + \sqrt{v_{1,t}}d\tilde{z}_{1,t} + \sqrt{v_{2,t}}d\tilde{z}_{2,t} \\
dv_{1,t} &= \tilde{\kappa}_1(\tilde{\theta}_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}(\rho_1 d\tilde{z}_{1,t} + \sqrt{1-\rho_1^2}d\tilde{B}_{1,t}) \\
dv_{2,t} &= \tilde{\kappa}_2(\tilde{\theta}_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}(\rho_2 d\tilde{z}_{2,t} + \sqrt{1-\rho_2^2}d\tilde{B}_{2,t})
\end{aligned}
\tag{A.2}
$$

Then, following Christoffersen et al. (2013), we show that the pricing kernel links the physical and risk neutral measures has the following exponential affine form.

$$
\frac{M_t}{M_0} = \big(\frac{S_t}{S_0}\big)^{\phi} \exp\Big[\delta t + \eta_1\int_0^t v_{1,s}ds + \eta_2\int_0^t v_{2,s}ds + \zeta_1(v_{1,t} - v_{1,0}) + \zeta_2(v_{2,t} - v_{2,0})\Big] \tag{A.3}
$$

Note that in the sprite of Cox et al. (1985) and Heston (1993) we assume that the market price of each variance risk factor is proportional to spot variance. Therefore, the risk neutral process in (A.2) can be defined as follows.

$$
\begin{aligned}
dS_t/S_t &= rdt + \sqrt{v_{1,t}}d\tilde{z}_{1,t} + \sqrt{v_{2,t}}d\tilde{z}_{2,t} \\
dv_{1,t} &= (\kappa_1(\theta_1 - v_{1,t}) - \lambda_1 v_1)dt + \sigma_1\sqrt{v_{1,t}}d\tilde{w}_{1,t} \\
dv_{2,t} &= (\kappa_2(\theta_2 - v_{2,t}) - \lambda_2 v_2)dt + \sigma_2\sqrt{v_{2,t}}d\tilde{w}_{2,t}
\end{aligned}
\tag{A.4}
$$

The log stock price process under physical measure and log pricing kernel process have the following dynamics respectively.

$$
d(\log(S_t)) = (r + \mu_1 v_{1,t} + \mu_2 v_{2,t} - \frac{1}{2}v_{1,t} - \frac{1}{2}v_{2,t})dt + \sqrt{v_{1,t}}dz_{1,t} + \sqrt{v_{2,t}}dz_{2,t} \tag{A.5}
$$

$$
d(\log(M_t)) = \phi \cdot d(\log(S_t)) + (\delta + \eta_1 v_{1,t} + \eta_2 v_{2,t})dt + \zeta_1 dv_{1,t} + \zeta_2 dv_{2,t} \tag{A.6}
$$

Replacing (A.5) and (A.1) into (A.6) we have:

$$
\begin{aligned}
d(\log(M_t)) =& \big[\phi(r + \mu_1 v_{1,t} + \mu_2 v_{2,t} - \frac{1}{2}v_{1,t} - \frac{1}{2}v_{2,t}) + \delta + \eta_1 v_{1,t} + \eta_2 v_{2,t} \\
&+ \zeta_1\kappa_1(\theta_1 - v_{1,t}) + \zeta_2\kappa_2(\theta_2 - v_{2,t})\big]dt \\
&+ \big[\phi\sqrt{v_{1,t}} + \zeta_1\rho_1\sigma_1\sqrt{v_{1,t}}\big]dz_{1,t} + \big[\phi\sqrt{v_{2,t}} + \zeta_2\rho_2\sigma_2\sqrt{v_{2,t}}\big]dz_{2,t} \\
&+ \big[\zeta_1\sigma_1\sqrt{v_{1,t}}\sqrt{1-\rho_1^2}\big]dB_{1,t} + \big[\zeta_2\sigma_2\sqrt{v_{2,t}}\sqrt{1-\rho_2^2}\big]dB_{2,t}.
\end{aligned}
\tag{A.7}
$$

As $dM_t/M_t = d(\log(M_t)) + \frac{1}{2}[d(\log(M_t))]^2$ we have

$$
\begin{aligned}
dM_t/M_t =& \big[\phi(r + \mu_1 v_{1,t} + \mu_2 v_{2,t} - \frac{1}{2}v_{1,t} - \frac{1}{2}v_{2,t}) + \delta + \eta_1 v_{1,t} + \eta_2 v_{2,t} \\
&+ \zeta_1\kappa_1(\theta_1 - v_{1,t}) + \zeta_2\kappa_2(\theta_2 - v_{2,t}) + \frac{1}{2}\phi^2(v_{1,t} + v_{2,t}) \\
&+ \phi(\zeta_1\rho_1\sigma_1 v_{1,t} + \zeta_2\rho_2\sigma_2 v_{2,t}) + \frac{1}{2}\zeta_1^2\sigma_1^2 v_{1,t}^2 + \frac{1}{2}\zeta_2^2\sigma_2^2 v_{2,t}^2\big]dt \\
&+ \big[\phi\sqrt{v_{1,t}} + \zeta_1\rho_1\sigma_1\sqrt{v_{1,t}}\big]dz_{1,t} + \big[\phi\sqrt{v_{2,t}} + \zeta_2\rho_2\sigma_2\sqrt{v_{2,t}}\big]dz_{2,t} \\
&+ \big[\zeta_1\sigma_1\sqrt{v_{1,t}}\sqrt{1-\rho_1^2}\big]dB_{1,t} + \big[\zeta_2\sigma_2\sqrt{v_{2,t}}\sqrt{1-\rho_2^2}\big]dB_{2,t}.
\end{aligned}
\tag{A.8}
$$

The first restriction on the pricing kernel is that the product of the money market account, $B_t = B_0\exp(rt)$, and the pricing kernel, M_t, should be a martingale under physical measure. Therefore, $E[d(B_t \cdot M_t)] = 0$ or $E[dM_t/M_t] = -rdt$.

$$\big[\phi(r + \mu_1 v_{1,t} + \mu_2 v_{2,t} - \frac{1}{2}v_{1,t} - \frac{1}{2}v_{2,t}) + \delta + \eta_1 v_{1,t} + \eta_2 v_{2,t} + \zeta_1\kappa_1(\theta_1 - v_{1,t}) + \zeta_2\kappa_2(\theta_2 - v_{2,t}) + \frac{1}{2}\phi^2(v_{1,t} + v_{2,t}) + \phi(\zeta_1\rho_1\sigma_1 v_{1,t} + \zeta_2\rho_2\sigma_2 v_{2,t}) + \frac{1}{2}\zeta_1^2\sigma_1^2 v_{1,t}^2 + \frac{1}{2}\zeta_2^2\sigma_2^2 v_{2,t}^2\big]dt = -rdt \tag{A.9}$$

As (A.9) holds for $v_{1,t} = v_{2,t} = 0$,

$$\delta = -r(\phi + 1) - \zeta_1\kappa_1\theta_1 - \zeta_2\kappa_2\theta_2. \tag{A.10}$$

(A.9) also holds for $v_{1,t} = v_{2,t} = \infty$.

$$\begin{aligned} \eta_1 &= -\phi\mu_1 + 1/2\phi + \zeta_1\kappa_1 - 1/2(\phi^2 + \zeta_1^2\sigma_1^2 + 2\phi\zeta_1\sigma_1\rho_1) \\ \eta_2 &= -\phi\mu_2 + 1/2\phi + \zeta_2\kappa_2 - 1/2(\phi^2 + \zeta_2^2\sigma_2^2 + 2\phi\zeta_2\sigma_2\rho_2) \end{aligned} \tag{A.11}$$

The second restriction on the pricing kernel is based on the fact that $[S_t.M_t]$ is also a martingale under physical measure. Therefore, $E[d(S_t \cdot M_t)] = 0$. As a result of this restriction we have

$$\begin{aligned} & v_{1,t}(\mu_1 + \phi + \zeta_1\sigma_1\rho_1) + v_{2,t}(\mu_2 + \phi + \zeta_2\sigma_2\rho_2) = 0, \\ & \phi = \frac{-1}{v_{1,t} + v_{2,t}}\big[(\mu_1 + \zeta_1\sigma_1\rho_1)v_{1,t} + (\mu_2 + \zeta_2\sigma_2\rho_2)v_{2,t}\big]. \end{aligned} \tag{A.12}$$

If we impose the restriction that $\mu_1 + \zeta_1\sigma_1\rho_1 \equiv \mu_2 + \zeta_2\sigma_2\rho_2$, then (A.12) can be simplified as follows.

$$\phi = -(\mu_1 + \zeta_1\sigma_1\rho_1) = -(\mu_2 + \zeta_2\sigma_2\rho_2) \tag{A.13}$$

We impose the third restriction on pricing kernel so that for any asset $U \equiv U(S, v_1, v_2, t)$, $[U(t).M_t]$ is also a martingale under P-distribution. Therefore, $E[d(U \cdot M_t)] = E[dU.M_t + U.dM_t + dU.dM_t] = 0$. Replacing M_t and dM_t into this equation we have the following restriction where $U_S = \partial U(S, v_1, v_2, t)/\partial S$, $U_{v_1} = \partial U(S, v_1, v_2, t)/\partial v_1$, and $U_{v_2} = \partial U(S, v_1, v_2, t)/\partial v_2$.

$$
\begin{aligned}
&-rU + U_t + U_S(r + \mu_1 v_{1,t} + \mu_2 v_{2,t})S + U_{v_{1,t}}\kappa_1(\theta_1 - v_{1,t}) + U_{v_{2,t}}\kappa_2(\theta_2 - v_{2,t}) \\
&+ \frac{1}{2}U_{SS}(v_{1,t} + v_{2,t}) + \frac{1}{2}U_{v_{1,t}v_{1,t}}\sigma_1^2 v_{1,t} + \frac{1}{2}U_{v_{2,t}v_{2,t}}\sigma_2^2 v_{2,t} + U_{Sv_{1,t}}\rho_1\sigma_1 v_{1,t} + U_{Sv_{2,t}}\rho_2\sigma_2 v_{2,t} \\
&+ (U_S S\sqrt{v_{1,t}} + U_{v_{1,t}}\rho_1\sigma_1\sqrt{v_{1,t}})(\phi\sqrt{v_{1,t}} + \zeta_1\rho_1\sigma_1\sqrt{v_{1,t}}) \\
&+ (U_S S\sqrt{v_{2,t}} + U_{v_{2,t}}\rho_2\sigma_2\sqrt{v_{2,t}})(\phi\sqrt{v_{2,t}} + \zeta_2\rho_2\sigma_2\sqrt{v_{2,t}}) \\
&+ U_{v_{1,t}}\zeta_1\sigma_1^2 v_{1,t}(1-\rho_1^2) + U_{v_{2,t}}\zeta_2\sigma_2^2 v_{2,t}(1-\rho_2^2) = 0
\end{aligned}
\tag{A.14}
$$

The last restriction is based on the fact that discounted price process should be a martingale under risk neutral measure. Therefore, for any asset, $U(S, v_1, v_2, t)$, whose payoff depends on the state variables $\{S, v_1, v_2\}$, U/B_t is a Q-martingale. This restriction implies that $E^Q[d(U/B_t)] = 0$ or equivalently $E^Q[d(U(S, v_1, v_2, t))] = rU(S, v_1, v_2, t)$.

$$
\begin{aligned}
&U_t + rSU_S + U_{v_{1,t}}(\kappa_1(\theta_1 - v_{1,t}) - \lambda_1 v_{1,t}) + U_{v_{2,t}}(\kappa_1(\theta_1 - v_{1,t}) - \lambda_2 v_{2,t}) + \frac{1}{2}U_{SS}(v_{1,t} + v_{1,t}) \\
&+ \frac{1}{2}U_{v_{1,t}v_{1,t}}\sigma_1^2 v_{1,t} + \frac{1}{2}U_{v_{2,t}v_{2,t}}\sigma_2^2 v_{2,t} + U_{Sv_{1,t}}\rho_1\sigma_1 v_{1,t} + U_{Sv_{2,t}}\rho_2\sigma_2 v_{2,t} = rU.
\end{aligned}
\tag{A.15}
$$

Replace (A.15) from the last restriction into (A.14) from the third restriction.

$$
\begin{aligned}
&U_S(\mu_1 v_{1,t} + \mu_2 v_{2,t})S + U_{v_{1,t}}\lambda_1 v_{1,t} + U_{v_{2,t}}\lambda_2 v_{2,t} \\
&+ (U_S S\sqrt{v_{1,t}} + U_{v_{1,t}}\rho_1\sigma_1\sqrt{v_{1,t}})(\phi\sqrt{v_{1,t}} + \zeta_1\rho_1\sigma_1\sqrt{v_{1,t}}) \\
&+ (U_S S\sqrt{v_{2,t}} + U_{v_{2,t}}\rho_2\sigma_2\sqrt{v_{2,t}})(\phi\sqrt{v_{2,t}} + \zeta_2\rho_2\sigma_2\sqrt{v_{2,t}}) \\
&+ U_{v_{1,t}}\zeta_1\sigma_1^2 v_{1,t}(1-\rho_1^2) + U_{v_{2,t}}\zeta_2\sigma_2^2 v_{2,t}(1-\rho_2^2) = 0
\end{aligned}
$$

$$
\begin{aligned}
&U_S(\mu_1 v_{1,t} + \mu_2 v_{2,t})S + U_{v_{1,t}}\lambda_1 v_{1,t} + U_{v_{2,t}}\lambda_2 v_{2,t} \\
&+ U_S S\phi v_{1,t} + U_S S\zeta_1\rho_1\sigma_1 v_{1,t} + U_{v_{1,t}}\rho_1\sigma_1\phi v_{1,t} + U_{v_{1,t}}\zeta_1\sigma_1^2 v_{1,t} \\
&+ U_S S\phi v_{2,t} + U_S S\zeta_2\rho_2\sigma_2 v_{2,t} + U_{v_{2,t}}\rho_2\sigma_2\phi v_{2,t} + U_{v_{2,t}}\zeta_2\sigma_2^2 v_{2,t} = 0
\end{aligned}
\tag{A.16}
$$

From the second restriction in (A.12) we know that $\mu_1 v_{1,t} + \mu_2 v_{2,t} = -\phi v_{1,t} - \zeta_1\rho_1\sigma_1 v_{1,t} - \phi v_{2,t} - \zeta_2\rho_2\sigma_2 v_{2,t}$. Therefore, we can further simplify (A.16).

$$
U_{v_{1,t}}\big(\rho_1\sigma_1\phi + \lambda_1 + \zeta_1\sigma_1^2\big)v_{1,t} + U_{v_{2,t}}\big(\rho_2\sigma_2\phi + \lambda_2 + \zeta_2\sigma_2^2\big)v_{2,t} = 0 \tag{A.17}
$$

One admissible solution for (A.17) would be:

$$\rho_1\sigma_1\phi + \lambda_1 + \zeta_1\sigma_1^2 = 0$$
$$\rho_2\sigma_2\phi + \lambda_2 + \zeta_2\sigma_2^2 = 0 \tag{A.18}$$

If we combine restrictions in (A.18) with those introduced in (A.13) and replace them back into (A.13) we have ϕ, ζ_1, and ζ_2.

$$\zeta_1 = \frac{\rho_1\sigma_1\mu_1 - \lambda_1}{\sigma_1^2(1-\rho_1^2)}$$
$$\zeta_2 = \frac{\rho_2\sigma_2\mu_2 - \lambda_2}{\sigma_2^2(1-\rho_2^2)} \tag{A.19}$$

$$\phi = -\mu_1 - \frac{\rho_1^2\sigma_1^2\mu_1 - \lambda_1\rho_1\sigma_1}{\sigma_1^2(1-\rho_1^2)} = -\mu_2 - \frac{\rho_2^2\sigma_2^2\mu_2 - \lambda_2\rho_2\sigma_2}{\sigma_2^2(1-\rho_2^2)} \tag{A.20}$$

Therefore, an admissible pricing kernel linking the P and Q dynamics in (A.1) and (A.2) is as follows.

$$\frac{dM_t}{M_t} = -rdt - \mu_1\sqrt{v_{1,t}}dz_{1,t} - \mu_2\sqrt{v_{2,t}}dz_{2,t} + \frac{\rho_1\sigma_1\mu_1 - \lambda_1}{\sigma_1^2(1-\rho_1^2)}dB_{1,t} + \frac{\rho_2\sigma_2\mu_2 - \lambda_2}{\sigma_2^2(1-\rho_2^2)}dB_{2,t} \tag{A.21}$$

This is the pricing kernel introduced in (1).

Now, we show that how physical shocks are linked to risk neutral shocks through equity premium $\{\mu_1, \mu_2\}$ and variance premium $\{\lambda_1, \lambda_2\}$ parameters.

$$d\tilde{z}_{1,t} = dz_{1,t} + (\psi_{1,t} + \rho_1\psi_{3,t})dt$$
$$d\tilde{z}_{2,t} = dz_{2,t} + (\psi_{2,t} + \rho_2\psi_{4,t})dt$$
$$d\tilde{w}_{1,t} = dw_{1,t} + (\psi_{3,t} + \rho_1\psi_{1,t})dt$$
$$d\tilde{w}_{2,t} = dw_{2,t} + (\psi_{4,t} + \rho_2\psi_{2,t})dt \tag{A.22}$$

Replace physical shocks in return dynamics (1) by risk neutral shocks introduced in (A.22).

$$dS_t/S_t = (r + \mu_1 v_{1,t} + \mu_2 v_{2,t})dt$$
$$+ \sqrt{v_{1,t}}d\tilde{z}_{1,t} - (\psi_{1,t} + \rho_1\psi_{3,t})\sqrt{v_{1,t}}dt + \sqrt{v_{2,t}}d\tilde{z}_{2,t} - (\psi_{2,t} + \rho_2\psi_{4,t})\sqrt{v_{2,t}}dt \tag{A.23}$$

As a result of risk neutralization in (A.23), the expected stock returns in (A.23) should be equal to the risk free rate of returns. Therefore, we have the following restriction.

$$(\mu_1 v_{1,t} + \mu_2 v_{2,t})dt = (\psi_{1,t} + \rho_1 \psi_{3,t})\sqrt{v_{1,t}}dt + (\psi_{2,t} + \rho_2 \psi_{4,t})\sqrt{v_{2,t}}dt \tag{A.24}$$

One possible solution of (A.24) is as follows.

$$\begin{aligned} \mu_1\sqrt{v_{1,t}} &= \psi_{1,t} + \rho_1\psi_{3,t} \\ \mu_2\sqrt{v_{2,t}} &= \psi_{2,t} + \rho_2\psi_{4,t} \end{aligned} \tag{A.25}$$

Similarly, we replace the proposed transformation in (A.22) into the dynamics of volatilities in (1).

$$\begin{aligned} dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}d\tilde{w}_{1,t} - \sigma_1\sqrt{v_{1,t}}(\psi_{3,t} + \rho_1\psi_{1,t})dt \\ dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}d\tilde{w}_{2,t} - \sigma_2\sqrt{v_{2,t}}(\psi_{4,t} + \rho_2\psi_{2,t})dt \end{aligned} \tag{A.26}$$

The risk-neutral variance dynamics in (A.26) should be equivalent to those in (A.4), where the market price of variance risk factors is proportional to spot variance. Therefore, we have following restrictions:

$$\begin{aligned} \sigma_1\sqrt{v_{1,t}}(\psi_{3,t} + \rho_1\psi_{1,t}) &= \lambda_1 v_{1,t} \\ \sigma_2\sqrt{v_{2,t}}(\psi_{4,t} + \rho_2\psi_{2,t}) &= \lambda_2 v_{2,t} \end{aligned} \tag{A.27}$$

Combining the restrictions in (A.25) and (A.27), we have the following results, which link the physical distribution (1) to the risk neutral distribution (4).

$$\begin{aligned} \psi_{1,t} &= \frac{\sigma_1\mu_1 - \rho_1\lambda_1}{\sigma_1(1-\rho_1^2)}\sqrt{v_{1,t}} \\ \psi_{2,t} &= \frac{\sigma_2\mu_2 - \rho_2\lambda_2}{\sigma_2(1-\rho_2^2)}\sqrt{v_{2,t}} \\ \psi_{3,t} &= \frac{\lambda_1 - \rho_1\sigma_1\mu_1}{\sigma_1(1-\rho_1^2)}\sqrt{v_{1,t}} \\ \psi_{4,t} &= \frac{\lambda_2 - \rho_2\sigma_2\mu_2}{\sigma_2(1-\rho_2^2)}\sqrt{v_{2,t}} \end{aligned} \tag{A.28}$$

B Risk Neutral Distribution

Risk neutral distribution in (4) can also be extracted by assuming the following standard stochastic discount factor, without explicit assumptions about the investor's variance preferences.

$$\frac{dM_t}{M_t} = -rdt - \psi_t' dW_t \ , \tag{B.1}$$

where $\psi_t \equiv [\psi_{1,t}, \psi_{2,t}, \psi_{3,t}, \psi_{4,t}]$ is the vector of market price of risk factors and $W_t \equiv [z_{1,t}, z_{2,t}, w_{1,t}, w_{2,t}]$ is the vector of innovations in market index return and variance components. Given the SDF in (B.1), the change-of-measure from P to Q distribution has the following exponential form.

$$\frac{dQ}{dP}(t) \equiv M_t \exp(rt) = \exp\Big[-\int_0^t \psi_u^{'} dW_u - \frac{1}{2}\int_0^t \psi_u^{'} d\langle W, W^{'}\rangle_u \psi_u \Big] \tag{B.2}$$

where $\langle W, W^{'}\rangle$ is the covariance operator.

We follow the notion of Doléans-Dade exponential (stochastic exponential) and define the stochastic exponential $\varepsilon(\cdot)$ as follow.

$$\varepsilon\Big(\int_0^t \vartheta_u^{'} dW_u \Big) \equiv \exp\Big[\int_0^t \vartheta_u^{'} dW_u - \frac{1}{2}\int_0^t \vartheta_u^{'} d\langle W, W^{'}\rangle_u \vartheta_u \Big] \tag{B.3}$$

Therefore, the change-of-measure (B.2) can be expressed in term of stochastic exponential as

$$\frac{dQ}{dP}(t) = \varepsilon\Big(\int_0^t -\psi_u^{'} dW_u \Big) \tag{B.4}$$

Applying Ito's lemma, we get the following dynamic for the log stock price process under physical measure.

$$\log\Big(\frac{S_t}{S_0}\Big) = (r + \mu_1 v_{1,t} + \mu_2 v_{2,t})t - \frac{1}{2}v_{1,t}t + \int_0^t \sqrt{v_{1,u}} dz_{1,u} - \frac{1}{2}v_{2,t} + \int_0^t \sqrt{v_{2,u}} dz_{2,u} \tag{B.5}$$

Given (B.5) and definition of stochastic exponential (B.3) we have

$$\frac{S_t}{S_0} = \exp\big[(r + \mu_1 v_{1,t} + \mu_2 v_{2,t})t\big] \varepsilon\Big(\int_0^t \sqrt{v_{1,u}} dz_{1,u} \Big) \varepsilon\Big(\int_0^t \sqrt{v_{2,u}} dz_{2,u} \Big) \tag{B.6}$$

To find the market prices of risk we impose the restriction that the product of the price of any traded asset and the pricing kernel under physical measure is a P-martingale. Given the change-of-measure (B.2), the following process, $N(t)$, should be a P-martingale.

$$N(t) \equiv \frac{S_t}{S_0}\frac{dQ}{dP}(t)\exp\left(-rt\right) \tag{B.7}$$

where

$$\begin{aligned} N(t) = & \exp\left[(\mu_1 v_{1,t} + \mu_2 v_{2,t})t\right] \\ & \varepsilon\Big(\int_0^t \sqrt{v_{1,u}}dz_{1,u}\Big)\varepsilon\Big(-\int_0^t \psi_{1,u}dz_{1,u} - \int_0^t \psi_{3,u}dw_{1,u}\Big) \\ & \varepsilon\Big(\int_0^t \sqrt{v_{2,u}}dz_{2,u}\Big)\varepsilon\Big(-\int_0^t \psi_{2,u}dz_{2,u} - \int_0^t \psi_{4,u}dw_{2,u}\Big) \end{aligned} \tag{B.8}$$

Using the properties of a stochastic exponential $\varepsilon(\cdot)$, $\varepsilon(X_t)\varepsilon(Y_t) = \varepsilon(X_t + Y_t)\exp(\langle X, Y\rangle_t)$ we can rewrite the process of $N(t)$ as follows.

$$\begin{aligned} N(t) = & \exp\left[(\mu_1 v_{1,t} + \mu_2 v_{2,t})t\right] \\ & \varepsilon\Big(\int_0^t \big(\sqrt{v_{1,u}} - \psi_{1,u}\big)dz_{1,u} - \int_0^t \psi_{3,u}dw_{1,u}\Big)\exp\Big[-\int_0^t \sqrt{v_{1,u}}(\psi_{1,u} + \rho_1\psi_{3,u})du\Big] \\ & \varepsilon\Big(\int_0^t \big(\sqrt{v_{2,u}} - \psi_{2,u}\big)dz_{2,u} - \int_0^t \psi_{4,u}dw_{2,u}\Big)\exp\Big[-\int_0^t \sqrt{v_{2,u}}(\psi_{2,u} + \rho_2\psi_{4,u})du\Big] \end{aligned} \tag{B.9}$$

From the definition of a stochastic exponential we know that $\varepsilon(\cdot)$ are P-martingales. Thus, the process $N(t)$ is a P-martingale when the following restriction holds.

$$\exp\left[(\mu_1 v_{1,t} + \mu_2 v_{2,t})t\right]\exp\Big[-\int_0^t \sqrt{v_{1,u}}(\psi_{1,u} + \rho_1\psi_{3,u})du\Big]\exp\Big[-\int_0^t \sqrt{v_{2,u}}(\psi_{2,u} + \rho_2\psi_{4,u})du\Big] = 1 \tag{B.10}$$

The restriction in (B.10) can be satisfied if

$$\begin{aligned} \mu_1 v_{1,t}t - \sqrt{v_{1,t}}(\psi_{1,t} + \rho_1\psi_{3,t})t = 0 \\ \mu_2 v_{2,t}t - \sqrt{v_{2,t}}(\psi_{3,t} + \rho_2\psi_{4,t})t = 0 \end{aligned} \tag{B.11}$$

To fully specify the market prices of risk we assume that market price of variance risk factors are proportional to spot volatilites, following Heston (1993).

$$
\begin{aligned}
(\psi_{3,t} + \rho_1 \psi_{1,t}) &= \frac{v_{1,t}}{\sigma_1 \sqrt{v_{1,t}}} \lambda_1 \\
(\psi_{4,t} + \rho_2 \psi_{2,t}) &= \frac{v_{2,t}}{\sigma_2 \sqrt{v_{2,t}}} \lambda_2
\end{aligned} \tag{B.12}
$$

Combining the restrictions in (B.11) and (B.12), we have the following market price of risk factors. Note that these prices are the same as those we find in Proposition (1).

$$
\begin{aligned}
\psi_{1,t} &= \frac{\sigma_1 \mu_1 - \rho_1 \lambda_1}{(1-\rho_1^2)} \frac{\sqrt{v_{1,t}}}{\sigma_1} \\
\psi_{2,t} &= \frac{\sigma_2 \mu_2 - \rho_2 \lambda_2}{(1-\rho_2^2)} \frac{\sqrt{v_{2,t}}}{\sigma_2} \\
\psi_{3,t} &= \frac{\lambda_1 - \rho_1 \sigma_1 \mu_1}{(1-\rho_1^2)} \frac{\sqrt{v_{1,t}}}{\sigma_1} \\
\psi_{4,t} &= \frac{\lambda_2 - \rho_2 \sigma_2 \mu_2}{(1-\rho_2^2)} \frac{\sqrt{v_{2,t}}}{\sigma_2}
\end{aligned} \tag{B.13}
$$

Given the market price of risk factors (B.13), we can apply Girsanov's theorem to find transform physical innovations in (1) to its risk neutral counterpart in (4).

$$
\begin{aligned}
d\tilde{z}_{1,t} &= dz_{1,t} + \psi_{1,t} dt + \rho_1 \psi_{3,t} dt \\
d\tilde{z}_{2,t} &= dz_{2,t} + \psi_{2,t} dt + \rho_2 \psi_{4,t} dt \\
d\tilde{w}_{1,t} &= dw_{1,t} + \psi_{3,t} dt + \rho_1 \psi_{1,t} dt \\
d\tilde{w}_{2,t} &= dw_{2,t} + \psi_{4,t} dt + \rho_2 \psi_{2,t} dt
\end{aligned} \tag{B.14}
$$

With some algebra we have the following transformations.

$$
\begin{aligned}
d\tilde{z}_{1,t} &= dz_{1,t} + \mu_1 \sqrt{v_{1,t}} dt \\
d\tilde{z}_{2,t} &= dz_{2,t} + \mu_2 \sqrt{v_{2,t}} dt \\
d\tilde{w}_{1,t} &= dw_{1,t} + (\lambda_1/\sigma_1) \sqrt{v_{1,t}} dt \\
d\tilde{w}_{2,t} &= dw_{2,t} + (\lambda_2/\sigma_2) \sqrt{v_{2,t}} dt
\end{aligned} \tag{B.15}
$$

Replacing $dz_{1,t}, dz_{2,t}, dw_{1,t}, dw_{2,t}$ from (B.15) into the physical dynamics in (1) and knowing that $\tilde{\kappa}_1 = \kappa_1 + \lambda_1$, $\tilde{\kappa}_2 = \kappa_2 + \lambda_2$, $\tilde{\theta}_1 = \frac{k_1 \theta_1}{k_1 + \lambda_1}$, $\tilde{\theta}_2 = \frac{k_2 \theta_2}{k_2 + \lambda_2}$ we obtain risk neutral return and variance dynamics.

$$
\begin{aligned}
dS_t/S_t &= (r + \mu_1 v_{1,t} + \mu_2 v_{2,t}) dt + \sqrt{v_{1,t}} dz_{1,t} + \sqrt{v_{2,t}} dz_{2,t} \\
&= (r + \mu_1 v_{1,t} + \mu_2 v_{2,t}) dt + \sqrt{v_{1,t}} (d\tilde{z}_{1,t} - \mu_1 \sqrt{v_{1,t}} dt) + \sqrt{v_{2,t}} (d\tilde{z}_{2,t} - \mu_2 \sqrt{v_{2,t}} dt) \\
&= r dt + \sqrt{v_{1,t}} d\tilde{z}_{1,t} + \sqrt{v_{2,t}} d\tilde{z}_{2,t}
\end{aligned} \tag{B.16}
$$

$$\begin{aligned} dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}(d\tilde{w}_{1,t} - (\lambda_1/\sigma_1)\sqrt{v_{1,t}}dt) \\ &= (\kappa_1\theta_1 - (\kappa_1 + \lambda_1)v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}d\tilde{w}_{1,t} \\ &= \tilde{\kappa}_1(\tilde{\theta}_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}d\tilde{w}_{1,t} \end{aligned} \tag{B.17}$$

$$\begin{aligned} dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}(d\tilde{w}_{2,t} - (\lambda_2/\sigma_2)\sqrt{v_{2,t}}dt) \\ &= (\kappa_2\theta_2 - (\kappa_2 + \lambda_2)v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}d\tilde{w}_{2,t} \\ &= \tilde{\kappa}_2(\tilde{\theta}_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}d\tilde{w}_{2,t} \end{aligned} \tag{B.18}$$

C Proof of Proposition 2

We show that the GARCH model under physical measure (14) is linked to the GARCH model under risk-neutral measure (21) with the proposed pricing kernel (19) by specifying a set of sufficient conditions (20), (22), and (23). We first impose Euler equation for the risk-free asset and subsequently impose Euler equation for the underlying asset to find this parameters mapping.

Given the pricing kernel (19), we have

$$\frac{M_t}{M_{t-1}} = (\frac{S_t}{S_{t-1}})^{\phi} \exp\Big[\delta + \eta_1 h_{1,t} + \eta_2 h_{2,t} + \zeta_1(h_{1,t+1} - h_{1,t}) + \zeta_2(h_{2,t+1} - h_{2,t})\Big] \tag{C.1}$$

Rewrite the physical GRACH dynamics (14) as follows.

$$\begin{aligned} &S_t/S_{t-1} = \exp\big[r + (\mu_1 - \frac{1}{2})h_{1,t} + (\mu_2 - \frac{1}{2})h_{2,t} + \sqrt{h_{1,t}}z_{1,t} + \sqrt{h_{2,t}}z_{2,t}\big] \\ &h_{1,t+1} - h_{1,t} = w_1 + (\beta_1 - 1)h_{1,t} + \alpha_1(z_{1,t} - \gamma_1\sqrt{h_{1,t}})^2 \\ &h_{2,t+1} - h_{2,t} = w_2 + (\beta_2 - 1)h_{2,t} + \alpha_2(z_{2,t} - \gamma_2\sqrt{h_{2,t}})^2 \end{aligned} \tag{C.2}$$

Substitute the dynamics (C.2) into (C.1)

$$\begin{aligned} \frac{M_t}{M_{t-1}} = \exp\Big[&r\phi + (\mu_1 - \frac{1}{2})\phi h_{1,t} + (\mu_2 - \frac{1}{2})\phi h_{2,t} + \sqrt{h_{1,t}}\phi z_{1,t} + \sqrt{h_{2,t}}\phi z_{2,t} \\ &+ \delta + \eta_1 h_{1,t} + \eta_2 h_{2,t} \\ &+ w_1\zeta_1 + (\beta_1 - 1)\zeta_1 h_{1,t} + \alpha_1\zeta_1(z_{1,t} - \gamma_1\sqrt{h_{1,t}})^2 \\ &+ w_2\zeta_2 + (\beta_2 - 1)\zeta_2 h_{2,t} + \alpha_2\zeta_2(z_{2,t} - \gamma_2\sqrt{h_{2,t}})^2\Big]. \end{aligned} \tag{C.3}$$

Expanding squares and collecting some terms yield the following expression for a one-day pricing kernel.

$$
\begin{aligned}
\frac{M_t}{M_{t-1}} = \exp\Big[& r\phi + \delta + w_1\zeta_1 + w_2\zeta_2 \\
& + \big((\mu_1 - \frac{1}{2})\phi + \eta_1 + (\beta_1 - 1)\zeta_1 + \alpha_1\gamma_1^2\zeta_1\big)\ h_{1,t} \\
& + \big((\mu_2 - \frac{1}{2})\phi + \eta_2 + (\beta_2 - 1)\zeta_2 + \alpha_2\gamma_2^2\zeta_2\big)\ h_{2,t} \\
& + (\phi - 2\alpha_1\gamma_1\zeta_1)\sqrt{h_{1,t}}z_{1,t} + (\alpha_1\zeta_1)z_{1,t}^2 \\
& + (\phi - 2\alpha_2\gamma_2\zeta_2)\sqrt{h_{2,t}}z_{2,t} + (\alpha_2\zeta_2)z_{2,t}^2\Big]
\end{aligned}
\tag{C.4}
$$

Before imposing the Euler equation, we introduce the expectations (C.5), where $z_{1,t}$ and $z_{2,t}$ follow a standard normal distribution.

$$
\begin{aligned}
\mathrm{E}\left[\exp(2a_1b_1z_{1,t} + a_1z_{1,t}^2)\right] &= \exp\left[-\frac{1}{2}\ln(1 - 2a_1) + \frac{2a_1^2b_1^2}{1 - 2a_1}\right] \\
\mathrm{E}\left[\exp(2a_2b_2z_{2,t} + a_2z_{2,t}^2)\right] &= \exp\left[-\frac{1}{2}\ln(1 - 2a_2) + \frac{2a_2^2b_2^2}{1 - 2a_2}\right]
\end{aligned}
\tag{C.5}
$$

where in our case

$$
\begin{aligned}
a_1 = \alpha_1\zeta_1\ ,\ b_1 &= \frac{\phi - 2\alpha_1\gamma_1\zeta_1}{2\alpha_1\zeta_1}\sqrt{h_{1,t}} \\
a_2 = \alpha_2\zeta_2\ ,\ b_2 &= \frac{\phi - 2\alpha_2\gamma_2\zeta_2}{2\alpha_2\zeta_2}\sqrt{h_{2,t}}
\end{aligned}
\tag{C.6}
$$

and thus

$$
\begin{aligned}
2a_1^2b_1^2 &= 2\alpha_1^2\zeta_1^2\big(\frac{\phi - 2\alpha_1\gamma_1\zeta_1}{2\alpha_1\zeta_1}\big)^2 h_{1,t} = \frac{1}{2}\big(\phi - 2\alpha_1\gamma_1\zeta_1\big)^2 h_{1,t} \\
2a_2^2b_2^2 &= 2\alpha_2^2\zeta_2^2\big(\frac{\phi - 2\alpha_2\gamma_2\zeta_2}{2\alpha_2\zeta_2}\big)^2 h_{2,t} = \frac{1}{2}\big(\phi - 2\alpha_2\gamma_2\zeta_2\big)^2 h_{2,t}
\end{aligned}
\tag{C.7}
$$

Therefore, conditional expectations of the last two lines of pricing kernel (C.4) may be simplified as follows.

$$
\begin{aligned}
\mathrm{E}_{t-1}\left[\exp\left[(\phi - 2\alpha_1\gamma_1\zeta_1)\sqrt{h_{1,t}}z_{1,t} + \alpha_1\zeta_1z_{1,t}^2\right]\right] &= \exp\left[-\frac{1}{2}\ln(1 - 2\alpha_1\zeta_1) + \frac{\phi - 2\alpha_1\gamma_1\zeta_1}{2(1 - 2\alpha_1\zeta_1)}h_{1,t}\right] \\
\mathrm{E}_{t-1}\left[\exp\left[(\phi - 2\alpha_2\gamma_2\zeta_2)\sqrt{h_{2,t}}z_{2,t} + \alpha_2\zeta_2z_{2,t}^2\right]\right] &= \exp\left[-\frac{1}{2}\ln(1 - 2\alpha_2\zeta_2) + \frac{\phi - 2\alpha_2\gamma_2\zeta_2}{2(1 - 2\alpha_2\zeta_2)}h_{2,t}\right]
\end{aligned}
\tag{C.8}
$$

We begin the proof by imposing the Euler equation for the risk-free asset.

$$\mathrm{E}_{t-1}\left[\frac{M_t}{M_{t-1}}\right] = \exp(-r) \tag{C.9}$$

Substituting (C.4) into (C.9), taking conditional expectation, and using the results (C.8) yield

$$\begin{aligned}
\mathrm{E}_{t-1}\left[\frac{M_t}{M_{t-1}}\right] = \exp\Big[& r\phi + \delta + w_1\zeta_1 + w_2\zeta_2 \\
& + \left((\mu_1 - \frac{1}{2})\phi + \eta_1 + (\beta_1 - 1)\zeta_1 + \alpha_1\gamma_1^2\zeta_1\right) h_{1,t} \\
& + \left((\mu_2 - \frac{1}{2})\phi + \eta_2 + (\beta_2 - 1)\zeta_2 + \alpha_2\gamma_2^2\zeta_2\right) h_{2,t} \\
& - \frac{1}{2}\ln(1 - 2\alpha_1\zeta_1) + \frac{\phi - 2\alpha_1\gamma_1\zeta_1}{2(1 - 2\alpha_1\zeta_1)} h_{1,t} \\
& - \frac{1}{2}\ln(1 - 2\alpha_2\zeta_2) + \frac{\phi - 2\alpha_2\gamma_2\zeta_2}{2(1 - 2\alpha_2\zeta_2)} h_{2,t}\Big] = \exp(-r)
\end{aligned} \tag{C.10}$$

Taking logs requires

$$\begin{aligned}
& (1+\phi)r + \delta + w_1\zeta_1 + w_2\zeta_2 - \frac{1}{2}\ln(1 - 2\alpha_1\zeta_1) - \frac{1}{2}\ln(1 - 2\alpha_2\zeta_2) \\
& + \Big[\left((\mu_1 - \frac{1}{2})\phi + \eta_1 + (\beta_1 - 1)\zeta_1 + \alpha_1\gamma_1^2\zeta_1\right) + \frac{\phi - 2\alpha_1\gamma_1\zeta_1}{2(1 - 2\alpha_1\zeta_1)}\Big] h_{1,t} \\
& + \Big[\left((\mu_2 - \frac{1}{2})\phi + \eta_2 + (\beta_2 - 1)\zeta_2 + \alpha_2\gamma_2^2\zeta_2\right) + \frac{\phi - 2\alpha_2\gamma_2\zeta_2}{2(1 - 2\alpha_2\zeta_2)}\Big] h_{2,t} = 0
\end{aligned} \tag{C.11}$$

Therefore, one possible solution of (C.11) can be defined as follows.

$$\begin{aligned}
\delta &= -(\phi + 1)r - \zeta_1 w_1 - \zeta_2 w_2 + \frac{1}{2}\ln(1 - 2\zeta_1\alpha_1) + \frac{1}{2}(1 - 2\zeta_2\alpha_2) \\
\eta_1 &= -(\mu_1 - \frac{1}{2})\phi - \zeta_1\alpha_1\gamma_1^2 + (1 - \beta_1)\zeta_1 - \frac{(\phi - 2\zeta_1\alpha_1\gamma_1)^2}{2(1 - 2\zeta_1\alpha_1)} \\
\eta_2 &= -(\mu_2 - \frac{1}{2})\phi - \zeta_1\alpha_2\gamma_2^2 + (1 - \beta_2)\zeta_2 - \frac{(\phi - 2\zeta_2\alpha_2\gamma_2)^2}{2(1 - 2\zeta_2\alpha_2)}
\end{aligned} \tag{C.12}$$

Then, we impose the Euler equation for the underlying index.

$$\mathrm{E}_{t-1}\left[\frac{S_t}{S_{t-1}} \times \frac{M_t}{M_{t-1}}\right] = 1 \tag{C.13}$$

where

$$\frac{M_t}{M_{t-1}} \times \frac{S_t}{S_{t-1}} = (\frac{S_t}{S_{t-1}})^{(\phi+1)} \exp\left[\delta + \eta_1 h_{1,t} + \eta_2 h_{2,t} + \zeta_1(h_{1,t+1} - h_{1,t}) + \zeta_2(h_{2,t+1} - h_{2,t})\right]. \tag{C.14}$$

Following the results in (C.10), we replace ϕ by $\phi + 1$ and we have

$$\begin{aligned} \mathrm{E}_{t-1}\left[\frac{M_t}{M_{t-1}} \times \frac{S_t}{S_{t-1}}\right] = \exp\Big[& r(\phi+1) + \delta + w_1\zeta_1 + w_2\zeta_2 \\ & + \big((\mu_1 - \frac{1}{2})(\phi+1) + \eta_1 + (\beta_1 - 1)\zeta_1 + \alpha_1\gamma_1^2\zeta_1\big)\ h_{1,t} \\ & + \big((\mu_2 - \frac{1}{2})(\phi+1) + \eta_2 + (\beta_2 - 1)\zeta_2 + \alpha_2\gamma_2^2\zeta_2\big)\ h_{2,t} \\ & - \frac{1}{2}\ln(1 - 2\alpha_1\zeta_1) + \frac{(\phi+1) - 2\alpha_1\gamma_1\zeta_1}{2(1 - 2\alpha_1\zeta_1)} h_{1,t} \\ & - \frac{1}{2}\ln(1 - 2\alpha_2\zeta_2) + \frac{(\phi+1) - 2\alpha_2\gamma_2\zeta_2}{2(1 - 2\alpha_2\zeta_2)} h_{2,t}\Big] = \exp(-r) \end{aligned} \tag{C.15}$$

Taking logs and substituting δ, η_1 and η_2 from (C.12) yield the following restriction.

$$(\mu_1 - \frac{1}{2}) + (\mu_2 - \frac{1}{2}) + \frac{1 + 2\phi - 4\alpha_1\gamma_1\zeta_1}{2(1 - 2\alpha_1\zeta_1)}\ h_{1,t} + \frac{1 - 2\phi - 4\alpha_2\gamma_2\zeta_2}{2(1 - 2\alpha_2\zeta_2)}\ h_{2,t} = 0 \tag{C.16}$$

Therefore, one admissible solution for the risk aversion parameter would be

$$\phi = -(\mu_1 - \frac{1}{2} + \gamma_1)(1 - 2\alpha_1\zeta_1) + \gamma_1 - \frac{1}{2} = -(\mu_2 - \frac{1}{2} + \gamma_2)(1 - 2\alpha_2\zeta_2) + \gamma_2 - \frac{1}{2} \tag{C.17}$$

To complete the proof, we need to specify how physical shocks $z_{1,t}$ and $z_{2,t}$ are transformed to risk-neutral shocks $z^*_{1,t}$ and $z^*_{2,t}$. We use the fact that the risk-neutral distribution is proportional to the physical distribution times pricing kernel. We also use the fact that $z_{1,t}$ and $z_{2,t}$ are independent.

$$f^*_{t-1}(S_t) = \frac{M_t}{\mathrm{E}_{t-1}[M_t]} \times f_{t-1}(S_t) \tag{C.18}$$

Using the proposed pricing kernel and physical dynamics and after some algebra, we find that the mean and variance may shift according to the following transformations.

$$\begin{aligned} z^*_{1,t} &= \sqrt{1 - 2\alpha_1\zeta_1}\big(z_{1,t} + (\mu_1 + \frac{\alpha_1\zeta_1}{1 - 2\alpha_1\zeta_1})\sqrt{h_{1,t}}\big) \\ z^*_{2,t} &= \sqrt{1 - 2\alpha_2\zeta_2}\big(z_{2,t} + (\mu_2 + \frac{\alpha_2\zeta_2}{1 - 2\alpha_2\zeta_2})\sqrt{h_{2,t}}\big) \end{aligned} \tag{C.19}$$

Note that the risk-neutral (21) dynamics can be derived by replacing the risk-neutral shocks (C.19) into the physical dynamics (14).

References

Adrian, T. and J. Rosenberg (2008). Stock returns and volatility: Pricing the short-run and long-run components of market risk. *The Journal of Finance 63*(6), 2997–3030.

Ait-Sahalia, Y. and R. Kimmel (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics 83*(2), 413–452.

Albrecher, H., P. Mayer, W. Schoutens, and J. Tistaert (2006). The little heston trap. *status: published 6*(05).

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*(3), 1047–1091.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of Finance 52*(5), 2003–2049.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies 9*(1), 69–107.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics 94*(1), 181–238.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. and H. Zhou (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics 109*(1), 33–65.

Bollerslev, T. and H. Zhou (2006). Volatility puzzles: a simple framework for gauging return-volatility regressions. *Journal of Econometrics 131*(1), 123–150.

Boloorforoosh, A. (2014). Is idiosyncratic volatility risk priced? evidence from the physical and risk-neutral distributions. *Concordia University Working Paper*.

Broadie, M., M. Chernov, and M. Johannes (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance 62*(3), 1453–1490.

Carr, P. and L. Wu (2007). Stochastic skew in currency options. *Journal of Financial Economics 86*(1), 213–247.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative models for stock price dynamics. *Journal of Econometrics 116*(1), 225–257.

Chernov, M. and E. Ghysels (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics 56*(3), 407–458.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2010). Option valuation with conditional heteroskedasticity and nonnormality. *Review of Financial Studies 23*(5), 2139–2183.

Christoffersen, P., S. Heston, and K. Jacobs (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science 55*(12), 1914–1932.

Christoffersen, P., S. Heston, and K. Jacobs (2013). Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies 26*(8), 1963–2006.

Christoffersen, P., K. Jacobs, and K. Mimouni (2010). Volatility dynamics for the s&p 500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies 23*(8), 3141–3189.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics 90*(3), 272–297.

Corradi, V. (2000). Reconsidering the continuous time limit of the garch (1,1) process. *Journal of Econometrics 96*(1), 145–153.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*(2), 385–407.

Dai, Q. and K. J. Singleton (2000). Specification analysis of affine term structure models. *The Journal of Finance 55*(5), 1943–1978.

Dai, Q. and K. J. Singleton (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics 63*(3), 415–441.

Derman, E. (1999). Regimes of volatility. *Risk 4*, 55–59.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical Finance 5*(1), 13–32.

Duan, J.-C., G. Gauthier, and J.-G. Simonato (1999). An analytical approximation for the garch option pricing model. *Journal of Computational Finance 2*(4), 75–116.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Egloff, D., M. Leippold, and L. Wu (2010). The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis 45*(5), 1279–1310.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. and G. Lee (1999). *A permanent and transitory component model of stock return volatility, in Robert F. Engle, and Halbert L. White, eds.: Cointegration, Causality, and Forecasting: A Festschrift in Honor of CliveW. J. Granger*. Oxford University Press.

Eraker, B. (2004). Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance 59*(3), 1367–1404.

Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics 54*(1), 173–182.

Feunou, B. and R. Tédongap (2012). A stochastic volatility model with conditional skewness. *Journal of Business & Economic Statistics 30*(4), 576–591.

Gordon, N. J., D. J. Salmond, and A. F. Smith (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, Volume 140, pp. 107–113. IET.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Heston, S. L. and S. Nandi (2000). A closed-form garch option valuation model. *Review of Financial Studies 13*(3), 585–625.

Huang, J.-z. and L. Wu (2004). Specification analysis of option pricing models based on time-changed lévy processes. *The Journal of Finance 59*(3), 1405–1440.

Johannes, M. and N. Polson (2009). Particle filtering. *Handbook of Financial Time Series*, 1015–1029.

Johannes, M. S., N. G. Polson, and J. R. Stroud (2009). Optimal filtering of jump diffusions: Extracting latent states from asset prices. *Review of Financial Studies 22*(7), 2759–2799.

Jones, C. S. (2003). The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics 116*(1), 181–224.

Khrapov, E. R. S. and E. Renault (2016). Affine option pricing model in discrete time. *Working Paper*.

Li, H., M. T. Wells, and L. Y. Cindy (2008). A bayesian analysis of return dynamics with lévy jumps. *Review of Financial Studies 21*(5), 2345–2378.

Mencía, J. and E. Sentana (2013). Valuation of vix derivatives. *Journal of Financial Economics 108*(2), 367–391.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics 63*(1), 3–50.

Pitt, M. K. (2002). Smooth particle filters for likelihood evaluation and maximisation. the warwick economics research paper series. *University of Warwick, Department of Economics*.

Santa-Clara, P. and S. Yan (2010). Crashes, volatility, and the equity premium: Lessons from s&p 500 options. *The Review of Economics and Statistics 92*(2), 435–451.

Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics 19*(2), 351–372.

Table 1: S&P 500 Index Call Option Data Characteristics by Moneyness and Maturity

Panel A: Number of call option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤0.92	152	3,371	12,690	8,782	24,995
0.92<S/K≤0.94	642	8,220	17,345	8,342	34,549
0.94<S/K≤0.96	4,033	14,436	18,557	8,096	45,122
0.96<S/K≤0.98	10,761	17,202	17,000	7,167	52,130
S/K>0.98	13,052	16,137	15,628	6,485	51,302
All	28,640	59,366	81,220	38,872	208,098

Panel B: Average price of call option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤0.92	13.6200	15.5478	23.0998	47.0797	24.8368
0.92<S/K≤0.94	11.7434	16.1440	26.2574	56.2993	27.6110
0.94<S/K≤0.96	9.9935	18.0151	34.2459	69.4400	32.9236
0.96<S/K≤0.98	11.5532	24.4015	44.6126	82.1867	40.6885
S/K>0.98	18.5235	35.5330	57.9296	95.6642	51.9126
All	13.0867	21.9283	37.2290	70.1340	35.5945

Panel C: Average implied volatility of call option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤0.92	0.4071	0.2299	0.1894	0.1791	0.2514
0.92<S/K≤0.94	0.3163	0.2034	0.1760	0.1831	0.2197
0.94<S/K≤0.96	0.2213	0.1792	0.1770	0.1881	0.1914
0.96<S/K≤0.98	0.1784	0.1741	0.1833	0.1958	0.1829
S/K>0.98	0.1715	0.1829	0.1900	0.2028	0.1868
All	0.2589	0.1939	0.1831	0.1898	0.2064

Panel D: Average delta of call option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤0.92	0.2316	0.2302	0.2724	0.3726	0.2767
0.92<S/K≤0.94	0.2329	0.2549	0.3121	0.4268	0.3067
0.94<S/K≤0.96	0.2381	0.2984	0.3832	0.4827	0.3506
0.96<S/K≤0.98	0.2996	0.3843	0.4608	0.5319	0.4191
S/K>0.98	0.4422	0.4976	0.5377	0.5771	0.5136
All	0.2889	0.3331	0.3932	0.4782	0.3733

Note to Table: This table reports the summary statistics of out-of-the-money S&P 500 call option contracts in our sample, from January 1, 1996 to December 31, 2011. The implied volatilities and the deltas are from the OptionMetrics volatility surface data set. S denotes the price of the S&P 500 index, K the option strike price, and DTM denotes the number of calendar days to maturity.

Table 2: S&P 500 Index Put Option Data Characteristics by Moneyness and Maturity

Panel A: Number of put option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤1.02	10,776	13,499	13,463	5,904	43,642
1.02<S/K≤1.04	7,163	10,951	12,018	5,008	35,140
1.04<S/K≤1.06	3,699	8,083	10,399	5,317	27,498
1.06<S/K≤1.08	1,248	5,334	8,105	3,908	18,595
S/K>1.08	385	3,173	5,591	3,588	12,737
All	23,271	41,040	49,576	23,725	137,612

Panel B: Average price of put option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤1.02	18.7121	30.3521	44.9423	63.5550	39.3904
1.02<S/K≤1.04	13.9689	25.4113	40.1731	59.5418	34.7738
1.04<S/K≤1.06	12.7334	21.7862	34.1231	55.3294	30.9930
1.06<S/K≤1.08	14.0224	20.8254	30.5229	44.3883	27.4397
S/K>1.08	16.1005	20.9994	30.9259	43.7921	27.9545
All	15.1075	23.8749	36.1375	53.3213	32.1103

Panel C: Average implied volatility of put option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤1.02	0.1929	0.1933	0.1992	0.2121	0.1994
1.02<S/K≤1.04	0.2194	0.2134	0.2158	0.2127	0.2153
1.04<S/K≤1.06	0.2646	0.2314	0.2233	0.2313	0.2376
1.06<S/K≤1.08	0.3342	0.2599	0.2367	0.2200	0.2627
S/K>1.08	0.4255	0.2904	0.2583	0.2343	0.3021
All	0.2873	0.2377	0.2266	0.2221	0.2434

Panel D: Average delta of put option contracts

	DTM≤30	30<DTM≤91	91<DTM≤182	DTM>182	All
S/K≤1.02	-0.3931	-0.3988	-0.3931	-0.3631	-0.3870
1.02<S/K≤1.04	-0.2860	-0.3221	-0.3403	-0.3334	-0.3204
1.04<S/K≤1.06	-0.2348	-0.2699	-0.2932	-0.3060	-0.2760
1.06<S/K≤1.08	-0.2194	-0.2395	-0.2579	-0.2612	-0.2445
S/K>1.08	-0.2175	-0.2209	-0.2431	-0.2547	-0.2341
All	-0.2702	-0.2902	-0.3055	-0.3037	-0.2924

Note to Table: This table reports the summary statistics of out-of-the-money S&P 500 put option contracts in our sample, from January 1, 1996 to December 31, 2011. The implied volatilities and delta are from the OptionMetrics volatility surface data set. S denotes the price of the S&P 500 index, K the option strike price, and DTM denotes the number of calendar days to maturity.

Table 3: Market Parameter Estimates

Panel A: Parameter Estimates (Physical) - Joint Estimation

κ_1	κ_2	θ_1	θ_2	σ_1	σ_2	ρ_1	ρ_2	λ_1	λ_2
1.4271	3.5874	0.0026	0.0171	0.0855	0.3496	-0.6918	-0.2173	-1.0798	-1.0355
9.38E-02	8.26E-02	1.12E-02	5.10E-03	8.93E-03	1.09E-02	3.47E-02	3.91E-02	5.55E-02	4.39E-02

Panel B: Parameter Estimates (Risk Neutral) - Options-based Estimation

$\tilde{\kappa}_1$	$\tilde{\kappa}_2$	$\tilde{\theta}_1$	$\tilde{\theta}_2$	σ_1	σ_2	ρ_1	ρ_2
0.2267	2.9137	0.0590	0.0100	0.0958	0.5678	-0.9135	-0.4934
4.73E-02	3.16E-02	6.01E-03	3.37E-03	9.75E-03	1.03E-02	2.85E-02	3.83E-02

Note to Table: This table reports the structural parameter estimates of the S&P 500 Index for the two-factor stochastic volatility model. The reported results in Panel A are from the joint estimation using the daily S&P 500 index returns and options data. Structural parameters in Panel B are estimated using only options data. In both panels, we use 10% OTM call and put options over the period 1996-2011. As in Proposition (1), $\tilde{\kappa}_1 = \kappa_1 + \lambda_1$, $\tilde{\kappa}_2 = \kappa_2 + \lambda_2$, $\tilde{\theta}_1 = \frac{k_1\theta_1}{k_1+\lambda_1}$, $\tilde{\theta}_2 = \frac{k_2\theta_2}{k_2+\lambda_2}$. Therefore, risk neutral parameters from joint estimation are $\tilde{\kappa}_1 = 0.3473$, $\tilde{\kappa}_2 = 2.5520$, $\tilde{\theta}_1 = 0.0106$, $\tilde{\theta}_2 = 0.0240$. Standard errors are reported below the parameter estimates and computed by the outer product of gradient matrix evaluated at the optimum parameter estimates.

Table 4: Goodness of Fit

	Option Based Estimation				Joint Estimation		
	Number of Obs.	Vega RMSE	IV RMSE	IVRMSE/ Avg. IV	Vega RMSE	IV RMSE	IVRMSE/ Avg. IV
Panel A: Goodness of Fit - Call Option Contracts							
DTM≤30	28,640	1.2956			2.7171		
30<DTM≤91	59,366	0.8695			2.5104		
91<DTM≤182	81,220	0.6913			2.3505		
DTM>182	38,872	0.8943			2.6032		
All	208,098	0.8846	0.9132	4.4244	2.5299	2.5637	12.4210
Panel B: Goodness of Fit - Put Option Contracts							
DTM≤30	23,271	1.6193			2.8857		
30<DTM≤91	41,040	1.0712			2.4509		
91<DTM≤182	49,576	0.8342			2.4941		
DTM>182	23,725	1.0440			2.5256		
All	137,612	1.1064	1.1167	4.5879	2.5877	2.6389	10.8418
Panel C: Goodness of Fit - All Option Contracts							
DTM≤30	51,911	1.4497			2.7946		
30<DTM≤91	100,406	0.9571			2.4835		
91<DTM≤182	130,796	0.7486			2.4180		
DTM>182	62,597	0.9538			2.5665		
All	345,710	0.9790	0.9992	4.4428	2.5566	2.5939	11.5335

Note to Table: This table reports in-sample goodness-of-fit for our two-factor stochastic volatility model over the entire sample, 1996 through 2011 for various maturities. We also report in-sample fit for calls and puts separately. All numbers are in percentage points. We compute the Vega-weighted root mean squared error (Vega RMSE) along with the implied volatility root mean squared error (IVRMSE). We also report the ratio of IVRMSE over the average implied volatility. To provide a basis for caparison the left panel reports pricing errors based on the option data and the right panel reports those of joint estimation.

Table 5: Subsample Parameter Estimates

κ_1	κ_2	θ_1	θ_2	σ_1	σ_2	ρ_1	ρ_2	λ_1	λ_2
Panel A: Joint Estimation:1996 - 2003									
1.2138	3.2780	0.0033	0.0195	0.0855	0.3220	-0.6514	-0.2985	-1.1008	-0.9755
9.62E-02	8.59E-02	1.19E-02	5.44E-03	1.02E-02	1.03E-02	3.29E-02	3.89E-02	5.92E-02	3.95E-02
Panel B: Joint Estimation (2003 - 2011)									
1.1274	4.2337	0.0069	0.0289	0.0793	0.4675	-0.5102	-0.3086	-1.0684	-1.0351
5.02E-02	3.85E-02	4.81E-03	3.03E-03	1.09E-02	1.33E-02	2.96E-02	3.71E-02		
Panel C: Options-based Estimation (1996-2003)									
0.1794	2.6176	0.0437	0.0104	0.0912	0.3732	-0.8891	-0.4434		
8.97E-02	8.01E-02	1.08E-02	4.76E-03	8.57E-03	1.16E-02	4.13E-02	3.47E-02	5.36E-02	5.27E-02
Panel D: Options-based Estimation (2003-2011)									
0.1117	3.4731	0.0623	0.0247	0.0837	0.6692	-0.7550	-0.6497		
4.14E-02	2.97E-02	6.13E-03	3.48E-03	8.84E-03	9.15E-03	2.73E-02	3.86E-02		

Note to Table: This table reports the structural parameter estimates of the S&P 500 Index for the two-factor stochastic volatility model over two subsample periods. The first subsample is from January 1996 to December 2003 and the second one is from January 2004 to December 2011. The point estimates in Panel A and Panel B are from the joint estimation using the daily S&P 500 index returns and options data. Entries in Panel C and Panel D are estimated using only options data. In both panels, we use OTM call and put options up to 10% moneyness over the period 1996-2011. Standard errors are reported below the parameter estimates and computed by the outer product of gradient matrix evaluated at the optimum parameter estimates.

Table 6: Subsample Goodness of Fit (1996-2003)

	Option Based Estimation				Joint Estimation		
	Number of Obs.	**Vega RMSE**	**IV RMSE**	**IVRMSE/ Avg. IV**	**Vega RMSE**	**IV RMSE**	**IVRMSE/ Avg. IV**
Panel A: Subsample Goodness of Fit (1996-2003) - Call Option Contracts							
DTM≤30	14,267	1.2355			2.9061		
30<DTM≤91	30,414	0.8397			2.8784		
91<DTM≤182	39,160	0.7194			2.7826		
DTM>182	18,237	0.7593			3.0274		
All	102,078	0.8514	0.8846	4.5041	2.8787	2.9137	12.8697
Panel B: Subsample Goodness of Fit (1996-2003) - Put Option Contracts							
DTM≤30	11,775	1.5167			3.3108		
30<DTM≤91	20,282	1.1038			2.9729		
91<DTM≤182	24,137	0.8742			2.9596		
DTM>182	11,528	1.0111			2.9025		
All	67,722	1.1006	1.1067	4.7416	3.0462	3.1389	11.9169
Panel C: Subsample Goodness of Fit (1996-2003) - All Option Contracts							
DTM≤30	26,042	1.3698			3.1091		
30<DTM≤91	50,696	0.9542			2.9218		
91<DTM≤182	63,297	0.7820			2.8691		
DTM>182	29,765	0.8655			2.9682		
All	169,800	0.9586	0.9792	4.5567	2.9592	3.0055	12.2725

Note to Table: This table reports in-sample goodness-of-fit for our two-factor stochastic volatility model over the entire sample, 1996 through 2011 for various maturities. We also report in-sample fit for calls and puts separately. All numbers are in percentage points. We compute the Vega-weighted root mean squared error (Vega RMSE) along with the implied volatility root mean squared error (IVRMSE). We also report the ration of IVRMSE over the average implied volatility. To provide a basis for caparison the left panel reports pricing errors based on the option data and the right panel reports those of joint estimation.

Table 7: Subsample Goodness of Fit (2004-2011)

	Option Based Estimation				**Joint Estimation**		
	Number of Obs.	**Vega RMSE**	**IV RMSE**	**IVRMSE/ Avg. IV**	**Vega RMSE**	**IV RMSE**	**IVRMSE/ Avg. IV**
Panel A: Subsample Goodness of Fit (2004-2011) - Call Option Contracts							
DTM≤30	14,373	1.3526			2.5715		
30<DTM≤91	28,952	0.8998			2.1570		
91<DTM≤182	42,060	0.6640			1.9298		
DTM>182	20,635	0.9985			2.0532		
All	106,020	0.9155	0.9471	4.1833	2.2014	2.3017	10.1665
Panel B: Subsample Goodness of Fit (2004-2011) - Put Option Contracts							
DTM≤30	11,496	1.7181			2.4266		
30<DTM≤91	20,758	1.0383			1.9112		
91<DTM≤182	25,439	0.7944			1.9656		
DTM>182	12,197	1.0741			2.0348		
All	69,890	1.1121	1.1437	4.3421	2.0802	2.1294	8.0843
Panel C: Subsample Goodness of Fit (2004-2011) - All Option Contracts							
DTM≤30	25,869	1.5259			2.5109		
30<DTM≤91	49,710	0.9601			2.0487		
91<DTM≤182	67,499	0.7159			1.9459		
DTM>182	32,832	1.0273			2.0445		
All	175,910	0.9982	1.0297	4.2046	2.1480	2.2348	9.1255

Note to Table: This table reports goodness-of-fit for our two-factor stochastic volatility model over the subsample from January 2004 through December 2011 for various maturities. We also report in-sample fit for calls and puts separately. All numbers are in percentage points. We compute vega-weighted root mean squared error (Vega RMSE) along with implied volatility root mean squared error (IVRMSE). We also report the ration of IVRMSE over the average implied volatility. To provide a basis for caparison the left panel reports pricing errors based on the option data and the right panel reports those of joint estimation.

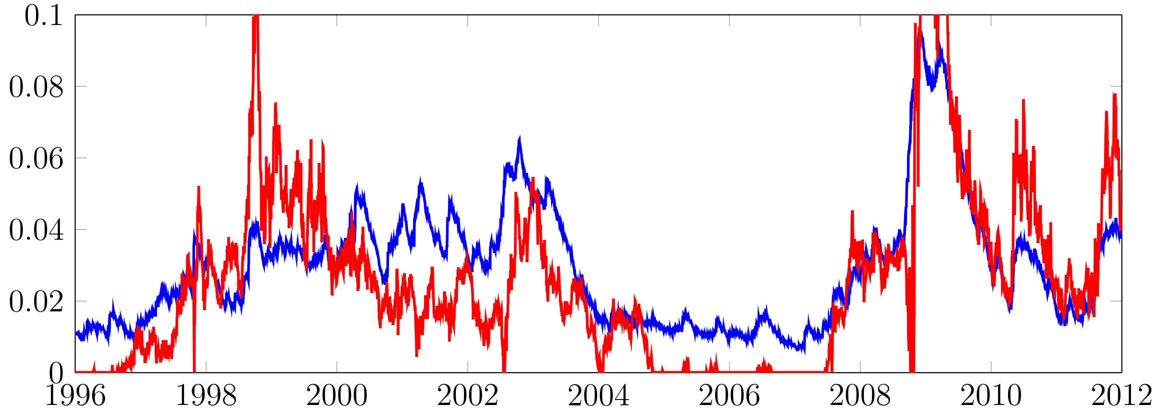
Table 8: Out of Sample Goodness of Fit (2004-2011)

	Option Based Estimation				Joint Estimation		
	Number of Obs.	Vega RMSE	IV RMSE	IVRMSE/ Avg. IV	Vega RMSE	IV RMSE	IVRMSE/ Avg. IV
Panel A: Out of Sample Goodness of Fit (2004-2011) - Call Option Contracts							
DTM≤30	14,373	1.4764			2.7853		
30<DTM≤91	28,952	0.9372			2.2801		
91<DTM≤182	42,060	0.6902			1.9978		
DTM>182	20,635	1.0797			2.1189		
All	106,020	0.9753	0.9985	4.4103	2.2201	2.3907	10.5596
Panel B: Out of Sample Goodness of Fit (2004-2011) - Put Option Contracts							
DTM≤30	11,496	1.8064			2.5780		
30<DTM≤91	20,758	1.1048			1.9984		
91<DTM≤182	25,439	0.8359			1.9856		
DTM>182	12,197	1.1153			2.1478		
All	69,890	1.1708	1.2142	4.6097	2.1259	2.2087	8.3853
Panel C: Out of Sample Goodness of Fit (2004-2011) - All Option Contracts							
DTM≤30	25,869	1.6313			2.6952		
30<DTM≤91	49,710	1.0105			2.1670		
91<DTM≤182	67,499	0.7485			1.9932		
DTM>182	32,832	1.0931			2.1297		
All	175,910	1.0573	1.0893	4.4480	2.1831	2.3201	9.4737

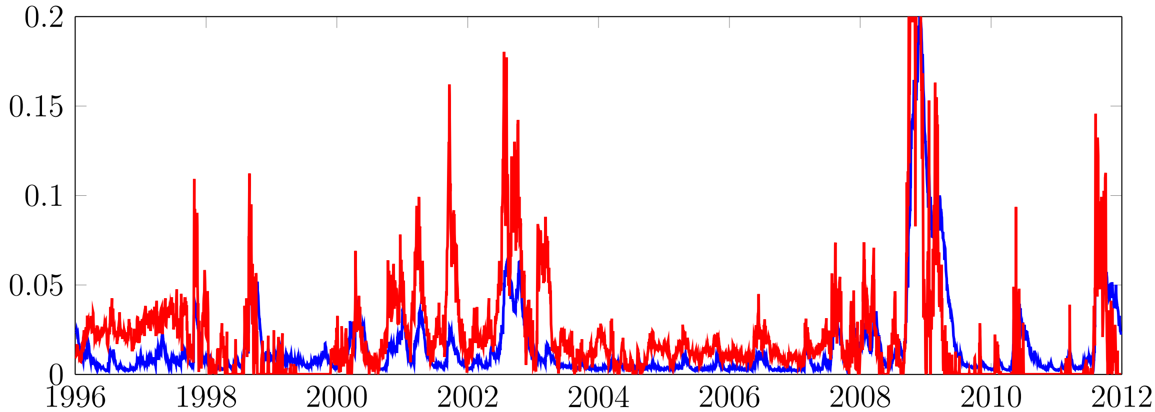
Note to Table: This table reports out-of-sample goodness-of-fit for our two-factor stochastic volatility model over the period from January 2004 through December 2011 for various maturities. We also report out-of-sample fit for calls and puts separately. All numbers are in percentage points. Out-of-sample daily spot persistent and transient variance components are filtered with Particle Filter method given the in-sample structural parameter estimates over the period January 1996 through December 2003. The Vega RMSE along with the IVRMSE are computed given in-sample structural parameters and filtered variance components. We also report the ratio of IVRMSE over the average implied volatility. To provide a basis for capariison the left panel reports pricing errors based on the option data and the right panel reports those of joint estimation.

Figure 1: The S&P 500 Index Spot Variance Components Paths

Panel A: Persistent Variance Component



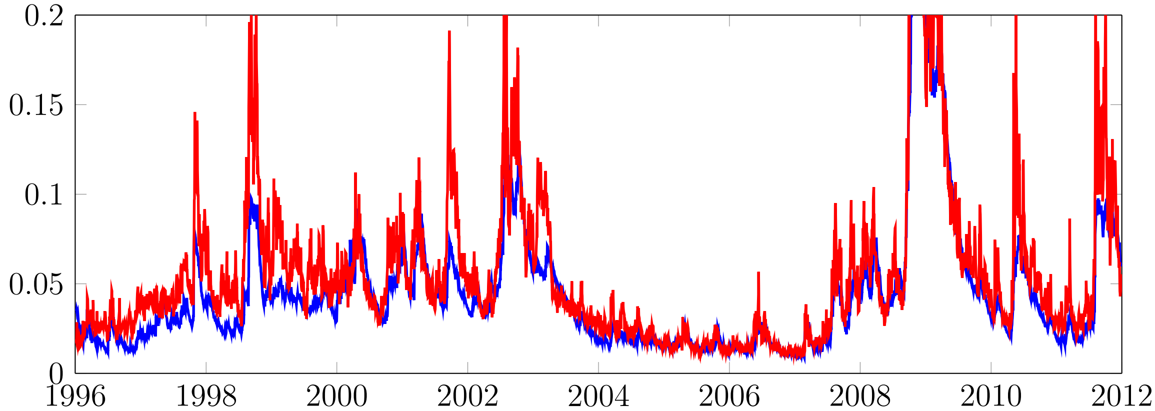
Panel B: Transient Variance Component



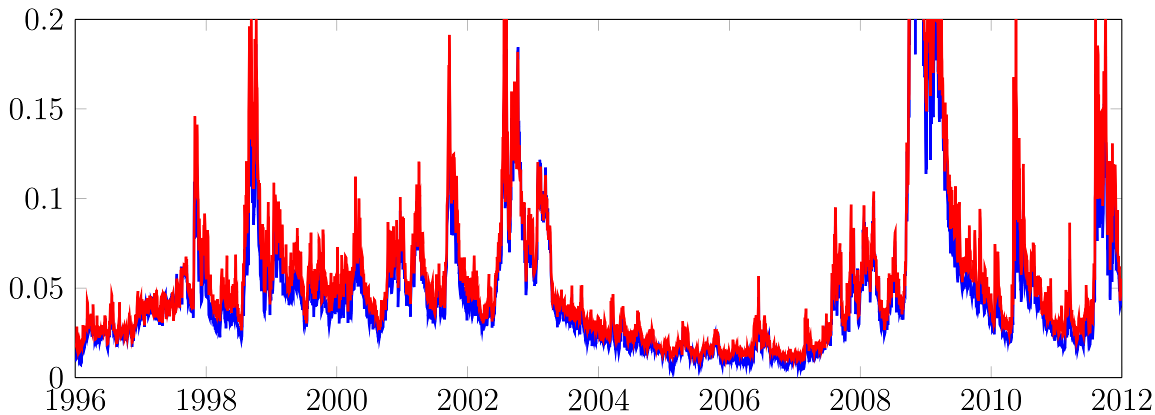
Note to Figure: We plot time series of risk-neutral spot variances for the S&P 500 index in the two-factor stochastic volatility model. Panel A shows time series of persistent variance component and Panel B shows time series of transient variance component. The blue plots are based on the Particle Filter method using data from both S&P 500 index and option markets (joint estimation). The red plots are filtered spot variances using data from S&P 500 option market only.

Figure 2: The S&P 500 Index Total Spot Variance Path Versus VIX

Panel A: Joint Estimation Versus VIX



Panel B: Option-Based Estimation Versus VIX



Note to Figure: We plot time series of risk-neutral total spot variance for the S&P 500 index by combining persistent and transient variance components of the two-factor stochastic volatility model. The blue plots in Panel A is based on the Particle Filter method using data from both S&P 500 index and option markets (joint estimation). The blue plot in Panel B is based on data from S&P 500 option market only. Red plots in both panels are time series of the VIX option implied volatility index.