Estimation for Non-Affine Stochastic Volatility Models and Volatility Risk Premiums with VIX

Mao-Wei Hung\textsuperscript{a}, Yi-Chen Ko\textsuperscript{b}, and Jr-Yan Wang\textsuperscript{c}

Abstract

Due to the unsatisfactory empirical performance of affine stochastic volatility models, it is an ongoing issue to explore more general stochastic volatility models. This paper develops a novel estimation method for a generalized stochastic volatility process with a nonlinear drift, the characteristic of constant elasticity of variance, and jumps. In contrast to all previous estimation methods, which employ time- and stock-index-independent links between the CBOE Volatility Index (VIX) and the corresponding variance level, the main idea of our estimation method is to discover the relationships between VIX and variance levels conditional on the inferred contemporary stock index return innovations. Our method more faithfully reflects real-world condition and yields reliable estimation results. Without such conditional relationships between VIX and variance levels, estimation results cannot be trusted, in particular the correlation between the stock index and variance level.

Keywords: Non-affine stochastic volatility model; Nonlinear drift; VIX; Volatility risk premium; Maximum likelihood estimation

\textit{JEL classification:} G10; G13

\textsuperscript{a} Department of International Business, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan. E-mail: mwhung@ntu.edu.tw.

\textsuperscript{b} Department of International Business, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan. E-mail: d97724012@ntu.edu.tw.

\textsuperscript{c} Corresponding author. Department of International Business, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan. E-mail: jryanwang@ntu.edu.tw; Tel: +886-2-33664987; Fax: +886-2-33652245.
Estimation for Non-Affine Stochastic Volatility Models and Volatility Risk Premiums with VIX

Abstract

Due to the unsatisfactory empirical performance of affine stochastic volatility models, it is an ongoing issue to explore more general stochastic volatility models. This paper develops a novel estimation method for a generalized stochastic volatility process with a nonlinear drift, the characteristic of constant elasticity of variance, and jumps. In contrast to all previous estimation methods, which employ time- and stock-index-independent links between the CBOE Volatility Index (VIX) and the corresponding variance level, the main idea of our estimation method is to discover the relationships between VIX and variance levels conditional on the inferred contemporary stock index return innovations. Our method more faithfully reflects real-world condition and yields reliable estimation results. Without such conditional relationships between VIX and variance levels, estimation results cannot be trusted, in particular the correlation between the stock index and variance level.

*Keywords*: Non-affine stochastic volatility model; Nonlinear drift; VIX; Volatility risk premium; Maximum likelihood estimation

*JEL classification*: G10; G13
1. Introduction

In this paper, we propose a novel estimation method for a general stochastic volatility model with nonlinear drift (NLD), constant elasticity of variance diffusion (CEV), and jumps. Hereafter, we term as NLD-CEV-J our general stochastic volatility model. Our method follows the stream of approaches that extract variance levels implicitly from VIX (CBOE Volatility Index) levels, whose squared value is theoretically a model-free quantity for the expected integrated variance over the following one month (see Britten-Jones and Neuberger (2000)). In contrast to previous methods that employ a fixed VIX-variance relationship, that is, independent of both time and asset price behavior, with our estimation method we propose using VIX-variance relationships conditional on the behavior of the innovations of contemporary asset returns (inferred from our estimation procedures) to imply unobservable variance levels. The parameters are next estimated by maximizing the log-likelihood of the joint density function of asset prices and implied variances. Based on the proposed method, we estimate not only parameter values but also the corresponding asset, volatility, and jump risk premiums in our NLD-CEV-J model.

Estimating parameters of stochastic volatility models is a difficult but important

---

1 The empirical performance of the affine stochastic volatility models in, for example, Heston (1993), Bates (1996), Bakshi et al. (1997) are not satisfactory. The performance of stochastic volatility models with CEV volatilities is investigated in, for example, Aït-Sahalia and Kimmel (2007), Duan and Yeh (2010), Chourdakis and Dotsis (2011), and Durham (2013). In addition to CEV volatility, Durham (2013) also includes jumps in his stochastic volatility model. NLD stochastic variance models are examined in Bakshi et al. (2006) and Chourdakis and Dotsis (2011), for instance. These all argue the necessity to include NLD, CEV, and/or jumps in stochastic volatility processes.
task, as variance levels are unobservable but significantly influence derivative prices.

In the literature, estimation methods for stochastic volatility models can be classified into two categories: one purely based on asset prices under the physical measure, and the other exploiting the information on volatilities implied under the risk-neutral measure. In the first category, Harvey et al. (1994) employ the quasi maximum likelihood estimation approach, and Duffie and Singleton (1993), Chacko and Viceira (2003), Gallant and Tauchen (1996), and Andersen et al. (1999) propose simulated, generalized, or efficient methods of moments approaches, respectively. Jacquier et al. (1994) and Eraker (2001) use the Markov chain Monte Carlo method. A more efficient approach, the particle filter method, is advocated in, for example, Christoffersen et al. (2010) and Malik and Pitt (2011). However, parameter values of stochastic volatility models estimated based on this stream of methods are values under the physical measure and cannot be used directly for pricing derivatives.

In the second category of estimation methods, a traditional approach is to develop analytic-form pricing formulas for derivatives under stochastic volatility assumptions (see Heston (1993), Bates (1996), and Bakshi et al. (1997)) and then calibrate the parameters by minimizing the distances between market prices and theoretical values of derivatives. However, the affine stochastic volatility model is necessary for developing analytic-form formulas. This constraint limits the usefulness of this approach. To estimate stochastic volatility models with NLD, CEV features, or even jumps, a new type of approach was recently developed in Aït-Sahalia and Kimmel
Duan and Yeh (2010), and Durham (2013). These models share several common features. First, they all focus on how to extract reliable implied variances from observable price levels (usually VIXs) in financial markets. Second, once the time series of implied variances are obtained, the parameters are obtained by maximizing the likelihood value based on the probability density function of implied variances and/or asset prices. Their methods for extracting variance levels are briefly discussed respectively as follows.

For a stochastic volatility model with linear drift (LD) and CEV features, Aït-Sahalia and Kimmel (2007) propose a model-free adjustment method for (but not limited to) Black-Scholes implied variances (approximated by VIX squared values in their empirical study) to capture the effect of mean reversion. Moreover, when calculating likelihood values, they develop a closed-form expansion to approximate the true joint transition density of the asset price and the variance level. Under the assumption of a LD-CEV stochastic volatility and jump (for the asset price) processes, Duan and Yeh (2010) obtain the relationship between the VIX and the variance level by proposing a closed-form function to evaluate the expected integrated variance and then use a bivariate normal density to approximate the true joint transition density of the asset price and the implied variance level. Since the closed-form relationship between the VIX and variance levels proposed by Duan and Yeh (2010) does not apply

---

2 Aït-Sahalia and Kimmel (2007) also examine the Heston’s (1993) affine stochastic volatility model. Since the option pricing formula is available in this case, they extract the variance level directly by calibrating at-the-money, short-dated option prices.
to stochastic volatility models with NLD or jumps, Durham (2013) proposes a simulated estimation (SE) method to evaluate the expectation of the integrated variance.\(^3\) He derives the relationship for a grid of variances and the corresponding theoretical VIX squared values by simulating the stochastic variance process and then uses cubic splines to interpolate the variance level implied from the VIX market value. Due to the flexibility of Monte Carlo simulation, Durham’s (2013) SE approach can be used to estimate parameters for virtually all types of stochastic volatility models, even with NLD and/or CEV features or jumps in asset prices or variances. The results of the above papers demonstrate the poor performance of the affine stochastic volatility model, and suggest NDL, CEV, or jumps as potential ways to improve fitting performance.

However, we argue that the implied variance level should be codetermined by the prevailing VIX value and the expected stock index behavior. Note first that the significant negative correlation between the stock index return and the variance is widely confirmed both in practice and in theory. Therefore, for a wide range of expected stock index return behavior, future variance levels are expected to vary accordingly and thus the relationship between the current VIX and variance levels may be distinct, reflecting the future comovement behavior of the stock index returns and variance levels. Second, empirical data also supports a strong correlation between VIX and the

\(^3\) On the other hand, Chourdakis and Dotsis (2011) employ the Markov chain process to approximate the expected integrated variance for a NLD-CEV variance process. By assuming the squared value of VIX follows the normal distribution centering around the obtained expectation of the integrated variance under the risk-neutral probability measure, they estimate the associated parameters by maximizing the likelihood value of the VIX time series.
stock index. We plot for example in Figure 1 the S&P 500 stock index and the corresponding VIX level from 1990 to 2016. Clearly, the VIX index is highly sensitive to changes in the S&P 500 index in a negative way: when the S&P 500 moves downward significantly, the VIX moves upward rapidly. During this period, the correlation coefficient between the log returns of the S&P 500 index and the change in VIX levels was -0.7906. However, when evaluating the expectation of the integrated variance, not only the analytic-form approaches in Aït-Sahalia and Kimmel (2007) and Duan and Yeh (2010) but also the Markov chain approximation in Chourdakis and Dotsis (2011) and the SE method in Durham (2013) ignore the influence of the contemporary stock index on VIX levels and assume a theoretical VIX value that is absolutely independent of the stock index returns. Since the prospects of the stock index and its variance level are intertwined, one cannot deny the possible effect of the expected stock index behavior on the current VIX, especially noting that what we can observe is the stock index rather than the variance level. Due to the above reasons, we believe that it is necessary to have a sounder estimation model which properly reflects the relationships among VIXs, stock index levels, and variance levels when extracting the variance level from the market VIX value.

To address the above problem, we modify Durham’s (2013) SE method by discovering the VIX-variance relationships conditional on the inferred behavior of the

---

4 A similar argument is found in Section 5.1 in Aït-Sahalia and Kimmel (2007). They point out that their method to evaluate the expectation of the integrated variance is merely an approximation since the instantaneous correlation between the variance and the asset price is unlikely to be zero.
index return innovations. Hereafter, we term our model the conditionally simulated estimation (CSE) method. In our model, to reconcile with real-world condition, we ensure the volatility innovations of each simulated path to be correlated with the inferred index return innovations when simulating the stochastic variance process. To be more specific, our CSE method obtains one VIX-variance relationship at each examined time point according to the pattern of the subsequent 21-trading-day index return innovations inferred from the previous iteration. Equipped with these conditional relationships between the VIX and variance levels, our CSE method is more capable of generating reliable estimation results, especially for estimating the correlation between the asset price and the variance. Lastly, since our method determines the theoretical VIX value using both the asset price and variance processes, our method is implemented in a complete framework, rather than the partial one (that is, one in which the theoretical VIX is determined solely based on the variance process) considered in most previous papers.

The remainder of this paper is structured as follows. Section 2 first introduces the general stochastic volatility model with NLD, CEV, and jumps examined in our paper and then proposes our novel CSE method for estimating parameters associated with the general stochastic volatility model. The simulation study, which evaluates the performance of our estimation method, is shown in Section 3. Section 4 concludes our contribution and presents the ongoing work for empirical studies.
2. Model specifications and estimation procedure

2.1 Proposed NLD-CEV-J stochastic volatility model

Suppose that the dynamics of the asset price ($S_t$) and the variance ($V_t$) of the proposed NLD-CEV-J model under the risk-neutral measure satisfy the following stochastic differential equations:

$$d \ln S_t = (r(t) - q(t) - 0.5V_t)dt + \sqrt{V_t}dW_{S,t}. \quad (1)$$

$$dV_t = \left(\alpha_0 + \bar{\alpha}_1V_t + \bar{\alpha}_2V_t^2 + \bar{\alpha}_3V_t^3 - \Lambda J \bar{\mu}_J\right)dt + \beta \Lambda \nu dW_{V,t} + \nu dN_t. \quad (2)$$

where $r(t)$ and $q(t)$ are deterministic (not stochastic) risk-free rate and dividend yield, respectively, $W_{S,t}$ and $W_{V,t}$ are standard geometric Brownian motions with \( \text{Corr}(dW_{S,t}, dW_{V,t}) = \rho \), $dN_t$ is a Poisson process with arrival intensity $\Lambda J$ and independent of $W_{S,t}$ and $W_{V,t}$, and the variance jump size $J_t$ follows a normal distribution $N(\hat{\mu}_J, \sigma_J^2)$, where $\hat{\mu}_J = \mu_J - \lambda_J$, $\mu_J$ is the mean of the variance jump size under the physical measure, and $\lambda_J$ represents the market price of jump risk. Here we follow Pan (2002) in consolidating the jump risk premium on jump-size uncertainty rather than on jump-timing uncertainty. To capture the mean-reverting phenomenon by

---

5 To focus our analysis, this paper concentrates on the effect caused by the proposed stochastic volatility process rather than that from introducing jumps in asset prices.

6 Another NLD stochastic variance process involving a $V_t^{-1}$ term is examined in Chourdakis and Dotsis (2011). However, this setting leads to a singularity point of extremely high drift levels as the variance approaches zero. That is, for assets with very small variances (i.e., almost no fluctuation in today's asset prices), the explosive drift term implies that there definitely will be a large positive rebound in the variance level tomorrow (i.e., asset prices must become very volatile tomorrow). In practice, it is rare to observe this consequence, and this implication also contradicts the well-known clustering feature in the time series of the variance.
a polynomial drift of $V_t$, an odd-degree polynomial function is needed. Thus, this paper considers a third-degree polynomial function in $V_t$, for which $\alpha_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$ are the coefficients in the polynomial drift term. As for the CEV volatility term $\beta_1 V_t^{\beta_2}$, the common assumptions of $\beta_1 > 0$ and $\beta_2 \geq 0$ are made. Lastly, depending on the value of $\beta_2$, the coefficients of $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$ are defined following Takamizawa (2008) and Chourdakis and Dotsis (2011) in several cases to capture the market price of risk of the CEV volatility term of the variance process: if $\beta_2 = 0$, it needs not consider the risk premium from the CEV volatility term, so all $\hat{\alpha}_i = \alpha_i$, where $\alpha_i$ represents the coefficients of the polynomial NLD under the physical measure; if $0 < \beta_2 \leq 1$, $\hat{\alpha}_1 = \alpha_1 - \beta_1 \lambda_1$, $\hat{\alpha}_2 = \alpha_2 - \beta_1 \lambda_2$, and $\hat{\alpha}_3 = \alpha_3$; if $1 < \beta_2 \leq 2$, $\hat{\alpha}_1 = \alpha_1$, $\hat{\alpha}_2 = \alpha_2 - \beta_1 \lambda_2$, and $\hat{\alpha}_3 = \alpha_3$; if $\beta_2 > 2$, $\hat{\alpha}_1 = \alpha_1$, $\hat{\alpha}_2 = \alpha_2$, and $\hat{\alpha}_3 = \alpha_3 - \beta_1 \lambda_3$.

Our NLD-CEV-J reduces to LD-CEV model as a special case when $\hat{\alpha}_2$, $\hat{\alpha}_3$, and $\Lambda_f$ are restricted to zero:

$$V_t = (\alpha_0 + \hat{\alpha}_1 V_t)dt + \beta_1 V_t^{\beta_2}dW_{V,t}, \quad (3)$$

where $\hat{\alpha}_1 = \alpha_1 - \beta_1 \lambda_1$ regardless of the value of $\beta_2$. In addition, by imposing the restriction $\beta_2 = 0.5$ or $\beta_2 = 1$ to Equations (2) or (3), our CEV volatility term represents SQR (Heston’s (1993) model) and GARCH models, respectively. Table 1 lists the parameter settings and shorthand notation for all the specifications of our NLD-CEV-J model. Here, LD-CEV, LD-GARCH, and LD-SQR models are identical to the CEV, GARCH, and SQR models examined in Aït-Sahalia and Kimmel (2007) and Duan and Yeh (2010). To streamline our analysis, we do not examine the case of $\beta_2 = 0$ in
2.2 Estimation method

By defining $Y_t = \ln S_t$ (thus $Y_{t+\Delta t} - Y_t$ denoting the asset return), the Euler’s method is employed to discretize Equations (1) and (2) as follows:

$$Y_{t+\Delta t} - Y_t = (r(t) - q(t) - 0.5V_t)\Delta t + \sqrt{V_t\Delta t}z_{S,t},$$  \hspace{1cm} (4)

$$V_{t+\Delta t} - V_t = (\alpha_0 + \hat{\alpha}_1 V_t + \hat{\alpha}_2 V_t^2 + \hat{\alpha}_3 V_t^3 - \Lambda_j \hat{\mu}_j)\Delta t + \beta_1 V_t^{\beta_2} \sqrt{\Delta t} z_{V,t} + J_t \sigma_t,$$ \hspace{1cm} (5)

where $\Delta t = 1/252$, $z_{i,t} \sim N(0,1)$ for $i = S, V$, and $\text{Corr}(z_{S,t}, z_{V,t}) = \rho$. To approximate the jump component, $dN_t$, we use a Bernoulli approximation, $\sigma_t \sim Ber(\Lambda_j \Delta t)$, i.e., $\sigma_t = 1$ if the jump occurs and $\sigma_t = 0$ otherwise.

Following Aït-Sahalia and Kimmel (2007), the theoretical link between the VIX and variance levels is expressed as follows:

$$\text{VIX}_t^2(\tau) = \frac{1}{\tau} E_t^Q \left( \int_t^{t+\tau} V_s \, ds \right),$$ \hspace{1cm} (6)

where $\tau = 21/252 = 21\Delta t$ by assuming that there are 21 trading days in a month.

Via evaluating the above equation, one can extract the time series of the unobservable variance levels based on the market values of VIX. For LD-CEV stochastic volatility model, Duan and Yeh (2010) derive the analytic solution of Equation (6) as

$$\text{VIX}_t^2(\tau) = -\frac{\alpha_0}{\hat{\alpha}_1} \left( 1 + \frac{1-\exp(\hat{\alpha}_1 \tau)}{\hat{\alpha}_1 \tau} \right) \frac{1-\exp(\hat{\alpha}_1 \tau)}{\hat{\alpha}_1 \tau} V_t,$$ \hspace{1cm} (7)

Bollerslev and Zhou (2002) also propose a similar result.
where $\hat{\alpha}_1 = \alpha_1 - \beta_1 \lambda_1$. However, the analytic solution of Equation (6) is not available under the proposed NLD-CEV-J stochastic volatility models. Using the Monte Carlo technique, Durham (2013) approximates Equation (6) as

$$\text{VIX}_t^2(\tau) = \frac{1}{M} \sum_{j=1}^{M} \left( \frac{1}{21} \sum_{i=1}^{21} \tilde{\nu}_i^j \right) \Delta t$$

(7)

where $M$ is the number of simulated paths, and $\{\tilde{\nu}_i^j\}$ is the set of simulated volatilities that will be introduced later. In essence, Equation (7) is the discrete version of Equation (6).

Before introducing our CSE method, we emphasize that Equation (6) is established based on only the variance process in Equations (2) or (3) but without taking into account the asset price in Equation (1). We believe that the link between the VIX and variance could vary greatly for different expected behavior of the time series of $S_t$. For example, for a scenario where the series of $\{S_t\}$ is expected to rise (decline) continuously for the following 21 days, the corresponding average of the series of $\{V_t\}$ is expected to decrease (increase) with time provided $S_t$ and $V_t$ are negatively correlated. As a result, when evaluating Equation (6), the resulting values of $\text{VIX}_t^2(\tau)$ could be very different in these two scenarios even given the same variance level today. These arguments motivate us to modify Durham’s (2013) SE method by introducing the conditional link between the VIX and variance on the inferred contemporary index return innovations according to the prior estimation results.

The two phases of our CSE estimation method are presented as follows:
Phase 1:

**Step A.** Generate pairs of \( \{ \sqrt{VIX^2_g(\tau)}, \tilde{V}_g \} \)

First, choose a grid of variance values, \( \hat{V}_g \), for \( g = 1, \ldots, G \), where \( G \) is the number of volatility grids. For each \( g \), set the initial variance \( \tilde{V}_0^j = \hat{V}_g \) for \( j = 1, \ldots, M \), where \( M \) is the number of simulated variance paths. Draw samples from \( \tilde{z}_t^j \sim N(0,1) \), \( \tilde{f}_t \sim N(\tilde{\mu}_f, \sigma_f^2) \), and \( \tilde{\alpha}_t \sim Ber(\Lambda_t \Delta t) \) to simulate \( \tilde{V}_t^j \) progressively until \( t + 21\Delta t \) through

\[
V_t^j - V_t^j = (a_0 + \tilde{\alpha}_1 V_t^j + \tilde{\alpha}_2 (V_t^j)^2 + \tilde{\alpha}_3 (V_t^j)^3 - \tilde{\Lambda}_t \tilde{\mu}_f) \Delta t + \tilde{\beta}_1 (V_t^j) \sqrt{\Delta t} \tilde{z}_t^j + \tilde{f}_t \tilde{\alpha}_t. \tag{8}
\]

Then the corresponding \( \sqrt{\bar{V}IX^2_g(\tau)} \) is obtained by evaluating Equation (7).

**Step B.** Extract \( V_t \) from market series of \( \sqrt{VIX^2_{\tilde{g}}(\tau)} \)

Given the collection of pairs \( \{ \sqrt{VIX^2_{\tilde{g}}(\tau)}, \tilde{V}_g \} \), the cubic spline technique is used to construct the curve relationship between the latent variance and the market value of \( \sqrt{VIX^2_{\tilde{g}}(\tau)} \). If \( \sqrt{VIX^2_{\tilde{g}}(\tau)} < \sqrt{VIX^2_{\tilde{g}}} \leq \sqrt{VIX^2_{\tilde{g}+1}(\tau)} \), the interpolated variance is

\[
\begin{align*}
V_t &= a_{g0} + a_{g1} \left( VIX^2_{\tilde{g}} - \sqrt{VIX^2_{\tilde{g}}(\tau)} \right) + a_{g2} \left( VIX^2_{\tilde{g}} - \sqrt{VIX^2_{\tilde{g}}(\tau)} \right)^2 \\
& \quad + a_{g3} \left( VIX^2_{\tilde{g}} - \sqrt{VIX^2_{\tilde{g}}(\tau)} \right)^3,
\end{align*}
\]

where \( a_{gk} \) are the polynomial coefficients of the interpolated curve bounded with range values \( (\sqrt{VIX^2_{\tilde{g}}(\tau)}, \sqrt{VIX^2_{\tilde{g}+1}(\tau)}) \), for \( g = 1,2, \ldots, G - 1 \) and \( k = 0, \ldots, 3 \).
**Step C.** Maximize log-likelihood value

Given $Y_{t+\Delta t}$ and $V_{t+\Delta t}$ conditional on $Y_t$ and $V_t$ based on the actual data, we calculate the joint likelihood of the observations of $Y_t$ and $VIX_t^2$ under the physical measure as follows:

$$L_{t+\Delta t} = \Lambda_j \Delta t \cdot f_2 \left( Y_{t+\Delta t}, V_{t+\Delta t}, \mu_Y, \mu_V, \sigma_Y^2, (\sigma_V^2)^2, \rho \right) \cdot \left| \frac{\partial V_{t+\Delta t}}{\partial VIX_{t+\Delta t}(\tau)} \right|$$

$$+ \left( 1 - \Lambda_j \Delta t \right) \cdot f_2 \left( Y_{t+\Delta t}, V_{t+\Delta t}, \mu_Y, \mu_V, \sigma_Y^2, \sigma_V^2, \rho \right) \cdot \left| \frac{\partial V_{t+\Delta t}}{\partial VIX_{t+\Delta t}(\tau)} \right|,$$

where $f_2(\cdot)$ is a bivariate normal density function, and

$$\mu_Y = Y_t + [r(t) - q(t) + (\lambda_S - 0.5) V_t] \Delta t,$$

$$\sigma_Y = \sqrt{V_t \Delta t},$$

$$\mu_V = V_t + (\alpha_0 + \alpha_1 V_t + \alpha_2 V_t^2 + \alpha_3 V_t^3 - \Lambda_j \mu_Y) \Delta t,$$

$$\sigma_V = \beta_1 V_t^{\beta_2} \sqrt{\Delta t},$$

$$\mu'_V = V_t + (\alpha_0 + \alpha_1 V_t + \alpha_2 V_t^2 + \alpha_3 V_t^3 - \Lambda_j \mu_Y) \Delta t + \mu_j,$$

$$\sigma'_V = \left( \beta_1 V_t^{\beta_2} \sqrt{\Delta t} \right)^2 + \sigma_j^2,$$

where $\lambda_S$ is the risk premium for the asset price. The notation of $| \cdot |$ represents the Jacobian transformation from $VIX_t^2$ to $V_t$. If $VIX_g^2(\tau) < VIX_{t+\Delta t}^2 \leq VIX_{g+1}^2(\tau)$, we calculate the Jacobian using the following equation:

$$\left| \frac{\partial V_{t+\Delta t}}{\partial VIX_{t+\Delta t}(\tau)} \right| = a_{g1} + 2a_{g2} \left( VIX_{t+\Delta t}^2 - VIX_g^2(\tau) \right) + 3a_{g2} \left( VIX_{t+\Delta t}^2 - VIX_g^2(\tau) \right)^2.$$

The parameters in NLD-CEV-J model are finally estimated by maximizing

$$\sum_{i=1}^{N} \ln (L_{t+i\Delta t}),$$

where $N$ is the number of observation time points. We denote as $LV_0$ the maximized
log-likelihood value obtained in Phase 1.

In Phase 2, we use the results of the implied time series of \( \{V_t\} \) and the estimated parameter values in Phase 1 as the initial condition. The three steps in Phase 2 are introduced as follows.

**Phase 2:**

**Step 1.** Generate conditional relationship between \( \sqrt{\text{VIX}^2_g(\tau)} \) and \( \tilde{V}_g \) at each \( t \)

Given the time series of \( V_t \), we can infer the index return innovations \( \tilde{z}_{S,t} \) based on Equation (4):

\[
\tilde{z}_{S,t} = \frac{Y_{t+\Delta t} - Y_t - (\tau(t) - q(t) - 0.5V_t)\Delta t}{\sqrt{V_t \Delta t}} \quad \text{for all } t. \tag{9}
\]

Next, we obtain \( \sqrt{\text{VIX}^2_g(\tau)} \) corresponding to each \( \tilde{V}_g \) in a similar way to that of step A in Phase 1, except that \( \tilde{z}_{V,t} \) is simulated as

\[
\tilde{z}_{V,t}^i = \rho \tilde{z}_{S,t} + \sqrt{1 - \rho^2} \epsilon_{V,t}^i, \tag{10}
\]

where \( \epsilon_{V,t}^i \sim N(0,1) \), for capturing the instantaneous correlation between \( z_{S,t} \) and \( z_{V,t} \).

**Step 2.** Reestimate parameter values

When \( \{\sqrt{\text{VIX}^2_g(\tau)}, \tilde{V}_g\} \) for every time point \( t \) are derived from step 1, we perform steps B to C in Phase 1 to obtain the sum of the log-likelihood values of the observations of \( Y_t \) and \( \text{VIX}_t^2 \) for all time points. A new set of estimated parameters is obtained by maximizing this log-likelihood; we denote as \( L_{V_k} \) the resulting highest log-likelihood.
Step 3. Repeat until convergence

We repeat steps 1 and 2 until $|LV_k - LV_{k-1}| < 10^{-6}$, yielding the stationary results of estimated parameters.

Note first that Phase 1 of our CSE method is essentially identical to the Durham’s (2013) SE method except that we approximate the Poisson distribution with the Bernoulli distribution. This technique can save computational time and perform satisfactorily when $\Delta t$ is small, such as the case of $\Delta t = 1/252$ in our method.

Second, our CSE method obviously incorporates Durham’s (2013) SE method as a special case. Note that Phase 2 is reduced equivalently to Phase 1 (Durham’s SE method) when $\rho = 0$ and $\bar{Z}_{S,t} = 0$ for all $t$ in Equation (10). However, neither of these two conditions are reasonable, and they may indeed cause unexpected estimation results. In contrast to Durham (2013) which considers a fixed set of $\{\text{VIX}_{g}^{2}(\tau), \bar{v}_{g}\}$ for all examined time points and are thus irrespective of the stock index behavior, our model carefully handles the correlation between the expected patterns of index return innovations and the variance levels and embeds information from the stock index in the conditional relationship between $\text{VIX}_{g}^{2}(\tau)$ and $\bar{v}_{g}$. Although our CSE method appears to be more time-consuming than Durham’s (2013) SE method, our model is more complete and more faithfully reflects the stylized fact that the asset price return and the variance are negatively related. It is also worth noting that this problem applies not only to Durham’s method but also to all previous methods which extract the variance level
from the VIX value. Since all relevant models in the literature miss this important relationship between index return and volatility innovations, model misspecification ensures unreliable estimation results, particularly for the correlation coefficient between the asset price and the variance level, as shown later in the next section.

Last, it should be emphasized that we do not intend to use the future information of spot returns in Equation (9) when calculating theoretical values of $\bar{\text{VIX}}_{g}(\tau)$ in our estimation method. Note first that since the task of this paper is to estimate the stochastic volatility model from an ex post viewpoint (rather than to calculate the real-time expected VIX value based on known information at every time point), it is normal to allow the proposed estimation method to endogenously differentiate or recognize various patterns in the examined time series. As a matter of fact, the original design of this paper tries to discover the relationships between the VIX and variance levels conditional on various time-series patterns of index return innovations. Consequently, at each time point, we can employ a specific relationship between the VIX and variance levels according to different expected patterns of future stock index returns. However, it is impossible to identify and examine all patterns of index return innovations exhaustively. Therefore, our estimation method turns to derive conditional VIX-variance relationships on every patterns of 21-day index return innovations that we infer via Equation (9) in the examined period. Of course, those inferred patterns may not reflect precisely the true patterns in the real world. Nevertheless, the step of inferring the patterns of 21-day index return innovations in Equation (9) can help us to
differentiate different expected scenarios in the examined time series and thus derive the corresponding conditional relationships between the VIX and variance levels.

3. Simulation study

In this section, we conduct two simulation experiments to evaluate the performance of the proposed CSE method under LD-CEV and NLD-CEV stochastic volatility models, respectively. Given the simulated time series of asset prices $S_t$, variances $V_t$, and the corresponding $\text{VIX}_t^2(\tau)$ according to the posited parameter values, we seek to evaluate the performance of the CSE method, Durham’s (2013) SE method, and the analytic-form approach in Duan and Yeh (2010) (represented by DY method hereafter) based only on the simulated time series of $S_t$ and $\text{VIX}_t^2(\tau)$.

To generate the needed time series, we first simulate $N$-day series of asset log returns $Y_t$ and volatilities $V_t$ under the physical measure:

$$Y_{t+\Delta t} - Y_t = (r(t) - q(t) + (\lambda_S - 0.5)V_t)\Delta t + \sqrt{V_t\Delta t}u_{S,t},$$

$$V_{t+\Delta t} - V_t = (\alpha_0 + \alpha_1V_t + \alpha_2V_t^2 + \alpha_3V_t^3 - \Lambda_j\mu_j)\Delta t + \beta_1\sqrt{V_t}\Delta t,\quad \theta_t, \quad \theta_t \sim Ber(\Lambda_j\Delta t),$$

where $\Delta t = 1/252$, $u_{i,t} \sim N(0,1)$ for $i = S, V$, $\text{Corr}(u_{S,t}, u_{V,t}) = \rho$, $\theta_t \sim Ber(\Lambda_j\Delta t)$, and $J_t \sim N(\mu_j, \sigma_j^2)$. Next, for each $V_t$ at time point $t$, we obtain $\text{VIX}_t^2(\tau)$ through Equation (7) by simulating $\tilde{V}_{t+i\Delta t}^j$ for $i = 1, 2, ..., 21$ and $j = 1, 2, ..., M$ under the risk-neutral measure. Note that when simulating $\tilde{V}_{t+i\Delta t}^j$ via Equation (8), we follow Equation (10) to simulate $\tilde{z}_{V,t}^j$ by conditioning on the
subsequent 21-day risk-neutral index return innovations $\tilde{z}_{S,t}$, which can be derived via Equation (9). However, to avoid the foresight bias,\(^8\) we do not use the obtained risk-neutral innovations $\tilde{z}_{S,t}$ directly when implementing Equation (10). Instead, for each 21-day window, the normally distributed $\tilde{z}_{S,t}$ are resampled by controlling their mean and standard deviation equal to the observed ones in that window, and those resampled $\tilde{z}_{S,t}$ are then employed to simulating $\tilde{z}_{\nu,t}^j$ in Equation (10). Finally, the simulated time series of $S_t$ and $\text{VIX}_t^2 = \text{VIX}_t^2(\tau)$ are used as inputs of our CSE, SE, and DY methods.

For LD-CEV process, most parameter values are inherited from SV0 model in Table 3 of Duan and Yeh (2010): $q(t) = -0.0788$, $\lambda_S = -2.0863$, $\alpha_0 = 0.0392$, $\alpha_1 = -0.8309$, $\lambda_1 = -8.3547$, $\beta_1 = 1.3873$, $\beta_2 = 0.8936$, and $\rho = -0.6916$. For NLD-CEV process, all parameters are the same as those for LD-CEV model except that $\alpha_0 = 0.6$, $\alpha_1 = -10$, $\alpha_2 = 30$, $\alpha_3 = -50$, $\lambda_1 = -2.5$, and $\lambda_2 = -20$. For both cases, we further assume $r(t) = \Lambda_f = \lambda_f = 0$, $S_0 = 100$, and $V_0 = 0.04$ to simplify our analysis. As for the number of observations, we examine $N = 2520$. When implementing our CSE method and Durham’s (2013) SE method, we choose $G = 10$, $M = 500$, and $\{\bar{V}_g\} = \{0.1^2, 0.15^2, 0.2^2, 0.25^2, 0.3^2, 0.4^2, 0.5^2, 0.6^2, 0.7^2, 0.8^2\}$.

Table 2 shows the estimation results for LD-CEV process based on the three

---

\(^8\) Recall that our CSE method utilizes the contemporary innovations $\tilde{z}_{S,t}$ in Equation (9) to differentiate different expected scenarios. Therefore, if here we use the innovations $\tilde{z}_{S,t}$ derived in Equation (9) to simulate $\tilde{z}_{\nu,t}^j$ and thus $\bar{V}_t^j$, it is equivalent to allowing our CSE method to have perfect foresight into the future.
examined methods. For the parameters other than $\rho$, all of these three methods can generate reasonable estimation results. However, since the proposed CSE method does take into account the influence of asset prices on the variance level, in comparison to SE and DY methods, it exhibits superior performance in generating accurate estimation for the correlation $\rho$. The estimated value of $\rho$ is -0.5757 in our CSE method and 0.0003 and 0.0004 in DY and SE methods, respectively. This is because the fixed links between the VIX and variance in DY and SE methods, which are time- and stock-index independent, could obscure the actual correlation between the asset prices and variances.

Figure 2 shows the difference between the selected VIX-volatility relationships in our CSE model and those in SE and DY methods. It can be observed that either SE or DY method employs a fixed relationship to infer the variance level from the market value of VIX at any time point. The nearly identical VIX-volatility relationships of these two models verify the correctness and feasibility of SE method developed by Durham (2013). In contrast, our CSE method uses different VIX-volatility relationships at different time points to derive implied variances, and the range of VIX-volatility relationships generated by our CSE method is bounded by the CSE-High and CSE-Low curves in Figure 2. The differences between the CSE-High and CSE-low curves and those generated by SE and DY methods are pronounced. For example, when VIX is 30%, the implied volatility based on both SE and DY methods is about 36%, but the implied volatility based on our CSE method ranges from 25% to 51%. We further plot the 21-day index return innovations (inferred by Equation (9)) corresponding to the
CSE-High and CSE-low relationships in Figure 3. It is obvious that the innovations trend rises (falls) for the CSE-High relationship (CSE-Low relationship). As discussed in Section 2, this analysis demonstrates our conjecture that when asset prices behave dissimilarly, the VIX-volatility relationship should be different.

As for NLD-CEV stochastic volatility model, the counterpart results are reported in Table 3 and Figures 4 and 5. We only compare our CSE method with SE method because DY method is not available under NLD-CEV stochastic volatility model. The results for NLD-CEV model are similar to those for LD-CEV model. Table 3 shows that our CSE method estimates the correlation coefficient $\rho$ accurately, whereas Durham’s SE method performs unsatisfactorily in estimating the correlation coefficient $\rho$. The VIX-volatility curves of our CSE and Durham’s SE methods are compared in Figure 4, where the range of VIX-volatility relationships generated by our CSE method is bounded by the CSE-High and CSE-Low curves. The differences between the CSE-High and CSE-low curves and those generated by SE and DY methods are still significant. For example, when VIX is 30%, the implied volatility based on SE and method is about 34%, but the implied volatility based on our CSE method ranges from 24% to 42%. The inferred 21-day index return innovations corresponding to the CSE-High and CSE-low relationships are presented in Figure 5. One can observe that the innovations trend rises (falls) for the CSE-High relationship (CSE-Low relationship).

According to the simulation experiments for LD-CEV and NLD-CEV stochastic volatility models in this section, we demonstrate that the proposed CSE method can
reconcile the real-world condition more appropriately and thus provide more reliable estimation results for stochastic volatility models, especially for the correlation between the stock index and variance level.

4. Conclusion and future work

In this paper, we propose the CSE method to estimate the stochastic volatility model with NLD, CEV, and jumps. In contrast to previous studies, our CSE method conditions the relationship between the VIX and variance levels on the inferred contemporary stock index return innovations, thus resulting in an estimation procedure that more faithfully reflects actual market behavior and yielding estimation results that are more meaningful. Simulation results show that the CSE method outperforms previous models which do not take into consideration the conditional relationships between the VIX and variance levels.

Our work in the near future is to examine the fitting performance of NLD-CEV-J model for the actual stock index and VIX data with our CSE method. The data sample covers the daily prices of S&P 500 index and VIX index over the period from January 2, 1990 to June 30, 2016. Table 4 shows the descriptive statistics of our empirical sample. The mean and standard deviation values are similar to Duan and Yeh’s (2010) results. The left-skewed and heavy-tail S&P 500 index return and right-skewed and heavy-tail VIX index are also consistent with those in Duan and Yeh (2010). However, the magnitude of skewness and kurtosis is clearly larger than theirs. This may be
attributed to the influence of the financial crisis from 2007 to 2009.

We plan to conduct horse-race comparisons for the fitting performance of all models nested in our NLD-CEV-J model. Furthermore, we will revisit the issue of volatility, asset, and jump risk premiums associated with the stochastic volatility model based on our more reasonable estimation method.

References


Table 1. All possible specifications nested in the proposed NLD-CEV-J model

<table>
<thead>
<tr>
<th>With/without jumps</th>
<th>$\beta_2 = 0.5$</th>
<th>$\beta_2 = 1$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear drift</td>
<td>LD-SQR-J/ LD-SQR</td>
<td>LD-GARCH-J/ LD-GARCH</td>
<td>LD-CEV-J/ LD-CEV</td>
</tr>
<tr>
<td>($\hat{a}_2 = \hat{a}_3 = 0$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\hat{a}_1 = \alpha_1 - \beta_1 \lambda_1$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonlinear drift</td>
<td>NLD-SQR-J/ NLD-SQR</td>
<td>NLD-GARCH-J/ NLD-GARCH</td>
<td></td>
</tr>
<tr>
<td>($\hat{a}_1 = \alpha_1 - \beta_1 \lambda_1$)</td>
<td></td>
<td>($\alpha_1 = \alpha_1$)</td>
<td>NLD-CEV1-J/ NLD-CEV1 (0 &lt; $\beta_2$ ≤ 1)</td>
</tr>
<tr>
<td>($\hat{a}_2 = \alpha_2 - \beta_1 \lambda_2$)</td>
<td></td>
<td>($\hat{a}_2 = \alpha_2 - \beta_1 \lambda_2$)</td>
<td>($\hat{a}_1 = \alpha_1 - \beta_1 \lambda_1$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($\hat{a}_2 = \alpha_2 - \beta_1 \lambda_2$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NLD-CEV2-J/ NLD-CEV2 (1 &lt; $\beta_2$ ≤ 2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($\hat{a}_2 = \alpha_2 - \beta_1 \lambda_2$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NLD-CEV3-J/ NLD-CEV3 ($\beta_2 &gt; 2$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($\hat{a}_3 = \alpha_3 - \beta_1 \lambda_3$)</td>
</tr>
<tr>
<td>Parameter</td>
<td>True value</td>
<td>DY method</td>
<td>SE method</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.0392</td>
<td>0.0198</td>
<td>0.0194</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0494e-3)</td>
<td>(0.0011)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.8309</td>
<td>-0.2437</td>
<td>-0.2202</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0259e-03)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-8.3547</td>
<td>-9.5808</td>
<td>-9.2903</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0003)</td>
<td>(0.0051)</td>
</tr>
<tr>
<td>$\lambda_S$</td>
<td>-2.0863</td>
<td>-1.8914</td>
<td>-1.8883</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0229e-03)</td>
<td>(0.0057)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>1.3873</td>
<td>1.1183</td>
<td>1.1152</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0326e-03)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.8936</td>
<td>0.9257</td>
<td>0.9245</td>
</tr>
<tr>
<td></td>
<td>(0.0011)</td>
<td>(0.0302e-03)</td>
<td>(0.0014)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.6916</td>
<td>0.0003</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.0218e-03)</td>
<td>(0.0011)</td>
</tr>
</tbody>
</table>

This table reports the estimated results and their standard errors (shown underneath in parentheses) for LD-CEV variance process based on Duan and Yeh (2010) (DY method), Durham’s (2013) SE method, and our CSE method. The other parameters are $q(t) = -0.0788$, $r(t) = \lambda_j = \Lambda_j = 0$, $S_0 = 100$, $V_0 = 0.04$, and $N = 2,520$. 

24
Table 3. Estimation results for NLD-CEV model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>SE method</th>
<th>CSE method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.6</td>
<td>0.1101</td>
<td>0.5156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0043)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-10</td>
<td>-0.4663</td>
<td>-8.1547</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0020)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>30</td>
<td>0.0602</td>
<td>23.1205</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0389)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>-50</td>
<td>-46.7257</td>
<td>-42.9582</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.3032)</td>
<td>(0.0020)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-2.5</td>
<td>-1.6282</td>
<td>-2.8565</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0888)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-20</td>
<td>-13.6425</td>
<td>-7.1706</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.5294)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>-2.0863</td>
<td>-2.0066</td>
<td>-1.8115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0359)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>1.3873</td>
<td>1.9167</td>
<td>1.7101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0817)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.8936</td>
<td>1.1806</td>
<td>0.9829</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0316)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.6916</td>
<td>0.0134</td>
<td>-0.6044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0012)</td>
<td>(0.0001)</td>
</tr>
</tbody>
</table>

This table reports the estimated results and their standard errors (shown underneath in parentheses) for NLD-CEV variance process based on Durham’s (2013) SE and our CSE methods. The other parameters are $q(t) = -0.0788$, $r(t) = \Lambda_j = \lambda_j = 0$, $S_0 = 100$, $V_0 = 0.04$, and $N = 2,520$. 

This table reports the estimated results and their standard errors (shown underneath in parentheses) for NLD-CEV variance process based on Durham’s (2013) SE and our CSE methods. The other parameters are $q(t) = -0.0788$, $r(t) = \Lambda_j = \lambda_j = 0$, $S_0 = 100$, $V_0 = 0.04$, and $N = 2,520$. 


Table 4. Descriptive statistics for S&P 500 return and VIX index

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P 500 index return</th>
<th>VIX index (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0003</td>
<td>19.7750</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0114</td>
<td>7.8479</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.2806</td>
<td>2.1006</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>11.4522</td>
<td>10.7307</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1042</td>
<td>80.8600</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.0947</td>
<td>9.3100</td>
</tr>
</tbody>
</table>

The data sample covers the daily prices of the S&P 500 index and the VIX index over the period from January 2, 1990 to June 30, 2016. All data are collected from CBOE website.
Figure 1. The S&P 500 index and VIX index from January 2, 1990 to June 30, 2016
Figure 2. VIX-volatility relationships of our CSE method, Durham’s (2013) SE method, and Duan and Yeh’s (2010) analytic approach for the LD-CEV stochastic volatility process.
Figure 3. Inferred stock return innovations for the CSE-High and CSE-Low cases under the LD-CEV stochastic volatility process.
Figure 4. VIX-volatility relationships between our CSE method and Durham’s (2013) SE method for the NLD-CEV stochastic volatility process.
Figure 5. Inferred stock return innovations for the CSE-High and CSE-Low cases under the NLD-CEV stochastic volatility process.