https://doi.org/10.1016/j.jfineco.2012.08.011

Diagnosing Affine Models of Options Pricing: Evidence from VIX

Gang Li and Chu Zhang*

December, 2011

*Hong Kong Baptist University and Hong Kong University of Science and Technology, respectively. We would like to thank the referee whose comments and suggestions substantially improved the paper. We would also like to thank Hai Lin, Jun Yu, and seminar participants at Asian Financial Management Association meeting, Fudan University, Hong Kong University, National University of Singapore, Shanghai University of Finance and Economics, and Singapore Management University for helpful comments on earlier versions. Chu Zhang acknowledges financial support from the Hong Kong RGC Competitive Earmarked Research Grant HKUST 647011. All remaining errors are ours.

Diagnosing Affine Models of Options Pricing: Evidence from VIX

Abstract

Affine jump-diffusion models have been the mainstream in options pricing because of their analytical tractability. Popular affine jump-diffusion models, however, are still unsatisfactory in describing the options data and the problem is often attributed to the diffusion term of the unobserved state variables. Using prices of variance-swaps (i.e., squared VIX) implied from options prices, we provide fresh evidence regarding the mis-specification of affine jump-diffusion models, as variance-swap prices are affine functions of the state variables in a broader class of models that do not restrict the diffusion term of the state variables. We apply the nonparametric methodology used by Aït-Sahalia (1996b), supplemented with bootstrap tests and other parametric tests, to the S&P 500 index options data from January 1996 to September 2008. We find that, while the affine diffusion term of the state variables may contribute to the mis-specification as the literature has suggested, the affine drift of the state variables, jump intensities, and risk premiums are also sources of mis-specification.

1. Introduction

Recent advances in modeling options prices are aimed at solving the problems of volatility smile and smirk, which refer to the phenomena that the implied volatility from the Black-Scholes formula is a smile-shaped function of the strike price before the 1987 stock market crash and a decreasing function after the crash. Since the problems stem from the constant volatility assumption of the Black-Scholes model, major advances are made along the line of stochastic volatility models, in which unobserved instantaneous volatility of the underlying security follows another stochastic process and serves as an additional state variable in pricing options. More recent developments add jump components to the processes of both the price of the underlying security and the state variables which are components of stochastic volatilities. The affine jump-diffusion models make headway towards resolving the pricing issue. Important milestones include the models of Heston (1993) and Duffie, Pan, and Singleton (2000).[1] However, empirical studies show that the existing models are still inadequate in fitting the observed options prices in cross-sections to various degrees. Bakshi, Cao, and Chen (1997) find that the Heston's model requires highly implausible parameters of the volatility-return correlation and the volatility-of-volatility. Bates (2000) extends the stochastic volatility/jump-diffusion model to a two-factor specification with time-varying jump intensity. Pan (2002) and Eraker (2004) report improvements of models with jumps in both the underlying price and the stochastic volatility processes in the time-series dimension. The fit of cross-sections of options data, however, is still unsatisfactory even with added jump components.

The existing evidence in the literature suggests that the rejection of affine jump-diffusion models of options pricing is due to the mis-specification of the diffusion term of stochastic volatility. For example, Duffie et al. (2000) suggest that the deficiency of certain

[1] Empirical analysis of equity index options using affine jump-diffusion models includes Bakshi, Cao, and Chen (1997), Bates (2000), Chernov and Ghysels (2000), Pan (2002), Eraker, Johannes, and Polson (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007), among others. For recent surveys on the options pricing literature, see Broadie and Detemple (2004) and Garcia, Ghysels and Renault (2010).

specific affine jump-diffusion options pricing models is that these models unnecessarily restrict the correlation between the state variables driving the underlying returns and the stochastic volatility. Jones (2003) finds that the square-root stochastic volatility model of Heston's type is incapable of generating realistic return behavior and concludes that the stochastic volatility models in the constant elasticity-of-variance class or with a time-varying leverage effect are more consistent with the underlying asset and options data. Christoffersen, Jacobs, and Mimouni (2010) find that a stochastic volatility model with a linear diffusion term fits options prices better than the square-root process, which implies that the conditional variance of the stochastic volatility is better modeled as a quadratic function of the state variables. To sum up, some authors identify the problem with specific affine models as the restrictiveness of the diffusion term of the state variables which can be solved by using less restrictive diffusion terms within the affine class of models, while others find that the entire affine class is inadequate because of the empirical evidence of the non-affine diffusion term of the state variables.

In this paper we address the following questions. Are diffusion terms of the state variables in specific affine models the only source of the problems in options pricing? Can the problem of specific models be resolved within the class of affine jump-diffusion models by having a more flexible diffusion term? The analysis we conduct in this paper is based on one observation that, in a much wider class of models than the affine jump-diffusion models, variance-swap prices are affine functions of state variables and inherit the properties of state variables. The wider class with this property, named as the semi-affine models in this paper, is the class that imposes no restriction on the diffusion term of the state variables. We examine the affine properties of the conditional mean and conditional variance of the variance-swap prices implied from options and reject them. Since the affine property of the conditional mean of variance-swap prices does not rely on the affine property of the diffusion term of the state variables, the rejection of the affine property of the conditional mean of variance-swap prices can be traced to the inappropriate affine specifications of either the conditional mean of state variables, jump intensities, or risk premiums.

The methodology we use in this paper begins with the nonparametric method used by Aït-Sahalia (1996b) and Stanton (1997) on short-term interest rate, followed by bootstrap tests and some parametric tests. The use of nonparametric methods allows us to address issues with general affine models, rather than specific affine models. We apply our methods to the S&P 500 index options from January 1996 to September 2008. The nonparametric estimation and testing results in this paper show that the conditional mean, conditional variance and conditional covariance of the variance-swap prices of the S&P 500 index exhibit strong non-affine properties. More specifically, the mean reversion of variance-swap prices is much faster and the volatility of variance-swap prices is much greater at the high levels of variance-swap prices than what affine functions imply. Parametric tests further confirm the results. Both nonparametric and parametric results suggest that the specifications of the affine drift of state variables, jump intensities, and risk premiums are all potential reasons for the rejection of affine jump-diffusion models. The problems of specific affine models cannot be resolved by more flexible diffusion terms not only within the affine class of models, but also within the semi-affine class of models. It should be noted, however, that our empirical evidence presented in this paper is limited to the case of S&P 500 index options, and does not necessarily extend to other data.

Our specification analysis of the affine jump-diffusion models is based on the dynamic features of the variance-swap prices inferred from cross-sections of option prices. This is in contract to the approach in the existing literature where option pricing models are fitted to the prices of individual options. Our approach has the following advantages. First, since state variables are unobserved, complicated econometric methods have been used in the literature to estimate the models, which make it difficult to identify which aspect of the affine jump-diffusion models causes problems. This is especially so when Heston's univariate model is extended to multivariate models. Our approach of examining the conditional mean and variance of transformed state variables in parametric and nonparametric analyses is straightforward to implement and avoids complicated econometric procedures. Second, the prices of certain long-maturity, deep in- or out-of-the-money options may contain errors due to liquidity reasons. The existence of these errors makes it difficult to know whether a model is rejected because of model mis-specification or because of data errors. Our approach of using variance-swap prices avoids this problem because prices of individual options are aggregated, so the impact of idiosyncratic errors is substantially reduced.

Our intended contribution of this paper is to provide evidence at a fairly general level that the mis-specification of the affine models goes beyond the diffusion term of the state variables, as the literature has been focused on. Since models outside the affine class are difficult to solve, such information can be valuable to theoretical modelers in directing their efforts towards finding better models. While our empirical results are limited to the S&P 500 index options only, the methodology can be used in other cases in which sufficient cross-sections of options data are available.

The rest of the paper is organized as follows. Section 2 presents the jump-diffusion models that we investigate and the properties of variance-swap prices under jump-diffusion models. Section 3 explains the construction of the S&P 500 index model-free variance-swap prices and provides the summary statistics. Section 4 presents the results for the nonparametric estimation of the conditional moments of model-free variance-swap prices. Section 5 discusses the nonparametric tests and presents the results. Section 6 conducts parametric tests. Section 7 discusses some robustness issues of the main results and Section 8 concludes the paper.

2. The options pricing framework and variance-swap prices

2.1. The options pricing framework

Suppose the log price of an underlying security, $s_t = \log S_t$, of a number of European options is driven by the stochastic process under the physical probability, P, as follows,

$$ds_t = \mu_s(x_t)dt + \sigma_s(x_t)dW_t + z_{st}dJ_t - \nu_s\lambda(x_t)dt, \tag{1}$$

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)dW_t + z_{xt}dJ_t - \nu_x\lambda(x_t)dt, \tag{2}$$

where x_t is a k-dimensional state variable, W_t is a standard n-vector Brownian motion with $n \geq k+1$, J_t is an m-vector counting process with jump intensity $\lambda(x_t)$ independent of W_t, $\mu_s(x_t)$ and $\mu_x(x_t)$ are the conditional mean of ds_t and dx_t, $\sigma_s\sigma_s{}'$ and $\sigma_x\sigma_x{}'$ are the conditional variance of the diffusive component of ds_t and dx_t, z_{st} is the conformable matrix of random jump sizes of ds_t with mean ν_s, z_{xt} is the conformable matrix of random jump sizes of dx_t with mean ν_x, both independent of W_t and J_t. Since the functional form of $\mu_s(x_t)$, $\sigma_s(x_t)$, $\mu_x(x_t)$, $\sigma_x(x_t)$ and $\lambda(x_t)$ and the distribution function of z_{st} and z_{xt} are all unspecified, except for the regularity conditions to guarantee the existence of the solution, s_t and x_t, this is a very general class of jump-diffusion models used in the options pricing literature.[2]

Suppose a riskfree asset exists with the riskfree rate being r_t which may depend on x_t. Since x_t is not assumed to be traded and the jump components of the processes cannot be hedged, there exists a risk-neutral probability, $\tilde{P}$, though not unique, under which $\tilde{W}_t = W_t + \int_0^t \phi(x_s)ds$ is a Brownian motion for an n-vector, $\phi(x_t)$, J_t has an intensity function, $\tilde{\lambda}(x_t)$, and distributions of jump sizes (z_{st}, z_{xt}) are potentially different with means $\tilde{\nu}_s$ and $\tilde{\nu}_x$. Under $\tilde{P}$, the log price of the underlying security and the state variables

[2] Our analysis is limited to the jump-diffusion models with finite activity jumps, which do not include infinite activity jump processes used in the literature.

evolve as

$$ds_t = \tilde{\mu}_s(x_t)dt + \sigma_s(x_t)d\tilde{W}_t + z_{st}dJ_t - \tilde{\nu}_s\tilde{\lambda}(x_t)dt, \tag{3}$$

$$dx_t = \tilde{\mu}_x(x_t)dt + \sigma_x(x_t)d\tilde{W}_t + z_{xt}dJ_t - \tilde{\nu}_x\tilde{\lambda}(x_t)dt, \tag{4}$$

where $\tilde{\mu}_s(x_t) = \mu_s(x_t) - \sigma_s(x_t)\phi(x_t) - \nu_s\lambda(x_t) + \tilde{\nu}_s\tilde{\lambda}(x_t)$ and $\tilde{\mu}_x(x_t) = \mu_x(x_t) - \sigma_x(x_t)\phi(x_t) - \nu_x\lambda(x_t) + \tilde{\nu}_x\tilde{\lambda}(x_t)$. For $\tilde{P}$ to be a risk-neutral probability,

$$\tilde{\mu}_s(x_t) = r_t - \frac{1}{2}\sigma_s(x_t)\sigma_s(x_t)' - [\tilde{E}_t(e^{z_{st}} - 1 - z_{st})]\tilde{\lambda}(x_t), \tag{5}$$

so that, by Ito's lemma, $\tilde{E}_t(dS_t/S_t) = r_t dt$ where $\tilde{E}_t$ is the expectation with respect to $\tilde{P}$ conditional on time t information. Options can be priced under the risk-neutral probability as discounted expected future payoffs.

For the model to be affine under the actual probability P, the following conditions are imposed.

- (i) $\mu_s(x_t)$ is affine in x_t.
- (ii) $\sigma_s(x_t)\sigma_s(x_t)'$ is affine in x_t.
- (iii) Each element of $\mu_x(x_t)$ is affine in x_t.
- (iv) Each element of $\sigma_x(x_t)\sigma_x(x_t)'$ is affine in x_t.
- (v) Each element of $\sigma_s(x_t)\sigma_x(x_t)'$ is affine in x_t.
- (vi) Each element of $\lambda(x_t)$ is affine in x_t.

For the model to be affine under the risk-neutral probability $\tilde{P}$, the following additional conditions are imposed.

- (vii) $\sigma_s(x_t)\phi(x_t)$ is affine in x_t.
- (viii) Each element of $\sigma_x(x_t)\phi(x_t)$ is affine in x_t.
- (ix) Each element of $\tilde{\lambda}(x_t)$ is affine in x_t.

Together with (i)-(vi), the additional conditions (vii)-(ix) guarantee that $\tilde{\mu}_s(x_t)$ and $\tilde{\mu}_x(x_t)$ are affine in x_t, so the model is affine under the risk-neutral probability, $\tilde{P}$. The advantage of affine jump-diffusion models is their analytical tractability. Heston (1993) shows the closed form expression of prices of European options for the simplest affine diffusion model. Duffie et al. (2000) extend it to general affine models and to other derivatives.

Strictly speaking, a closed form option pricing formula can be obtained as long as the model is affine only under the risk-neutral probability, $\tilde{P}$. That is, one may directly specify that $\tilde{\mu}_s(x_t)$, $\tilde{\mu}_x(x_t)$, $\sigma_s(x_t)\sigma_s(x_t)'$, $\sigma_x(x_t)\sigma_x(x_t)'$, $\sigma_s(x_t)\sigma_x(x_t)'$ and $\tilde{\lambda}(x_t)$ are affine without specifying that $\mu_s(x_t)$, $\mu_x(x_t)$, and $\lambda(x_t)$ are affine. The empirical results we present in this paper have nothing to say about this alternative approach.[3] Since the paper is motivated by the fact that affine models under $\tilde{P}$ face challenges and the ultimate goal of this line of research is to find out the sources of mis-specification, considerations of models that have no restrictions under actual probability P defeat the purpose.

The thrust of this paper is based on the observation that the prices of a class of derivatives known as variance-swaps are affine functions of x_t in models less restrictive than the affine models defined above. A variance-swap is a forward contract on the realized quadratic variation of s_t over a fixed time horizon. It will be shown below that the variance-swap price, which is the expectation of the quadratic variation under the risk-neutral probability, is a function of the first moment of the state variable x_t alone. No restrictions on $\sigma_x(x_t)$ are required for the variance-swap price to be affine in x_t. This observation leads us to define the class of semi-affine jump-diffusion models as follows.

Definition. The semi-affine class of jump diffusion models used in this paper is defined by conditions (ii), (iii), (vi), (viii), and (ix).

[3] In the literature of the term structure of interest rates, there are models that are non-affine under the actual probability, but affine under the risk-neutral probability with a contrived specification of risk premiums. We are not aware of any such models in options pricing.

The condition (ii) can be viewed as a definition of state variables in this class of models and it does not impose any material restrictions. The key ingredients of a semi-affine model are, therefore, the affine drift of the state variables under probability P, the affine jump intensities under both P and $\tilde{P}$, and the risk premiums associated with the state variables. The class of semi-affine models is obviously broader than the class of affine models because it is defined with less restrictive conditions.

The key difference between affine models and semi-affine models is that the crucial conditions (iv) and (v) of the affine models are not imposed on the semi-affine models which leave $\sigma_x(x_t)$ unrestricted. This statement, however, requires further clarification because $\sigma_x(x_t)$ still appears in (viii). Since $\phi(x_t)$ is unrestricted, for any specification of $\sigma_x(x_t)$ that is bounded away from zero as in all affine models, there always exists a $\phi(x_t)$ such that $\sigma_x(x_t)\phi(x_t)$ is an affine function of x_t. It is in this sense that semi-affine models do not put restriction on $\sigma_x(x_t)$.[4]

This feature of the semi-affine models distinguishes them from affine models and plays a crucial role in this paper. As many authors, such as Jones (2003) and Christoffersen et al. (2010), attribute the unsatisfactory performance of specific affine models to the affine restriction on the diffusion term of the state variables, i.e., condition (iv), examinations of semi-affine models help discover if the affine restriction of the diffusion term of the state variables is the only problem the affine models have.

2.2. Variance-swap prices

Since state variables in the jump-diffusion options pricing models cannot be observed directly, econometric inferences have to rely on variables that track the unobserved state variables. Variance-swap prices are good candidates for this purpose. A variance-swap is a forward contract determined at t between two parties to exchange at $t+\tau$ a value

[4] A related discussion can be found in Cheridito, Filipovič, and Kimmel (2007) in the context of affine models of interest-rates.

$\tilde{V}_{\tau,t}$ and the realized quadratic variation of s_u between t and $t+\tau$, $\frac{1}{\tau}\int_t^{t+\tau}\langle s_u, s_u\rangle du$, where $\langle\cdot,\cdot\rangle$ indicates quadratic variation and the multiplier $\frac{1}{\tau}$ reflects the convention of annualization when the unit of t is a year. The value $\tilde{V}_{\tau,t}$ is known as the variance-swap price. Since a variance-swap has no value at its inception, t, it must be true that, theoretically, $\tilde{V}_{\tau,t} = \tilde{E}_t\frac{1}{\tau}\int_t^{t+\tau}\langle s_u, s_u\rangle du$. Since $\tilde{V}_{\tau,t}$ is the expectation conditional on the information at t, it must be a measurable function of x_t. The empirical part of this paper is based on the following propositions. The proofs of these results are presented in the Appendix.

Proposition 1. Under the assumptions of semi-affine jump-diffusion models, the prices of variance-swaps are affine functions of the state variables, x_t.

Therefore, under the assumptions of semi-affine models, variance-swap prices inherit the affine properties of the state variables under both the actual probability P and risk-neutral probability $\tilde{P}$. This facilitates a test of the affine properties of the state variables using the properties of variance-swap prices under P. In the following sections, we examine the conditional means and conditional variances of the variance-swap prices under probability P as a test of affine properties of the state variables. Since the semi-affine jump-diffusion models do not require that the squared diffusion term of the state variables be affine in the state variables, we are able to tell whether this requirement is the only restriction responsible for the unsatisfactory performance of certain specific affine models in fitting options prices.

Similar results have appeared in the literature, so the proposition is not new. For example, Duan and Yeh (2010) find a similar result for a one-factor affine model and Egloff, Leippold, and Wu (2010) find the same result for a multi-factor affine model without jumps. The proofs of the results are more-or-less the same. However, none of them emphasize that the result can be obtained in semi-affine models such as the one defined in this paper. To our knowledge, no attempt has been made for the purpose of model diagnostics, using the result that variance-swap prices inherit the properties of the state variables in semi-affine models.

Variance-swap prices can be obtained from the actual market for variance-swaps or can be created synthetically using prices of liquid options on the underlying. For our purpose of testing the semi-affine models, the latter approach has an obvious advantage. This is so because variance-swaps are traded over-the-counter, without a central clearing house. Price data from any specific dealer may contain all kinds of errors, due to lack of liquidity, behavior biases of specific traders, or simply human errors, which make the inferences based on these actual prices less reliable. On the other hand, the test of affine properties conducted in this paper does not require observations of actual variance-swap prices. All it requires is that, theoretically, variance-swap prices are affine functions of the unobserved state variables.

We construct variance-swap prices as portfolios of out-of-the-money calls and puts, using the approach demonstrated by Carr and Madan (1998) and Demeterfi et al. (1999) as follows.

$$V_{\tau,t} = \frac{2}{\tau} e^{r\tau} \left(\int_0^{F_{\tau,t}} \frac{1}{K^2} p_{\tau,t}(K) dK + \int_{F_{\tau,t}}^{\infty} \frac{1}{K^2} c_{\tau,t}(K) dK \right), \tag{6}$$

where r is the riskfree rate and is assumed to be constant, $F_{\tau,t}$ is the $t+\tau$-forward price of S_t, and $c_{\tau,t}(K)$ and $p_{\tau,t}(K)$ are prices of European calls and puts with strike price K and maturity at $t+\tau$. The resultant synthetic variance-swap price is known as the model-free variance-swap price. Under the assumption that the stochastic process for the price of the underlying security is continuous, i.e., there is no jump component, $V_{\tau,t} = \tilde{V}_{\tau,t}$. As such, the variance-swap prices are replicated by portfolios of out-of-the-money calls and puts. In practice, calls and puts of all strikes are not available and the forward price of S_t involves estimating dividends paid from t to $t+\tau$, so approximations are involved. We adopt the same approach to calculating the VIX as used by the Chicago Board Options Exchange (CBOE). The approach includes implying the forward price from the prices of near-the-money options according to the put-call parity, approximating the integration with a numerical integration scheme over the range of available strikes, estimating the price of the variance-swap of a fixed maturity by linearly interpolating prices of variance-swaps of adjacent maturities, and using the bid-ask average prices of the options as the true prices. All of these ingredients may introduce some approximation errors. Since there are so many factors which are essentially unrelated to each other, and the effects of most of them on the sign of the approximation error cannot be determined, the approximation is unlikely to yield systematic biases.[5]

With jump components, however, the replicating strategy in Eq. (6) is no longer exact, i.e., $V_{\tau,t} \neq \tilde{V}_{\tau,t}$. The errors induced by jumps in the replication, $V_{\tau,t} - \tilde{V}_{\tau,t}$, are small on average for the S&P 500 index. Carr and Wu (2009) provide the order of the approximation error of Eq. (6) due to jumps and conclude that the approximation error is less than 1% of the average variance level. While the errors are small on average, they can be large at times. In general, the error can be unbounded. Nevertheless, the next proposition justifies our use of model-free variance-swap prices in the analysis.

Proposition 2. Suppose the riskfree rate is a constant. Denote the drift of x_t as $\tilde{\mu}_x(x_t) = \Gamma(\theta - x_t)$ and the jump intensity as $\tilde{\lambda}(x_t) = \tilde{\lambda}_0 + \tilde{\Lambda} x_t$. The approximation error of $V_{\tau,t}$ as $\tilde{V}_{\tau,t}$ is also an affine function of the state variables under the assumptions of semi-affine models, as follows,

$$
\begin{aligned}
& V_{\tau,t} - \tilde{V}_{\tau,t} \\
= \; & \frac{2}{\tau} \tilde{E} \left(e^{z_{st}} - 1 - z_{st} - z_{st}^2/2 \right) \tilde{E}_t \left[\int_t^{t+\tau} (\tilde{\lambda}_0 + \tilde{\Lambda} x_u) du \right] \\
= \; & 2 \left[\frac{\tilde{E} z_{st}^3}{3!} + \frac{\tilde{E} z_{st}^4}{4!} + \cdots \right] \left[\tilde{\lambda}_0 + \tilde{\Lambda}\theta + \tilde{\Lambda} (\tau\Gamma)^{-1} (I_k - e^{-\tau\Gamma})(x_t - \theta) \right]. \qquad (7)
\end{aligned}
$$

According to the proposition, the first factor of the approximation error due to jumps is determined by the shape of the jump size distribution in the process of the underlying

[5] As a robustness check, we adopt the method proposed by Jiang and Tian (2007) to reduce the errors arising from the unavailability of options of all strikes. The variance-swap prices calculated using this alternative approach are quantitatively similar to those from the standard approach and the conclusions on testing the affine properties using the two approaches are the same.

asset. If the jump size distribution is more skewed or with fatter tails, then the absolute value of the approximation error will be greater. The second factor of the approximation error is determined by the expected jump intensity from the present to the maturity of the variance-swap conditioned on the information at the time. Since the jump intensity is an affine function of the state variable, $\tilde{\lambda}_0+\tilde{\Lambda}x_u$, the unconditional mean of the approximation error, $\tilde{E}[V_{\tau,t}-\tilde{V}_{\tau,t}]$, will be greater if, loosely speaking, $\tilde{\lambda}_0$ and $\tilde{\Lambda}$ are greater, and the mean of x_u, θ, is greater. The sign of the unexpected approximation error, $V_{\tau,t}-\tilde{V}_{\tau,t}-\tilde{E}[V_{\tau,t}-\tilde{V}_{\tau,t}]$, is positive when $x_t > \theta$. The absolute value of the unexpected approximation error is greater if the persistence of x_t is greater (i.e., Γ is smaller).[6]

It should be realized that the approximation error caused by the existence of jump components can be unbounded because it is an affine function of the state variables, which are unbounded for most non-trivial affine processes.[7] The importance of Proposition 2, however, is that the approximation error does not affect the use of model-free variance-swap prices in testing the semi-affine properties of the state variables. If the conditional means of model-free variance-swap prices are found to be non-affine, it must be caused by the assumptions of semi-affine properties, rather than the approximation error due to jumps. This is so because the approximation error itself is affine in the state variables and will not generate non-affine distortions if the state variables follow a semi-affine process.

From Eq. (5), the riskfree rate is a function of the state variables. The assumption of a constant riskfree rate in the proposition is made mainly for simplicity, without which, the use of the model-free variance-swap price formula as the theoretical variance-swap price

[6]Jiang and Oomen (2008) provide a test for the presence of jumps based on the replication error due to jumps for the payoff of the realized quadratic variation. The replication involves a continuously rebalanced delta hedging strategy for a short position in two log contracts. The annualized replication error is given by $\frac{1}{\tau}\sum_{i=J_t+1}^{J_{t+\tau}} 2(e^{z_{st_i}}-1-z_{st_i}-z_{st_i}^2/2)$, where t_i is the time for the ith jump. The expression given in Proposition 2 can be obtained alternatively by taking risk-neutral expectation of the above expression under the assumptions of semi-affine models.

[7]For example, a variable following the square-root process, the simplest non-trivial affine process without jumps, has a conditional distribution of the non-central χ^2 type, and therefore is unbounded. Adding and mixing an independent jump component to it will not change that even if the jump size distribution is bounded.

may contain additional approximation errors. Fortunately, the assumption of a constant riskfree rate is innocuous as far as the empirical performance of option pricing models is concerned, as most options are relatively short term, with time-to-expiration less than two years typically. Bakshi et al. (1997), for example, find that a stochastic riskfree rate is not helpful for improving the options pricing performance.

3. Data description

We use daily data of the S&P 500 index options from January 1996 to September 2008. The options written on the S&P 500 index are the most actively traded European-style contracts, and the S&P 500 index options and the S&P 500 futures options have been the focus of recent empirical options studies. The best known model-free variance-swap price for the S&P 500 index is the squared VIX, which has a maturity of one month and is constructed by CBOE. Data for VIX are downloaded from the CBOE website. A few recent studies show that options pricing models should contain multiple unobserved state variables.[8] We construct longer-term variance-swap prices of the S&P 500 index using the same CBOE VIX calculation method (Chicago Board Options Exchange, 2003). The options data needed are from OptionMetrics. The daily interest rate data are from the U.S. Treasury Department's website.

The prices of a model-free variance-swap with a fixed maturity are calculated as follows. For a given day, the variance-swap prices with maturities equal to those of available options contracts are calculated first. The variance-swap price with a given maturity is then calculated by interpolating two variance-swap prices with maturities closest to the given maturity. According to CBOE, the VIX is constructed using options of maturities

[8] Christoffersen, Heston, and Jacobs (2009) find that a two-factor stochastic volatility model provides more flexibility in modeling the time-series variation in the smirk and the volatility term structure than single-factor stochastic volatility models. Li and Zhang (2010) use a nonparametric approach to arrive at the conclusion that in addition to the price of the underlying security, exactly two state variables are required for pricing S&P 500 index options. Christoffersen et al. (2008) propose a GARCH options pricing model with long-run and short-run volatility factors that outperforms the one-factor options pricing model of Heston and Nandi (2000), especially for pricing long maturity options.

no greater than 39 days. We choose options with the longest maturities to construct the long-maturity variance-swap price. To ensure that at least two maturities are available to calculate a fixed maturity variance-swap price for the entire sample, the options with two shortest maturities of no less than 350 days are used to calculate the 18-month variance-swap price. The 1-month and 18-month variance-swap prices are denoted as V_1 and V_{18}, respectively.

Fig. 1 plots V_1 and V_{18}. In general, V_1 and V_{18} move up and down together. Both of the short-maturity and the long-maturity variance-swap prices are relatively low in 1996 and during 2004-2006 and relatively high for the rest of the years in the sample. There are also clear contrasts between V_1 and V_{18}. V_1 is more volatile and has several spikes during the sample period. V_{18} is more stable and more persistent.

Fig. 1 Here

Table 1 reports the summary statistics of V_1, V_{18} and their first differences, ΔV_1 and ΔV_{18}. It is shown that V_1 is slightly higher than is V_{18} on average. V_1 is also more volatile and more positively skewed than is V_{18}. The time-series of V_1 and V_{18} are only modestly persistent, as the autocorrelations of V_1 and V_{18} decay quickly when compared with daily interest-rate data.[9] The p-values of the augmented Dickey-Fuller unit root test of V_1 and V_{18} are low, suggesting that the levels of V_1 and V_{18} are stationary. ΔV_1 has a larger standard deviation than does ΔV_{18} and they are about equally positively skewed.

Table 1 Here

In the next few sections, we estimate the conditional mean and variance-covariance of the variance-swap price changes and test the affine properties of these quantities. The

[9] For example, the autocorrelations for the daily 1-month constant maturity treasury yield for the same sample period are $\rho_1 = 0.9990$, $\rho_2 = 0.9976$, $\rho_3 = 0.9964$, $\rho_5 = 0.9946$, $\rho_{10} = 0.9903$, $\rho_{20} = 0.9838$ and $\rho_{30} = 0.9780$, much more persistent than V_1 and V_{18}.

main focus is on the conditional mean because semi-affine models do not impose restrictions on variances and covariances. We begin with a nonparametric method because there are no obvious alternatives to the affine jump-diffusion models.[10] The method we use is an extension of the method used by Aït-Sahalia (1996b) and Stanton (1997) on the conditional mean and conditional variance functions of the short-term interest-rate model.

The method, when applied to interest rates, draws certain criticism, Pristker (1998) argues that Aït-Sahalia's test does not perform well in finite samples and it over-rejects the null of affine conditional mean because the interest rate data are highly persistent. Chapman and Pearson (2000) argue that due to the truncation of the upper limit of finite samples, the kernel regression estimation of the conditional mean is downward biased at the upper end. The criticism does not apply squarely to the case of variance-swap prices we study in this paper for two reasons. First, variance-swap prices are much less persistent than the interest rates. Because of this, we don't need to invoke the result in Aït-Sahalia and Park (2012) who derive the asymptotic distribution of the test for integrated data. Second, our method analyzes transition density, instead of stationary density, of the process. The transition density based nonparametric tests do not rely as much on stationarity and the transition density contains additional information in the data that is unavailable in the stationary density. Hong and Li (2005) and Chen and Gao (2004) show that transition density based tests have good size and power properties.

We use a bootstrap method to improve the finite sample performance of the nonparametric test. The bootstrap method has been employed in various studies in the literature and has been shown to have good finite-sample performances. Corradi and Swanson (2005) show that bootstrap methods provide accurate standard errors even for highly persistent data. We also conduct a simulation study which shows that the bootstrap

[10]Nonparametric methods are applied to options pricing in Hutchinson, Lo, and Poggio (1994), Aït-Sahalia (1996a), Aït-Sahalia and Lo (1998), Broadie et al. (2000a, 2000b), Aït-Sahalia, Wang, and Yared (2001), Aït-Sahalia and Duarte (2003), and Li and Zhao (2009), among others.

test performs well for persistent data. To supplement the nonparametric method, we also adopt a parametric approach with specific non-affine alternatives gleaned from the nonparametric estimation.

4. Nonparametric estimation

In this section, we estimate the conditional moments of variance-swap prices nonparametrically and examine their affine properties. Specifically, we estimate the conditional moments of the daily changes in the variance-swap prices, ΔV_1 and ΔV_{18}, as

$$\Delta V_{1,t+1} = \mu_1(V_{1,t}, V_{18,t}) + \eta_{1,t+1}, \tag{8}$$

$$\Delta V_{18,t+1} = \mu_{18}(V_{1,t}, V_{18,t}) + \eta_{18,t+1}, \tag{9}$$

and

$$\hat{\eta}^2_{1,t+1} = \sigma^2_1(V_{1,t}, V_{18,t}) + \xi_{1,t+1}, \tag{10}$$

$$\hat{\eta}^2_{18,t+1} = \sigma^2_{18}(V_{1,t}, V_{18,t}) + \xi_{18,t+1}, \tag{11}$$

$$\hat{\eta}_{1,t+1}\hat{\eta}_{18,t+1} = \sigma_{1,18}(V_{1,t}, V_{18,t}) + \xi_{1,18,t+1}. \tag{12}$$

where $\Delta V_{\tau_j,t+1} = V_{\tau_j,t+1} - V_{\tau_j,t}$ for $\tau_j = 1$ and 18, and $\hat{\eta}_{1,t+1}$ and $\hat{\eta}_{18,t+1}$ are fitted residuals from Eq. (8) and Eq. (9).

The conditional moments are fitted using both the local constant (Nadaraya-Watson) and local linear kernel estimators. For the former,

$$\widehat{\mu_{\tau_j}}(V_1, V_{18}) = \frac{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})\Delta V_{\tau_j,t+1}}{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})} \tag{13}$$

$$\widehat{\sigma^2_{\tau_j}}(V_1, V_{18}) = \frac{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})\hat{\eta}^2_{\tau_j,t+1}}{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})} \tag{14}$$

$$\widehat{\sigma_{1,18}}(V_1, V_{18}) = \frac{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})\hat{\eta}_{1,t+1}\hat{\eta}_{18,t+1}}{\sum_{t=1}^{T} \phi(\frac{V_{1,t}-V_1}{h_{V_1}})\phi(\frac{V_{18,t}-V_{18}}{h_{V_{18}}})} \tag{15}$$

where $\tau_j = 1, 18$, ϕ is a kernel function, h_w is the bandwidth for an explanatory variable, w, and T is the number of observations. We choose the second-order Gaussian kernel with $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$. The optimal bandwidth, h_w, is determined by the cross-validation method for each conditional moment and for each explanatory variable. Using the cross-validation method, a vector of bandwidth h is chosen to minimize the objective function

$$\text{CV}(h) = \frac{1}{T} \sum_{t=1}^{T} [z_t - \hat{m}_{-t,h}(u_t)]^2 \nu(u_t), \tag{16}$$

where $\hat{m}_{-t,h}(u_t)$ is the kernel estimator of z_t without using the observation z_t, and $\nu(u_t)$ is a weighting function for the observation u_t. The role of ν is to reduce the boundary biases in the bandwidth selection by reducing the weight of the extreme levels of u_t. ν is one if each component of u_t is between the 2.5th and 97.5th percentile and zero otherwise. We also estimate an independent case in which $\mu_1(V_{1,t}, V_{18,t})$ and $\sigma_1^2(V_{1,t}, V_{18,t})$ only depend on $V_{1,t}$, $\mu_{18}(V_{1,t}, V_{18,t})$ and $\sigma_{18}^2(V_{1,t}, V_{18,t})$ only depend on $V_{18,t}$, and $\sigma_{1,18}(V_{1,t}, V_{18,t}) = 0$. For the case of one-dimensional u_t, ν is one if u_t is between the 5th and 95th percentile and zero otherwise.

The estimation result of the conditional mean and the conditional variance for the independent case is shown in Fig. 2. The solid line shows the mean estimate, and the dashed lines cover the 90% confidence interval.[11] The estimated conditional mean of ΔV_1 is a concave function of V_1. From the low to medium level of V_1, the conditional mean is close to zero. For the high level of V_1, the conditional mean is negative and the speed of mean reversion is relatively fast. The unconditional mean of V_1 over the sample period is 0.047, at which level, the conditional mean of ΔV_1 is slightly above zero. This suggests that V_1 has a tendency to move even higher at the mean level. Similar to ΔV_1,

[11] The confidence interval is calculated using Kunsch's (1989) block bootstrap method to account for the time-series dependence of the observations.

the conditional mean of ΔV_{18} is concave in V_{18}. The non-affine conditional mean of the variance-swap prices for the independent case rejects the affine restrictions on the semi-affine jump-diffusion model. The conditional variance of ΔV_1 and ΔV_{18} is convex, which suggests that it increases faster in V_1 and V_{18} than what the affine process dictates.

The local constant kernel regression tends to estimate a flatter surface near the boundary of the domain when the slopes of the surface are actually steep because the average is taken for observations from inside the boundary only. The local linear kernel regression is known for its better performance in this situation. The results based on local linear kernel regressions show that the fitted conditional mean functions are quantitatively similar to those based on local constant kernel regressions for low variance-swap price levels, as the slopes of the functions are flat. But for high variance-swap price levels, the non-affine properties of the conditional means of ΔV_1 and ΔV_{18} are even more pronounced when local linear kernel regressions are used in the estimation. On the other hand, the non-affine properties of the conditional variances estimated by local linear kernel regressions are less pronounced because the means are fitted better. The results are not reported here to conserve the space.

Fig. 2 Here

The patterns of the condition mean and conditional variance of the variance-swap prices can be compared with those of short-term interest rates estimated using nonparametric methods by Aït-Sahalia (1996b) and Stanton (1997). The similarity is that, for both short-term interest rates and variance-swap prices, the conditional mean and the conditional variance of their changes are found to be non-affine. The conditional means are concave and the conditional variances are convex. There are some small differences, however. The conditional mean of the change in short-term interest rates exhibits a certain degree of curvature throughout the entire range of the short-term interest rates. The conditional mean of variance-swap prices, however, is more like two connected straight lines with different slopes: it is almost flat in the low range of the variance-swap prices, but it abruptly changes to a very steep line in the high range. The conditional variance of the short-term interest rates is slightly U-shaped, while the conditional variance of the variance-swap prices is monotonically increasing.

For the dependent case, the conditional mean of ΔV_1 is a function of V_1 and V_{18}. Its estimate is presented conditional on V_{18} at the low, medium and high levels in the left panels of Fig. 3. Conditional on the low and medium levels of V_{18}, V_1 shows little tendency to revert to the mean. Conditional on the high level of V_{18}, the conditional mean is positive for some regions of V_1, and it becomes negative when V_1 is also very high. The confidence interval is wider conditional on the high level of V_{18} than conditional on the low level of V_{18}. The conditional mean of ΔV_{18} as a function of V_1 and V_{18} is shown in the right panels of Fig. 3. The magnitude of the conditional mean of ΔV_{18} is smaller than that of ΔV_1, suggesting that V_{18} is more persistent. The strongest mean reversion of V_{18} also occurs conditional on the high level of V_1. The non-affine conditional mean of ΔV_1 and ΔV_{18} of the dependent case confirms the findings of the univariate case that the affine restrictions on the semi-affine jump-diffusion model are rejected.

Fig. 3 Here

The estimated conditional variance for the dependent case is shown in Fig. 4. In the left panels, the estimated conditional variance of ΔV_1 is shown as a function of V_1 conditional on the low, medium and high levels of V_{18}. The right panels are for the estimated conditional variance of ΔV_{18}. We note the difference in scales for different panels. For both ΔV_1 and ΔV_{18}, the conditional variances are convex functions of their levels. The non-affine property is stronger for ΔV_1 than for ΔV_{18}, and more so when conditional on low levels of V_1 and V_{18} than conditional on high levels. It is also shown that ΔV_1 has a higher variability than does ΔV_{18}.

Fig. 4 Here

The estimated conditional covariance for the dependent case is shown in Fig. 5. In the left panels, the estimated conditional covariance between ΔV_1 and ΔV_{18} is shown as a function of V_1 for the low, medium and high levels of V_{18}. The right panels show the estimated conditional covariance as a function of V_{18} for the different levels of V_1. The conditional covariance is increasing in V_1 and V_{18} in general. It is a convex function of V_1 and V_{18} at low levels of V_1 and V_{18}. However, at high levels of V_1 and V_{18}, the conditional covariance is a concave function of V_1 and V_{18}, and decreases at the extreme high levels of V_1 and V_{18}. The results suggest that both the conditional variance of ΔV_1 and ΔV_{18} and their conditional covariance are inconsistent with the affine process since they increase at much faster rates in the levels of V_1 and V_{18} than what the affine process suggests.

Fig. 5 Here

5. Nonparametric tests

In this section, we conduct rigorous nonparametric tests of the affine properties of the conditional mean, conditional variance, and conditional covariance of ΔV_1 and ΔV_{18}. We consider four tests for the independent case, in which the conditional mean and conditional variance of ΔV_1 are functions of V_1 only, the conditional mean and conditional variance of ΔV_{18} are functions of V_{18} only, and the conditional covariance of $(\Delta V_1, \Delta V_{18})$ is zero. Specifically, we test the null hypotheses $\mu_1(V_1) = a + bV_1$, $\sigma_1^2(V_1) = a + bV_1$, $\mu_{18}(V_{18}) = a + bV_{18}$, and $\sigma_{18}^2(V_{18}) = a + bV_{18}$ against unrestricted alternatives. For the dependent case, we consider fifteen tests, in which each of the following five moments, μ_1, μ_{18}, σ_1^2, σ_{18}^2, $\sigma_{1,18}$, takes the following functional forms under the null hypothesis: $a + b_1V_1 + b_2V_{18}$, $g(V_1) + bV_{18}$, or $bV_1 + g(V_{18})$, where $g(\cdot)$ is unrestricted. The totally unrestricted moments under the alternative and the unrestricted $g(\cdot)$ are estimated nonparametrically.

We use the nonparametric test developed in Fan and Li (1996) and Zheng (1996) to test the parametric or semiparametric function forms against the nonparametric alternatives, supplemented by the so-called two-point wild bootstrap method to approximate the null distribution of the statistic to achieve more accurate finite sample results. Under the null hypothesis, the test statistic is asymptotically distributed as a standard normal variate under certain regularity conditions. A number of studies, such as Li and Wang (1998), Li (2005), Li and Racine (2007, Chapter 12), and Gu, Li, and Liu (2007), show that the test of Fan and Li (1996) and Zheng (1996) has good finite sample performance when used in combination with the bootstrap method in various applications. Some of the additional advantages of the bootstrap method are that it allows for heteroskedasticity, it works well with serially correlated data, and the result is insensitive to the choice of the bandwidth. We illustrate the methods of testing an affine model and a partially affine model in turn.

5.1. Testing affine models

For the case of an affine model, suppose we test the null hypothesis that the conditional mean of ΔV_1 is affine in (V_1, V_{18}). That is

$$H_0 : E[\Delta V_1 | V_1, V_{18}] = a + b_1 V_1 + b_2 V_{18}. \tag{17}$$

A statistic can be constructed as $I = E[\varepsilon E(\varepsilon|V_1, V_{18})]$, where $\varepsilon = \Delta V_1 - a - b_1 V_1 - b_2 V_{18}$ is the residual under the null hypothesis. By the law of iterated expectations, $I = E[E^2(\varepsilon|V_1, V_{18})] \geq 0$. The equality holds if and only if the null hypothesis is true. Thus, I serves as a proper statistic for consistently testing the null hypothesis. A density weighted sample analogue of I is $I_T = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{t+1} E(\hat{\varepsilon}_{t+1}|V_{1,t}, V_{18,t}) f(V_{1,t}, V_{18,t})$, where $\hat{\varepsilon}_{t+1} = \Delta V_{1,t+1} - \hat{a} - \hat{b}_1 V_{1,t} - \hat{b}_2 V_{18,t}$ is the OLS regression residual and $f(V_{1,t}, V_{18,t})$ is the joint density of $(V_{1,t}, V_{18,t})$.[12] Both $E(\hat{\varepsilon}_{t+1}|V_{1,t}, V_{18,t})$ and $f(V_{1,t}, V_{18,t})$ are estimated nonparametrically. $\hat{I}_T$, which is I_T standardized by a consistent estimator of its standard error, follows the standard normal distribution asymptotically under the null. The null

[12] Fan and Li (1996) indicate that using the density-weighted version overcomes the random denominator problem in the kernel estimation of $E(\hat{\varepsilon}_{t+1}|V_{1,t}, V_{18,t})$ and simplifies the asymptotic analysis of I_T.

hypothesis is rejected if $\tilde{I}_T$ is greater than a threshold, say, 1.645 at the 5% significance level.

Using the residual $\hat{\varepsilon}_{t+1}$, the two-point wild bootstrap samples under the null hypothesis can be constructed as,

$$\Delta V^*_{1,t+1} = \hat{a} + \hat{b}_1 V_{1,t} + \hat{b}_2 V_{18,t} + \varepsilon^*_{t+1}, \tag{18}$$

where $\varepsilon^*_{t+1} = \frac{1-\sqrt{5}}{2}\hat{\varepsilon}_{t+1}$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$, and $\varepsilon^*_{t+1} = \frac{1+\sqrt{5}}{2}\hat{\varepsilon}_{t+1}$ with probability $\frac{-1+\sqrt{5}}{2\sqrt{5}}$. The new errors have the following property: $E^*(\varepsilon^*_{t+1}) = 0$, $E^*(\varepsilon^{*2}_{t+1}) = \hat{\varepsilon}^2_{t+1}$ and $E^*(\varepsilon^{*3}_{t+1}) = \hat{\varepsilon}^3_{t+1}$, where E^* indicates the expected value in the simulation. Then, the bootstrap samples are used to compute the test statistic $\hat{I}^*_T$ in the same way $\hat{I}_T$ is computed. The empirical distribution of $\hat{I}^*_T$ under the null hypothesis can be obtained from many bootstrap samples. In practice, we construct 100 bootstrap samples. When $\hat{I}_T$ is greater than the 95th percentile of the empirical distribution of $\hat{I}^*_T$, the bootstrap test rejects the null hypothesis at the 5% significance level and the bootstrap p-value is 5%.

5.2. Estimating and testing partially affine models

To estimate a partially affine model, consider $\mu_1(V_1, V_{18}) = g(V_1) + bV_{18}$ as an example. The regression is

$$\Delta V_{1,t+1} = g(V_{1,t}) + bV_{18,t} + \varepsilon_{t+1}, \tag{19}$$

where $\Delta V_{1,t+1} = V_{1,t+1} - V_{1,t}$ and $g(.)$ is unspecified. Taking the expectation of Eq. (19) conditional on $V_{1,t}$ gives

$$E(\Delta V_{1,t+1}|V_{1,t}) = g(V_{1,t}) + bE(V_{18,t}|V_{1,t}). \tag{20}$$

Subtracting Eq. (20) from Eq. (19) and multiplying by the density of $V_{1,t}$, $f(V_{1,t})$, yields

$$[\Delta V_{1,t+1} - E(\Delta V_{1,t+1}|V_{1,t})]f(V_{1,t}) = b[V_{18,t} - E(V_{18,t}|V_{1,t})]f(V_{1,t}) + \varepsilon_{t+1}f(V_{1,t}). \tag{21}$$

$E(\Delta V_{1,t+1}|V_{1,t})$, $E(V_{18,t}|V_{1,t})$ and $f(V_{1,t})$ are estimated nonparametrically. Given the estimates of the conditional expectations and the density function, b is estimated by regressing $[\Delta V_{1,t+1} - E(\Delta V_{1,t+1}|V_{1,t})]f(V_{1,t})$ on $[V_{18,t} - E(V_{18,t}|V_{1,t})]f(V_{1,t})$ using OLS. With the estimated $\hat{b}$, $g(V_{1,t})$ is estimated by regressing $\Delta V_{1,t+1} - \hat{b}V_{18,t}$ on $V_{1,t}$ nonparametrically, or is simply estimated by $E(\Delta V_{1,t+1}|V_{1,t}) - \hat{b}E(V_{18,t}|V_{1,t})$ as suggested by Eq. (20).[13]

To test the null hypothesis $E(\Delta V_1|V_1, V_{18}) = g(V_1) + bV_{18}$ against $E(\Delta V_1|V_1, V_{18}) = g_1(V_1, V_{18})$ is equivalent to testing $E(\Delta V_1 - bV_{18}|V_1) = g(V_1)$ against $E(\Delta V_1 - bV_{18}|V_1, V_{18}) = g_2(V_1, V_{18})$, where $g_1(V_1, V_{18})$ and $g_2(V_1, V_{18})$ are unrestricted. This suggests that given b, testing partially affine models against nonparametric alternatives is equivalent to testing omitted variables. A density-weighted version of the test statistic for omitted variables or partially affine models is $I = E[\varepsilon f(V_1)E(\varepsilon f(V_1)|V_1, V_{18})f(V_1, V_{18})]$, where ε is the residual from the null hypothesis. For our case of testing partially affine models against nonparametric alternatives, $\varepsilon = \Delta V_1 - g(V_1) - bV_{18}$. The sample analogue of I is $I_T = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{t+1} f(V_{1,t})E(\hat{\varepsilon}_{t+1} f(V_{1,t})|V_{1,t}, V_{18,t})f(V_{1,t}, V_{18,t})$, where $\hat{\varepsilon}_{t+1} = \Delta V_{1,t+1} - \hat{g}(V_{1,t}) - \hat{b}V_{18,t}$, $\hat{g}(V_{1,t})$ is the nonparametric estimator of $g(V_{1,t})$, and $f(V_{1,t})$, $E(\hat{\varepsilon}_{t+1} f(V_{1,t})|V_{1,t}, V_{18,t})$ and $f(V_{1,t}, V_{18,t})$ are estimated nonparametrically.

The bootstrap samples under the null hypothesis can be constructed as,

$$\Delta V^*_{1,t+1} = \hat{g}(V_{1,t}) + \varepsilon^*_{t+1}, \qquad (22)$$

where $\varepsilon^*_{t+1} = \frac{1-\sqrt{5}}{2}\hat{\varepsilon}_{t+1}$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$, and $\varepsilon^*_{t+1} = \frac{1+\sqrt{5}}{2}\hat{\varepsilon}_{t+1}$ with probability $\frac{-1+\sqrt{5}}{2\sqrt{5}}$. The linear term $\hat{b}V_{18,t}$ is unnecessary since the test can be regarded as an omitted variables test and the estimation of the linear term can be avoid. The test statistic computed from the bootstrap samples is $I^*_T = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}^*_{t+1} f(V_{1,t})E(\hat{\varepsilon}^*_{t+1} f(V_{1,t})|V_{1,t}, V_{18,t})f(V_{1,t}, V_{18,t})$, where $\hat{\varepsilon}^*_{t+1} = \Delta V^*_{1,t+1} - \hat{g}^*(V_{1,t})$ and $\hat{g}^*(V_{1,t})$ is the nonparametric estimator of $E(\Delta V^*_{1,t+1}|V_{1,t})$. Statistical inference is then made by comparing the standardized test statistic $\hat{I}_T$ from

[13]The reason for multiplying by $f(V_{1,t})$ is to avoid the technical difficulties in deriving the asymptotic distribution of $\hat{b}$ arising from the random denominator problem in the kernel estimation of $E(\Delta V_{1,t+1}|V_{1,t})$ and $E(V_{18,t}|V_{1,t})$, indicated in Li and Racine (2007, p.224). $\hat{b}$ is a consistent estimator of b.

the original sample with a distribution of the standardized test statistics $\hat{I}_T^*$ from the bootstrap samples.

5.3. Testing results

The results of the nonparametric tests shown in Table 2 are in line with the impression from Fig. 2-5. Panel A reports results based on local constant kernel regressions in the estimation. For the independent case, the conditional mean of ΔV_1 is marginally significant, while the affine properties of other conditional moments are strongly rejected. For the dependent case, we find strong non-affine properties in the conditional means of ΔV_1 and ΔV_{18}, the conditional variances of ΔV_1 and ΔV_{18}, and their conditional covariance, evidenced by the very low p-values. The results also show that the nonparametric components of V_1 and V_{18} capture the non-affine property in the conditional moments. The partially affine model allowing for the nonparametric component of $g(V_1)$ captures the non-affine property in the conditional mean and conditional variance of ΔV_1. The nonparametric component of $g(V_{18})$ captures the non-affine property not only in the conditional mean and conditional variance of ΔV_{18}, but also in the conditional mean of ΔV_1 and the conditional covariance. Panel B reports results based on local linear kernel regressions. The p-values for testing the affine properties of the conditional means of ΔV_1 and ΔV_{18} are virtually identical to those based on local constant kernel regressions. The non-affine properties of the conditional variances and conditional covariance are slightly weaker than those based on local constant kernel regressions because, as we mentioned earlier, local linear kernel regressions fit the conditional means better at the upper end of V_1 and V_{18} and the estimated values of second moments are lower at the upper end.

Table 2 here

The non-affine conditional means of the variance-swap prices indicate that either the affine jump intensity of the underlying price/volatility process, the drift or the risk premium of the volatility process of the semi-affine jump-diffusion model is mis-specified. Therefore, the non-affine diffusion of the volatility process is not the only reason for the rejection of the affine jump-diffusion model of options pricing.

5.4. Finite sample performance of the bootstrap test

In this subsection, we conduct a simulation analysis of the finite sample performance of the bootstrap testing method. The data are simulated to capture the important features of the variance-swap prices so that the performance of the test on persistent time-series data can be examined. The data are generated under the null hypothesis of either affine or semi-affine models. We explain the independent case first.

For the independent case, one variance-swap suffices. Let a time-series of $\check{V}_{1,t+1}$ be generated by the following model,

$$\check{V}_{1,t+1} = \phi_0 + (\phi_1 + 1)\check{V}_{1,t} + \sqrt{\phi_2 \check{V}_{1,t}^{\phi_3}}\varepsilon_{t+1}, \tag{23}$$

where ε_{t+1} is drawn from the standard normal distribution independently across t. The starting value, $\check{V}_{1,1}$ is equal to 0.02, which is approximately the level of the actual 1-month variance-swap price at the beginning of the sample period. The parameters are estimated by the quasi maximum likelihood method using the entire sample of the actual 1-month variance-swap prices. The estimates are $\phi_0 = 0.0004$, $\phi_1 = -0.0048$, $\phi_2 = 0.019$ and $\phi_3 = 2.1$. The parameters determined this way correspond to a semi-affine model. We also consider an affine model in which $\phi_3 = 1$, so the conditional variance of $\Delta\check{V}_1$ is affine. The remaining parameters in the affine model are estimated as $\phi_0 = 0.00066$, $\phi_1 = -0.013$ and $\phi_2 = 0.00075$.

For the semi-affine model with $\phi_3 \neq 1$, we test the null hypothesis that the conditional mean of $\Delta\check{V}_1$ is affine. For the case of $\phi_3 = 1$, we further test the affine property of the conditional variance. We construct 500 time-series samples from each of the above models, and use the bootstrap method described previously to test the null hypothesis at a certain significance level. From the 500 samples, we calculate the rejection rate. A test with the rejection rate close to the significance level is considered as a good test. We also vary the sample size and bandwidth to examine the sensitivity of the test to these variables. We consider the sample sizes of 1000, 2000, and 3208. The latter is the size of the actual sample used in this study. We consider the case of the optimal bandwidth h^*, an under-smoothed case with $h^*/1.5$ and an over-smoothed case with $1.5h^*$.

Table 3 reports results based on local constant kernel regressions in the estimation. The percentage rejection rates out of the 500 tests on the simulated samples for the independent case are reported in the left panel of Table 3. For the affine case with $\phi_3 = 1$, the rejections rates for both of the conditional mean and conditional variance are in line with the significance levels. For the semi-affine case with $\phi_3 \neq 1$, the test tends to over-reject the null for smaller samples. For the simulated data of the same sample size as the actual sample used in this study, the performance is good.

Table 3 Here

For the dependent case, the data are generated by the model,

$$\check{V}_{1,t+1} = \psi_{0,1} + (\psi_{1,1} + 1)\check{V}_{1,t} + \psi_{2,1}\check{V}_{18,t} + \sqrt{\psi_{3,1}\check{V}_{1,t}^{\psi_{4,1}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,1}\check{V}_{18,t}^{\psi_{6,1}}}\varepsilon_{2,t+1} \quad (24)$$

$$\check{V}_{18,t+1} = \psi_{0,18} + \psi_{1,18}\check{V}_{1,t} + (\psi_{2,18} + 1)\check{V}_{18,t} + \sqrt{\psi_{3,18}\check{V}_{1,t}^{\psi_{4,18}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,18}\check{V}_{18,t}^{\psi_{6,18}}}\varepsilon_{2,t+1}, \quad (25)$$

where $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are drawn from the standard normal distribution independent of each other and across t. The parameters are estimated by the quasi maximum likelihood method using both the 1-month and 18-month variance-swap prices. The parameters estimated for Eq. (24) are $\psi_{0,1} = 0.00011$, $\psi_{1,1} = -0.025$, $\psi_{2,1} = 0.025$, $\psi_{3,1} = 0.037$, $\psi_{4,1} = 2.4$, $\psi_{5,1} = 0.000048$, $\psi_{6,1} = 0.77$; the parameters estimated for Eq. (25) are $\psi_{0,18} = 0.00045$, $\psi_{1,18} = 0.011$, $\psi_{2,18} = -0.022$, $\psi_{3,18} = 0.000078$, $\psi_{4,18} = 2.7$, $\psi_{5,18} = 0.0033$ and $\psi_{6,18} = 2.1$. The parameters for the affine case with $\psi_{4,1} = \psi_{6,1} = \psi_{4,18} = \psi_{6,18} = 1$ are $\psi_{0,1} = 0.00011$, $\psi_{1,1} = -0.034$, $\psi_{2,1} = 0.034$, $\psi_{3,1} = 0.0007$, $\psi_{5,1} = 0.000049$,

$\psi_{0,18} = 0.00034$, $\psi_{1,18} = 0.0087$, $\psi_{2,18} = -0.016$, $\psi_{3,18} = 0.000005$ and $\psi_{5,18} = 0.00012$. The results of testing performance for the dependent case are in the right panel of Table 3. They are similar to those of the independent cases.

Table 4 reports results based on local linear kernel regressions. The rejection rates are also very close to nominal sizes of the tests. Overall, the results in Table 3 and 4 show that the bootstrap test performs quite well for persistent time-series data. They also indicate that the rejection rates are not sensitive to the bandwidth choices in both the independent and dependent cases.

Table 4 Here

5.5. Alternative choices of testing variables

We also consider the tests based on the first difference of the slope and curvature of the term structure of variance-swap prices. The slope and curvature are defined as $V_{18} - V_1$ and $(V_1 + V_{18})/2 - V_9$, respectively, where V_9 is the 9-month variance-swap price. Under the null hypothesis that the conditional means of the variance-swap prices are affine in the variance-swap prices, so are the conditional means of the slope and curvature since they are linear combinations of the conditional means of the variance-swap prices.

The advantage of using the slope and curvature for testing purpose is that they are less persistent than the variance-swap prices are, so the nonparametric tests of affine properties on the conditional means of the slope and curvature have better finite sample properties than those for the variance-swap prices themselves.[14] On the other hand, the tests based on the slope and curvature lose power to a certain extent because the non-affine patterns in the conditional means of the variance-swap prices tend to cancel with each other in the slope and curvature, as they involve differences.

[14]The autocorrelations for the daily observations of $V_{18} - V_1$ are $\rho_1 = 0.9436$, $\rho_2 = 0.9003$, $\rho_3 = 0.8668$, $\rho_5 = 0.8062$, $\rho_{10} = 0.7094$, $\rho_{20} = 0.5747$ and $\rho_{30} = 0.4710$. For $(V_1 + V_{18})/2 - V_9$, the autocorrelations are $\rho_1 = 0.9121$, $\rho_2 = 0.8546$, $\rho_3 = 0.8096$, $\rho_5 = 0.7428$, $\rho_{10} = 0.6240$, $\rho_{20} = 0.4765$ and $\rho_{30} = 0.3644$.

The results based on the slope and curvature are qualitatively the same as those on variance-swap prices themselves. The p-values for the tests tend to be higher, however. They are not reported here to save the space.[15]

6. Parametric tests

To supplement the tests against nonparametric alternatives, we adopt a parametric approach with specific non-affine alternatives. Tests against carefully designed parametric alternatives are more powerful. We use the nonparametric estimation results in Section 4 as a guide to specify the non-affine parametric alternatives. The non-affine property in the conditional mean is captured by the squared and reciprocal terms quite well. For the independent case, the conditional mean function is specified as

$$\mu_{\tau_j}(V_{\tau_j}) = \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{\tau_j} + \alpha_{2,\tau_j} V_{\tau_j}^2 + \alpha_{3,\tau_j}(1/V_{\tau_j}), \tag{26}$$

for $\tau_j = 1, 18$. The affine property of the conditional mean is rejected if $\alpha_{2,\tau_j} \neq 0$ or $\alpha_{3,\tau_j} \neq 0$. Aït-Sahalia (1996b) also considers this specification for the non-affine conditional mean of interest rate models. For the dependent case, the specification is

$$\mu_{\tau_j}(V_1, V_{18}) = \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_1 + \alpha_{2,\tau_j} V_{18} + \alpha_{3,\tau_j} V_1^2 + \alpha_{4,\tau_j} V_{18}^2 + \alpha_{5,\tau_j}(1/V_1) + \alpha_{6,\tau_j}(1/V_{18}), \tag{27}$$

for $\tau_j = 1, 18$. Similarly, the affine property of the conditional mean is rejected if $\alpha_{3,\tau_j} \neq 0$, or $\alpha_{4,\tau_j} \neq 0$, or $\alpha_{5,\tau_j} \neq 0$, or $\alpha_{6,\tau_j} \neq 0$. As the explanatory variables are simple transformations of V_τs, they are highly correlated. To reduce the multicollinearity problem, we orthogonalize the explanatory variables as in Chapman and Pearson (2000).

To illustrate the orthogonalization and estimation procedure, consider the dependent case as an example. V_{18} is regressed on a constant and V_1, and the regression residual is the orthogonalized V_{18}, denoted as $\widetilde{V_{18}}$. Likewise, V_1^2 is regressed on a constant, V_1 and $\widetilde{V_{18}}$ and the regression residual is the orthogonalized V_1^2, denoted as $\widetilde{V_1^2}$. Other variables, $\widetilde{V_{18}^2}$,

[15] We thank the referee for suggesting such tests. The results are available upon request.

$(\widetilde{1/V_1})$ and $(\widetilde{1/V_{18}})$ are defined similarly. To account for the possible spurious non-affine property of the conditional mean at the upper end of the variance-swap prices, where variance of residuals tends to be large, similar to the approach in Chapman and Pearson (2000), we use weighted least squares to estimate the regression,

$$\begin{aligned}\Delta V_{\tau_j,t+1} &= \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{1,t} + \alpha_{2,\tau_j} \widetilde{V_{18,t}} + \alpha_{3,\tau_j} \widetilde{V_{1,t}^2} \\ &\quad + \alpha_{4,\tau_j} \widetilde{V_{18,t}^2} + \alpha_{5,\tau_j} (\widetilde{1/V_{1,t}}) + \alpha_{6,\tau_j} (\widetilde{1/V_{18,t}}) + \epsilon_{\tau_j,t+1}, \end{aligned} \quad (28)$$

where $\Delta V_{\tau_j,t+1} = V_{\tau_j,t+1} - V_{\tau_j,t}$ for $\tau_j = 1, 18$. The weight is the reciprocal of the variance of $\Delta V_{\tau_j,t+1}$ estimated as a function of $V_{1,t}$ and $V_{18,t}$ nonparametrically.

The conditional variance and conditional covariance increase exponentially with the levels of the variance-swap prices, which suggests a power function to capture such non-affine property. For the independent case, the conditional variance function is specified as

$$\sigma^2_{\tau_j}(V_{\tau_j}) = (\beta_{0,\tau_j} + \beta_{1,\tau_j} V_{\tau_j,t})^{\gamma_{\tau_j}}, \quad (29)$$

for $\tau_j = 1, 18$. For the dependent case, the conditional variance function is specified as

$$\sigma^2_{\tau_j}(V_1, V_{18}) = (\beta_{0,\tau_j} + \beta_{1,\tau_j} V_{1,t} + \beta_{2,\tau_j} V_{18,t})^{\gamma_{\tau_j}}, \quad (30)$$

for $\tau_j = 1, 18$ and the specification for the conditional covariance is

$$\sigma_{1,18}(V_1, V_{18}) = (\beta_{0,1,18} + \beta_{1,1,18} V_{1,t} + \beta_{2,1,18} V_{18,t})^{\gamma_{1,18}}. \quad (31)$$

The affine property of the conditional variance or conditional covariance is rejected if $\gamma \neq 1$.

The parameters in the conditional variance and conditional covariance functions are estimated using nonlinear least squares. As shown in Fig. 2-5, the variance of residuals of conditional variance and conditional covariance tend to increase with the level of variance-swap prices, we similarly weight the observations by the reciprocal of the nonparametric estimate of the variance of dependent variables before applying the nonlinear least squares estimation. Fig. 2-5 also show that β_0s in Eq. (29)-(31) are close to zero, and the estimates of all β_0s in Eq. (29)-(31) are statistically indifferent from zero. Without loss of generality, we restrict $\beta_0 = 0$ when estimating the models.[16]

The results are shown in Table 5 for the independent case and in Table 6 for the dependent case. The t-statistics reported are adjusted for heteroskedasticity and 24 lags of autocorrelation using Newey and West (1987). For the independent case, $\alpha_{2,1}$, $\alpha_{2,18}$ and $\alpha_{3,18}$, the coefficient estimates on the non-affine terms of the conditional mean functions, are statistically significant. γ is significantly greater than one for both of ΔV_1 and ΔV_{18}, indicating that the conditional variances are disproportionally large at high levels of V_1 and V_{18}. The results suggest that both of the conditional means and conditional variances are non-affine in the level of variance-swap prices.

Table 5 Here

For the dependent case, $\alpha_{3,1}$ and $\alpha_{6,1}$ are statistically significant, which evidences the non-affine property of the conditional mean of ΔV_1. $\alpha_{4,18}$ and $\alpha_{6,18}$, the coefficient estimates on the non-affine terms of the conditional mean function of ΔV_{18}, are also significant. The estimated γ for the conditional variance and conditional covariance functions of ΔV_1 and ΔV_{18} is significantly greater than 1. Overall, the results of the parametric tests are in line with the impression from the nonparametric estimation in Fig. 2-5, and consistent with those of the nonparametric tests in the previous section.

Table 6 Here

7. Robustness to omitted state variables

The non-affine property of the conditional means, conditional variances and conditional covariances of variance-swap prices is identified from the analysis of the joint dynamics

[16] Results also suggest that restricting $\beta_0 = 0$ increases the estimation precision of other parameters in the models since their standard errors are reduced substantially.

up to two variances-swap prices. Some recent studies show that options pricing models with two unobserved state variables significantly outperform those with one state variable. However, the two state variables models may still be inadequate to capture the time-series and cross-sectional dynamics of options prices. This gives rise to the concern that the non-affine property may potentially be due to missing variables in the model specifications.

To study the robustness of the results to missing variables, we include a medium-maturity variance-swap in the specifications. Using the same model-free variance-swap construction method, we calculate the prices of a 9-month variance-swap, denoted as V_9. The options with two shortest maturities of no less than 189 days are used to calculate the 9-month variance-swap prices. Similarly, under the null, each of the nine conditional moments, μ_1, μ_9, μ_{18}, σ_1^2, σ_9^2, σ_{18}^2, $\sigma_{1,9}$, $\sigma_{1,18}$, $\sigma_{9,18}$, takes hypothesized functional forms, and is tested against nonparametric alternatives. The following functional forms for the null hypothesis are considered: $a+b_1V_1+b_2V_9+b_3V_{18}$, $g(V_1)+b_1V_9+b_2V_{18}$, $b_1V_1+g(V_9)+b_2V_{18}$, and $b_1V_1+b_2V_9+g(V_{18})$. The first functional form is affine in all three variance-swap prices, whereas the rest contain a nonparametric component in one of (V_1, V_9, V_{18}). There are 36 tests totally.

The p-values of the tests are reported in Table 7. Panel A reports results based on local constant kernel regressions in the estimation. The affine property is rejected for all the conditional means, conditional variances and conditional covariances of ΔV_1, ΔV_9 and ΔV_{18}, as evidenced by the low p-values of the affine models. The p-values are much higher for the models with a nonparametric component. The results show that the non-affine property in the conditional mean and conditional variance of a variance-swap price is captured by the nonparametric component of itself, whereas the non-affine property in the conditional covariance is mostly captured by the nonparametric component of one of the two variance-swap prices involved. In some cases, the non-affine property of the conditional mean is also captured by the nonparametric components of other variance-swap prices in addition to that of itself. Results based on local linear kernel regressions are reported in Panel B. Again, the affine properties of all the conditional moments are rejected. Overall, the results suggest that the non-affine property of the conditional means, conditional variances and conditional covariances of variance-swap prices is robust to missing variables in the specifications in Section 5.

Table 7 Here

8. Conclusion

The affine jump-diffusion models are popular in the options pricing literature because they are tractable to provide closed-form solutions of option prices. The evidence in the existing literature suggests that the problem of the affine jump-diffusion models is with the diffusion term of the state variables. In this paper, we focus on the affine restrictions on the other aspects of the affine jump-diffusion models, namely the drift term, jump intensities, and the risk premiums associated with the state variables. In the semi-affine models, variance-swap prices are affine functions of the unobserved state variables and inherit the affine properties of the unobserved state variables. Testing these affine restrictions on the state variables is tantamount to testing the affine properties of the variance-swap prices. We use both nonparametric and parametric methods and find strong non-affine properties of the conditional mean, conditional variance and conditional covariance of the model-free variance-swap prices.

The non-affine conditional mean of the change in variance-swap prices indicates that the semi-affine jump-diffusion models are problematic. This suggests that the affine diffusion term is not the only problem with the affine jump-diffusion models, as the semi-affine models do not restrict the diffusion term of the state variables. The drift term of the state variables, the jump intensities, and the risk premiums are all likely to be non-affine. The exclusive attention in the literature on the diffusion term of the state variables is too narrowly focused. The finding of non-affine conditional variance of the variance-swap prices is interesting by itself, but in no way implies that the conditional variance of the state variables is non-affine when the variance-swap prices themselves are not affine in the state variables.

The empirical results presented in this paper are helpful in directing future theoretical research on modeling options prices. To that end, our further empirical research aims at finding out the appropriate dynamic features of the stochastic processes governing the underlying assets and the state variables, in terms of the functional forms of the drift terms, diffusion terms, jump intensities, and risk premiums.

Appendix

Under the assumptions of semi-affine models, ds_t and dx_t evolve under risk-neutral probability, $\tilde{P}$, as

$$ds_t = \tilde{\mu}_s(x_t)dt + \sigma_s(x_t)d\tilde{W}_t + z_{st}dJ_t - \tilde{\nu}_s(\tilde{\lambda}_0 + \tilde{\Lambda}x_t)dt, \tag{32}$$

$$dx_t = \Gamma(\theta - x_t)dt + \sigma_x(x_t)d\tilde{W}_t + z_{xt}dJ_t - \tilde{\nu}_x(\tilde{\lambda}_0 + \tilde{\Lambda}x_t)dt, \tag{33}$$

where

$$\tilde{\mu}_s(x_t) = r_t - \frac{1}{2}(\sigma_0 + \sigma_1' x_t) - \tilde{E}(e^{z_{st}} - 1 - z_{st})(\tilde{\lambda}_0 + \tilde{\Lambda}x_t), \tag{34}$$

$$\sigma_s(x_t)\sigma_s(x_t)' = \sigma_0 + \sigma_1' x_t. \tag{35}$$

Proof of Proposition 1. The annualized quadratic variation of s_u between t and $t+\tau$ is

$$\frac{1}{\tau}\int_t^{t+\tau} \langle s_u, s_u \rangle du = \frac{1}{\tau}\int_t^{t+\tau} (\sigma_0 + \sigma_1' x_u)du + \frac{1}{\tau}\sum_{i=J_t+1}^{J_{t+\tau}} z_{st_i}^2, \tag{36}$$

where t_i is the time for the ith jump. Taking expectation of the both sides under the risk-neutral probability $\tilde{P}$ gives the variance-swap price $\tilde{V}_{\tau,t}$ as

$$\tilde{V}_{\tau,t} = \frac{1}{\tau}\tilde{E}_t\left[\int_t^{t+\tau} (\sigma_0 + \sigma_1' x_u)du\right] + \frac{1}{\tau}\tilde{E}(z_{st}^2)\tilde{E}_t\left[\int_t^{t+\tau} (\tilde{\lambda}_0 + \tilde{\Lambda}x_u)du\right]. \tag{37}$$

The terms $\frac{1}{\tau}\tilde{E}_t[\int_t^{t+\tau}(\sigma_0 + \sigma_1' x_u)du]$ and $\frac{1}{\tau}\tilde{E}_t[\int_t^{t+\tau}(\tilde{\lambda}_0 + \tilde{\Lambda}x_u)du]$ can be calculated as follows. From Eq. (33),

$$de^{\Gamma t}x_t = e^{\Gamma t}\Gamma\theta dt + e^{\Gamma t}\sigma_x(x_t)d\tilde{W}_t + e^{\Gamma t}z_{xt}dJ_t - e^{\Gamma t}\tilde{\nu}_x(\tilde{\lambda}_0 + \tilde{\Lambda}x_t)dt. \tag{38}$$

Integrating on both sides from t to $t+\tau$ and taking expectation gives,

$$e^{\Gamma(t+\tau)}\tilde{E}_t[x_{t+\tau}] = e^{\Gamma t}x_t + \int_t^{t+\tau} e^{\Gamma u}\Gamma\theta du = e^{\Gamma t}x_t + \left[e^{\Gamma(t+\tau)} - e^{\Gamma t}\right]\theta, \tag{39}$$

where the diffusion and compensated jump terms are dropped out because their expectations are zero. It follows that

$$\tilde{E}_t[x_{t+\tau}] = e^{-\Gamma\tau}x_t + (I_k - e^{-\Gamma\tau})\theta, \tag{40}$$

where I_k is the identity matrix. As a result,

$$\frac{1}{\tau}\tilde{E}_t\left[\int_t^{t+\tau}(\sigma_0+\sigma_1' x_u)du\right] = \frac{1}{\tau}\int_t^{t+\tau}[\sigma_0+\sigma_1'(e^{-\Gamma(u-t)}x_t+(I_k-e^{-\Gamma(u-t)})\theta)]du$$
$$= \sigma_0+\sigma_1'\theta+\sigma_1'(\tau\Gamma)^{-1}(I_k-e^{-\tau\Gamma})(x_t-\theta), \quad (41)$$

and similarly,

$$\frac{1}{\tau}\tilde{E}_t\left[\int_t^{t+\tau}(\tilde{\lambda}_0+\tilde{\Lambda}x_u)du\right] = \tilde{\lambda}_0+\tilde{\Lambda}\theta+\tilde{\Lambda}(\tau\Gamma)^{-1}(I_k-e^{-\tau\Gamma})(x_t-\theta). \quad (42)$$

The variance-swap price $\tilde{V}_{\tau,t}$ is then given by

$$\tilde{V}_{\tau,t} = \sigma_0+\sigma_1'\theta+\sigma_1'(\tau\Gamma)^{-1}(I_k-e^{-\tau\Gamma})(x_t-\theta)$$
$$+(\tilde{E}z_{st}^2)\left[\tilde{\lambda}_0+\tilde{\Lambda}\theta+\tilde{\Lambda}(\tau\Gamma)^{-1}(I_k-e^{-\tau\Gamma})(x_t-\theta)\right], \quad (43)$$

which is affine in x_t.

Proof of Proposition 2. As shown by Carr and Madan (1998), any twice-continuous differentiable payoff function, $H(u)$, can be written as

$$H(u)=H(\overline{u})+(u-\overline{u})H_u(\overline{u})+\int_{\overline{u}}^{\infty}H_{uu}(K)(u-K)^+dK+\int_0^{\overline{u}}H_{uu}(K)(K-u)^+dK. \quad (44)$$

Let $H(S_{t+\tau})=\ln(S_{t+\tau}/S_t)$, then,

$$\ln\left(\frac{S_{t+\tau}}{S_t}\right) = \ln\left(\frac{\overline{S}}{S_t}\right)+\frac{S_{t+\tau}-\overline{S}}{\overline{S}}$$
$$-\left[\int_{\overline{S}}^{\infty}\frac{1}{K^2}(S_{t+\tau}-K)^+dK+\int_0^{\overline{S}}\frac{1}{K^2}(K-S_{t+\tau})^+dK\right], \quad (45)$$

where $a^+=\max(a,0)$. Taking expectation on both sides and letting $\overline{S}=F_{\tau,t}=S_te^{r\tau}$, where r is the riskfree rate and $F_{\tau,t}$ is the time t forward price with maturity at $t+\tau$, gives

$$\tilde{E}_t\left[\ln\left(\frac{S_{t+\tau}}{S_t}\right)\right]=r\tau-e^{r\tau}\left[\int_{F_{\tau,t}}^{\infty}\frac{1}{K^2}c_{\tau,t}(K)dK+\int_0^{F_{\tau,t}}\frac{1}{K^2}p_{\tau,t}(K)dK\right], \quad (46)$$

where $c_{\tau,t}(K)=e^{-r\tau}\tilde{E}_t(S_{t+\tau}-K)^+$ and $p_{\tau,t}(K)=e^{-r\tau}\tilde{E}_t(K-S_{t+\tau})^+$ are prices of calls and puts, respectively, with maturity $t+\tau$ and strike price K. The second term multiplied by

$2/\tau$ is the model-free formula for variance-swap price, $V_{\tau,t}$. We now calculate $\tilde{E}_t[\ln(\frac{S_{t+\tau}}{S_t})]$. Under the process Eq. (32)-(35),

$$\begin{aligned}
& \tilde{E}_t\left[\ln\left(\frac{S_{t+\tau}}{S_t}\right)\right] = \tilde{E}_t[s_{t+\tau} - s_t] \\
= \;& r\tau - \tilde{E}_t\left[\int_t^{t+\tau} \frac{\sigma_0 + \sigma_1' x_u}{2} du\right] - \tilde{E}(e^{z_{st}} - 1 - z_{st})\tilde{E}_t\left[\int_t^{t+\tau} (\tilde{\lambda}_0 + \tilde{\Lambda} x_u) du\right] \\
= \;& r\tau - \frac{\tau}{2}[\sigma_0 + \sigma_1'\theta + \sigma_1'(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta)] \\
& - \tau\tilde{E}(e^{z_{st}} - 1 - z_{st})\left[\tilde{\lambda}_0 + \tilde{\Lambda}\theta + \tilde{\Lambda}(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta)\right]. \qquad (47)
\end{aligned}$$

Using Eq. (46) and Eq. (47), the model-free variance-swap price can be written as

$$\begin{aligned}
V_{\tau,t} &= \frac{2}{\tau} e^{r\tau}\left[\int_{F_{\tau,t}}^{\infty} \frac{1}{K^2} c_{\tau,t}(K) dK + \int_0^{F_{\tau,t}} \frac{1}{K^2} p_{\tau,t}(K) dK\right] \\
&= 2r - \frac{2}{\tau}\tilde{E}_t\left[\ln\left(\frac{S_{t+\tau}}{S_t}\right)\right] \\
&= \sigma_0 + \sigma_1'\theta + \sigma_1'(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta) \\
&\quad + 2\tilde{E}(e^{z_{st}} - 1 - z_{st})\left[\tilde{\lambda}_0 + \tilde{\Lambda}\theta + \tilde{\Lambda}(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta)\right]. \qquad (48)
\end{aligned}$$

From Eq. (43) and Eq. (48), the approximation error of the model-free formula of the variance-swap price is

$$\begin{aligned}
& V_{\tau,t} - \tilde{V}_{\tau,t} \\
= \;& 2\tilde{E}\left(e^{z_{st}} - 1 - z_{st} - z_{st}^2/2\right)\left[\tilde{\lambda}_0 + \tilde{\Lambda}\theta + \tilde{\Lambda}(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta)\right] \\
= \;& 2\left[\frac{\tilde{E}z_{st}^3}{3!} + \frac{\tilde{E}z_{st}^4}{4!} + \cdots\right]\left[\tilde{\lambda}_0 + \tilde{\Lambda}\theta + \tilde{\Lambda}(\tau\Gamma)^{-1}(I_k - e^{-\tau\Gamma})(x_t - \theta)\right]. \qquad (49)
\end{aligned}$$

The last step follows from expanding $e^{z_{st}}$ around $z_{st} = 0$ using Taylor series expansion and taking expectation. The last two equations show that the approximation error is affine in x_t and that the leading term in the approximation error is proportional to the third moment of the jump size, whereas from Eq. (43) the second term in $\tilde{V}_{\tau,t}$ is proportional to the second moment of the jump size.

References

Aït-Sahalia, Y., 1996a. Nonparametric pricing of interest rate derivative securities. Econometrica 64, 527-560.

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

Aït-Sahalia, Y., Duarte, J., 2003. Nonparametric option pricing under shape restrictions. Journal of Econometrics 116, 9-47.

Aït-Sahalia, Y., Lo, A. W., 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance 53, 499-547.

Aït-Sahalia, Y., Park, J. Y., 2012. Stationary-based specification tests for diffusions when the process is nonstationary. Journal of Econometrics, Forthcoming.

Aït-Sahalia, Y., Wang, Y., Yared, F., 2001. Do option markets correctly price the probabilities of movement of the underlying asset?. Journal of Econometrics 102, 67-110.

Bakshi, G., Cao. C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003-2049.

Bates, D. S., 2000. Post-'87 crash fears in the S&P 500 futures option market. Journal of Econometrics 94, 181-238.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-659.

Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. Journal of Finance 62, 1453-1490.

Broadie, M., Detemple, J., 2004. Option pricing: Valuation models and applications. Management Science 50, 1145-1177.

Broadie, M., Detemple, J., Ghysels, E., Torres, O., 2000a. American options with stochastic dividends and volatility: A nonparametric investigation. Journal of Econometrics 94, 53-92.

Broadie, M., Detemple, J., Ghysels, E., Torres, O., 2000b. Nonparametric estimation of American options' exercise boundaries and call prices. Journal of Economic Dynamics and Control 24, 1829-1857.

Carr, P., Madan, D., 1998. Towards a theory of volatility trading. In Jarrow, R. (Eds.), Risk Book on Volatility. Risk, New York, pp. 417-427.

Carr, P., Wu, L., 2009. Variance risk premiums. Review of Financial Studies 22, 1311-1341.

Chapman, D. A., Pearson, N. D., 2000. Is the short rate drift actually nonlinear?. Journal of Finance 55, 355-388.

Chen, S. X., Gao, J., 2004. On the use of the kernel method for specification test of diffusion models. Unpublished working paper. Iowa State University.

Cheridito, P., Filipoviȁ, D., Kimmel, R., 2007. Market price of risk specifications for affine models: theory and evidence. Journal of Financial Economics 83, 123-170.

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

Chicago Board Options Exchange, 2003. VIX-CBOE volatility index. http://www.cboe.com/micro/vix/vixwhite.pdf.

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

Christoffersen, P., Jacobs, K., Mimouni, K., 2010. Volatility dynamics for the S&P 500: Evidence from realized volatility, daily returns, and option prices. Review of Financial Studies 23, 3141-3189.

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

Corradi, V., Swanson, N. R., 2005. Bootstrap specification tests for diffusion processes. Journal of Econometrics 124, 117-148.

Demeterfi, K., Derman, E., Kamal, M., Zou, J., 1999. A guide to volatility and variance swaps. Journal of Derivatives 6, 9-32.

Duan, J., Yeh, C., 2010. Jump and volatility risk premiums implied by VIX. Journal of Economic Dynamics and Control 34, 2232-2244.

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

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

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 59, 1367-1403.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269-1300.

Fan, Y., Li, Q., 1996. Consistent model specification tests: Omitted variables and semiparametric functional forms. Econometrica 64, 865-890.

Garcia, R., Ghysels, E., Renault, E., 2010. The econometrics of option pricing. In Aït-Sahalia, Y., Hansen., L. P. (Eds.), Handbook of Financial Econometrics, Vol. 1. North-Holland, Amsterdam, pp. 479-552.

Gu, J., Li, D., Liu, D., 2007. Bootstrap non-parametric significance test. Journal of Nonparametric Statistics 19, 215-230.

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

Heston, S. L., Nandi, S., 2000. A closed-form GARCH option valuation model. Review of Financial Studies 13, 585-625.

Hong, Y., Li, H., 2005. Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. Review of Financial Studies 18, 37-84.

Hutchinson, J. M., Lo, A. W., Poggio, T., 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. Journal of Finance 49, 851-889.

Jiang, G. J., Oomen, R. C. A., 2008. Testing for jumps when asset prices are observed with noise - a “swap variance” approach. Journal of Econometrics 144, 352-370.

Jiang, G. J., Tian, Y. S., 2007. Extracting model-free volatility from option prices: An examination of the VIX index. Journal of Derivatives 15(2), 35-60.

Jones, C. S., 2003. The dynamics of stochastic volatility: Evidence from underlying and option markets. Journal of Econometrics 116, 181-224.

Kunsch, H. R., 1989. The jackknife and the bootstrap for general stationary observations. Annals of Statistics 17, 1217-1241.

Li, D., 2005. A simple bootstrap test for time series regression models. Journal of Nonparametric Statistics 17, 513-520.

Li, G., Zhang, C., 2010. On the number of state variables in options pricing. Management Science 55, 2058-2075.

Li, H., Zhao, F., 2009. Nonparametric estimation of state-price densities implicit in interest rate cap prices. Review of Financial Studies 22, 4335-4376.

Li, Q., Racine, J. S., 2007. Nonparametric econometrics. Princeton University Press, New Jersey.

Li, Q., Wang, S., 1998. A simple consistent bootstrap test for a parametric regression function. Journal of Econometrics 1998, 145-165.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

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

Pritsker, M., 1998. Nonparametric density estimation and tests of continuous time interest rate models. Review of Financial Studies 11, 449-487.

Stanton, R., 1997. A nonparametric model of term structure dynamics and the market price of interest rate risk. Journal of Finance 52, 1973-2002.

Zheng, X., 1996. A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics 75, 263-289.

Table 1
Summary statistics of the V_τs and ΔV_τs

This table presents the mean, standard deviation, skewness, and the 1st, 5th, 50th, 95th, and 99th percentiles of the empirical distribution of the daily observations of the 1-month variance-swap price $V_{1,t}$, the 18-month variance-swap price $V_{18,t}$, and their first differences, $\Delta V_{1,t+1}$ and $\Delta V_{18,t+1}$, where $\Delta V_{1,t+1} = V_{1,t+1} - V_{1,t}$ and $\Delta V_{18,t+1} = V_{18,t+1} - V_{18,t}$. It also reports the autocorrelations of $V_{1,t}$ and $V_{18,t}$, as well as the p-values of the augmented Dickey-Fuller test for unit root. The sample period is from January 1996 to September 2008.

A. Summary statistics and tests of V_τs								
	Mean	Std	Skew	1P	5P	50P	95P	99P
V_1	0.0470	0.0306	1.6284	0.0109	0.0133	0.0410	0.1061	0.1581
V_{18}	0.0446	0.0195	0.9317	0.0179	0.0206	0.0430	0.0812	0.0977
	Autocorrelations							ADF
	ρ_1	ρ_2	ρ_3	ρ_5	ρ_{10}	ρ_{20}	ρ_{30}	p-value
V_1	0.9697	0.9449	0.9259	0.8918	0.8303	0.7215	0.6313	0.0011
V_{18}	0.9895	0.9835	0.9767	0.9630	0.9315	0.8743	0.8340	0.0270
B. Summary statistics of ΔV_τs								
	Mean	Std	Skew	1P	5P	50P	95P	99P
ΔV_1	0.0000	0.0075	0.8567	-0.0207	-0.0100	-0.0001	0.0109	0.0233
ΔV_{18}	0.0000	0.0028	0.8531	-0.0067	-0.0030	0.0000	0.0031	0.0072

Table 2
Nonparametric bootstrap tests of the affine property of the conditional mean, variance, and covariance of $(\Delta V_1, \Delta V_{18})$

This table reports the p-values of bootstrap tests of the affine property of the conditional mean, conditional variance, and conditional covariance of the first differences of the 1-month and 18-month variance-swap prices, ΔV_1 and ΔV_{18}. The first column indicates the functional forms under the null hypothesis that are affine in at least one component of (V_1, V_{18}). The functional forms under alternative hypothesis are unrestricted and estimated nonparametrically. μ_1 is the conditional mean of ΔV_1, μ_{18} is the conditional mean of ΔV_{18}, σ_1^2 is the conditional variance of ΔV_1, σ_{18}^2 is the conditional variance of ΔV_{18}, and $\sigma_{1,18}$ is the conditional covariance between ΔV_1 and ΔV_{18}. The test is based on 100 bootstrap samples. Panel A reports results based on local constant kernel regressions and Panel B reports results based on local linear kernel regressions. The sample period is from January 1996 to September 2008.

A. Local constant kernel regressions

	Independent case				
	$\mu_1(V_1)$	$\mu_{18}(V_{18})$	$\sigma_1^2(V_1)$	$\sigma_{18}^2(V_{18})$	
$a + bV_1$	0.05		0.00		
$a + bV_{18}$		0.00		0.00	
	Dependent case				
	$\mu_1(V_1, V_{18})$	$\mu_{18}(V_1, V_{18})$	$\sigma_1^2(V_1, V_{18})$	$\sigma_{18}^2(V_1, V_{18})$	$\sigma_{1,18}(V_1, V_{18})$
$a + b_1V_1 + b_2V_{18}$	0.03	0.00	0.00	0.04	0.00
$g(V_1) + bV_{18}$	0.28	0.00	0.47	0.00	0.00
$bV_1 + g(V_{18})$	0.60	0.24	0.02	0.09	0.44

B. Local linear kernel regressions

	Independent case				
	$\mu_1(V_1)$	$\mu_{18}(V_{18})$	$\sigma_1^2(V_1)$	$\sigma_{18}^2(V_{18})$	
$a + bV_1$	0.05		0.00		
$a + bV_{18}$		0.00		0.01	
	Dependent case				
	$\mu_1(V_1, V_{18})$	$\mu_{18}(V_1, V_{18})$	$\sigma_1^2(V_1, V_{18})$	$\sigma_{18}^2(V_1, V_{18})$	$\sigma_{1,18}(V_1, V_{18})$
$a + b_1V_1 + b_2V_{18}$	0.03	0.00	0.00	0.02	0.03
$g(V_1) + bV_{18}$	0.22	0.00	0.78	0.01	0.04
$bV_1 + g(V_{18})$	0.29	0.12	0.04	0.12	0.30

Table 3
Rejection rates of the bootstrap test with simulated data (local constant)

This table shows the percentage rejection rates of the bootstrap test of the null hypothesis of affine property for 500 simulated data series. The local constant kernel regression method is used in the estimation. The left panels are for the independent case where the data are simulated from the model

$$\check{V}_{1,t+1} = \phi_0 + (\phi_1 + 1)\check{V}_{1,t} + \sqrt{\phi_2 \check{V}_{1,t}^{\phi_3}}\varepsilon_{t+1}.$$

The right panels are for the dependent case where the data are generated by the models

$$\begin{aligned} \check{V}_{1,t+1} &= \psi_{0,1} + (\psi_{1,1} + 1)\check{V}_{1,t} + \psi_{2,1}\check{V}_{18,t} + \sqrt{\psi_{3,1}\check{V}_{1,t}^{\psi_{4,1}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,1}\check{V}_{18,t}^{\psi_{6,1}}}\varepsilon_{2,t+1} \\ \check{V}_{18,t+1} &= \psi_{0,18} + \psi_{1,18}\check{V}_{1,t} + (\psi_{2,18} + 1)\check{V}_{18,t} + \sqrt{\psi_{3,18}\check{V}_{1,t}^{\psi_{4,18}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,18}\check{V}_{18,t}^{\psi_{6,18}}}\varepsilon_{2,t+1}. \end{aligned}$$

The results are reported for the optimal bandwidth h^*, under-smoothed bandwidth, $h^*/1.5$, and over-smoothed bandwidth, $1.5h^*$, for the sample sizes of 1000, 2000 and 3208, and for 5% and 10% significance levels.

	Independent case						Dependent case					
	$H_0: \mu_1(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 = 1$)						$H_0: \mu_1(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1} = 1,\ \psi_{6,1} = 1,\ \psi_{4,18} = 1,\ \psi_{6,18} = 1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	6.4	11.0	6.2	11.8	5.4	12.2	5.6	10.8	4.6	10.8	6.0	11.0
2000	4.6	10.6	5.6	11.0	5.0	11.2	5.6	10.6	6.0	11.4	6.4	11.6
3208	3.6	10.2	3.4	10.6	4.4	9.8	4.6	9.8	4.2	8.6	5.2	9.6
	$H_0: \sigma_1^2(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 = 1$)						$H_0: \sigma_1^2(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1} = 1,\ \psi_{6,1} = 1,\ \psi_{4,18} = 1,\ \psi_{6,18} = 1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	7.0	10.0	6.6	10.4	6.8	12.0	5.0	9.4	5.2	9.8	6.4	10.2
2000	6.8	11.8	7.0	11.8	7.2	12.6	5.4	10.8	5.4	10.4	6.2	11.0
3208	5.8	11.0	6.4	11.4	6.8	11.8	4.8	8.6	4.4	9.8	4.4	11.2
	$H_0: \mu_1(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 \neq 1$)						$H_0: \mu_1(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1} \neq 1,\ \psi_{6,1} \neq 1,\ \psi_{4,18} \neq 1,\ \psi_{6,18} \neq 1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	7.2	15.0	9.0	16.4	9.6	17.6	3.8	9.8	4.2	10.8	5.4	12.2
2000	6.8	14.2	7.6	16.2	9.0	16.8	5.2	10.2	5.8	11.0	7.4	12.8
3208	6.2	12.8	6.0	13.4	7.2	13.2	4.2	10.4	4.8	11.2	6.6	12.6

Table 4
Rejection rates of the bootstrap test with simulated data (local linear)

This table shows the percentage rejection rates of the bootstrap test of the null hypothesis of affine property for 500 simulated data series. The local linear kernel regression method is used in the estimation. The left panels are for the independent case where the data are simulated from the model

$$\check{V}_{1,t+1} = \phi_0 + (\phi_1 + 1)\check{V}_{1,t} + \sqrt{\phi_2 \check{V}_{1,t}^{\phi_3}}\varepsilon_{t+1}.$$

The right panels are for the dependent case where the data are generated by the models

$$\begin{aligned}\check{V}_{1,t+1} &= \psi_{0,1} + (\psi_{1,1}+1)\check{V}_{1,t} + \psi_{2,1}\check{V}_{18,t} + \sqrt{\psi_{3,1}\check{V}_{1,t}^{\psi_{4,1}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,1}\check{V}_{18,t}^{\psi_{6,1}}}\varepsilon_{2,t+1} \\ \check{V}_{18,t+1} &= \psi_{0,18} + \psi_{1,18}\check{V}_{1,t} + (\psi_{2,18}+1)\check{V}_{18,t} + \sqrt{\psi_{3,18}\check{V}_{1,t}^{\psi_{4,18}}}\varepsilon_{1,t+1} + \sqrt{\psi_{5,18}\check{V}_{18,t}^{\psi_{6,18}}}\varepsilon_{2,t+1}.\end{aligned}$$

The results are reported for the optimal bandwidth h^*, under-smoothed bandwidth, $h^*/1.5$, and over-smoothed bandwidth, $1.5h^*$, for the sample sizes of 1000, 2000 and 3208, and for 5% and 10% significance levels.

	Independent case						Dependent case					
	$H_0: \mu_1(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 = 1$)						$H_0: \mu_1(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1}=1$, $\psi_{6,1}=1$, $\psi_{4,18}=1$, $\psi_{6,18}=1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	7.6	12.6	8.2	13.2	8.4	13.0	5.6	9.2	5.0	8.6	5.6	11.2
2000	5.6	13.4	5.8	11.8	5.6	11.6	5.8	12.4	7.4	13.6	6.2	13.0
3208	5.2	11.0	5.6	11.2	5.4	11.0	5.4	10.6	6.0	11.2	5.8	11.6
	$H_0: \sigma_1^2(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 = 1$)						$H_0: \sigma_1^2(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1}=1$, $\psi_{6,1}=1$, $\psi_{4,18}=1$, $\psi_{6,18}=1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	7.8	15.0	8.2	14.4	9.4	15.6	6.0	10.4	6.0	10.2	6.2	10.6
2000	7.6	11.8	8.6	12.6	7.6	12.6	4.8	9.4	4.6	10.2	5.6	10.6
3208	6.8	11.4	6.6	11.8	6.2	11.8	5.2	9.8	5.6	10.0	5.6	10.4
	$H_0: \mu_1(\check{V}_1) = a + b\check{V}_1$ ($\phi_3 \neq 1$)						$H_0: \mu_1(\check{V}_1, \check{V}_{18}) = a + b_1\check{V}_1 + b_2\check{V}_{18}$ ($\psi_{4,1}\neq 1$, $\psi_{6,1}\neq 1$, $\psi_{4,18}\neq 1$, $\psi_{6,18}\neq 1$)					
	$h^*/1.5$		h^*		$1.5h^*$		$h^*/1.5$		h^*		$1.5h^*$	
T	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1000	6.2	11.4	6.0	11.0	6.6	11.0	4.0	10.4	4.4	9.8	4.4	12.4
2000	6.6	13.4	7.2	11.8	7.0	12.4	6.0	10.4	7.0	11.4	7.2	13.4
3208	6.0	11.6	6.4	12.0	6.8	11.8	4.8	10.2	6.2	10.6	6.8	11.2

Table 5
Parametric tests of the affine property of the conditional mean and variance (independent case)
The parameters for the non-affine conditional mean function

$$\Delta V_{\tau_j,t+1} = \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{\tau_j,t} + \alpha_{2,\tau_j} \widetilde{V^2_{\tau_j,t}} + \alpha_{3,\tau_j} \widetilde{1/V_{\tau_j,t}} + \epsilon_{\tau_j,t+1}, \quad \tau_j = 1, 18,$$

and the non-affine conditional variance function

$$\hat{\eta}^2_{\tau_j,t+1} = (\beta_{1,\tau_j} V_{\tau_j,t})^{\gamma_{\tau_j}} + \varsigma_{\tau_j,t+1}, \quad \tau_j = 1, 18$$

are estimated by weighted least squares, where the weight is the reciprocal of the nonparametric estimate of the variance of dependent variables, and $\hat{\eta}_{\tau_j,t+1}$ is the fitted residual from the nonparametric regression of the conditional mean function

$$\Delta V_{\tau_j,t+1} = \mu_{\tau_j}(V_{\tau_j,t}) + \eta_{\tau_j,t+1}, \quad \tau_j = 1, 18.$$

$\widetilde{V^2_{\tau_j,t}}$ is the residual of $V^2_{\tau_j,t}$ regressed on $V_{\tau_j,t}$ and a constant. $\widetilde{1/V_{\tau_j,t}}$ is the residual of $1/V_{\tau_j,t}$ regressed on $V_{\tau_j,t}$, $\widetilde{V^2_{\tau_j,t}}$ and a constant. The t-statistics adjusted for heteroskedasticity and 24 lags of autocorrelation using Newey and West (1987) are reported in parentheses. The sample period is from January 1996 to September 2008.

A. Conditional mean of ΔV_1			
$\alpha_{0,1} \times 10^3$	$\alpha_{1,1}$	$\alpha_{2,1}$	$\alpha_{3,1} \times 10^5$
1.5817	-0.0344	-0.5433	1.2103
(6.23)	(-4.97)	(-4.73)	(1.70)
B. Conditional mean of ΔV_{18}			
$\alpha_{0,18} \times 10^3$	$\alpha_{1,18}$	$\alpha_{2,18}$	$\alpha_{3,18} \times 10^5$
0.4303	-0.0090	-0.3847	3.9728
(2.61)	(-2.08)	(-2.24)	(2.49)
C. Conditional variance of ΔV_1			
$\beta_{1,1}$	$\gamma_1 - 1$		
0.1756	1.1643		
(4.48)	(11.08)		
D. Conditional variance of ΔV_{18}			
$\beta_{1,18}$	$\gamma_{18} - 1$		
0.0986	1.2557		
(2.13)	(6.44)		

Table 6
Parametric tests of the affine property of the conditional mean, variance, and covariance (dependent case)
The parameters for the non-affine conditional mean function

$$\begin{aligned}\Delta V_{\tau_j,t+1} &= \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{1,t} + \alpha_{2,\tau_j} \widetilde{V_{18,t}} + \alpha_{3,\tau_j} \widetilde{V_{1,t}^2} \\ &\quad + \alpha_{4,\tau_j} \widetilde{V_{18,t}^2} + \alpha_{5,\tau_j} (\widetilde{1/V_{1,t}}) + \alpha_{6,\tau_j} (\widetilde{1/V_{18,t}}) + \epsilon_{\tau_j,t+1}, \quad \tau_j = 1, 18,\end{aligned}$$

and the non-affine conditional variance and covariance functions

$$\begin{aligned}\hat{\eta}^2_{\tau_j,t+1} &= (\beta_{1,\tau_j} V_{1,t} + \beta_{2,\tau_j} V_{18,t})^{\gamma_{\tau_j}} + \varsigma_{\tau_j,t+1}, \\ \hat{\eta}_{1,t+1}\hat{\eta}_{18,t+1} &= (\beta_{1,1,18} V_{1,t} + \beta_{2,1,18} V_{18,t})^{\gamma_{1,18}} + \varsigma_{1,18,t+1}, \quad \tau_j = 1, 18\end{aligned}$$

are estimated by weighted least squares, where the weight is the reciprocal of the nonparametric estimate of the variance of dependent variables, and $\hat{\eta}_{\tau_j,t+1}$ is the fitted residual from the nonparametric regression of the conditional mean function

$$\Delta V_{\tau_j,t+1} = \mu_{\tau_j}(V_{1,t}, V_{18,t}) + \eta_{\tau_j,t+1}, \quad \tau_j = 1, 18.$$

$\widetilde{V_{18,t}}$ is the residual of $V_{18,t}$ regressed on $V_{1,t}$ and a constant. $\widetilde{V_{1,t}^2}$ is the residual of $V_{1,t}^2$ regressed on $V_{1,t}$, $\widetilde{V_{18,t}}$ and a constant. Other variables, $\widetilde{V_{18,t}^2}$, $(\widetilde{1/V_{1,t}})$ and $(\widetilde{1/V_{18,t}})$ are defined similarly. The t-statistics adjusted for heteroskedasticity and 24 lags of autocorrelation using Newey and West (1987) are reported in parentheses. The sample period is from January 1996 to September 2008.

A. Conditional mean of ΔV_1

$\alpha_{0,1} \times 10^3$	$\alpha_{1,1}$	$\alpha_{2,1}$	$\alpha_{3,1}$	$\alpha_{4,1}$	$\alpha_{5,1} \times 10^5$	$\alpha_{6,1} \times 10^5$
1.3237	-0.0280	0.0314	-0.3438	-0.5890	1.2278	7.6973
(5.59)	(-4.27)	(3.05)	(-2.87)	(-1.77)	(1.35)	(2.27)

B. Conditional mean of ΔV_{18}

$\alpha_{0,18} \times 10^3$	$\alpha_{1,18}$	$\alpha_{2,18}$	$\alpha_{3,18}$	$\alpha_{4,18}$	$\alpha_{5,18} \times 10^5$	$\alpha_{6,18} \times 10^5$
0.1247	-0.0023	-0.0149	-0.0889	-0.3178	0.4660	3.8034
(1.47)	(-1.05)	(-3.27)	(-1.88)	(-2.15)	(1.40)	(2.40)

C. Conditional variance of ΔV_1

$\beta_{1,1}$	$\beta_{2,1}$	$\gamma_1 - 1$
0.1252	-0.0228	0.9441
(3.56)	(-2.45)	(6.78)

D. Conditional variance of ΔV_{18}

$\beta_{1,18}$	$\beta_{2,18}$	$\gamma_{18} - 1$
0.0073	0.0451	1.0247
(1.78)	(0.97)	(3.48)

E. Conditional covariance between ΔV_1 and ΔV_{18}

$\beta_{1,1,18}$	$\beta_{2,1,18}$	$\gamma_{1,18} - 1$
0.0359	0.0302	1.0987
(2.27)	(1.35)	(5.36)

Table 7

Nonparametric tests of the affine property of the conditional mean, variance, and covariance for $(\Delta V_1, \Delta V_9, \Delta V_{18})$

This table reports the p-values of bootstrap tests of the affine property of the conditional mean, conditional variance, and conditional covariance of the first differences of the 1-month, 9-month, and 18-month variance-swap prices, ΔV_1, ΔV_9 and ΔV_{18}. The first column indicates the functional forms under the null hypothesis in which at most one component of (V_1, V_9, V_{18}) is nonparametric. The alternative hypothesis is nonparametric. μ_1 is for the conditional mean of ΔV_1, μ_9 for the conditional mean of ΔV_9, and μ_{18} for the conditional mean of ΔV_{18}. σ_1^2 is for the conditional variance of ΔV_1, σ_9^2 for the conditional variance of ΔV_9, and σ_{18}^2 for the conditional variance of ΔV_{18}. $\sigma_{1,9}$ is for the conditional covariance between ΔV_1 and ΔV_9, $\sigma_{1,18}$ for the conditional covariance between ΔV_1 and ΔV_{18}, and $\sigma_{9,18}$ for the conditional covariance between ΔV_9 and ΔV_{18}. The test is based on 100 bootstrap samples. Panel A reports results based on local constant kernel regressions and Panel B reports results based on local linear kernel regressions. The sample period is from January 1996 to September 2008.

A. Local constant kernel regressions						
	Conditional mean					
	μ_1	μ_9	μ_{18}			
$a + b_1V_1 + b_2V_9 + b_3V_{18}$	0.01	0.04	0.01			
$g(V_1) + b_1V_9 + b_2V_{18}$	0.46	0.57	0.00			
$b_1V_1 + g(V_9) + b_2V_{18}$	0.75	0.26	0.05			
$b_1V_1 + b_2V_9 + g(V_{18})$	0.57	0.74	0.05			
	Conditional variance and conditional covariance					
	σ_1^2	σ_9^2	σ_{18}^2	$\sigma_{1,9}$	$\sigma_{1,18}$	$\sigma_{9,18}$
$a + b_1V_1 + b_2V_9 + b_3V_{18}$	0.00	0.00	0.01	0.00	0.00	0.00
$g(V_1) + b_1V_9 + b_2V_{18}$	0.56	0.42	0.00	0.12	0.01	0.00
$b_1V_1 + g(V_9) + b_2V_{18}$	0.00	0.07	0.00	0.18	0.29	0.39
$b_1V_1 + b_2V_9 + g(V_{18})$	0.00	0.00	0.15	0.03	0.42	0.67
B. Local linear kernel regressions						
	Conditional mean					
	μ_1	μ_9	μ_{18}			
$a + b_1V_1 + b_2V_9 + b_3V_{18}$	0.02	0.05	0.01			
$g(V_1) + b_1V_9 + b_2V_{18}$	0.64	0.46	0.00			
$b_1V_1 + g(V_9) + b_2V_{18}$	0.92	0.38	0.11			
$b_1V_1 + b_2V_9 + g(V_{18})$	0.55	0.61	0.17			
	Conditional variance and conditional covariance					
	σ_1^2	σ_9^2	σ_{18}^2	$\sigma_{1,9}$	$\sigma_{1,18}$	$\sigma_{9,18}$
$a + b_1V_1 + b_2V_9 + b_3V_{18}$	0.00	0.00	0.01	0.00	0.01	0.01
$g(V_1) + b_1V_9 + b_2V_{18}$	0.85	0.38	0.00	0.36	0.00	0.02
$b_1V_1 + g(V_9) + b_2V_{18}$	0.02	0.08	0.01	0.06	0.29	0.48
$b_1V_1 + b_2V_9 + g(V_{18})$	0.00	0.00	0.16	0.00	0.34	0.61

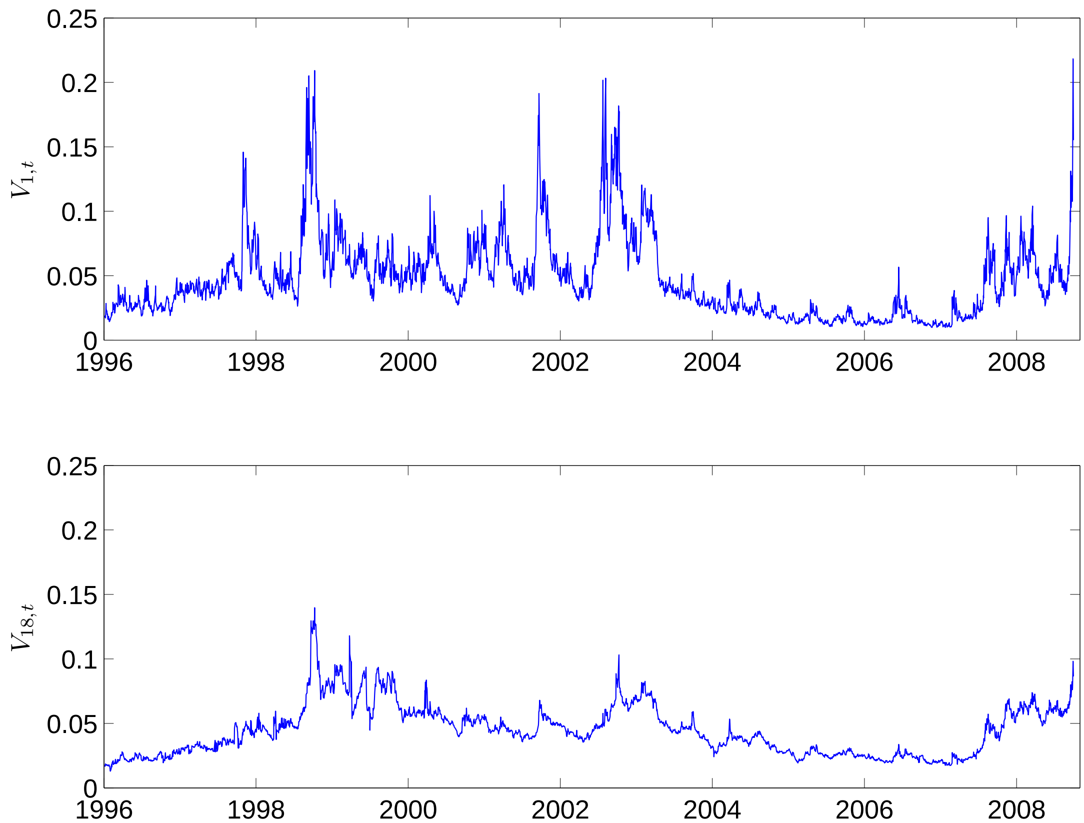

Fig. 1. Time-series plots of variance-swap prices

This figure shows the daily observations of the 1-month model-free variance-swap price V_1 and the 18-month model-free variance-swap price V_{18} from January 1996 to September 2008.

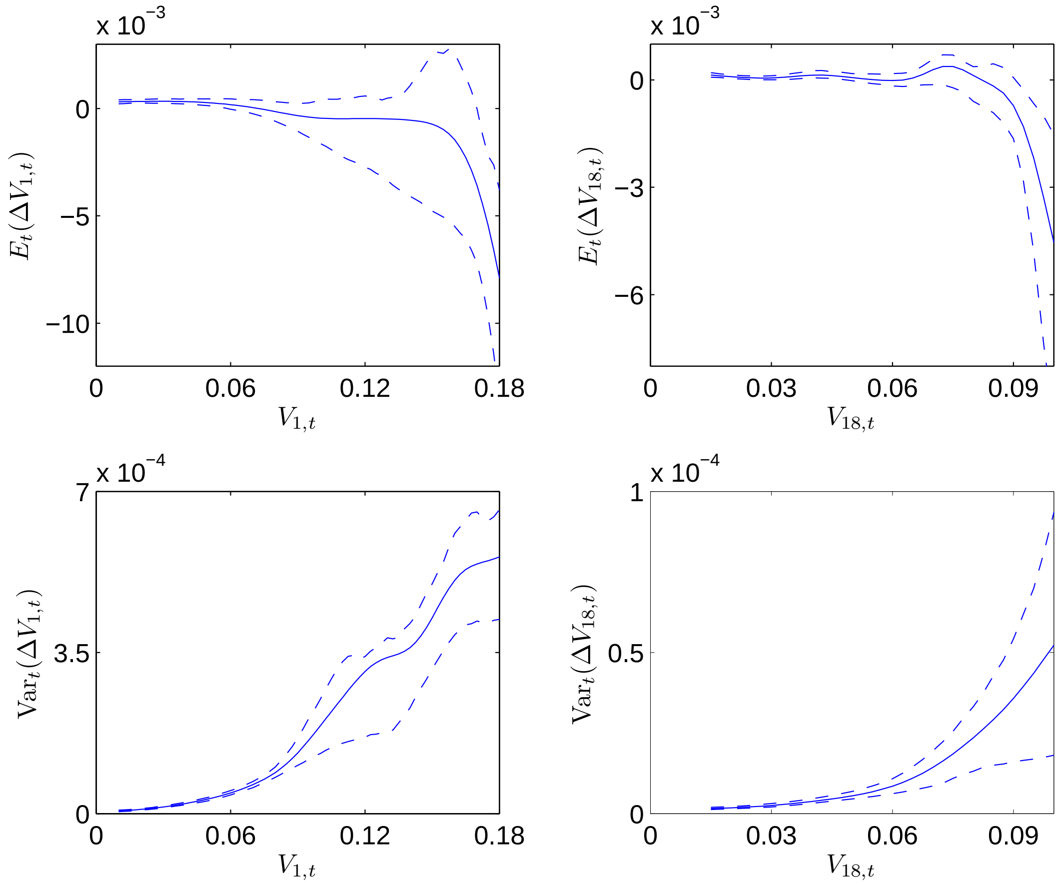

Fig. 2. Estimated conditional mean and conditional variance (independent case)

This figure shows the nonparametric estimation of the conditional mean and conditional variance for the first differences of the 1-month and 18-month variance-swap prices, ΔV_1 and ΔV_{18}. The solid line is the fitted curve. The dashed lines cover the 90% confidence interval. The two panels on the left are for ΔV_1 and the two panels on the right are for ΔV_{18}. The sample period is from January 1996 to September 2008.

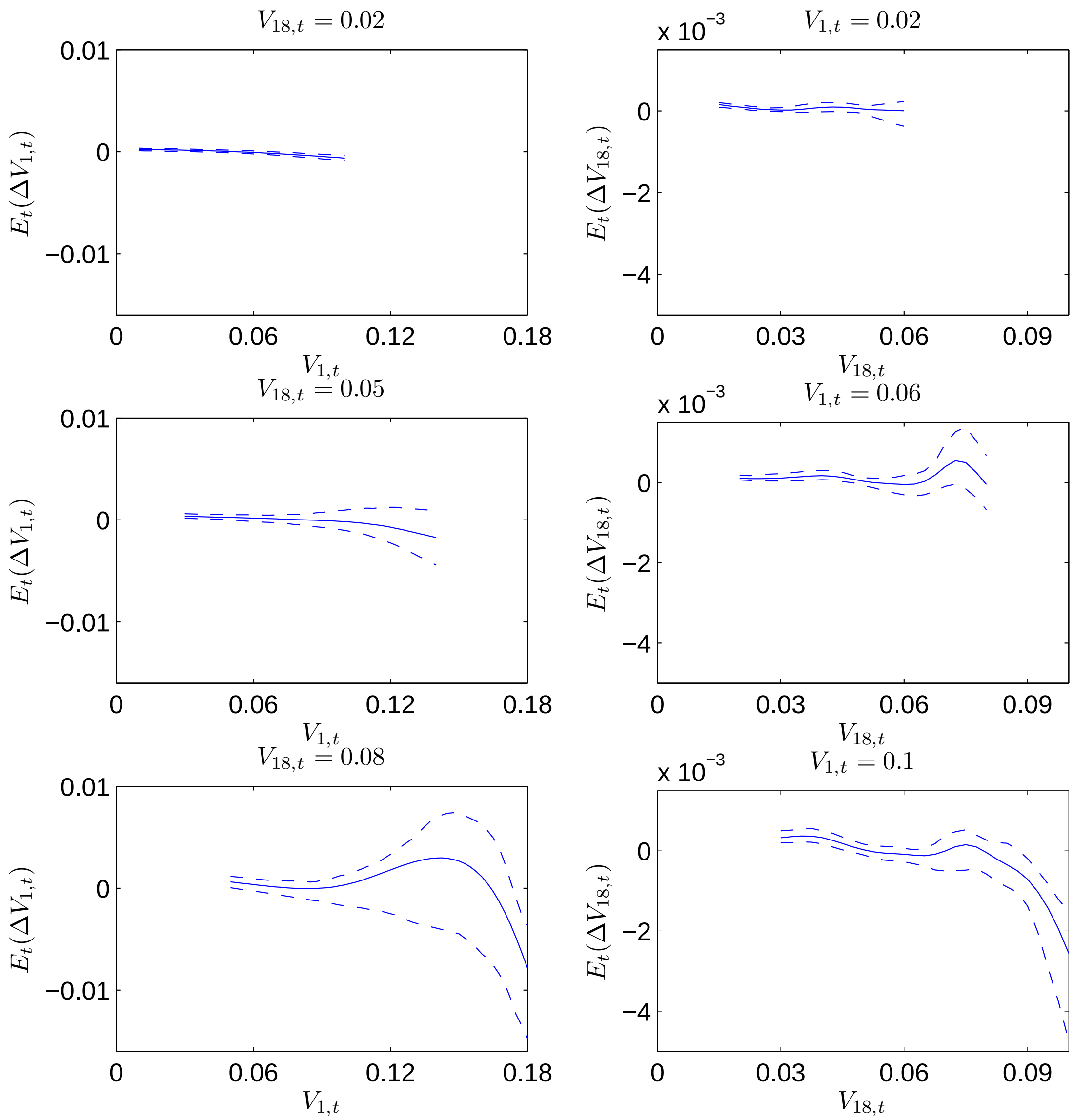

Fig. 3. Estimated conditional mean (dependent case)

This figure shows the nonparametric estimation of the conditional mean for the first differences of the 1-month and 18-month variance-swap prices, ΔV_1 and ΔV_{18}. The left panels are for the conditional mean of ΔV_1 as a function of V_1 for different levels of V_{18}, and the right panels are for the conditional mean of ΔV_{18} as a function of V_{18} for different levels of V_1. The solid line is the fitted curve. The dashed lines cover the 90% confidence interval. The sample period is from January 1996 to September 2008.

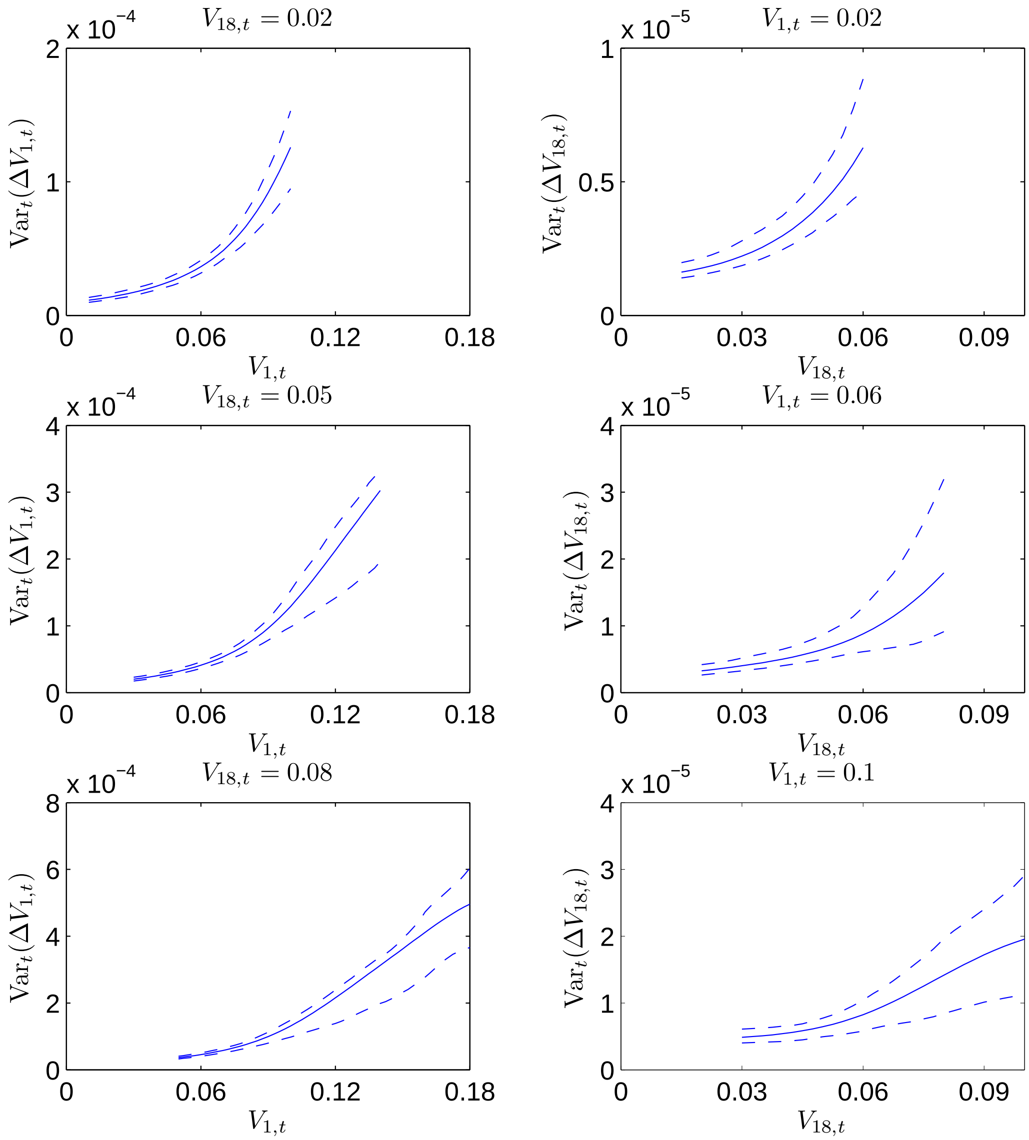

Fig. 4. Estimated conditional variance (dependent case)

This figure shows the nonparametric estimation of the conditional variance for the first differences of the 1-month and 18-month variance-swap prices, ΔV_1 and ΔV_{18}. The left panels are for the conditional variance of ΔV_1 as a function of V_1 for different levels of V_{18}, and the right panels are for the conditional variance of ΔV_{18} as a function of V_{18} for different levels of V_1. The solid line is the fitted curve. The dashed lines cover the 90% confidence interval. The sample period is from January 1996 to September 2008.

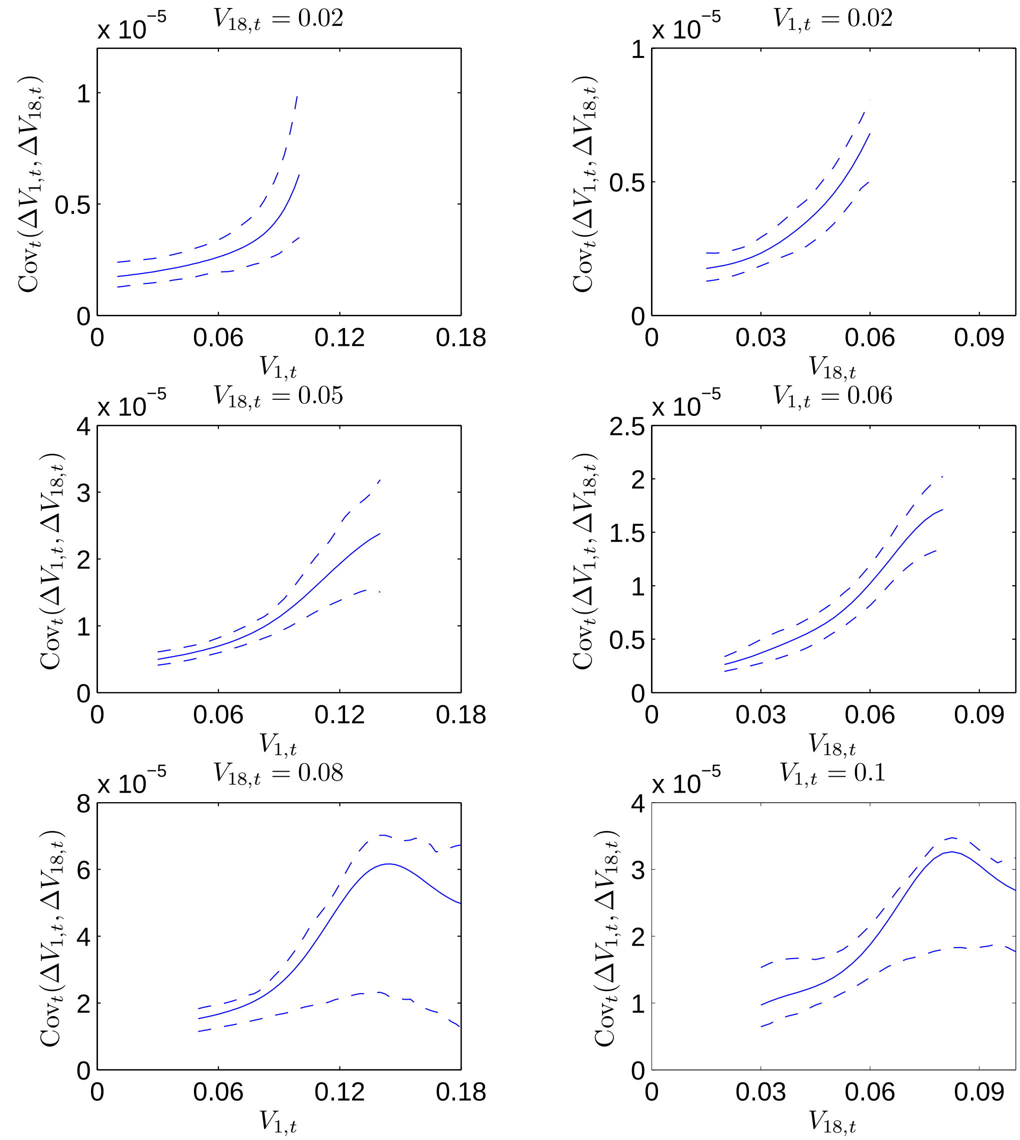

Fig. 5. Estimated conditional covariance (dependent case)

This figure shows the nonparametric estimation of the conditional covariance between the first differences of the 1-month and 18-month variance-swap prices, ΔV_1 and ΔV_{18}. The left panels show the conditional covariance as a function of V_1 for different levels of V_{18} and the right panels show the conditional covariance as a function of V_{18} for different levels of V_1. The solid line is the fitted curve. The dashed lines cover the 90% confidence interval. The sample period is from January 1996 to September 2008.