A new look at the continuous-time dynamics of international volatility indices amid the recent market turmoil

August 21, 2009

ABSTRACT

Volatility indices have been designed for many markets as gauges to measure investors' fear of market crash. The recent market turmoil has produced historically high volatility levels, in some cases around four times higher than their previous average levels. These levels are not anticipated by continuous-time models calibrated using old market data. We take a new look at the behavior of various volatility indices by including the recent market turmoil into the data. We re-estimate various continuous-time models for different volatility indices with focus on the structure of the drift and diffusion functions. Two methodologies are utilized: the maximum likelihood estimation, and Aït-Sahalia's parametric specification test. While the results from the parametric specification test advocate strongly for specifying more flexible drift and diffusion functions, nonlinear drift structure often only adds negligible benefit in terms of the likelihood function value. Our results call for caution against finite sample bias when adopting a particular model or fixing a particular parameter vector.

Keywords: Volatility Indices, Continuous-time Dynamics, Maximum Likelihood Estimation, Parametric Specification Test

JEL Classification: C60, G12, G13

I. Introduction

While the modern world financial markets have evolved for at least a few hundreds of years, the markets are still far from being complete. A complete market means that investors are able to hedge against any kinds of risk, such as traditional market risk, liquidity risk and credit risk, as well as nontraditional ones such as labor income risk, longevity risk, climate risk, etc. The process of financial engineering during the past half century can be thought of as an explosive effort by investors of trying to complete market.[1] A very significant boost in the process of completion is in 1973, when CBOE introduced first batch of exchange-traded stock options. Following that, there have been a plethora of new financial products both on exchanges and over the counter. These include single stock futures, credit derivatives, weather derivatives, freight derivatives, housing derivatives, economic derivatives, leveraged ETFs, catastrophe bonds, survival derivatives, etc. Merton Miller asked: "(c)an any twenty-year period in recorded history have witnessed even a tenth as much financial innovation?" The astonishing rapid pace of financial innovation has met with mixed reactions. While Warren Buffet famously labeled derivatives as "financial weapons of massive destruction," former Fed chair Allen Greenspan spoke highly of derivatives to a Senate Banking Committee in 2003: "What we have found over the years in the marketplace is that derivatives have been an extraordinarily useful vehicle to transfer risk from those who shouldn't be taking it to those who are willing to and are capable of doing so." The recent market turmoil has added fuel to some people's resentment to derivatives, a common phrase one often reads is "Math + Market = Mayhem." Nonetheless, the consensus of the market is that derivatives are going to stay and to develop further. In this regard, the recent market turmoil might turn out to be a good breathing point for financial engineers and regulators to take a critical look at the past achievements and the challenges ahead.

Adding to the arsenal of financial innovations is the introduction of volatility derivatives, such as volatility futures, variance swaps, volatility options, etc. These products allow investor to hedge against and speculate on pure volatility risks. Most of the time, these products are built upon some volatility indices, the most well-known one being the CBOE VIX index introduced in 1993. This index has been mimicked by many other markets both inside the U.S. and outside. For example, CBOE also computes the volatility index VXN for Nasdaq 100 (NDX) and the index VXD for Dow Jones Industrial Average (DJIA). Outside the U.S., there is a growing list of volatility indices being introduced. For example, Deutsche Börse back calculated volatility index

[1] Of course, financial engineering serves some other purposes too, such as minimizing transaction costs, alleviating agency problems or information asymmetry, etc. For articles on financial innovation, see Miller (1986, 1991, 1992) and Tufano (2003).

VDAX (as well as VDAX-NEW) based on the DAX from year 1992, while MONEP publishes VX1, VX6, and other related volatility indices based on CAC 40 since 1997. Two other examples are the VSTOXX and VSMI which will be studied in detail in this paper.

Derivative products are intrinsically related to volatility. Recently, direct bets on volatility have been introduced both on exchanges and over the counter. For example, the CBOE started trading VIX futures in March 2004. These products are welcome by market participants because they not only allow one to direct bet on the movement of volatility, but also allow one to isolate and hedge the volatility component in other derivative products. See Grünbichler and Longstaff (1996) on pricing volatility derivatives, Brenner, Ou, and Zhang (2006) on hedging volatility risk, and Windcliff, Forsyth and Vetzal (2006) on both. To reliably price volatility products and to calculate the optimal hedge positions, a good understanding of the dynamics of the volatility indices is essential. There are a few choices researchers can make. One could use a discrete-time process such as GARCH to model the volatility indices, or use a continuous-time process. In this paper, we focus on the second approach. Two earlier research papers are worth mentioning. In Bakshi, Ju, and Ou-Yang (2006), the authors employ maximum likelihood estimation to render a rank-ordering of various continuous-time models. A single volatility index (VXO) is used in this study.[2] Their results show that there is substantiate variance dynamics with nonlinear mean-reversion. Also, their results support the presence of a nonlinear diffusion coefficient structure. The combined specification of nonlinear drift and diffusion provides superior performance relative to its nested variants. Dotsis, Psychoyios, and Skiadopoulos (2007) further consider the addition of jumps in modeling the volatility dynamics. In addition to VIX, they also consider the VXO index, as well as several European volatility indices. However, because of the limitation of tractability, they have to restrict themselves to affine drift functions. This allows them to use a maximum likelihood estimation technique. With this restriction, their results show that it is necessary to add jumps to capture the evolution of implied volatility indices while (linear) mean-reversion is of second-order importance.

In this paper, we reexamine the continuous-time dynamics of various volatility indices. We have several motivations. First, most studies on volatility dynamics are completed before year 2007 and miss the recent market turmoil. Bakshi, Ju, and Ou-Yang (2006) use a sample period from July, 1, 1988 to January 10, 2000. Dotsis, Psychoyios, and Skiadopoulos (2007) use a sample period from October 14, 1997 to March 24, 2004. The year 2007 marks the dramatic downturn

[2] CBOE changed its methodology of computing VIX in 2003 and renamed the old VIX index to VXO. My private communications with the authors clarifies that the VIX index referred to in Bakshi, Ju, and Ou-Yang (2006) is actually the VXO index.

of the world financial markets and the subsequent dramatic increase of market volatility in 2008. For example, the maximum VIX level for the sample period in the study of Dotsis, Psychoyios, and Skiadopoulos (2007) is a little over 40 while it reached above 80 in 2008. This is largely unexpected from the estimations in previous studies. For example, using the parameter estimates from Bakshi, Ju, and Ou-Yang (2006), the probability that VIX reaches such a high level is extremely low. Thus, it is useful to reexamine the volatility dynamics and to understand what have changed and what have not. For example, is it still important to incorporate nonlinearity in the volatility dynamics? Second, most previous studies have focused on the maximum likelihood estimation technique. Li (2009) studies diffusion processes whose diffusion functions are damped. Through Monte Carlo simulation, the author shows that the maximum likelihood estimation is often susceptible to finite sample bias. For example, it often misidentifies the damped diffusion with nonlinear drift. Thus, in addition to maximum likelihood estimation, we also use a parametric specification test developed in Aït-Sahalia (1996). This test is based on matching the invariant density of the volatility process with the one implied by the data. Thus, one can roughly view this estimation as matching various sample moments altogether at the same time, somewhat similar to the method of generalized moments. Loosely speaking, the maximum likelihood estimation cares more about the local behavior such as the transition probability next step, while the parametric specification test cares more about global behavior. With large sample size, the two estimation methods have to give similar results because both are consistent. Thus, it is useful to see whether these two estimation techniques give similar results on the volatility dynamics. Conflicting results indicate the presence of finite sample bias and calls for caution when one wants to adopt a particular model. Third, we use two European volatility indices not studied in Bakshi, Ju, and Ou-Yang (2006) and Dotsis, Psychoyios, and Skiadopoulos (2007), namely, the VSTOXX and VSMI indices. In addition, we also briefly look at the volatility term structure by examining a constant one-year maturity VIX-1Y index. Finally, as a minor difference, we study the volatility process directly instead of its squared variance process.

Several important results come out of our study. First, we perform two identical maximum likelihood estimation on the VIX index with and without including the two most recent years. We find that the inclusion of the recent market turmoil has a much larger effect on the estimated parameters in the drift function than those in the diffusion function. With the inclusion of the recent data, the mean-reversion in the drift function becomes weaker and less significant. However, the diffusion function is almost unchanged. Second, our maximum likelihood estimation shows that nonlinearity parameters in the drift function are often insignificant or only marginally significant if we include most recent data in the analysis. Judging from the likelihood ratios, it often suggests not including the nonlinear parameters. Also, our results show that with the exception of the VSMI index, a CEV-type diffusion function is sufficient. This result is similar to the one obtained in Durham (2003) for interest rates, where he find that it is often optimal to specify a constant drift function if one's objective is to maximize the likelihood function. Our final result from a parametric specification test, however, contradicts the result obtained from the maximum likelihood estimation. The estimation unanimously rejects all models with only linear drift, and for the VSTOXX and VSMI indices, it even rejects the nonlinear drift and nonlinear diffusion model and calls for more complicated specifications. The reason seems to be that the observed data suggests a bimodal or multi-modal invariant density which a simple parametric specification is not able to deliver.

These results continue the debate on whether the existence or nonexistence of nonlinearity in the drift function estimated from various different methods is due to finite sample bias. A short list on papers examining this issue includes Aït-Sahalia (1996, 1999), Chapman and Pearson (2000), Durham (2003), Jones (2003), Li, Pearson and Poteshman (2004), and Takamizawa (2008). That maximum likelihood and parametric specification tests yield conflicting conclusions with regard to the inclusion of nonlinearity in the drift function indicates the existence of finite-sample bias in either or both of these two methods. Without additional data, it is difficult to draw the final conclusion on this issue. A joint estimation together with observations on VIX derivative prices might offer more insights on this issue, although the need to specify volatility risk-premium might make this estimation difficult.

The paper is organized as follows. Section II describes the data we use. Section III studies the continuous dynamics of various indices using the maximum likelihood estimation technique. Section IV studies the continuous dynamics of various indices using a parametric specification test. Section V concludes.

II. Data description

The data employed in this paper include time series for four volatility indices and three stock market indices. The four volatility indices are denoted as VIX, VIX-1Y, VSTOXX and VSMI. The first two volatility indices are for the U.S. market and cover both short-term (30-day) and long-term (1-year) horizons. The last two indices are for the European markets. The selection of these four indices allows us to examine both short-term and long-term volatility indices, and both domestic and foreign markets.

The VIX index is the volatility index introduced by the Chicago Board Options Exchange in 1993 based on the paper by Whaley (1993). See also Whaley (2000). The original volatility index (now with ticker symbol VXO) was based on model-specific implied volatilities of S&P 100 index options, while the new VIX index is a model independent measure of future volatility based on S&P 500 index options. Both can be regarded as measures of market's expectation of the stock market volatility with a 30-day constant maturity. A good analysis of the theoretical underpinnings of the two indices is carried out in Carr and Wu (2006). According to CBOE, "(s)ince its introduction in 1993, VIX has been considered by many to be the world's premier barometer of investor sentiment and market volatility." Since the VIX index is back calculated to year 1990, we use the time period January 2, 1990 to April 15, 2009, a total of 4859 daily observations.

Our second volatility index VIX-1Y is a volatility index very similar to VIX but with a constant horizon of one year. It is based on the VIX term structure data published by CBOE, which date back to year 1992. On each business day, the VIX term structure data consist of about 8 implied volatility indices with maturities up to about three years. Figure 1 plots the volatility term structure as a function of the calendar time. A few things are easy to notice. First, there is a dramatic increase in volatility levels during the recent years. Second, for different calendar days, the volatility term structure could be either downward-sloping or upward-sloping. The downward-sloping feature is particularly prominent during the recent market turmoil. Finally, the short-term volatility indices seem to be more volatile than long-term volatilities, although there is still considerable variability for long-term volatilities.

Since the maturities in Figure 1 change day by day and are not constants, we construct a constant one-year maturity volatility index (VIX-1Y) based on the same methodology that CBOE uses during their final step of computing the VIX index. Specifically, we do an interpolation of the total variance on each day using the two maturities nearest to one year to get the VIX-1Y. The time-period of VIX-1Y thus covers from January 2, 1992 to April 15, 2009, a total of 4342 daily observations. We use this VIX-1Y index as a proxy of the future volatility level implied by the market for one-year horizon.

Our last two indices cover the European markets. The VSTOXX index is based on Dow Jones EURO STOXX 50 option prices and is designed to reflect the market expectations of near-term (30-day) volatility of the Eurozone markets. The Dow Jones EURO STOXX 50 Index is a blue-chip index in the Eurozone covering Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain. The VSMI index is a similar volatility index based on the SMI index of the Six Swiss Exchange. The SMI index is Switzerland's key equity index representing 20 blue-chips. Both of these two indices start on January 4, 1999, with 2617 daily observations for VSTOXX and 2618 daily observations for VSMI.

Table 1 presents the summary statistics of the four volatility indices as well as their first-order differences. The means of the four indices are all very close to 20, except that the VSTOXX index has a mean of 25. This is a little bit surprising given that the STOXX index is for 50 blue-chip companies while the S&P 500 is a much broad market index. The VSTOXX index also behaves somewhat differently from the other three in terms of median and mode. For the maximum, the VIX-1Y index has a considerably smaller maximum of 55.53 while all the other three indices have a maximum larger than 80. The standard deviations of the indices are much larger than those reported in other studies using data before 2008, for example, Dotsis, Psychoyios, and Skiadopoulos (2007). All four indices have positive skewness and somewhat large kurtosis. For the first-order differences, the means are all close to 0 with standard deviations from 0.70 to 1.84. Judging from the relatively small standard deviations, the minimum and maximum values of the changes are huge in magnitude. For example, for the changes in VSTOXX, the highest one-day negative change is 13.98, while the highest one-day positive change is 22.64. All these large changes are realized in years 2008 and 2009 amid the recent market turmoil. The skewnesses of changes of the four indices behave differently, with two of them positive and two of them negative. The most striking feature is that the changes of all four indices display huge kurtosis, with the kurtosis of ΔVIX-1Y being 78.44. The first-order autocorrelations of changes of the indices are all slightly negative, except for changes of the VSMI index, which has a positive autocorrelation of 0.10.

In addition to these four volatility indices, we also collect daily data for the underlying market indices, namely, the S&P 500 index in the U.S., the DJ STOXX index, and the SMI index. The S&P 500 index data is from Yahoo Finance historical data, while the other two indices are obtained from STOXX Ltd. and Six Swiss Exchange, respectively. Figures 2 to 4 plot the volatility indices together with their underlying market indices. All three market indices experienced an "M-shape" during the sample periods, with a strong run leading to the first peak around the middle of year 2000. Following the burst of the internet bubble, stock indices picked up again starting around the end of year 2002. However, starting from the middle of year 2007, we see a big continuous drop of the market indices as a result of the subprime mortgage crisis and the following credit crunch. Both of these two market downturns have had significant impact on the behavior of the volatility indices. For example, the VIX shoot up to around 80% during the recent market turmoil. In addition, the Asian currency crisis in 1997 and the Russian sovereign bond default in 1998 also had strong impacts on the VIX and VIX-1Y.

Table 2 reports the correlation structure of the volatility indices and their underlying market indices (Panel A), as well as that of their first-order differences (Panel B). As we see, all four volatility indices are strongly correlated, with the two European volatility indices most strongly correlated with a coefficient of 0.96. All the market indices are also strongly positively correlated. However, all volatility indices are negatively correlated with the market indices. This is the reason why investors have treated these volatility indices as the market "fear gauges" for possible crashes. The behavior for the first-order differences is similar, although the correlations of volatility index changes are much smaller than the correlations of the volatility indices themselves. The two European indices, however, are still strongly positively correlated with coefficient 0.7095.

III. Maximum Likelihood Estimation of Volatility Indices

We consider a one-dimensional diffusion process for a state variable X_t:

$$\mathrm{d}X_t = \mu(X_t;\theta)\mathrm{d}t + \sigma(X_t;\theta)\mathrm{d}W_t, \tag{1}$$

where $\mu(X_t;\theta)$ and $\sigma(X_t;\theta)$ are the drift and diffusion coefficients respectively, and θ is the parameter vector. Different from Bakshi, Ju, and Ou-Yang (2006), we use the state variable X_t to model the volatility indices themselves instead of the squared process. The values of the indices are divided by 100 so that they are expressed in percentage terms. We will often suppress the argument θ.

The processes we want to consider are the same ones as in Bakshi, Ju, and Ou-Yang (2006). They are a class of nested models with the following drift and diffusion functions proposed in Aït-Sahalia (1996):

$$\mu(x;\theta) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_{-1} x^{-1}, \tag{2}$$

$$\sigma(x;\theta) = \sqrt{\beta_1 x + \beta_2 x^{\beta^3}}. \tag{3}$$

Thus, for the full model, $\theta = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$. We have chosen not to use a constant term β_0 in $\sigma(x;\theta)$ following Bakshi, Ju, and Ou-Yang (2006). Similar to their findings, we find that β_0 is often not significant and does not affect the likelihood function much. Different models are different restrictions of the parameter vector θ as listed below:

$$\text{AFF}: \quad \alpha_2 = \alpha_{-1} = 0, \quad \beta_2 = 0, \ \beta_3 \text{ irrelevant} \tag{4}$$

$$\text{CEV-CD}: \quad \alpha_1 = \alpha_2 = \alpha_{-1} = 0, \quad \beta_1 = 0 \tag{5}$$

$$\text{CEV-LD}: \quad \alpha_2 = \alpha_{-1} = 0, \quad \beta_1 = 0 \tag{6}$$

$$\text{CEV-ND}: \quad \beta_1 = 0 \tag{7}$$

$$\text{SEV-CD}: \quad \alpha_1 = \alpha_2 = \alpha_{-1} = 0 \tag{8}$$

$$\text{SEV-LD}: \quad \alpha_2 = \alpha_{-1} = 0 \tag{9}$$

$$\text{SEV-ND}: \quad \text{no restrictions} \tag{10}$$

Here in naming the models, AFF stands for the "affine" model which is often referred to as the Cox-Ingersoll-Ross (CIR) model in finance. See Cox, Ingersoll and Ross (1985). The shorthand notations CD, LD, and ND refer to constant drift, linear drift, and nonlinear drift, respectively. The notations CEV and SEV refer to constant elasticity of variance and stochastic elasticity of variance, respectively.

It is important to notice that the class of models considered in Dotsis, Psychoyios, and Skiadopoulos (2007) and the class of models considered in this paper are not subsets of each other. Dotsis, Psychoyios, and Skiadopoulos (2007) have focused more on adding jumps to the volatility dynamics. The cost is this makes them only able to examine a handful of special models with very special drift and diffusion functions. In this paper, we take Bakshi, Ju, and Ou-Yang (2006)'s approach by considering more flexible drift and diffusion functions at the cost of excluding the jump structure. As Dotsis, Psychoyios, and Skiadopoulos (2007) have rightly pointed out, while it would be useful to study jump diffusion models with flexible drift and diffusion functions, technically this is fairly difficult. Thus, in this paper we restrict ourselves with the class of models above.

Let $p_X(\Delta, x|x_0)$ be the transition density of the process X_t after time period Δ from the current level of x_0 to future level x. To use the maximum likelihood estimation, one needs the transition density $p_X(\Delta, x|x_0)$. Except for a few special cases, explicit expressions for the transition densities do not exist. Fortunately, Aït-Sahalia (1999, 2002) has designed efficient and accurate analytical approximations. The method is based on a series expansion method by treating Δ as small. More specifically, it first develops the part of the transition density that is singular in Δ and then uses a Taylor series expansion for the analytical part. The expansion coefficients were given iteratively in the original papers of Aït-Sahalia. The following proposition in Li (2009) simplifies the expansion by solving the iterations to the second order.[3] A brief proof is contained in Li (2009).

Proposition 1 *Consider the process* $\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$ *with* $\sigma(\cdot) > 0$ *except possibly at the boundaries. Define* $\hat{\mu}(x) \equiv \frac{\mu(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x)$ *and* $\lambda(x) \equiv -\frac{1}{2}(\hat{\mu}^2(x) + \hat{\mu}'(x)\sigma(x))$.
1. The approximate transition density $p_X^{(K)}(\Delta, x|x_0)$ *for the process* X_t *to order* $K = 2$ *in* Δ *is given by*

$$p_X^{(2)}(\Delta, x|x_0) = p_X^{(0)}(\Delta, x|x_0)(1 + c_1(x|x_0)\Delta + c_2(x|x_0)\Delta^2/2), \tag{11}$$

where

$$p_X^{(0)}(\Delta, x|x_0) = \frac{1}{\sqrt{2\pi\Delta}}\sqrt{\frac{\sigma(x_0)}{\sigma^3(x)}}\exp\left(\int_{x_0}^{x}\frac{\mu(u)}{\sigma^2(u)}\mathrm{d}u - \frac{1}{2\Delta}\left(\int_{x_0}^{x}\frac{1}{\sigma(u)}\mathrm{d}u\right)^2\right), \tag{12}$$

and for $x \neq x_0$,

$$c_1(x|x_0) = \frac{\int_{x_0}^{x}\lambda(u)/\sigma(u)\mathrm{d}u}{\int_{x_0}^{x}1/\sigma(u)\mathrm{d}u}, \qquad c_2(x|x_0) = c_1(x|x_0)^2 + \frac{\lambda(x) + \lambda(x_0) - 2c_1(x|x_0)}{\left(\int_{x_0}^{x}1/\sigma(u)\mathrm{d}u\right)^2}, \tag{13}$$

and for $x = x_0$, $c_1(x_0|x_0) = \lambda(x_0)$ *and* $c_2(x_0|x_0) = \lambda(x_0)^2 + (\sigma(\sigma\lambda')')(x_0)/6$.
2. The series approximation of $\log p_X(\Delta, x|x_0)$ *to order* $K = 2$ *in* Δ *is given by*

$$\log p_X^{(2)}(\Delta, x|x_0) = \log p_X^{(0)}(\Delta, x|x_0) + C_1(x|x_0)\Delta + \frac{1}{2}C_2(x|x_0)\Delta^2, \tag{14}$$

where $C_1(x|x_0) = c_1(x|x_0)$ *and* $C_2(x|x_0) = c_2(x|x_0) - c_1(x|x_0)^2$. *Writing out explicitly, we have*

$$C_1(x|x_0) = \frac{\int_{x_0}^{x}\lambda(u)/\sigma(u)du}{\int_{x_0}^{x}1/\sigma(u)du}, \quad C_2(x|x_0) = \frac{\lambda(x) + \lambda(x_0) - 2c_1(x|x_0)}{\left(\int_{x_0}^{x}1/\sigma(u)du\right)^2}. \tag{15}$$

The above proposition allows the fast computation of the transition densities. To implement the second-order approximation, one just needs to compute three integrals, namely, $\int \mu/\sigma^2$, $\int \lambda/\sigma$, and $\int 1/\sigma$. For many models, these integrals can be computed explicitly. For example, for the AFF model $\mathrm{d}X_t = (\alpha_0 + \alpha_1 X_t)\mathrm{d}t + \sqrt{\beta_1 X_t}\mathrm{d}W_t$ above, applying the above proposition

[3] Bakshi, Ju, and Ou-Yang (2006) solve the iteration to the fourth order for diffusion processes with unit diffusion functions. However, we use a second-order expansion because it is already extremely accurate. Another paper worth mentioning is Bakshi and Ju (2005).

gives the following second-order approximation:

$$p_X^{(0)}(\Delta, x|x_0) = \frac{1}{\sqrt{2\pi\Delta\beta_1 x}} \left(\frac{x}{x_0}\right)^{\alpha_0/\beta_1 - 1/4} \exp\left(\frac{\alpha_1(x - x_0)}{\beta_1} - \frac{2(\sqrt{x} - \sqrt{x_0})^2}{\beta_1\Delta}\right), \tag{16}$$

$$C_1(x|x_0) = -\frac{\alpha_1^2\left(6\alpha_0/\alpha_1 + x + \sqrt{xx_0} + x_0\right)}{6\beta_1} + \frac{16\alpha_0 - 3\beta_1}{32\sqrt{xx_0}} - \frac{\alpha_0^2}{2\beta_1\sqrt{xx_0}}, \tag{17}$$

$$C_2(x|x_0) = \frac{48\alpha_0\beta_1 - 9{\beta_1}^2 - 16\left(3\alpha_0^2 + \alpha_1^2 xx_0\right)}{384xx_0}. \tag{18}$$

In situations where explicit expressions are not available, these three integrals can be very efficiently computed using numerical integration such as quadrature method. This is the approach taken in this paper. The accuracy of Aït-Shalia's density expansion approximation has been established and verified in many papers, including Jensen and Poulsen (2002), Bakshi, Ju, and Ou-Yang (2006), and Li (2009). In particular, the uniform error of a second-order approximation is usually in the order of 10^{-4}.

Given discretely observed data x_i for $i = 0, 1, \cdots, n$ with fixed intervals Δ, the maximum likelihood estimation searches the optimal parameter values by maximizing the following log-likelihood function

$$\max_\theta \mathcal{L}[\theta] = \sum_{i=1}^{n} \log p_X(\Delta, x_i|x_{i-1}). \tag{19}$$

We use a simplex method for this maximization. This is a slow but very robust optimization algorithm. For an introduction of this method, see Press, Flannery, Teukolsky, and Vetterling (1992).

For all the four time series (VIX, VIX-1Y, VSTOXX and VSMI) considered in this paper, we collect data until April 15, 2009. Maximum likelihood estimation is performed on all four indices. For the VIX index, we also consider two different sample periods. The first sample period is from January 2, 1990 to December 31, 2007, while the second one covers the whole sample period, namely, from January 2, 1990 to April 15, 2009. This allows us to get a rough understanding of how the results change, if any, by including the most recent two years in the data. During the most recent two years, the market volatility level has been much higher than previous years. It is useful to know whether this has any effect on the estimated drift and diffusion functions. For example, does the inclusion of the most recent data affect the nonlinearity in the drift functions? Does the diffusion function (volatility of volatility index) become more volatile if we include the most recent data?

Table 3 reports the results of the maximum likelihood estimation on the VIX index for the shorter sample period. There are a total of 927 weekly observations (Wednesdays) from January

2, 1990 to December 31, 2007. We use weekly observations to save computational time and our results do not change much if we use daily observations. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L}-d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order two nested models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' easy reference, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

From Table 3 we see that the full model (SEV-ND) has the largest log-likelihood value, 2393.84. However, the LR test statistic between the CEV-ND and SEV-ND models is only 0.006, which is smaller than 3.84, the critical value of $\chi^2(1)$. Thus, the result favors CEV-ND over SEV-ND. Similar results hold for the LR statistics between the SEV-LD and CEV-LD pair, and the SEV-CD and CEV-CD pair. These results are consistent with the fact that β_1 are insignificant in all three SEV models judging from the large asymptotic standard errors. The AFF model is clearly inadequate if we look at the LR statistic between the AFF and SEV-LD models. Among the CEV models, the LR statistic between the CEV-CD and CEV-ND is 28.24 while the LR statistic between CEV-LD and CEV-ND is 15.78. Since both exceed the critical value of $\chi^2(3)$, the results favor the CEV-ND over either CEV-CD or CEV-LD. Judging from Akaike's information criterion, the CEV-ND model also has the lowest value. Therefore, the CEV-ND model seems to be the most favorable model among all seven models considered. The estimated drift function for the CEV-ND model is $\mu(x) = -10.0343 + 55.6598x - 99.4579x^2 + 0.5920/x$, with all four parameters statistically significant. Thus, similar to Bakshi, Ju, and Ou-Yang (2006), the nonlinearity in the drift function is important in capturing the strong mean-reverting behavior near both boundaries of the volatility process. On the other hand, a diffusion function with constant elasticity given by $\sigma(x) = \sqrt{2.0802x^{2.7365}}$ is sufficient.

The stationary density of the process X_t with drift and diffusion functions $\mu(x)$ and $\sigma(x)$ is given by (see Karlin and Taylor 1981)

$$\pi(x;\theta) = \frac{\xi(\theta)}{\sigma^2(x;\theta)} \exp\left(\int_0^x \frac{2\mu(u;\theta)}{\sigma^2(u;\theta)} \mathrm{d}u\right), \tag{20}$$

where $\xi(\theta)$ is a normalization constant so that $\pi(x;\theta)$ integrates to 1. For the CEV-ND model estimated in Table 3, under the stationary density, there is a 0.003 probability that the the volatility level is above 0.4, 5×10^{-7} probability above 0.6, and around 10^{-11} probability above 0.8. Thus, using purely data before year 2007, the probability that the process would ever reach a level of 0.8 is extremely low and the sudden increase in VIX in 2008 was unexpected at the end of 2007 purely from the VIX data.

Table 4 reports the results from the maximum likelihood estimation of the VIX for the full sample period from January 2, 1990 to April 15, 2009. The general pattern is very similar to that of Table 3. Again, the affine model AFF is insufficient. The SEV models all lose out to their CEV counterparts judging from the LR statistics. However, judging from the LR statistic between the CEV-LD and CEV-ND models, the CEV-LD model is weakly preferred over CEV-ND at 5% significance level. The LR statistic is $2 \times (2511.80 - 2509.07) = 5.46$ while the 5% critical value of $\chi^2(2)$ is 6.00. This is consistent with the fact that the nonlinear mean-reverting parameters α_2 and α_{-1} are not as significant as in Table 3. This behavior is different from the one implied by the shorter sample period ending in 2007, where the nonlinear mean-reverting parameters are very important. The strength of nonlinear mean-reverting in the CEV-ND model is also much weaker for the whole sample period compared with the shorter sample period. On the other hand, the CEV-LD model is quite strongly preferred over the CEV-CD model for the full sample period, indicating that although there is no strong evidence for nonlinear mean-reverting, linear mean-reverting is still very important in describing the data. However, the AIC gives a conflicting recommendation which ranks CEV-ND first, although the difference in AIC's is extremely small. These results together imply a marginal need for prescribing a nonlinear drift function. Nonetheless, by computing the LR statistics between the AFF model and its nesting SEV-LD model, we see that it is important to specify a more flexible diffusion function than the square-root function in the AFF model.

It is interesting to look at the stationary density with parameters estimated from the full sample period. Since the log-likelihood $\mathcal{L}$ and AIC give somewhat different rank-ordering, we consider both CEV-LD and CEV-ND models. For the CEV-LD model, under the stationary density, the probabilities that the volatility level is above 0.4, 0.6 and 0.8 are approximately 0.04, 0.01, and 0.004, respectively, while for the CEV-ND model, these three probabilities are 0.04, 0.003, and 0.0002, respectively. The corresponding empirical probabilities from the observed VIX time series are 0.032, 0.003, and 0.0002, respectively. Judging purely from these three tail probabilities, the CEV-ND seems to be doing a better job than the CEV-LD model.

Below we examine the results for the VIX-1Y, VSTOXX, and VSMI indices. Table 5 reports the results of the maximum likelihood estimation on the VIX-1Y index for the full sample period. There are a total of 891 weekly observations (Wednesdays) from January 2, 1992 to April 15,

2009. Overall, all models indicate a weaker mean-reversion strength and a smaller variance of volatility for the VIX-1Y index than the VIX. Same as the situation for the VIX and VIX-1Y indices, the AFF model is insufficient. This can be seen by computing the LR statistic between the AFF model and its nesting model SEV-LD. Also, all the SEV models again lose out to their CEV counterparts. Among the CEV models, the story is a little bit different from that of the VIX index. The CEV-CD model is weakly preferred over the CEV-LD and CEV-ND models judging from their respective LR statistics. However, the AIC criterion ranks CEV-ND first.

Table 6 reports the results of the maximum likelihood estimation on the VSTOXX index. There are a total of 528 weekly observations (Wednesdays) from January 4, 1994 to April 15, 2009. The AFF model is still insufficient. The SEV models still lose out to there CEV counterparts. However, judging from both the LR statistics and the AIC criterion, the CEV-CD model is preferred over the other two CEV models.

Table 7 reports the results of the maximum likelihood estimation on the VSMI index. There are a total of 528 weekly observations (Wednesdays) from January 4, 1994 to April 15, 2009. The AFF model is still insufficient. However, now the parameter β_1 in the diffusion function becomes very significant for all the SEV models. As a result, the SEV models now outperform their CEV counterparts. Among the SEV models, the SEV-CD model is preferred over the other two models judging from the LR statistics. However, the AIC criterion ranks SEV-LD first.

In summary, while there are individual differences for different indices, the maximum likelihood estimation for all four indices suggests that there is no need or only a marginal need to prescribe a nonlinear drift function. This result is somewhat similar to Durham (2003), where the author finds that to model the short-term interest rate, adding additional flexibility to the drift function beyond a constant term provides negligible benefit in terms of the value of the likelihood function. See also Li (2009) for a discussion. On the other hand, it is important to specify a more flexible diffusion function than that of the AFF model. Except for the VSMI index, a CEV diffusion function is sufficient for all other three indices.

IV. Specification Tests of Volatility Models

The maximum likelihood estimation above shows that it is only marginally useful to stipulate a nonlinear drift function if we include the more recent data into the analysis. However, as Li (2009) shows, maximum likelihood estimation is subject to finite sample bias, especially when applied to parametric models. For example, Monte Carlo experiment in Li (2009) shows that when the true data-generating model is linear drift but with damped diffusion function, maximum likelihood estimation tends to mistake the damped diffusion with nonlinear drift function when the time series is of finite sample size. One possible indication of finite sample bias is that the tail probabilities under the stationary density implied by the estimated parameters deviate significantly from those implied by the sample data. Therefore, in the following, we take a different approach, namely, the specification test for diffusion processes originally proposed in Aït-Sahalia (1996). This method was also used in Li (2009) to detect the existence of a damped diffusion function.

The idea of the approach is fairly intuitive. For a given parameter vector, we could compute the theoretical stationary density (or the marginal density as it was referred to as in Aït-Sahalia (1996)) as given by equation (20). At the optimal parameter values, this theoretical density should be close to the stationary density implied by the data if the data is truly generated by the proposed model. Thus, the difference of these two densities could be used as a criterion for estimating the parameters. The actual procedure involves a few steps, as we now describe below. We do not repeat all the details below such as the assumptions on the processes, the optimal choice of the bandwidth parameters, etc. Readers are referred to the original paper by Aït-Sahalia (1996).

1. Estimate the empirical stationary density $\hat{\pi}(X_i)$ nonparametrically for each realized value X_i in the data. More specifically, we use

$$\hat{\pi}(X_i) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{h_N}K\left(\frac{X_i - X_j}{h_N}\right), \tag{21}$$

   where $K(x) = \exp(-x^2/2)/\sqrt{2\pi}$ is the normal density kernel function, and h_N is the optimally chosen bandwidth depending on the sample size N.[4]

2. For each of the models we consider in equations (4) to (10), estimate the vector θ by minimizing the statistic

$$\hat{M} = \hat{M}(\theta) \equiv \frac{1}{N}\sum_{i=1}^{N}\hat{M}_i \tag{22}$$

   where

$$\hat{M}_i \equiv N h_N \big(\pi(X_i;\theta) - \hat{\pi}(X_i)\big)^2, \tag{23}$$

[4]More information on the kernel estimation technique can be found in Silverman (1986) and Wand and Jones (1994).

and $\pi(x,\theta)$ is the model implied marginal density given by equation (20). Let the true marginal density of X be $\pi_0(X)$ and let $\mathbb{E}$ denote the expectation under this true marginal density. Then $\hat{M}$ is the empirical counterpart of the following distance measure

$$M \equiv \min_{\theta} \mathbb{E}\Big[\big(\pi(X;\theta) - \pi_0(X)\big)^2\Big]. \tag{24}$$

3. To test whether a particular parametric model should be accepted or rejected, we compare $\hat{M}$ with the critical value c_α. It is shown in Aït-Sahalia (1996b) that the test statistic $\hat{M}$ is asymptotically distributed as

$$h_N^{-1/2}\big(\hat{M} - E_M\big) \xrightarrow{d} N(0, V_M), \tag{25}$$

where

$$E_M \equiv \frac{1}{2\sqrt{\pi}} \int_0^\infty \pi_0^2(x)\ \mathrm{d}x, \tag{26}$$

$$V_M \equiv \frac{1}{\sqrt{2\pi}} \int_0^\infty \pi_0^4(x)\ \mathrm{d}x. \tag{27}$$

That is, at a given significance level α, a particular parametric model cannot describe data well if $\hat{M} \geq c_\alpha \equiv \hat{E}_M + h_N^{1/2} z_\alpha \hat{V}_M^{1/2}$. If $\alpha = 5\%$, then $z_\alpha \approx 1.645$.[5]

In doing Step 1 above, it is useful to take a look at the histograms of the daily observations of the four volatility indices. This is plotted in Figure 5. For each of the four indices, daily observations from the full sample period are used. Therefore, the numbers of observations are 4859, 4342, 2617, and 2618 for the VIX, VIX-1Y, VSTOXX and VSMI indices, respectively. For each of the indices, the kernel estimated density $\hat{\pi}(X_i)$ can be thought of as a kernel smoothing of the histogram of the daily observations after a normalization. It seems that the empirical densities for all four indices are multi-modal with quite heavy right tails. This is especially true for the VIX-1Y, VSTOXX and VSMI indices, where the multimodality feature seems to be very prominent. It is also very important to notice that somewhat contrary to people's common reaction, this multimodality feature is not at all due to the recent market turmoil. A similar graph plotted using data before year 2007 or before year 2008 yields quantitatively very similar histograms. The recent market turmoil only makes the right tail heavier and does not really contribute to the multimodality feature.

In our actual implementation, we will fix the values of α_0 for all the models considered and all four indices. The fixed values of α_0 are those in Tables 4 to 7. The reason for doing

[5]Notice that there is a typo in equation (14) of Aït-Sahalia (1996).

this is as follows. A careful analysis of the implied marginal density $\pi(X;\theta)$ in equation (20) reveals that if we quadruple the drift function $\mu(X;\theta)$ and double the diffusion function $\sigma(X;\theta)$, then the marginal density does not change for any value of X. Therefore, depending on the particular functional forms of $\mu(X;\theta)$ and $\sigma(X;\theta)$, there could be an infinite set of parameters which minimize $\hat{M}$. For example, in the AFF model, without restricting the values of some parameters, if $(\alpha_0,\, \alpha_1, \beta_1)$ minimizes $\hat{M}$, then $(4\alpha_0, 4\alpha_1, 4\beta_1)$ would also minimize $\hat{M}$. Therefore, there is a need to fix some of the parameters in the drift or diffusion in order to get rid of the indeterminateness. This identification issue is also explained in Li (2009). It is important to understand that this indeterminateness does not in any way affect the validity of Aït-Sahalia (1996)'s specification test.

We now take a look at the results from the specification tests. While we examine all four indices, we will focus more on the VIX index. Table 8 presents the results from the specification tests of the continuous-time parametric models for the VIX index for the sample period from January 2, 1990 to April 15, 2009. It reports the estimated parameters, the test statistic $\hat{M}$, and the corresponding critical value c_α. All seven models are rejected by the specification tests, except for the CEV-ND and SEV-ND models. Also, the drift parameters for each of the models are quite close to the ones in Table 4 from the maximum likelihood estimation. However, for the five rejected CEV and SEV models, the parameters for the diffusion functions are quite different from those in Table 4. This reflects the fact that for these models, the optimization could not find a parameter vector that can fit the empirical marginal density. For these rejected models, the parameters estimated from the maximum likelihood function will give even larger values of $\hat{M}$, indicating that models without nonlinear drift are not able to match the theoretical and empirical stationary densities, even though the previous maximum likelihood estimation implies that nonlinear drift function is only of marginal importance. Since the two approaches give somewhat different results, a modeler has to exert caution if he/she has to choose between a model with linear drift and another one with nonlinear drift. Model risk has to be taken into account if these models in turn are used to price volatility derivatives.

Table 8 makes it clear that the nonlinear drift function is extremely important for the theoretical marginal density to match its empirical counterpart. The drift and diffusion functions from the AFF and SEV-ND models are plotted in Figures 6 and 7. The drift and diffusion functions for the other accepted model, CEV-ND, are very similar to those of the SEV-ND model and omitted. As we see in Figure 6, while the drift functions from the two models are very close in the volatility index range from 0.1 to 0.35, the drift function of the SEV-ND model is much larger for lower VIX levels and much more negative for higher VIX levels, resulting in stronger mean-reversion near the boundaries. For the diffusion function, Figure 7 shows that the diffusion function grows much faster in the SEV-ND model than in the AFF model, resulting in larger volatility when the volatility index level is high.

To get a better idea of how the parametric specification test works, it is useful to take a look at the empirical marginal density and the model implied densities from the optimized parameter vectors in Table 8. Two models are selected for this purpose, namely, the AFF model and the SEV-ND model. Figure 8 plots the goodness of fit for the AFF model. The empirical stationary density estimated from nonparametric kernel estimation (in gray) as well as the theoretical densities (in black) computed from the optimized parameter vector are plotted on the same graph. As we see, the AFF model is rejected by the specification test because the distance between these two densities is fairly large for much of the range of index values. Figure 9 plots the same two densities but now with the SEV-ND model. It is clear that the SEV-ND model fits better than the AFF model, and the test statistics shows that the SEV-ND model cannot be rejected at 5% significance level.

We now take a look at the specification test results for the other three volatility indices, namely, VIX-1Y, VSTOXX and VSMI. Table 9 presents the results from the specification tests of the continuous-time parametric models for the VIX-1Y index for the sample period from January 2, 1992 to April 15, 2009. As we see, all seven models are rejected by the specification tests, except for the nonlinear drift and stochastic variance (SEV-ND) model. Comparing with the parameters in Table 5 from the maximum likelihood estimation, we see that the drift parameters have not changed much for the SEV-ND model, but the point estimates of the diffusion parameters have changed quite a lot. Graphic analysis shows that the diffusion function from the specification test is much larger for volatility values above 0.4 than that from the maximum likelihood estimation. This is not that surprising as a large diffusion function is needed in order to generate high volatility values so as to match the empirical marginal density.

Table 10 presents the results from the specification tests of the continuous-time parametric models for the VSTOXX index for the sample period from January 4, 1999 to April 15, 2009. All seven models are now strongly rejected by the specification tests. The inability to match the empirical distribution of the VSTOXX index is clearly shown in Figures 10 and 11 for the AFF and SEV-ND models, respectively. As we see, the empirical distribution is strongly multimodal. At the optimized parameter vectors, the theoretical marginal densities from both the AFF and SEV-ND models are unimodal and cannot closely match the empirical distribution.

Finally, Table 11 presents the results from the specification tests of the continuous-time parametric models for the VSMI index for the sample period from January 4, 1999 to April

15, 2009. Similar to the results for the VSTOXX index, all seven models are rejected by the specification tests. Again, graphic analysis shows that none of the models are able to match the multimodal empirical density.

To summarize, the results of the parametric specification test advocate strongly for specifying more flexible drift and diffusion functions. In particular, a nonlinear drift structure is a minimum requirement in order for the theoretical marginal densities to possibly match the empirical densities. This result is at odds with that from the maximum likelihood function, where we see that nonlinear drift structure often only adds negligible benefit in terms of the likelihood function value. We take this as a strong indication of finite sample bias since both methods are asymptotically consistent. Li (2009) discusses the finite sample bias of the maximum likelihood estimation in detail. Our results call for caution against finite sample bias when adopting a particular model or fixing a particular parameter vector in modeling the dynamics of volatility indices. Methods such as Bayesian analysis with suitable priors or robust estimation with suitable criteria are probably useful in guiding the model and parameter section processes.

V. Conclusion

In this paper, we take a new look at the continuous-time dynamics of four international volatility indices: VIX, VIX-1Y, VSTOXX, and VSMI. Two methodologies are taken, namely the maximum likelihood estimation and the parametric specification test. With more recent data, the results from maximum likelihood estimation show that there is a marginal need to specify a nonlinear drift. However, the parametric specification test rejects all models with linear drift functions. That these two methodologies give somewhat conflicting results about the drift function indicate the presence of finite sample bias and the need to exert caution when a modeler has to make a choice among alternative models.

Several future (and probably difficult) directions might be taken. So far we have only considered diffusion models with no jumps. It is particularly interesting to examine whether the inclusion of jumps will get rid of the need of nonlinear drift functions in matching the theoretical and empirical densities. For example, with very restricted drift and diffusion functions, Dotsis, Psychoyios, and Skiadopoulos (2007) show that it is important to add jumps. Does this result still hold if we allow for more flexible drift and diffusion functions? To satisfactorily address this question requires the calculation of the theoretical densities for diffusion models with jumps, which at current stage poses technical difficulties. Another direction one can take is to estimate the volatility index dynamics jointly with the underlying index. This is appealing because it is well-known that the volatility index and its underlying stock index are negatively correlated. Unfortunately, this direction is much more difficult than it seems because the volatility index dynamics does not evolve freely, but rather it is dictated through a complicated mechanism by the stock index process under the physical measure as well as the unobserved market price of risk process. Estimating the stock index and volatility index under the physical measure jointly might not be consistent because there might not exist reasonable market price of risk process connecting the dynamics of stock and volatility indices. Finally, it is possible that the same finite sample bias is present in Dotsis, Psychoyios, and Skiadopoulos's (2007) analysis on jump diffusion models. One way to examine this is to extend Aït-Sahalia's (1996) parametric specification test to jump diffusion models, which is unfortunately quite difficult technically. We hope that these issues could be address in future research.

References

Aït-Sahalia, Y., 1996. Testing continuous-time models of the spot interest rate. *Review of Financial Studies* 9, 385–426.

Aït-Sahalia, Y., 1999. Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance* 54, 1361–1395.

Aït-Sahalia, Y., 2002. Maximum-likelihood estimation of discretely-sampled diffusions: A closed-form approximation approach. *Econometrica* 70, 223–262.

Bakshi, G., Ju N., 2005. A refinement to Aït-Sahalia's (2002) "maximum likelihood estimation of discretely sampled diffusions: A closed form approximation approach". *Journal of Business* 78, 2037–2052.

Bakshi, G., Ju N., and Ou-Yang H., 2006. Estimation of continuous-time models with an application to equity volatility dynamics. *Journal of Financial Economics* 82, 227–249.

Brenner, M., Ou, E.Y., Zhang, J.E., 2006. Hedging volatility risk. *Journal of Banking and Finance* 30, 811–821.

Carr, P., Wu L., 2006. A tale of two indices. *Journal of Derivatives*, Spring, 13–29.

Chapman, D., Pearson, N.D., 2000. Is the short rate drift actually nonlinear? *Journal of Finance* 55, 355–399.

Cox, J.C., Ingersoll J.E. Jr., and Ross S.A., 1985. A theory of the term structure of interest rates. *Econometrica*, Vol. 53, No. 2, 385–407.

Dotsis, G., Psychoyios D., and Skiadopoulos G., 2007. An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking and Finance* 31, 3584–3603.

Durham, G.B., 2003. Likelihood-based specification analysis of continuous-time models of the short-term interest rate. *Journal of Financial Economics* 70, 463–487.

Grünbichler, A., Longstaff, F., 1996. Valuing futures and options on volatility. *Journal of Banking and Finance* 20, 985–1001.

Jensen, B., Poulsen R., 2002. Transition densities of diffusion processes: Numerical comparison of approximation techniques. *Journal of Derivatives* 9, 18–32.

Jones, C.S., 2003. Nonlinear mean reversion in the short-term interest rate. *Review of Financial Studies* 16, 793–843.

Karlin, S., and Taylor H.M., 1981. A Second Course in Stochastic Processes. Academic Press, New York.

Li, M., Pearson N.D., and Poteshman A.M., 2004. Conditional estimation of diffusion processes. *Journal of Financial Economics* 74, 31–66.

Miller, M. H., 1986. Financial innovation: The last twenty years and the next. *Journal of Financial and Quantitative Analysis* 21(4), 459–471.

Miller, M.H., 1991. Financial Innovation and Market Volatility. Blackwell, Cambridge, MA.

Miller, M.H., 1992. Financial innovation: Achievements and prospects. *Journal of Applied Corporate Finance* 4, 4–11.

Press, W. H., Flannery B.P., Teukolsky S.A., Vetterling W.T., 1992. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, Cambridge.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC.

Takamizawa, H., 2008. Is nonlinear drift implied by the short end of the term structure? *Review of Financial Studies* 21, 311–346.

Tufano, P., 2003. Financial Innovation, in *Handbook of the Economics of Finance* (Volume 1a: Corporate Finance), edited by G. Constantinides, M. Harris and Rene Stulz, Elsevier, 307–336.

Wand, M., Jones C., 1994. Kernel Smoothing. Chapman & Hall/CRC.

Whaley, R.E., 1993. Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives* 1, 71–84.

Whaley, R.E., 2000. The investor fear gauge. *Journal of Portfolio Management* 26, 12–17.

Windcliff, H., Forsyth, P.A., Vetzal, K.R., 2006. Pricing methods and hedging strategies for volatility derivatives. *Journal of Banking and Finance* 30, 409–431.

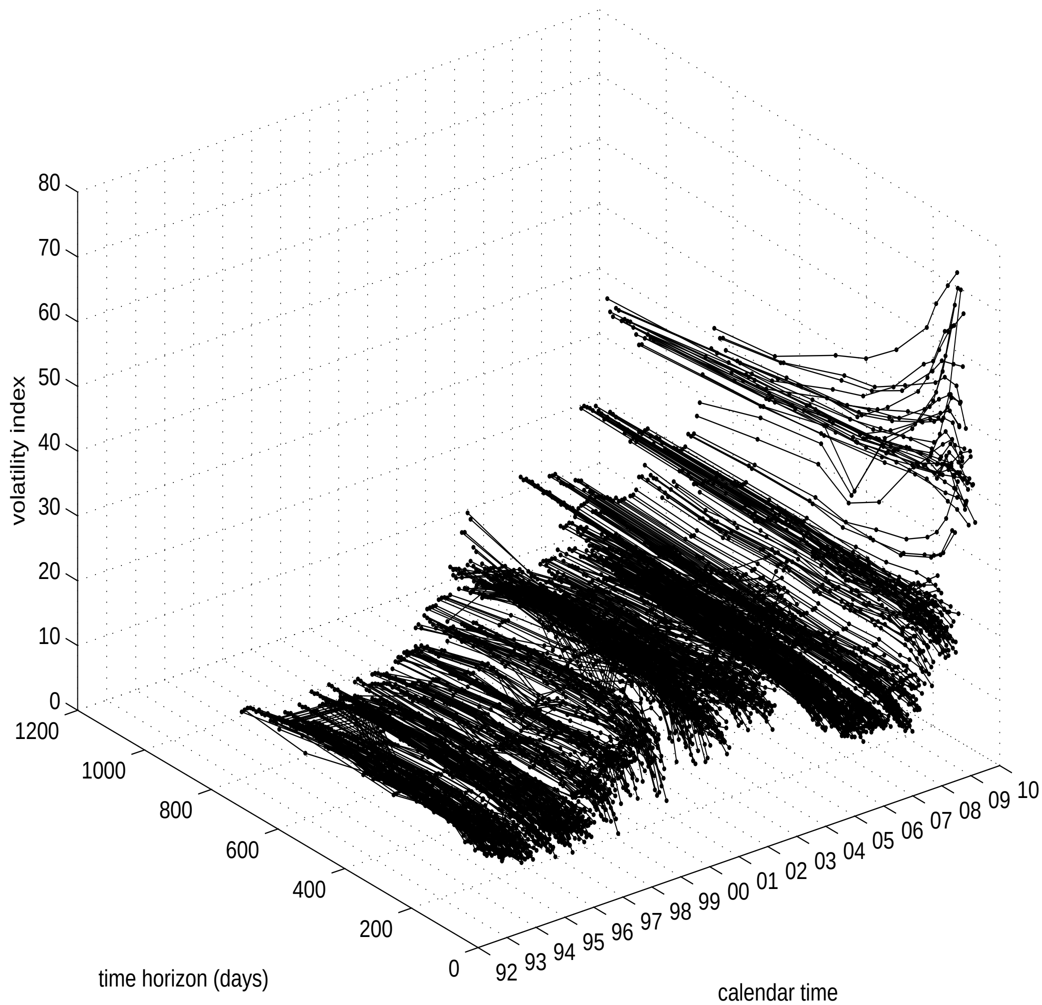

Figure 1. The CBOE volatility indices term structure. There is a dramatic increase in volatility levels during the recent years. For different calendar days, the volatility term structure could be either downward-sloping or upward-sloping. The short-term volatility indices seem to be more volatile than long-term volatilities, although there is still considerable variability for long-term volatilities.

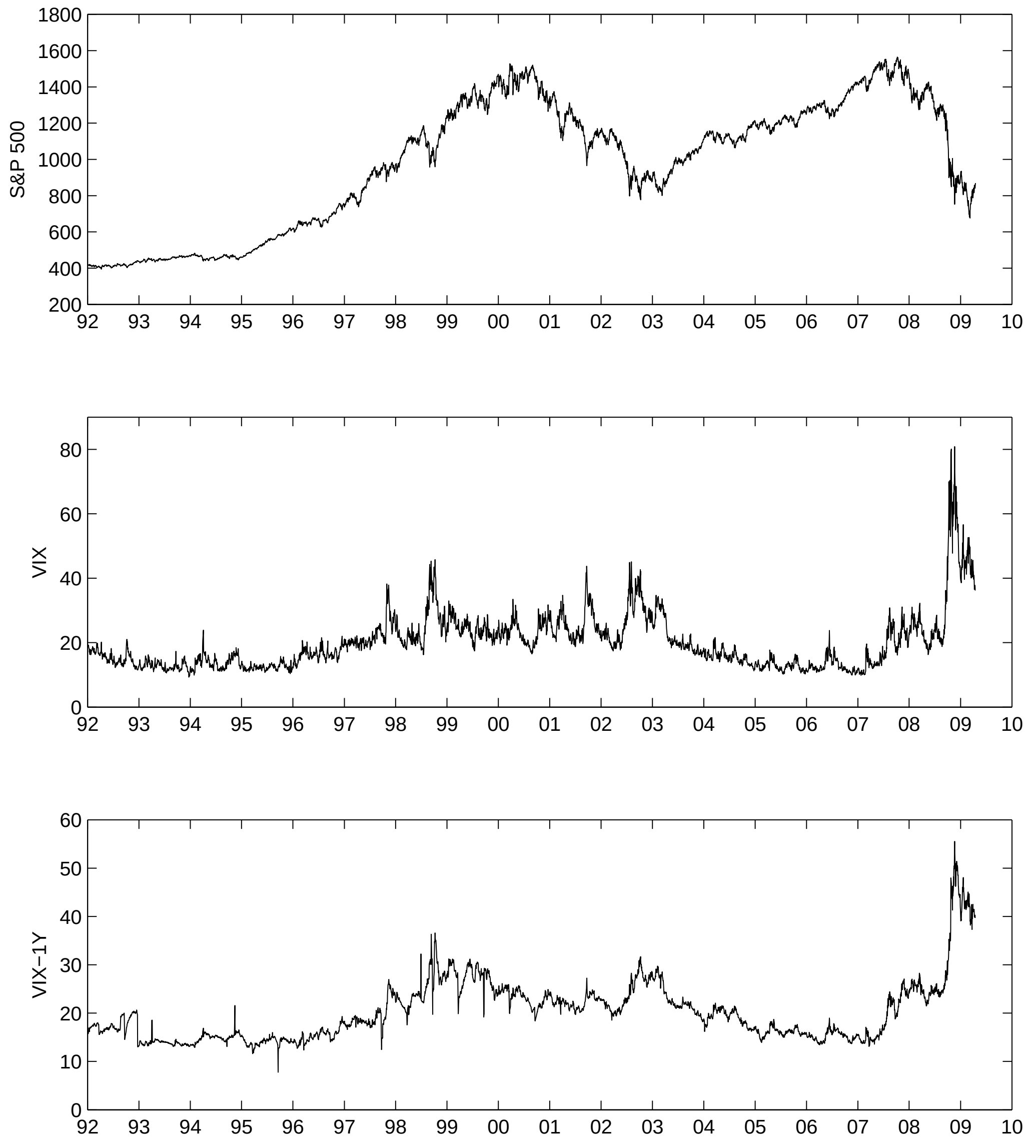

Figure 2. The S&P 500 index, the CBOE volatility index VIX, and the VIX-1Y index. The index experienced a few downturns during this sample period, including the Asian current crisis, the Russian default, the internet bubble burst, and more recently the subprime mortgage crisis. All these events have led the volatility indices to surge, with the recent market turmoil having the strongest impact.

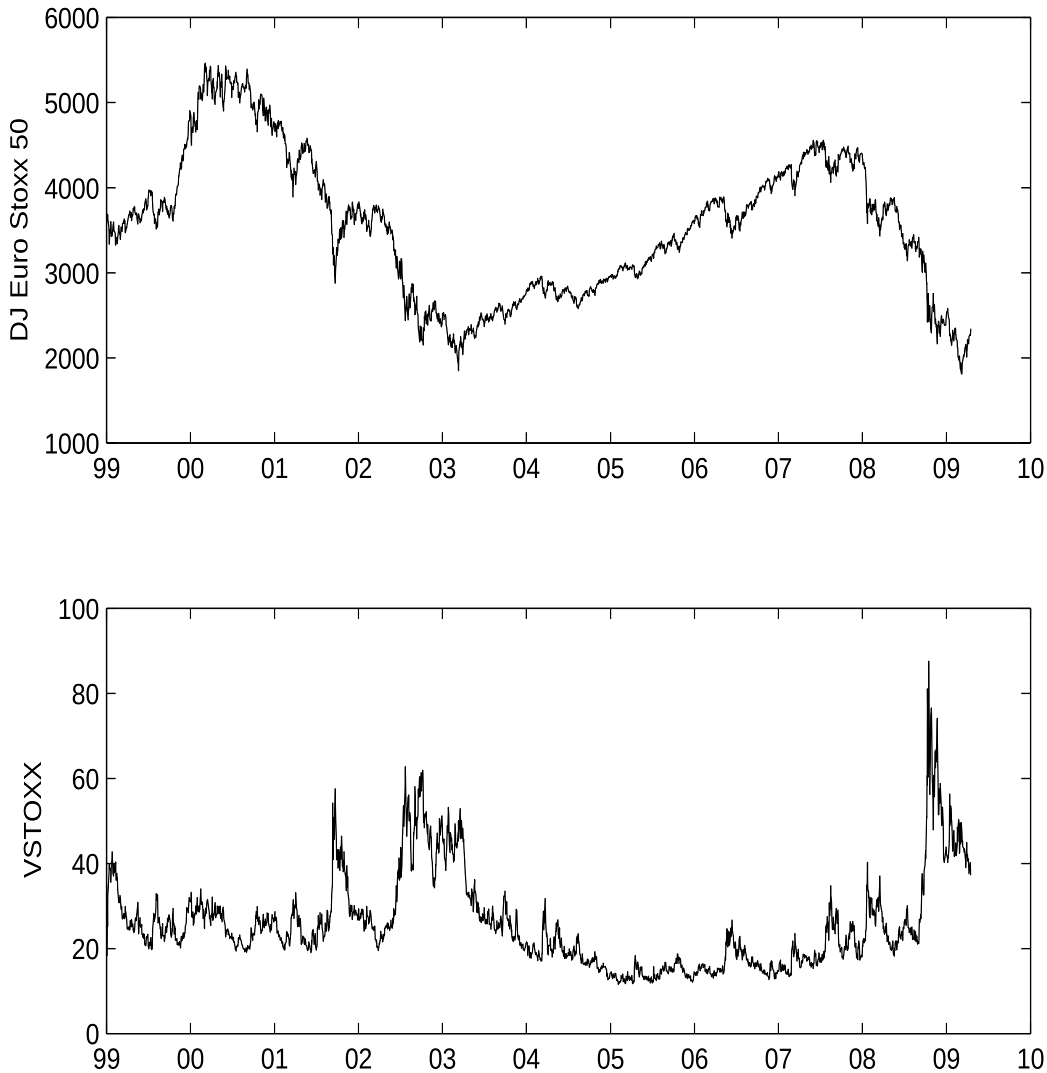

Figure 3. The Dow Jones Euro STOXX 50 index and the volatility index VSTOXX. The index experienced two major downturns during this sample period, namely, the internet bubble burst and more recently the subprime mortgage crisis. Both events have led the volatility indices to surge, especially for the recent market turmoil.

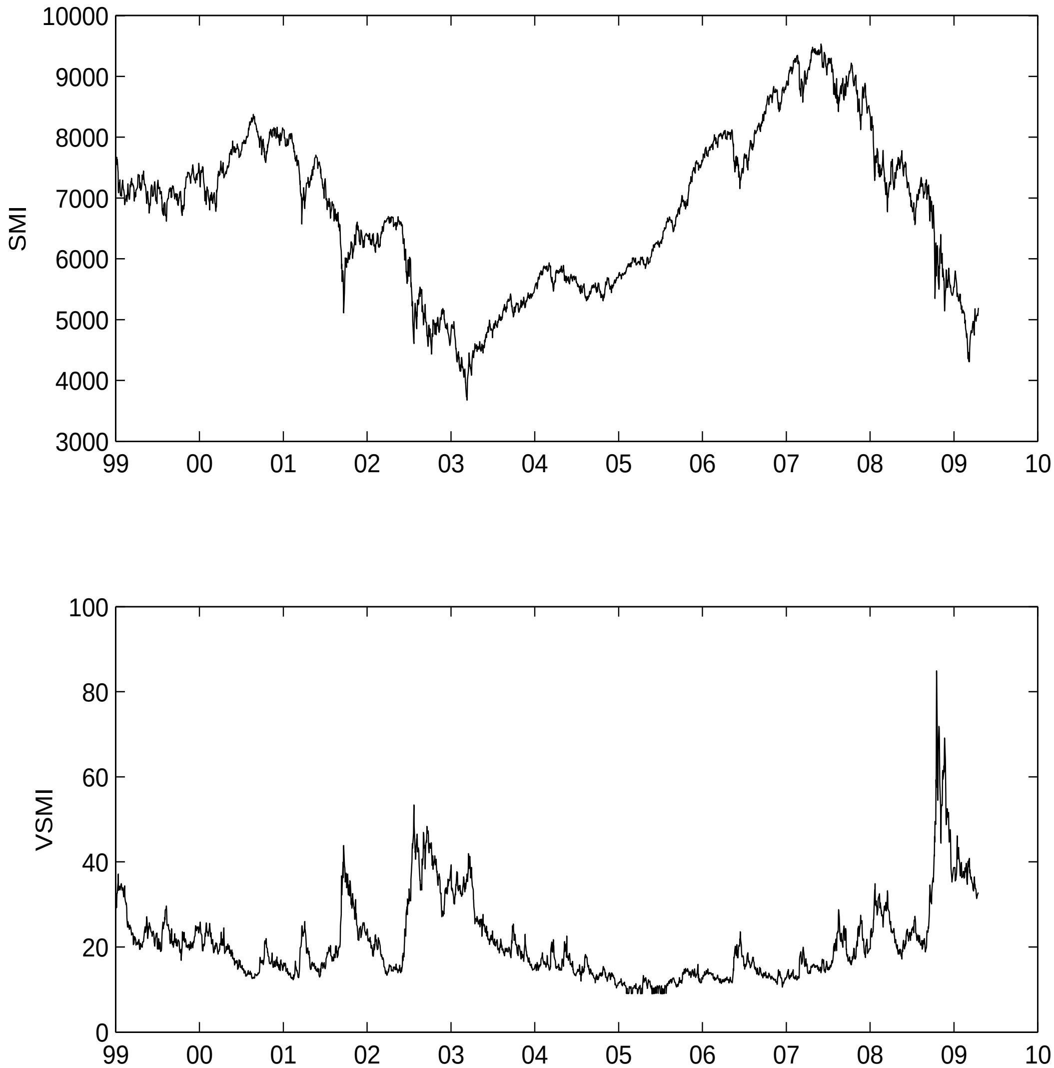

Figure 4. The Swiss market index SMI and the volatility index VSMI. The index experienced two major downturns during this sample period, namely, the internet bubble burst and more recently the subprime mortgage crisis. Both events have led the volatility indices to surge, especially for the recent market turmoil.

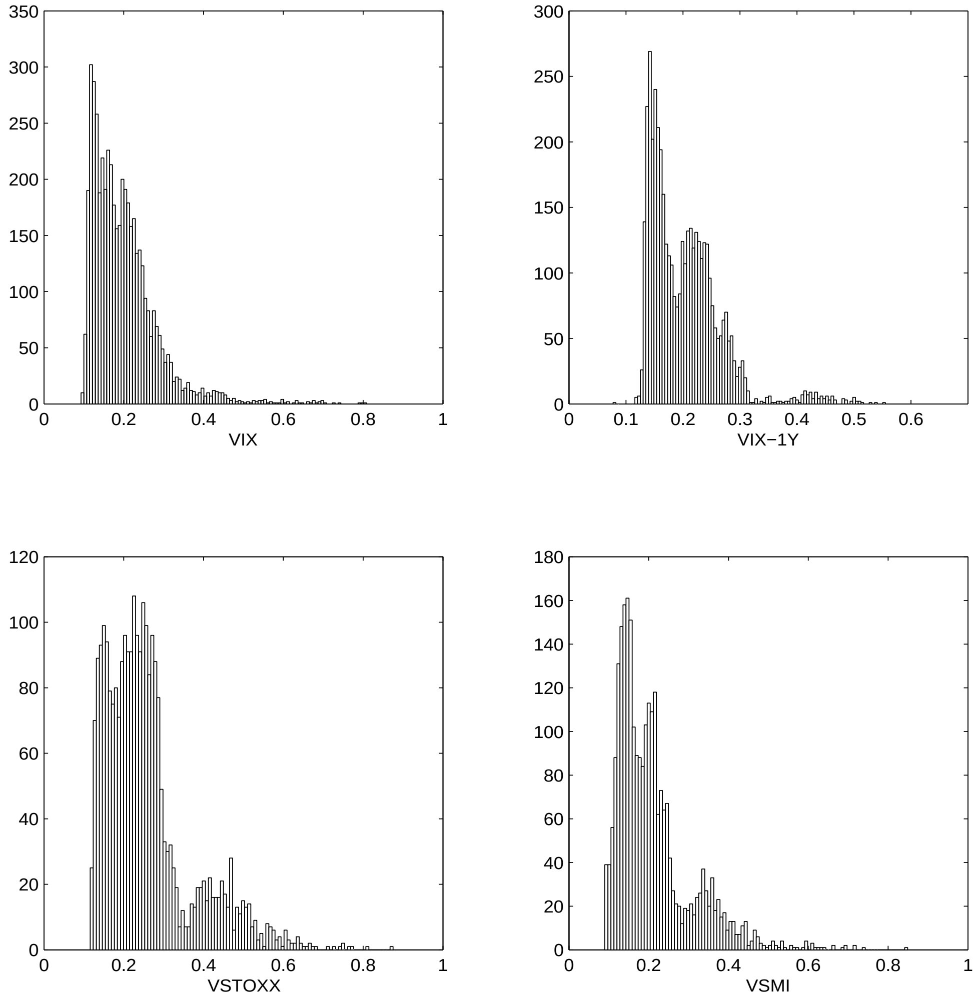

Figure 5. Histograms of the daily-observed volatility indices. The numbers of observations are 4859, 4342, 2617, and 2618 for the VIX, VIX-1Y, VSTOXX and VSMI indices, respectively. The empirical densities for all four indices are multi-modal with quite heavy right tails. This is especially true for the VIX-1Y, VSTOXX and VSMI indices, where the multimodality feature seems to be very prominent. Separate analysis reveals that the recent market turmoil only makes the right tail heavier and is not the source for the multimodality feature.

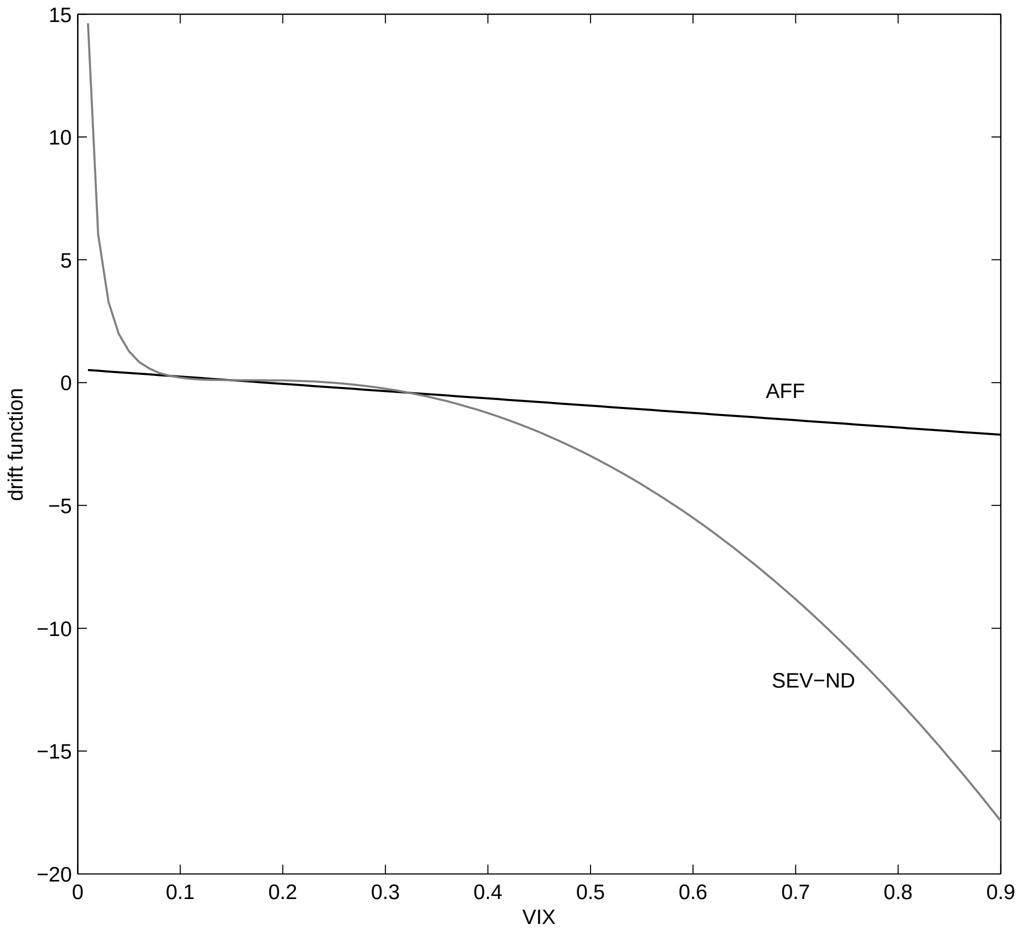

Figure 6. The drift functions in the AFF and SEV-ND models for the VIX index estimated from the parametric specification test. The optimized parameters are in Table 8. The drift function for the CEV-ND model (not plotted) is very similar to that of the SEV-ND model.

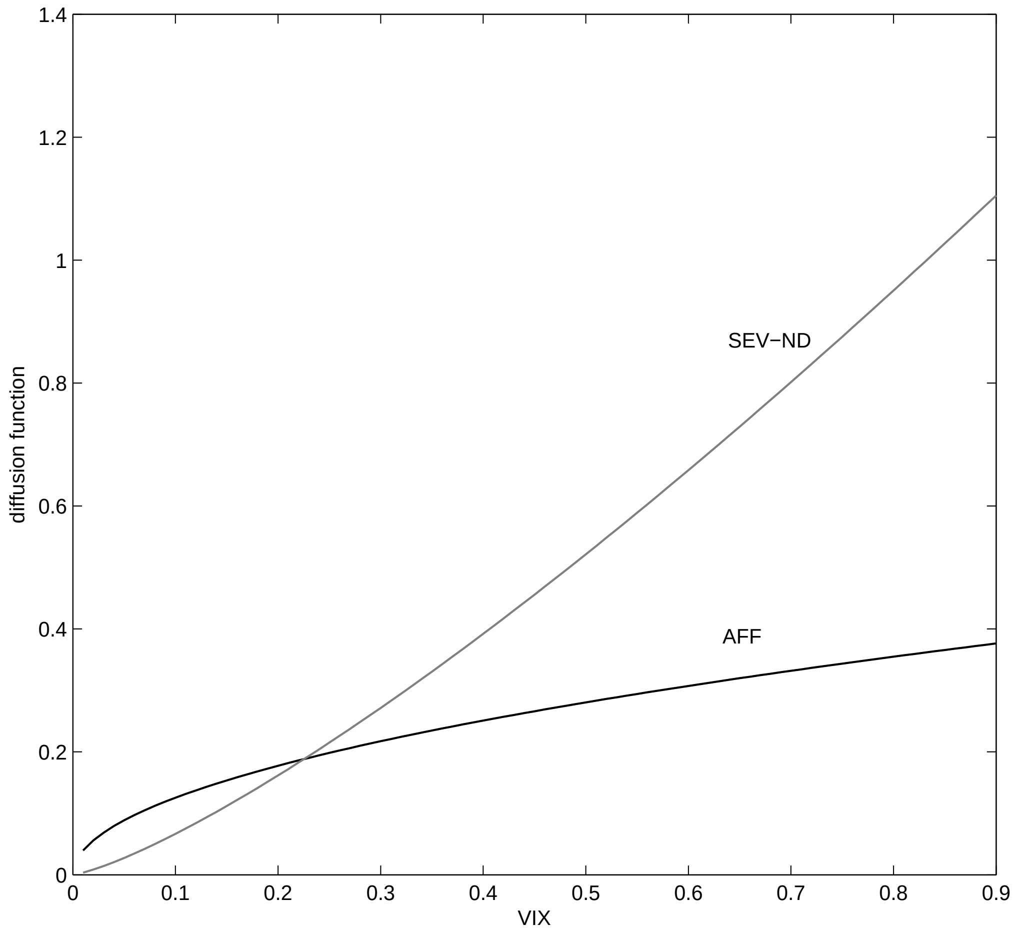

Figure 7. The diffusion functions in the AFF and SEV-ND models for the VIX index estimated from the parametric specification test. The optimized parameters are in Table 8. The diffusion function for the CEV-ND model (not plotted) is very similar to that of the SEV-ND model.

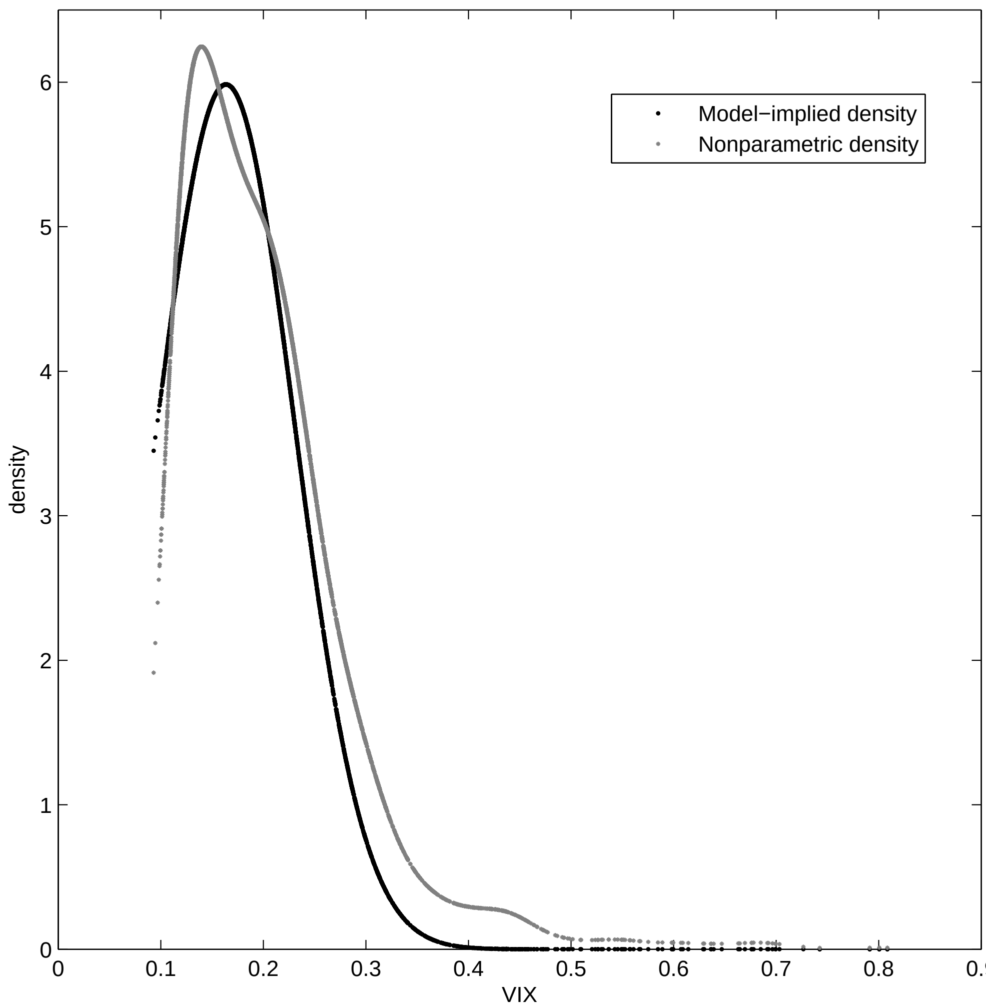

Figure 8. Goodness of fit of the AFF model on the VIX index by Aït-Sahalia's parametric specification test. The empirical stationary density estimated from nonparametric kernel estimation (in gray) as well as the theoretical densities (in black) computed from the optimized parameter vector are plotted. The AFF model is rejected by the specification test because the distance between these two densities is fairly large.

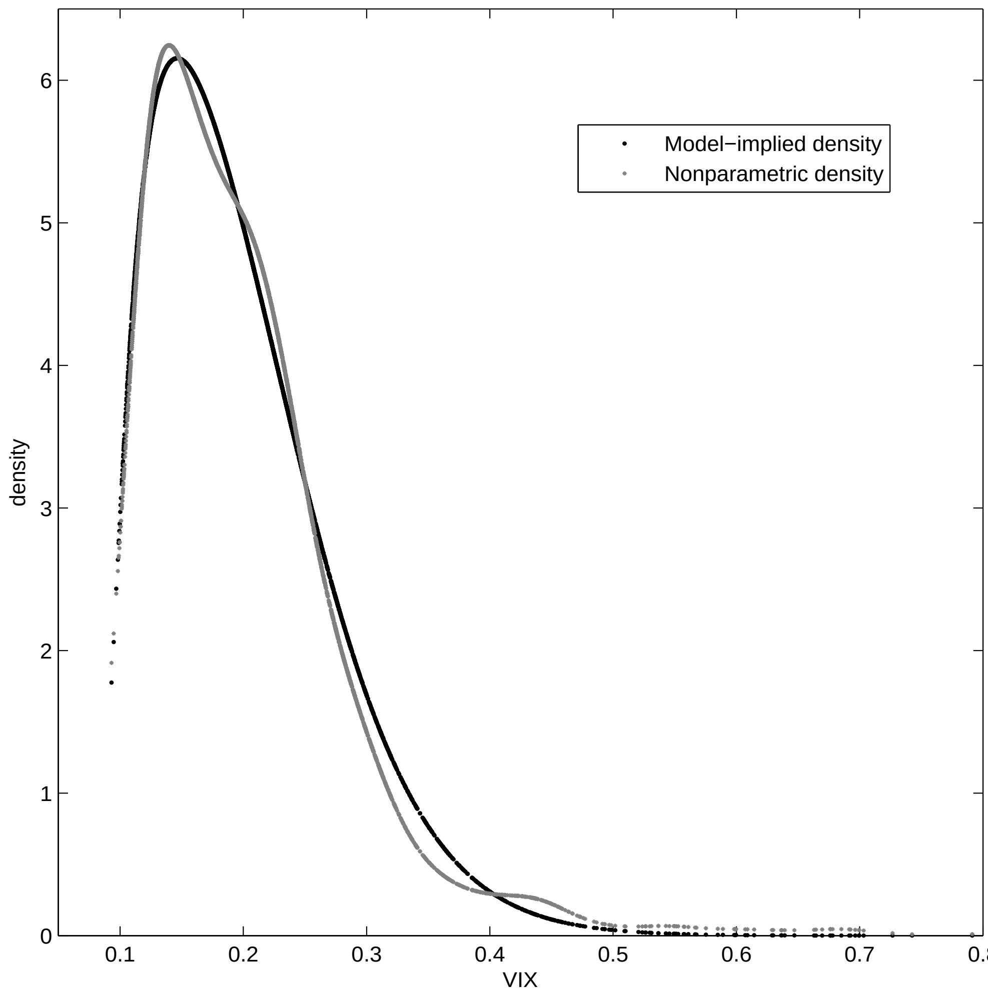

Figure 9. Goodness of fit of the SEV-ND model on the VIX index by Aït-Sahalia's parametric specification test. The empirical stationary density estimated from nonparametric kernel estimation (in gray) as well as the theoretical densities (in black) computed from the optimized parameter vector are plotted. The SEV-ND model cannot be rejected at 5% significance level by the specification test because the distance between these two densities is smaller than the critical value.

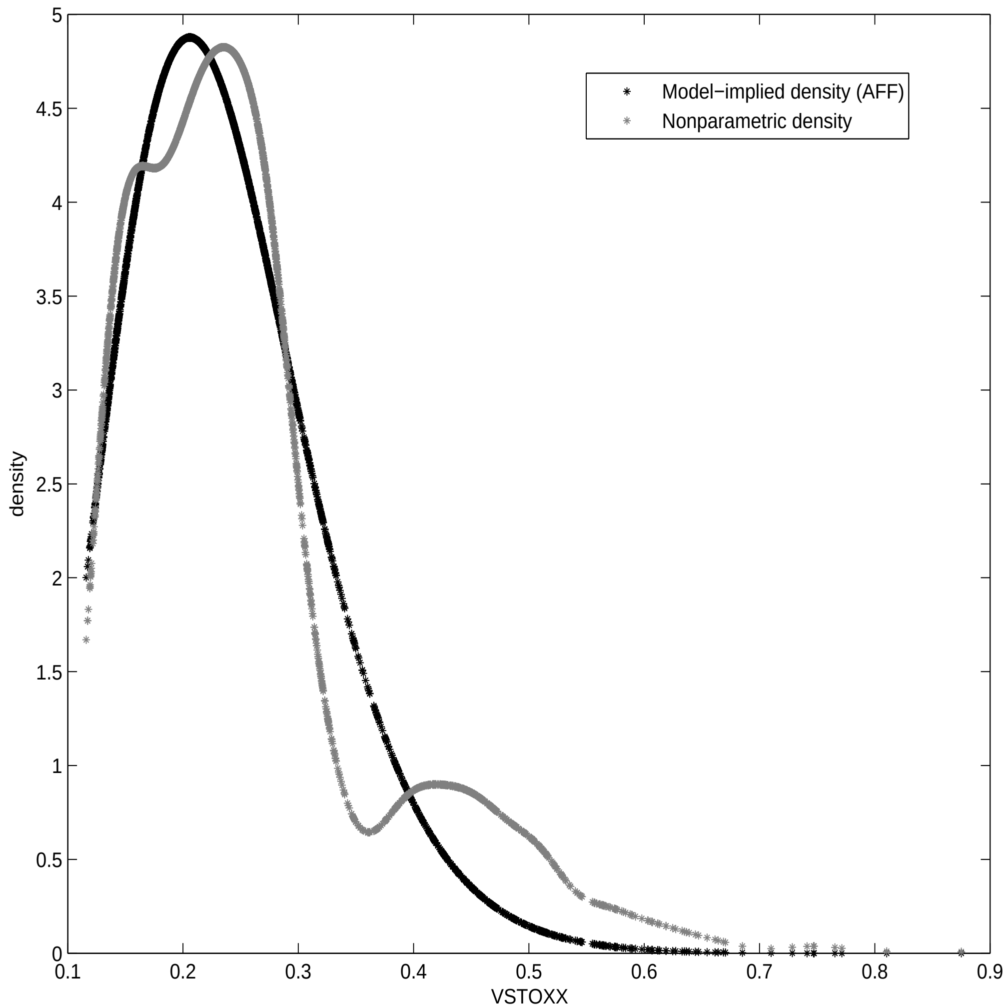

Figure 10. Goodness of fit of the AFF model on the VSTOXX index by Aït-Sahalia's parametric specification test. The empirical stationary density estimated from nonparametric kernel estimation (in gray) as well as the theoretical densities (in black) computed from the optimized parameter vector are plotted. The AFF model is rejected by the specification test because the distance between these two densities is fairly large.

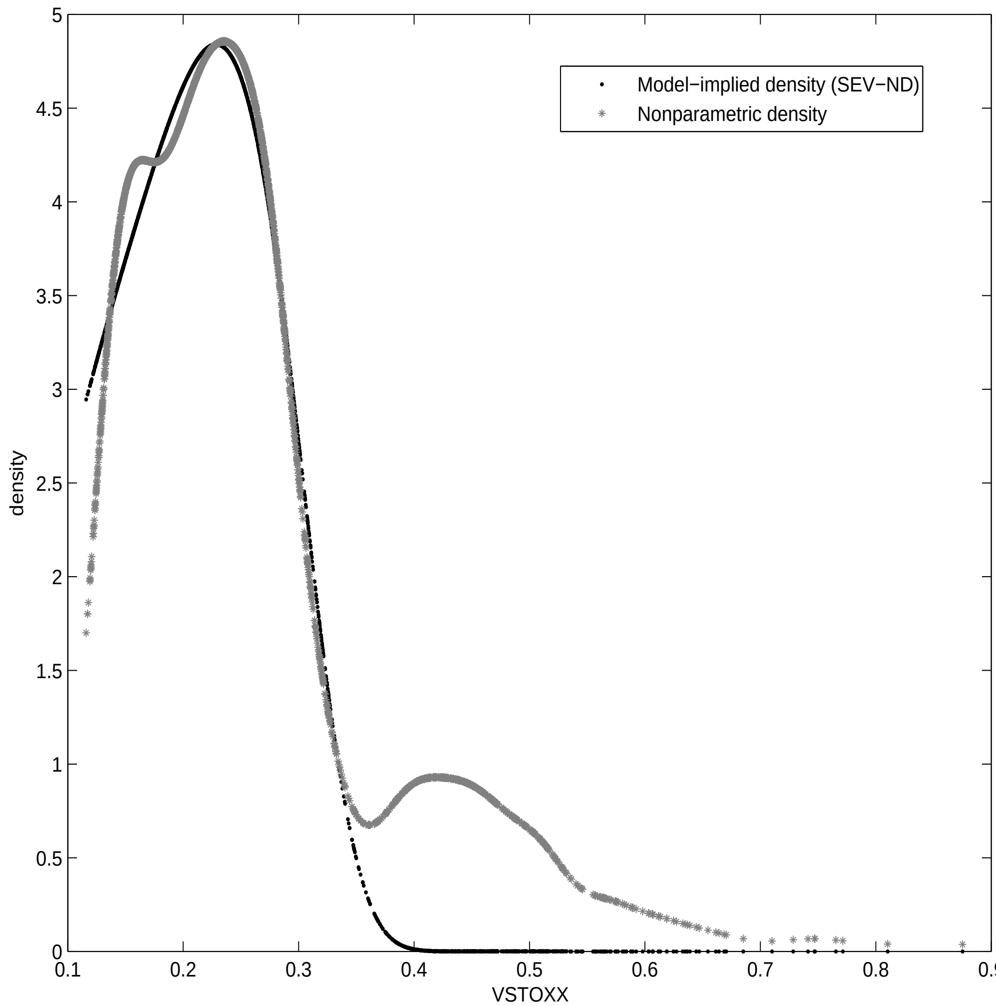

Figure 11. Goodness of fit of the SEV-ND model on the VSTOXX index by Aït-Sahalia's parametric specification test. The empirical stationary density estimated from nonparametric kernel estimation (in gray) as well as the theoretical densities (in black) computed from the optimized parameter vector are plotted. The SEV-ND model is rejected by the specification test because the distance between these two densities is fairly large.

Table 1
Summary Statistics of the Volatility Indices

This table presents the summary statistics of the four volatility indices (VIX, VIX-1Y, VSTOXX and VSMI) as well as their first-order differences.

	VIX	VIX-1Y	VSTOXX	VSMI	ΔVIX	ΔVIX-1Y	ΔVSTOXX	ΔVSMI
Start date	1/2/90	1/2/92	1/4/99	1/4/99				
Obs.	4859	4342	2617	2618	4858	4341	2616	2617
Mean	20.05	20.26	25.83	21.00	0.00	0.01	0.01	0.00
Median	18.45	19.18	23.40	18.64	−0.04	0.00	−0.09	0.03
Mode	11.65	15.78	28.93	19.52	0.00	−0.11	−0.01	0.00
Min	9.31	7.76	11.60	9.01	−17.36	−11.48	−13.98	−16.33
Max	80.86	55.53	87.51	84.90	16.54	10.24	22.64	16.41
Std. dev.	8.37	6.27	11.10	9.57	1.48	0.70	1.84	1.31
Skewness	2.15	1.66	1.46	1.74	0.41	−1.09	1.91	−0.39
Kurtosis	10.71	7.29	5.36	7.10	22.69	78.44	30.82	31.20
Autocorr.	0.98	0.99	0.99	0.99	−0.09	−0.13	−0.03	0.10

Table 2

Correlation Structure of the Volatility and Market Indices and Their Changes

This table presents the correlation structure of the volatility indices and the underlying market indices (Panel A) as well as that of their first-order differences (Panel B).

Panel A

	VIX	VIX-1Y	VSTOXX	VSMI	S&P 500	STOXX	SMI
VIX	1.0000						
VIX-1Y	0.9064	1.0000					
VSTOXX	0.9096	0.7973	1.0000				
VSMI	0.9048	0.8192	0.9609	1.0000			
S&P 500	−0.4929	−0.4407	−0.6257	−0.5654	1.0000		
STOXX	−0.2614	−0.2790	−0.3993	−0.4216	0.8751	1.0000	
SMI	−0.3484	−0.3653	−0.4982	−0.4332	0.9035	0.8394	1.0000

Panel B

	ΔVIX	ΔVIX-1Y	ΔVSTOXX	ΔVSMI	ΔS&P 500	ΔSTOXX	ΔSMI
ΔVIX	1.0000						
ΔVIX-1Y	0.6859	1.0000					
ΔVSTOXX	0.5306	0.3871	1.0000				
ΔVSMI	0.3946	0.2802	0.7095	1.0000			
Δ S&P 500	−0.8000	−0.6310	−0.3971	−0.2746	1.0000		
ΔSTOXX	−0.4659	−0.3755	−0.7167	−0.5451	0.5324	1.0000	
ΔSMI	−0.4418	−0.3468	−0.6753	−0.5461	0.4860	0.7949	1.0000

Table 3
Maximum Likelihood Estimation of the Continuous-time Dynamics of VIX (1990-2007)

This table reports the results of the maximum likelihood estimation on the VIX index for the shorter sample period. There are a total of 927 weekly observations (Wednesdays) from January 2, 1990 to December 31, 2007. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L} - d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order different models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' convenience, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\mathcal{L}$	AIC
AFF	0.6248	−3.3031			0.1218			2315.78	−4.9898
	(0.12)	(0.59)			(0.004)				
CEV-CD	0.1063					2.4201	2.8409	2379.69	−5.1277
	(0.03)					(0.51)	(0.12)		
CEV-LD	0.4526	−2.3463				1.9922	2.7202	2385.92	−5.1390
	(0.11)	(0.69)				(0.41)	(0.12)		
CEV-ND	−10.0343	55.6598	−99.4579	0.5920		2.0802	2.7365	2393.81	−5.1517
	(2.89)	(16.09)	(28.19)	(0.16)		(0.44)	(0.12)		
SEV-CD	0.1097				−0.0248	1.6669	2.4826	2379.89	−5.1260
	(0.03)				(0.04)	(0.78)	(0.44)		
SEV-LD	0.4813	−2.5205			−0.1015	0.9550	1.8725	2387.13	−5.1394
	(0.10)	(0.65)			(0.09)	(0.24)	(0.40)		
SEV-ND	−10.5288	58.2701	−103.7152	0.6211	0.0102	2.5423	2.9223	2393.84	−5.1496
	(3.39)	(18.72)	(32.53)	(0.19)	(0.03)	(1.58)	(0.54)		

Table 4
Maximum Likelihood Estimation of the Continuous-time Dynamics of VIX (1990-2009)

This table reports the results of the maximum likelihood estimation on the VIX index for the full sample period. There are a total of 995 weekly observations (Wednesdays) from January 2, 1990 to April 15, 2009. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L} - d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order different models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' convenience, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\mathcal{L}$	AIC
AFF	0.5412	−2.6587			0.1356			2406.86	−4.8319
	(0.10)	(0.42)			(0.004)				
CEV-CD	0.1094					2.1629	2.7824	2504.21	−5.0276
	(0.03)					(0.40)	(0.11)		
CEV-LD	0.3952	−1.9084				1.9062	2.7008	2509.07	−5.0353
	(0.10)	(0.64)				(0.34)	(0.10)		
CEV-ND	−3.1308	14.7547	−23.0396	0.2232		1.8935	2.6942	2511.80	−5.0368
	(1.57)	(7.54)	(11.23)	(0.10)		(0.36)	(0.11)		
SEV-CD	0.1130				−0.0228	1.6346	2.4854	2504.50	−5.0261
	(0.03)				(0.03)	(0.57)	(0.34)		
SEV-LD	0.4084	−1.9826			−0.0466	1.2491	2.2151	2509.85	−5.0349
	(0.10)	(0.62)			(0.04)	(0.36)	(0.32)		
SEV-ND	−2.8237	13.2936	−20.9981	0.2038	−0.0126	1.6248	2.5292	2511.86	−5.0349
	(1.61)	(7.72)	(11.31)	(0.10)	(0.03)	(0.63)	(0.38)		

Table 5
Maximum Likelihood Estimation of the Continuous-time Dynamics of VIX-1Y (1992-2009)

This table reports the results of the maximum likelihood estimation on the VIX-1Y index for the full sample period. There are a total of 891 weekly observations (Wednesdays) from January 2, 1992 to April 15, 2009. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L} - d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order different models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' convenience, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\mathcal{L}$	AIC
AFF	0.1731	−0.7864			0.0313			2772.52	−6.2167
	(0.07)	(0.28)			(0.0005)				
CEV-CD	0.0360					0.7027	3.0055	2845.17	−6.3797
	(0.02)					(0.10)	(0.08)		
CEV-LD	0.1500	−0.6795				0.6763	2.9788	2846.65	−6.3808
	(0.08)	(0.42)				(0.10)	(0.08)		
CEV-ND	−4.2749	18.6341	−25.7796	0.3139		0.6543	2.9553	2849.40	−6.3825
	(2.09)	(9.43)	(13.49)	(0.15)		(0.09)	(0.09)		
SEV-CD	0.0373				−0.0077	0.4705	2.5879	2845.41	−6.3780
	(0.02)				(0.009)	(0.21)	(0.44)		
SEV-LD	0.1567	−0.7133			−0.0114	0.4052	2.4277	2847.05	−6.3795
	(0.07)	(0.41)			(0.01)	(0.16)	(0.43)		
SEV-ND	−4.0265	17.4892	−24.1382	0.2970	−0.0052	0.4944	2.6644	2849.50	−6.3805
	(2.11)	(9.52)	(13.59)	(0.15)	(0.009)	(0.24)	(0.48)		

Table 6
Maximum Likelihood Estimation of the Continuous-time Dynamics of VSTOXX (1999-2009)

This table reports the results of the maximum likelihood estimation on the VSTOXX index for the full sample period. There are a total of 528 weekly observations (Wednesdays) from January 4, 1999 to April 15, 2009. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L}-d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order different models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' convenience, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\mathcal{L}$	AIC
AFF	0.7046	−2.6743			0.1982			1111.19	−4.1977
	(0.20)	(0.63)			(0.008)				
CEV-CD	0.1368					2.8479	3.0907	1191.52	−4.5020
	(0.05)					(0.60)	(0.14)		
CEV-LD	0.3352	−1.1035				2.6362	3.0339	1192.46	−4.5017
	(0.16)	(0.84)				(0.55)	(0.14)		
CEV-ND	−4.7998	20.9366	−28.1429	0.3628		2.7028	3.0477	1194.20	−4.5008
	(3.33)	(13.69)	(17.16)	(0.25)		(0.59)	(0.14)		
SEV-CD	0.1387				−0.0142	2.4065	2.8892	1191.64	−4.4986
	(0.05)				(0.03)	(0.91)	(0.39)		
SEV-LD	0.3446	−1.1536			−0.0219	2.0906	2.7489	1192.68	−4.4988
	(0.16)	(0.86)			(0.03)	(0.76)	(0.40)		
SEV-ND	−4.4961	19.7775	−26.8699	0.3393	−0.0099	2.4111	2.9092	1194.25	−4.4972
	(3.28)	(13.53)	(16.88)	(0.24)	(0.03)	(0.98)	(0.43)		

Table 7
Maximum Likelihood Estimation of the Continuous-time Dynamics of VSMI (1999-2009)

This table reports the results of the maximum likelihood estimation on the VSMI index for the full sample period. There are a total of 528 weekly observations (Wednesdays) from January 4, 1999 to April 15, 2009. We report the parameter estimates together with their standard errors in parentheses. Also reported are the optimized log-likelihood function $\mathcal{L}$ and Akaike's information criterion (AIC) which is computed as $-2/n(\mathcal{L} - d)$, where n is the number of observations and d is the dimension of the parameter vector. A more properly specified model has a smaller AIC value. To rank-order different models, likelihood ratio (LR) test statistic $-2(L_R - L_U)$ is used, when L_R and L_U refer to the likelihood functions of the restricted and unrestricted models respectively. The statistic is asymptotically chi-square distributed, with degree of freedom given by the number of restrictions. For readers' convenience, the 5% critical values of $\chi^2(df)$ with df equaling 1, 2, 3, and 4 are 3.84, 6.00, 7.82, and 9.50, respectively.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\mathcal{L}$	AIC
AFF	0.5167	−2.4761			0.1477			1244.21	−4.7016
	(0.14)	(0.51)			(0.005)				
CEV-CD	0.0995					1.7777	2.6971	1318.27	−4.9821
	(0.04)					(0.32)	(0.11)		
CEV-LD	0.3345	−1.5735				1.6451	2.6453	1320.43	−4.9865
	(0.11)	(0.71)				(0.31)	(0.11)		
CEV-ND	−0.9826	5.2398	−10.2385	0.0761		1.6462	2.6449	1320.76	−4.9802
	(1.65)	(8.53)	(12.72)	(0.09)		(0.32)	(0.12)		
SEV-CD	0.0886				0.0399	5.6161	3.8210	1322.83	−4.9956
	(0.04)				(0.01)	(2.97)	(0.46)		
SEV-LD	0.3086	−1.4314			0.0400	4.8189	3.7134	1324.44	−4.9979
	(0.13)	(0.82)			(0.01)	(2.46)	(0.45)		
SEV-ND	−2.1939	10.5882	−16.6976	0.1549	0.0435	5.4572	3.8400	1325.49	−4.9943
	(2.03)	(10.56)	(16.53)	(0.12)	(0.01)	(3.08)	(0.49)		

Table 8
Specification Test of the Continuous-time Parametric Models for VIX

This table presents the results from the specification tests of the continuous-time parametric models for the VIX index for the sample period from January 2, 1990 to April 15, 2009. Daily observations are used. The table reports the estimated parameters, the test statistic $\hat{M}$, and the corresponding critical value c_α. All seven models are rejected by the specification tests, except for the nonlinear drift CEV-ND and SEV-ND models.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\hat{M}$	c_α
AFF	0.5412	−2.9603			0.1574			9.20	2.83 reject
CEV-CD	0.1094					10.3108	3.6900	22.10	2.76 reject
CEV-LD	0.3952	−2.0832				0.2097	1.6462	8.43	2.83 reject
CEV-ND	−3.1308	19.7732	−40.3713	0.1757		1.5982	2.5554	2.40	2.93 cannot reject
SEV-CD	0.1130				0.0392	35.2988	4.6121	17.90	2.77 reject
SEV-LD	0.4084	−2.1351			−0.0665	0.2956	1.2690	8.26	2.83 reject
SEV-ND	−2.8237	17.8777	−36.4733	0.1577	0.00004	1.5547	2.6052	2.38	2.92 cannot reject

Table 9
Specification Test of the Continuous-time Parametric Models for VIX-1Y

This table presents the results from the specification tests of the continuous-time parametric models for the VIX-1Y index for the sample period from January 2, 1992 to April 15, 2009. Daily observations are used. The table reports the estimated parameters, the test statistic $\hat{M}$, and the corresponding critical value c_α. All seven models are rejected by the specification tests, except for the nonlinear drift and stochastic variance (SEV-ND) model.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\hat{M}$	c_α
AFF	0.1731	−0.9288			0.0328			44.51	3.54 reject
CEV-CD	0.0360					8.5115	4.4354	43.83	3.56 reject
CEV-LD	0.1500	−0.7564				0.1309	1.8898	38.18	3.58 reject
CEV-ND	−4.2749	22.5402	−39.0526	0.2700		0.2961	2.4163	6.70	3.82 reject
SEV-CD	0.0373				−0.00429	5.6025	4.0922	42.89	3.57 reject
SEV-LD	0.1567	−0.7261			−0.2904	0.4385	1.1832	29.50	3.60 reject
SEV-ND	−4.0265	20.0997	−32.1571	0.2621	0.0141	171.6217	7.4841	3.68	3.79 cannot reject

Table 10
Specification Test of the Continuous-time Parametric Models for VSTOXX

This table presents the results from the specification tests of the continuous-time parametric models for the VSTOXX index for the sample period from January 4, 1999 to April 15, 2009. Daily observations are used. The table reports the estimated parameters, the test statistic $\hat{M}$, and the corresponding critical value c_α. All seven models are rejected by the specification tests.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\hat{M}$	c_α
AFF	0.7046	−2.9663			0.1879			8.17	2.08 reject
CEV-CD	0.1368					6.2837	3.6552	23.97	1.93 reject
CEV-LD	0.3352	−2.1778				0.00835	−1.1930	6.53	2.08 reject
CEV-ND	−4.7998	25.0685	−49.4444	0.3088		0.0176	−1.1818	6.49	2.11 reject
SEV-CD	0.1387				−0.0052	6.4326	3.8034	7.40	2.10 reject
SEV-LD	0.3446	−2.2644			0.00994	0.00769	−1.2650	6.50	2.09 reject
SEV-ND	−4.4961	36.2891	−89.5761	0.0978	−0.0020	0.0232	−1.3033	5.95	2.10 reject

Table 11
Specification Test of the Continuous-time Parametric Models for VSMI

This table presents the results from the specification tests of the continuous-time parametric models for the VSMI index for the sample period from January 4, 1999 to April 15, 2009. Daily observations are used. The table reports the estimated parameters, the test statistic $\hat{M}$, and the corresponding critical value c_α. All seven models are rejected by the specification tests.

Model	α_0	α_1	α_2	α_{-1}	β_1	β_2	β_3	$\hat{M}$	c_α
AFF	0.5167	−2.8351			0.1482			10.30	2.79 reject
CEV-CD	0.0995					9.5208	3.7116	5.22	2.73 reject
CEV-LD	0.3345	−1.5508				1.1092	2.3505	4.14	2.77 reject
CEV-ND	−0.9826	6.0559	−17.1243	0.1117		3.4474	2.3931	3.95	2.78 reject
SEV-CD	0.0886				0.0119	12.8161	4.0331	4.56	2.75 reject
SEV-LD	0.3086	−1.4499			−0.0853	0.5850	1.6810	3.97	2.78 reject
SEV-ND	−2.1939	14.3730	−43.1347	0.2650	−0.0061	9.4519	2.4125	3.90	2.78 reject