

On the Number and Dynamic Features of State Variables in Options Pricing

Gang Li and Chu Zhang*

August 2008

*University of Macao and Hong Kong University of Science and Technology. We would like to thank Jinchuan Duan, Nengjiu Ju, Sophie Ni, Liuren Wu, and participants of Quantitative Methods in Finance 2008 conference for helpful comments on an earlier version. Chu Zhang acknowledges the financial support from the RGC Competitive Earmarked Research Grant HKUST646007. All remaining errors are ours. Address correspondence to Chu Zhang via E-mail: czhang@ust.hk.

On the Number and Dynamic Features of State Variables in Options Pricing

Abstract

In this paper, we investigate the number of state variables required for options pricing and the dynamic features that govern the evolution of these state variables. We adopt a nonparametric regression technique with the use of model-free implied variance-swap prices of various maturities as proxies for the transformed state variables. The methodology is applied to the prices of S&P500 index options during the period 1996-2005. We find that, in addition to the index value itself, two state variables, approximated by a short-term variance-swap price and a long-term variance-swap price, are adequate for pricing the index options and fitting the data well in time-series and cross-sections. The drift and the squared diffusion of the state variables exhibit strong non-linearity, which explains why popular models in the literature are unable to fit the data satisfactorily.

The volatility smile and the volatility smirk with respect to the well-known Black-Scholes formula have triggered major developments in the options pricing literature in recent years. The volatility smile refers to the observation that, before the 1987 stock market crash, the implied volatility from the Black-Scholes formula, which assumes a constant volatility, is a smile-shaped function of the strike price. After the crash, the implied volatility is shaped like a smirk with the low-strike options prices much higher than the high-strike options prices. The volatility smile and smirk are inconsistent with the assumptions of constant volatility and lognormal return distribution in the Black-Scholes options pricing model.

There have been tremendous theoretical advances in resolving the volatility smile/smirk issue. Since the problem is related to the volatility assumption, all the specific models begin with alternative assumptions about the volatility. The stochastic volatility models assume that the unobserved instantaneous volatility of the underlying security follows another stochastic process and serves as an additional state variable, or factor, in pricing options. Early stochastic volatility models include Hull and White (1987), Wiggins (1987), Scott (1987), and Stein and Stein (1990). Heston (1993) makes an important technical contribution to this literature by providing a closed-form options pricing formula. Stochastic volatility adds fatter tails to the normal distribution of the log price of the underlying security, required for producing the smile. Its negative correlation with the log price contributes the left skewness needed for explaining the smirk. In more recent studies such as Bates (2000) and Pan (2002), stochastic volatility models with jumps in the underlying security are advocated to enhance the effect of the stochastic volatility. Similar to the idea of stochastic volatility, another line of research in the discrete-time framework develops options pricing models in which the underlying security price follows processes with generalized autoregressive conditional heteroskedasticity (GARCH). Duan (1995) and Heston and Nandi (2000) represent the major advances in that direction. Both stochastic volatility models and GARCH models are theoretically capable of generating options prices that are consistent with the observed market prices.

Empirical studies confirm that stochastic volatility models and GARCH options pricing models make headway towards resolving the pricing issue. However, existing models are still inadequate, to various degrees, in fitting the observed options prices well. The typical problems for stochastic volatility models are, observed by Bakshi, Cao and Chen (1997), for example, that the models require highly implausible parameters of the volatility-return correlation and the volatility-of-volatility. The existing stochastic volatility models with jumps are particularly incapable of matching long-term option prices, as indicated by Bates (2000).

There can be two sources of model mis-specifications. One is the omitted state variables. It is conceivable that the instantaneous volatility of the underlying security can be driven by a multitude of state variables. If that is the case, models using the stochastic volatility as the only state variable in addition to the price of the underlying security are then misspecified. Recent research has shown the benefit of increasing the number of state variables. Chernov et al. (2003) study the dynamics of the Dow Jones index and find evidence that favors two state variables, which separate the tail effect and the volatility persistence effect, over the one-factor model. Christoffersen et al. (2006) 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. It is not clear, however, whether more state variables are better. In the extant literature, state variables beyond the price of the underlying security itself are mostly unobserved and each additional state variable unavoidably introduces a slew of parameters, making the identification and estimation of the unobserved state variables and parameters a very complicated procedure. As a result, no consensus has been reached on how many state variables are ideal for options pricing and what these state variables are.

The other source of model mis-specification is the functional form of the processes for the price of the underlying security and of the risk premiums associated with state vari-

ables, i.e., factor premiums. Bates (2000) finds strong evidence against the hypothesized square root diffusion processes driving instantaneous volatility in extensions of Heston's (1993) model, because it cannot account for the large and typically positive volatility shocks implied in the prices of long-term options. Jones (2003) also finds that Heston's (1993) square root stochastic volatility model is incapable of generating realistic return behavior. Chernov et al. (2003) considers ten models of various state variables and functional forms, but, in conclusion, they are unwilling to declare an overall winner. The two sources of model mis-specification, i.e., the uncertain number of state variables and the functional form of the underlying process and factor premiums, are obviously related. With omitted state variables, the functional form can never be correct. With an incorrect specification of the functional form, determining the correct number of state variables is difficult and the result is unreliable.

In this paper, we address the issue of determining the number of state variables for options pricing with respect to a given underlying security and the dynamic features of the state variables. While the goal is rather limited, we view the correct specification of the number of state variables as the top priority in modeling options prices, without which correct functional forms for the underlying process and factor premiums are out of the question. But since an incorrect functional form will also confound the determination of the number of state variables, we resolve the issue by using a nonparametric approach with state variables approximated by model-free implied variance-swap prices of various maturities. The nonparametric approach avoids mis-specification of the processes of the underlying security, the state variables, and the risk premiums associated with state variables. A feature that sets our approach apart from others in the literature is our use of observed proxies of state variables, in our case, the model-free variance-swap prices. More specifically, we fit the options prices nonparametrically as functions of the prices of variance-swaps of various maturities. Using the fitted options price functions, we then test how many variance-swap prices are sufficient to capture the dynamics of the options prices and what the identities (in terms of maturity) of these variance-swaps are.

Our approach of using variance-swap prices as proxies of state variables requires some explanation. The variance-swap prices we use are implied from the prices of European calls and puts. This approach disqualifies the resulting nonparametric functions we fit as a structural options pricing model even if the functional form is known because options prices are needed as input to the fitted functions. For our purpose of determining the number of state variables and characterizing these state variables, this is not a problem. The approach is analogous to the one in the literature of empirical studies of stock markets in which systematic factors of mimicking portfolio returns are constructed using returns on individual stocks and these factors are then used to determine the alphas and betas of individual stocks.¹

We apply the methodology to the analysis of S&P 500 index options, using weekly data from January 1996 to December 2005. We construct model-free variance-swap prices from options prices for various maturities, using the same approach the so-called VIX of Chicago Board Options Exchange (CBOE) is constructed. We find that two variance-swap prices are sufficient to capture the dynamics of the state variables in the time-series and cross-sections of the option prices. The two state variables characterize the expected volatility and the volatility-of-volatility under risk-neutral measures over short and long terms. The result suggests that, for S&P index options, there will be little gain to extend the number of state variables beyond two. Some of the models in the extant literature, Bates (2000) for example, use two state variables in addition to the price of the underlying security and are found unsatisfactory. The undesired features of these models, according to our analysis, come from the inappropriate functional forms in the model, rather than the inadequacy of the number of state variables. The nonparametric estimation results in this paper suggest that the drift and the squared diffusion of state variables exhibit strong nonlinearity. The class of affine jump-diffusion options pricing models, such as in Duffie et al. (2000), are insufficient to capture such dynamics.

¹While the nonparametric functions we fit are not a structural options pricing model, the fitted nonparametric function can nevertheless be useful in determining the relative pricing of options and hedging positions in various applications.

Bondarenko (2007) develops a methodology which constructs the so-called power contracts from option prices to track state variables needed in affine jump-diffusion processes. He uses principal components analysis to conclude that more than one latent factor is needed to explain option data. Since his paper does not address the issue of dynamic features of the state variables, he recommends using more state variables that follow affine jump diffusion processes. In our paper, in addition to using principal components as the first-step exploration, we develop a formal nonparametric test to reach the conclusion that exactly two latent factors are adequate to capture the dynamics of time-series and cross-sectional option prices. We also estimate and test the dynamics of the state variables nonparametrically and show the inadequacy of the affine structure. Therefore, our results are more informative about the property of the state variables in option prices.

The rest of the paper is organized as follows. Section 1 presents the methodology. Section 2 describes the S&P 500 index options price data for the application, presents preliminary analysis of the cross-sectional options prices, and explains the construction of variance-swap prices. Section 3 discusses the nonparametric test used for determining the number of required state variables and reports the results. Section 4 discusses the dynamic features of the two state variables. Section 5 presents some robustness checks of the main results and Section 6 concludes the paper.

1. Methodology

1.1. The Set-up

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

$$ds_t = \mu_s(x_t)dt + \sigma_s(x_t)d\omega_t + z(x_t)dJ_t - \mu_z(x_t)\lambda(x_t)dt, \quad (1)$$

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)d\omega_t + y(x_t)dJ_t - \mu_y(x_t)\lambda(x_t)dt, \quad (2)$$

where x_t is a k -dimensional state variable, ω_t is a standard vector Brownian motion with dimension greater than or equal to k , J_t is a vector counting process with jump intensity $\lambda(x_t)$, μ_s and μ_x are the conditional mean of ds and dx , $\sigma_s\sigma'_s$ and $\sigma_x\sigma'_x$ are the conditional variance of the diffusive component of ds_t and dx_t , z and y are the conformable matrices of jump sizes of ds_t and dx_t with means μ_z and μ_y , respectively, independent of ω_t and J_t .

Since x_t is not assumed to be tradable and the jump components of the processes cannot be hedged, there exists a k -dimensional risk premium for the purpose of option pricing. A risk-neutral probability, Q , exists correspondingly, though it is not unique. Let $\psi(S_{t+\tau})$ be the payoff of an European option at expiration date $t + \tau$. Its theoretical value at t is determined by $\psi_t = e^{-r\tau} E_t^Q[\psi(S_{t+\tau})]$. In particular, prices of calls and puts are functions of (x_t, S_t, K, τ) , where K is the strike price and τ is the time to maturity. We restrict our attention to the case in which, given x_t , the prices of European call and put options are homogeneous functions of degree one of the underlying price, S_t , and the strike price, K . As a result, options prices are functions of K/S_t , or, equivalently, of K/F where F is the $t + \tau$ futures price of the underlying security.

1.2. Variance-swap Prices

A variance swap is a forward contract between two parties to exchange at $t + \tau$ a value $V_{\tau,t}^2$ pre-determined at 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. The value $V_{\tau,t}^2$ is known as the variance-swap price. Since a variance swap has no value at its inception, t , it must be true that, theoretically, $V_{\tau,t}^2 = E_t^Q \frac{1}{\tau} \int_t^{t+\tau} \langle s_u, s_u \rangle du$. Since $V_{\tau,t}^2$ is the expectation conditioned on the time t information, it must be a measurable function of (x_t, τ) .

1.3. An Example

The following is an example in which variance-swap prices are affine functions of the state variables. Suppose, under the risk-neutral probability,

$$ds_t = \mu(x_t)dt + \sum_{j=1}^k \sqrt{\sigma_{0j} + \sigma_{1j}x_{jt}}d\omega_{jt} + z_t dJ_t - \mu_z(\lambda_0 + \lambda'_1 x_t)dt, \quad (3)$$

$$dx_t = K(\theta - x_t)dt + \Omega(x_t)d\omega_t^* + y_t dJ_t^* - l(x_t)\mu_y dt, \quad (4)$$

where $\omega_t = (\omega_{1t}, \dots, \omega_{kt})$ is a k -vector of independent standard Brownian motions, ω_t^* is a k -vector of independent standard Brownian motions, possibly correlated with ω_t , K is a $k \times k$ matrix, J_t is independent of ω_t , the jump size of J_t has mean μ_z and variance σ_z^2 , J_t^* has an intensity $l(x_t)$ and jump size with mean μ_y , $\mu(x_t)$ satisfies $E^Q[dS_t/S_t] = rdt$ where r is the instantaneous riskfree rate, $\Omega(x_t)$ and $l(x_t)$ are unspecified except for the usual regularity conditions. This process of $(s_t, x_t)'$ is more general than the model in Egloff et al. (2007) which does not have jumps in the process of the underlying asset price. It is also more general than the affine jump-diffusion process discussed in Duffie et al (2000) in that $\Omega(x_t)$ and $l(x_t)$ are unspecified.

It is not difficult to deduce that

$$V_{\tau,t}^2 = \sigma_0 + \sigma'_1 \theta + \sigma'_1 (\tau K)^{-1} (I_k - e^{-\tau K})(x_t - \theta) + [\lambda_0 + \lambda'_1 \theta + \lambda'_1 (\tau K)^{-1} (I_k - e^{-\tau K})(x_t - \theta)] (\mu_z^2 + \sigma_z^2), \quad (5)$$

where I_k is the identity matrix, $\sigma_0 = \sum_{j=1}^k \sigma_{0j}$ and $\sigma_1 = (\sigma_{11}, \dots, \sigma_{1k})'$. In this example, a variance-swap price is an affine function of the underlying state variables. Let $\tau_1 < \tau_2 < \dots < \tau_k$. From k equations for $V_{\tau_j,t}^2$, $j = 1, \dots, k$, it is easy to see that x_{jt} s can be solved as affine functions of $V_{\tau_j,t}^2$ s.

More generally, the relationship between $V_{\tau_j,t}^2$ and x_{jt} s may not be affine. But, under the mild condition that the Jacobian between $V_{\tau_j,t}^2$ s and x_{jt} s is nonzero, the unobserved x_{jt} s can be solved from $V_{\tau_j,t}^2$ s and, therefore, options prices can be written as (nonlinear) functions of variance-swap prices. It should be noted that, theoretically, variance swaps

do not have any unique role in tracking the unobserved state variables. Many other contracts can serve the same purpose.² The reason we adopt variance swaps is their practical popularity. Duan and Yeh (2007), for example, also use variance-swap price to infer the jump and volatility risk premiums.

1.4. Approximation of the Pricing Relationship

Theoretically, prices of European options are functions of s_t and x_t , so for any set of more than $k + 1$ options, their prices are functionally dependent. Practically, however, the observed options prices differ from their theoretical ones. The differences can be measurement errors from nonsynchronicity of options prices and their underlying and microstructural errors due to the bid-ask bounce and the discrete tick size. As a result, inferring the number of state variables from observed options prices becomes a matter of statistical inferencing.

In the empirical work below, we use model-free variance-swap prices inferred from options prices. Carr and Madan (1998) and Demeterfi et al. (1999) establish under the assumption of diffusion that the theoretical price of a variance swap equals

$$V_{\tau,t}^2 = \frac{2}{\tau} e^{r\tau} \left(\int_0^{\bar{K}} \frac{1}{K^2} p(K) dK + \int_{\bar{K}}^{\infty} \frac{1}{K^2} c(K) dK \right), \quad (6)$$

where \bar{K} is the $t + \tau$ -forward price of S_t and $c(K)$ and $p(K)$ are prices of European calls and puts with strike price K . In practice, \bar{K} is chosen as the strike at which the absolute value of the difference between prices of the calls and puts is minimized, so the variance swap is replicated by out-of-the-money calls and puts. In a slightly different setting, Jiang and Tian (2005) extend the result to a jump-diffusion model. With jump components, however, the replicating strategy in (6) is no longer exact. Carr and Wu (2008) give the order of the approximation error of (6).

The expression of (6) gives rise to a concern that using variance-swap prices to track

²Bondarenko (2007), for example, uses the so-called power contracts to track the state variables.

unobserved state variables may capture the variation in options prices across maturity, but not across moneyness because the variation across moneyness is integrated out. The example in the last subsection, in which variance-swap prices track the state variables perfectly, show that this concern is not theoretically valid. For processes different from the example, the corresponding variance-swap prices may not be able to track all the state variables, but it is not because the variation across moneyness is integrated out. For an empirical study where the true process governing the underlying security price is unknown, it is an empirical issue whether variance-swap prices can capture the state variables of the option prices. We pay special attention in our empirical work that the constructed proxies of state variables do not leave any systematic components in the fitting residuals.

In the empirical work below, we use the implied volatility of an option from the Black-Scholes formula, instead of the options price itself, as the dependent variable, and the square root of variance swap prices as explanatory variables. Suppressing the dependence of observations on time t , the function we fit is of the form

$$\sigma_i = \sigma(K_i/F_i, T_i, V_{\tau_1}, \dots, V_{\tau_k}) + \varepsilon_i, \quad (7)$$

where σ_i is the implied volatility, K_i/F_i is the moneyness, and T_i is the time-to-maturity of option i . The function σ is fitted with a nonparametric approach. Since, given other parameters, the Black-Scholes price of an option is a strictly increasing function of its implied volatility, using options prices and their implied volatilities is equivalent. However, having volatility measures on both sides of the equation eases the difficulty in nonparametric fitting and makes the typical homoskedasticity assumption about ε_i more appropriate. We use out-of-the-money calls and puts only because out-of-the-money options are more liquid and the implied volatility of an in-the-money call (put) is the same as that of the out-of-the-money put (call) with the same strike price.

2. Data and Preliminary Analysis

2.1. Data Description

We use weekly data on the S&P 500 index call and put options from 1996 to 2005. 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 options data are provided by OptionMetrics. OptionMetrics includes daily closing options prices, open interests, and trading volume for all US index options and equity options since 1996. For fitting options prices, we use weekly observations on Wednesdays to reduce the burden of nonparametric fitting. Wednesdays are chosen because there are the least number of holidays on Wednesdays. If Wednesday is a holiday, we choose Thursday. If both are holidays, we choose Tuesday, and so on. Options with the value less than $\frac{3}{8}$, or with the time-to-expiration less than seven days are excluded due to liquidity reasons. The S&P 500 index level is from Datastream. Daily interest rate data are from the U.S. Treasury Department's website. The VIX index is from CBOE's website.

To calculate the implied volatility, we use the forward price of the underlying security implied from options prices instead of the spot price of the underlying security. By doing so, we reduce the potential errors introduced by estimating the future dividends and the riskfree rate. We use the call-put option pair with the strike price for which the difference between call and put option prices is the smallest, and apply the put-call parity to infer the forward price, F , as $F = (c(K^*) - p(K^*))e^{r\tau} + K^*$, where $K^* = \operatorname{argmin}|c(K) - p(K)|$ and r is the riskfree rate.

Table 1 reports the summary statistics of the implied volatilities from the out-of-the-money options, i.e., put options with $K/F \leq 1.00$ and call options with $K/F > 1.00$. The total sample size is 105,574. We divide the entire sample into 6 moneyness and 6 maturity bins, creating 36 intersection groups in all. The maturity groups' cutoff points

are chosen to be consistent with the construction of the variance-swap prices explained later in this paper. The cutoff point for the shortest maturity group is 39 days, which is used by the CBOE for constructing the square root of the one-month variance-swap price, VIX. Since there are more short-term options contracts than long-term options contracts, to balance between missing values and maturity coverage in each group, we set the other cutoff points as 69 days, 189 days, 279 days and 369 days, which correspond to 2-month, 6-month, 9-month and 12-month options, respectively. Table 1 shows the number of options contracts and the average implied volatility of each group. For the long-maturity groups, more options belong to the deep in-the-money and deep out-of-the-money groups. Across moneyness, the implied volatility generally decreases with K/F . For the short-maturity groups, the implied volatility exhibits a smile pattern, whereas for the long-maturity groups, the implied volatility exhibits a smirk pattern. On average, the implied volatility of short-maturity options is slightly higher than that of long-maturity options.

Table 1 Here

2.2. Principal Components Analysis of Implied Volatilities

As a first-step exploration, the weekly average implied volatilities in all moneyness-maturity groups are examined using principal components analysis. The variance explained by the 36 principal components is shown in Table 2.³ The first principal component explains about 91% of the total variance in the implied volatility surface, the second principal component explains about 5.5%, and the third one explains additional 1.4%. The remaining principal components explain very little.

Table 2 Here

³About half of the time, at least one observation is missing for one of the 36 groups. We replace a missing observation with the average of the group with the nearest shorter maturity and the same moneyness.

Figure 1 shows the factor loadings of the first three principal components. Each surface represents the 36 factor loadings for each of the three principal components. For the first principal component, the surface is flat, indicating that the level is the most important factor in determining the variation in the implied volatility surface. The implied volatility for different moneyness and maturity tends to move up and down together. The second principal component is a slope factor, representing large-strike-minus-small-strike and short-maturity-minus-long-maturity differences. As we can see, the slopes of the implied volatility across moneyness and across maturity are negatively related, meaning that a common factor determines the slopes of the implied volatility across moneyness and across maturity simultaneously. The loadings on the second principal component suggest when the implied volatility across moneyness is more negatively skewed, the long-term implied volatility is higher than the short-term implied volatility. The loading on the third principal component shows a curvature pattern for the short maturity and medium maturity options. However, it only explains 1.4% of the total variation in the implied volatility surface.

Figure 1 Here

Figure 2 plots the time-series of the first three principal components. The first principal component representing the average implied volatility shows a familiar pattern documented in the literature. The second principal component, the slope factor, reaches its lowest level in late 1999, indicating that the options market expected higher long-term risk than short-term risk. Indeed, the S&P 500 index reached its peak about one year later and tumbled following the burst of the internet bubble. The slope factor then went up and peaked around mid-2002, right before the stock market recovery in 2003. The time-series plot of the first and second principal components contain important information regarding the market expectations of the risk in stock returns. The first two principal components are rather persistent and the third principal component is less so.

Figure 2 Here

2.3. Variance-swap Prices

CBOE calculates the square root of the one-month variance-swap price, VIX. We calculate the long-maturity VIXs according to the CBOE VIX construction method (Chicago Board Options Exchange, 2003). For a given day, the squared VIX with maturities equal to those of available options contracts are calculated first. The VIX with a given maturity is then calculated by interpolating two squared VIXs with maturities closest to the given maturity. More than 90% of the time, there are options with seven or eight different maturities. We construct five VIXs with 2-month, 6-month, 9-month, 12 month and 18-month maturities, in addition to the 1-month VIX provided by CBOE. We denote the VIXs for the various maturities as V_{τ_j} for $\tau_j = 1, 2, 6, 9, 12, 18$ measured in months for convenience. The cutoff points are the same as in the previous section except for the longest maturity group: the 2-month V_2 is constructed using options of the two shortest maturities of no less than 39 days; the 6-month V_6 uses the two shortest maturities options of no less than 69 days; similarly, there are no less than 189 days for the 9-month V_9 and no less than 279 days for the 12-month V_{12} . We use options with two shortest maturities of no less than 350 days to calculate the 18-month V_{18} to ensure that at least two maturities are available.

Table 3 reports the summary statistics of V_{τ} s. The calculation is based on daily observations from 1996 to 2005. It is shown that short-maturity V_{τ} s are slightly higher than long-maturity V_{τ} s. Short-maturity V_{τ} s are more volatile, more positively skewed than are long-maturity V_{τ} s, and they have higher excess kurtosis and larger ranges than do long-term V_{τ} s. The principal components analysis shows that the first two principal components explain nearly 99% of the variance in all the V_{τ} s. The eigenvectors indicate that the first principal component is a level factor and the second one is a slope factor.

Table 3 Here

Figure 3 plots the time-series of the first three principal components of the constructed V_{τ} s. The patterns of the first two principal components resemble those of the principal components of the average implied volatilities in Figure 2. The correlation between the first principal components of the average implied volatilities and of the V_{τ} s is 0.9921. The correlation between the second principal components of the two samples is 0.8925. The results hint that the first two principal components extracted from the variance-swap prices can explain most of the variance in the implied volatility surface. The third principal components are rather idiosyncratic between the two samples, with a correlation of -0.1129.

Figure 3 Here

3. The Number of State Variables

In this section, we conduct a more rigorous analysis of the number of state variables using a nonparametric approach. We fit the implied volatility surface using the square root of the constructed variance-swap prices and adopt a bootstrap procedure to test the number of state variables in the time-series and cross-section of the options prices. The performance of the bootstrap test is also studied using simulation.

3.1. Nonparametric Estimation

The implied volatility surface is fitted using the classical Nadaraya-Waston kernel estimator:

$$\hat{\sigma}(K/F, T, V_{\tau_1}, \dots, V_{\tau_k}) = \frac{\sum_{i=1}^N \phi\left(\frac{K_i/F_i - K/F}{h_{K/F}}\right) \phi\left(\frac{T_i - T}{h_T}\right) \prod_{j=1}^k \phi\left(\frac{V_{\tau_j, i} - V_{\tau_j}}{h_{V_{\tau_j}}}\right) \sigma_i}{\sum_{i=1}^N \phi\left(\frac{K_i/F_i - K/F}{h_{K/F}}\right) \phi\left(\frac{T_i - T}{h_T}\right) \prod_{j=1}^k \phi\left(\frac{V_{\tau_j, i} - V_{\tau_j}}{h_{V_{\tau_j}}}\right)}, \quad (8)$$

where ϕ is a kernel function, h_w is the bandwidth for an explanatory variable, w , and N is the number of observations. We choose the fourth-order Gaussian kernel with $\phi(z) =$

$\frac{1}{\sqrt{2\pi}}e^{-z^2/2}(\frac{3}{2} - \frac{z^2}{2})$ following Ait-Sahalia and Lo (1998) who use the Nadaraya-Waston kernel estimator combined with the fourth-order Gaussian kernel to estimate the implied volatility surface and demonstrate good performance. The optimal bandwidth, h_w , for an explanatory variable, w , is

$$h_w = c \cdot \text{SD}_w \left(\frac{1}{N} \right)^{\frac{1}{k+10}}, \quad (9)$$

where SD_w indicates the standard deviation of the variable w .⁴ We choose $c = 1$ as in the usual practice.

Models with various variance-swap prices as proxies for state variables are labeled as $M_k(\tau_1, \dots, \tau_k)$, where k is the number of state variables (in addition to the price of the underlying security). We calculate the mean squared errors (MSE) of the model to gauge its absolute performance and calculate R^2 , defined as

$$R^2 = 1 - \frac{\text{MSE}(M_{k+1}(\tau_1, \dots, \tau_k, \tau_{k+1}))}{\text{MSE}(M_k(\tau_1, \dots, \tau_k))}, \quad (10)$$

to gauge the improvement of the model with an addition of a state variable, $V_{\tau_{k+1}}$. Like the adjusted R^2 in linear regressions, the R^2 used here penalizes more explanatory variables with a larger bandwidth, which gives a larger MSE. Panel A of Table 4 reports the MSEs for models M_0 , M_1 , M_2 and M_3 and the corresponding R^2 s. M_0 is the model without any additional state variables to K/F . $M_1(1)$ is the model with one additional state variable, V_1 . In model $M_2(1, 18)$, we have V_1 and V_{18} as the state variables. By adding either V_6 , V_9 or V_{12} to the $M_2(1, 18)$ model, we have an M_3 model.

The results in the table show that the MSE drops from 2.3037×10^{-3} to 0.3894×10^{-3} when the square root of the shortest-maturity variance-swap price, V_1 , is added to the M_0 model. The R^2 is 0.8310. When the square root of the longest-maturity variance-swap price, V_{18} , enters the model as the second state variable, the MSE further drops to 0.2070×10^{-3} . The second state variable explains additional 46.83% of the variation $M_1(1)$ fails to explain. However, when either of the medium maturity variance-swap prices enters

⁴The formula for the power of $1/N$ in the optimal bandwidth is the reciprocal of two times the order of the kernel plus the number of explanatory variables, $2 \cdot 4 + (k + 2)$ in our case.

the model as the third state variable, the MSE decreases only marginally. The R^2 for the third state variable ranges from 0.0290 to 0.0604 for different medium-maturity variance-swap prices. The results suggest that given the short- and long-maturity variance-swap prices, the medium-maturity variance-swap prices do not improve the fitting significantly.

Table 4 Here

The performance of each model is also illustrated graphically. Figure 4 shows the weekly average residuals (the solid line), the fifth percentile, and the 95th percentile of the residual distribution (the shaded area) for the M_0 model. The upper panel is for the entire sample, the middle panel is for the options with maturities no greater than 180 days, and the lower panel is for the options with maturities greater than 180 days. The average residuals exhibit a persistent pattern and resemble the time-series plot of the V_1 . The fifth and 95th percentiles of the distribution cover a rather wide range. The long-term and short-term average residuals show a similar pattern, except that the distribution of the residuals from the short-term options is much wider, suggesting that the short-term options are more volatile than the long-term ones. The residuals shown in Figure 4 indicate that there are omitted state variables in the model.

Figure 4 Here

Figure 5 shows the residuals for the $M_1(1)$ model, in which the square root of the short-term variance-swap price, V_1 , is added to the model. There is a clear contrast between Figure 5 and Figure 4. The variation in the average residuals in Figure 5 is much less pronounced than that in Figure 4, and the distribution of the residuals is much narrower. The distribution from different maturities shows that the improvement is mainly from the short-maturity options. While there is also an improvement for the long-maturity options, the average residual is still persistent, suggesting that another long-term variance-swap price may explain the residual. Overall, the model with one state variable in addition to

the underlying security price significantly improves the model with the underlying security price as the only state variable.

Figure 5 Here

Figure 6 shows the residuals of $M_2(1, 18)$, which contains the short-term and the long-term variance-swap prices. By comparing Figure 6 with Figure 5, we can see that the overall fitting is further improved. Although most of the improvement is concentrated in the long-term options, the residuals in the short-term options are reduced in magnitude as well. The average residual from the $M_2(1, 18)$ model does not show obvious trends for both of the short-term and long-term options, suggesting that a model with two state variables in addition to the underlying security price captures most of the variation in the options prices cross-sectionally and over time.⁵

Figure 6 Here

Additional evidence about the number of state variables in the options prices can be gleaned from the principal components of the residuals from various models. Table 5 reports, for each model discussed in Table 4, the variances and the proportional variances of the principal components of the average residuals of 36 moneyness/maturity groups, defined in the previous section. To conserve space, only those for the first six out of 36 principal components are reported. For M_0 , the first principal component of the average residuals accounts for 92% of the total variance, clearly indicating that a common factor remains. For $M_1(1)$, that proportion drops to 68%, a substantial reduction. The difference in the proportional variance between the first and the second principal components is still obvious, however. For the remaining models in the table, the proportional variance of the first principal component is further reduced, and the difference in the proportional

⁵This can be interpreted as evidence that, in the case of S&P 500 index options, variance-swap prices are able to track all the state variables of the underlying asset and the risk premiums.

variance between the first and the second principal components is much less pronounced. This adds to the previous evidence that a two-factor model is adequate in explaining the dynamics of the cross-sectional option prices.

Table 5 Here

3.2. Nonparametric Testing

We use a bootstrap procedure to test the null hypothesis nonparametrically:

$$H_0 : \sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_k}) = \sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_{k+1}}) \quad (11)$$

against the alternative

$$H_1 : \sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_k}) \neq \sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_{k+1}}), \quad (12)$$

where $\sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_k})$ is the restricted model, and $\sigma(K/F, T, V_{\tau_1}, \dots, V_{\tau_{k+1}})$ is the unrestricted model. We start with $k = 0$ and increase the number of state variables until the null hypothesis cannot be rejected.

There are a few nonparametric tests of omitted variables in the literature that do not rely on simulation, for example, Fan and Li (1996), Lavergne and Vuong (2000) and Ait-Sahalia et al. (2001). These papers develop test statistics based on the fitted residuals from restricted and unrestricted models. Under the null hypothesis, the statistic is asymptotically distributed as a standard normal variate under certain regularity conditions. These tests work well when the number of explanatory variables is small, but not so well when the number of state variables is large or moderate. In addition, as these tests depend on the bandwidth for both the restricted and the unrestricted models in the nonparametric regression, the result is sensitive to the choice of the bandwidth.

To avoid the potential problems arising from such tests, we use the so-called two-point wild bootstrap procedure for the nonparametric testing. Applications of the bootstrap

approach in various nonparametric studies have demonstrated its good finite sample properties. See, for example, Li and Wang (1998) and Li and Racine (2007, Chapter 12). The bootstrap procedure is implemented as follows. We fit the implied volatility, σ_i , using the restricted model and calculate the residuals from the model as

$$\varepsilon_i = \sigma_i - \hat{\sigma}_i(K_i/F_i, T_i, V_{\tau_1}, \dots, V_{\tau_k}). \quad (13)$$

Using the residuals from the restricted model, we construct the bootstrap samples under the null hypothesis,

$$\sigma_i^* = \hat{\sigma}_i(K_i/F_i, T_i, V_{\tau_1}, \dots, V_{\tau_k}) + \varepsilon_i^*, \quad (14)$$

where $\varepsilon_i^* = \frac{1-\sqrt{5}}{2}\varepsilon_i$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$, and $\varepsilon_i^* = \frac{1+\sqrt{5}}{2}\varepsilon_i$ with probability $\frac{-1+\sqrt{5}}{2\sqrt{5}}$. The new errors have the following property: $E^*(\varepsilon_i^*) = 0$, $E^*(\varepsilon_i^{*2}) = \varepsilon_i^2$ and $E^*(\varepsilon_i^{*3}) = \varepsilon_i^3$, where E^* indicates the expected value in the simulation.⁶ We calculate the R^2 from many bootstrap samples and the empirical distribution of R^2 under the null hypothesis can be obtained. Comparing the R^2 from the original sample with the empirical distribution of R^2 from the bootstrap samples, we are able to make inferences about the validity of the null hypothesis. When the R^2 calculated from the original sample is higher than 95% of the empirical distribution of R^2 from the bootstrap samples, the bootstrap test rejects the null hypothesis at the 5% confidence level and the bootstrap p-value is 5%.

The last column in Table 4 reports the p-values for the bootstrap test of the number of state variables in the options prices. The p-value is calculated from 100 bootstrap samples. Not surprisingly, the model with one additional state variable significantly improves upon the model with the underlying security price as the only state variable. The second state variable further improves the one-factor variable model, and the bootstrap p-value suggests that the second state variable is important for explaining the variation in the cross-sectional options prices. When any of the additional medium-maturity variance-swap prices is added, the bootstrap test suggests that the null hypothesis of the two-factor model cannot be rejected.

⁶It should be noted that the two-point wild bootstrap approach allows heteroscedasticity.

We also conduct an analysis of the finite sample performance of the bootstrap non-parametric test. First, we randomly select a sample from our original data set, fit the implied volatility in that sample using one of the models, M_0 , $M_1(1)$ or $M_2(1, 18)$, and calculate the root mean squared error (RMSE). The fitted values plus the simulated random errors are considered as the data set generated from that particular model. The random errors are simulated from the normal distribution with a mean of zero and the standard deviation equal to the RMSE. A useless independent variable is another simulated random variable, which is a standard normal random variable and independent of the random errors. Thus, we have constructed a sample from our data set for which the null hypothesis is true. Then, we fit the simulated data using the restricted model (the true model) and the unrestricted model (the model with the additional useless explanatory variable), and calculate the MSE and R^2 . The fitted errors from the restricted model are resampled according to the bootstrap procedure described previously to construct the bootstrap samples. Using the bootstrap samples, we construct the empirical distribution of R^2 . The R^2 calculated from the simulated data is compared with the empirical distribution of R^2 from the bootstrap samples. If the R^2 calculated from the simulated data is higher than a threshold of the empirical distribution of R^2 , we reject the null hypothesis. The entire procedure is repeated many times, and we calculate the rejection rate. A test with the property that one minus the rejection rate is close to the threshold is considered as a good test. For example, if we reject the null hypothesis when the R^2 calculated from the simulated data is higher than the 95th percentile of the empirical distribution of R^2 from the bootstrap samples, a good test should have a rejection rate of 5%. A test with a rejection rate higher (lower) than 5% indicates that the test tends to over-reject (under-reject) the null hypothesis.

To study the performance of the test, we consider model M_p with p true additional state variables and one useless explanatory variables. The implied volatility is fitted using K/F and T in M_0 , using K/F , T and V_1 in $M_1(1)$, or using K/F , T , V_1 and V_{18} , in $M_2(1, 18)$. The performance of the test with different sample sizes is also compared.

We consider sample sizes of 500, 1000 and 2000. The number of bootstrap samples to construct the empirical distribution of R^2 is 100, and number of tests to calculate the rejection rate is also 100. To study how sensitive the result is to the bandwidth choice, we conduct the simulation study for different values of bandwidth. We vary the constant c in choosing the bandwidth for different cases. The constant, c , for both the restricted and the unrestricted models is the same, and we consider the cases of $c = 0.85, 1$ and 1.2 .

Table 6 shows the simulation results of the bootstrap test. For M_0 , the test performance is quite good. The results are similar with different sample sizes. More importantly, the results are not sensitive to a rather wide range of bandwidth choices. However, the test tends to over-reject for the case of 95% threshold and under-reject the case of 90% threshold. For $M_1(1)$, the performance is worse than in the previous case. A smaller bandwidth tends to over-reject the null hypothesis, and a larger bandwidth tends to under-reject the null hypothesis. For the case of $c = 1$, the result is still reasonably good. The test performance is improved with larger sample size. For $M_2(1, 18)$, the over-rejection and under-rejection problem is more obvious, and the help from a larger sample is not as significant as in the previous case. Nevertheless, for $c = 1$, the result is still satisfactory.

Table 6 Here

4. Dynamic Features of State Variables

After showing that two transformed state variables are sufficient to capture the variation in the cross-sectional options prices, we go one step further to examine the dynamic features of the transformed state variables. We study the conditional moments of V_1^2 and V_{18}^2 . More specifically, we estimate

$$V_{1,t+1}^2 - V_{1,t}^2 = \mu_1(V_{1,t}^2, V_{18,t}^2) + \eta_{1,t+1}, \quad (15)$$

$$V_{18,t+1}^2 - V_{18,t}^2 = \mu_{18}(V_{1,t}^2, V_{18,t}^2) + \eta_{18,t+1}. \quad (16)$$

We then estimate

$$\hat{\eta}_{1,t+1}^2 = \sigma_1^2(V_{1,t}^2, V_{18,t}^2) + \xi_{1,t+1}, \quad (17)$$

$$\hat{\eta}_{18,t+1}^2 = \sigma_{18}^2(V_{1,t}^2, V_{18,t}^2) + \xi_{18,t+1}, \quad (18)$$

$$\hat{\eta}_{1,t+1}\hat{\eta}_{18,t+1} = \sigma_{1,18}(V_{1,t}^2, V_{18,t}^2) + \xi_{1,18,t+1}, \quad (19)$$

where $\hat{\eta}_{1,t+1}$ and $\hat{\eta}_{18,t+1}$ are the residuals from (15)-(16). We also estimate a univariate version in which $\mu_1(V_{1,t}^2, V_{18,t}^2)$ and $\sigma_1^2(V_{1,t}^2, V_{18,t}^2)$ only depend on $V_{1,t}^2$, $\mu_{18}(V_{1,t}^2, V_{18,t}^2)$ and $\sigma_{18}^2(V_{1,t}^2, V_{18,t}^2)$ only depend on $V_{18,t}^2$, and $\sigma_{1,18}(V_{1,t}^2, V_{18,t}^2) = 0$. We use the Nadaraya-Waston estimation with the fourth order Gaussian kernel and use (9) with $c = 2$ for the bandwidth selection.⁷ We use daily observations to reduce the adverse effect that may arise from the discretization of continuous-time models.

The estimation result of the drift and the squared diffusion of V_1^2 in the univariate model is shown in the left panels of Figure 7. The drift is concave in V_1^2 , which is inconsistent with the most advocated stochastic volatility models in the literature, such as Heston's (1993) model, in which the drift is affine in the state variables. When the level of V_1^2 is high, the drift is negative and mean reversion is fast. When the level of V_1^2 is low, although the drift is positive, the magnitude is very small, meaning that the speed of mean reversion is much slower. There is a clear pattern of non-affine drift in V_1^2 . The unconditional mean of V_1^2 over the sample period is 0.0505, and the drift of V_1^2 at its mean is well above zero. This suggests that V_1^2 has a tendency to move even higher at the mean level. The squared diffusion of V_1^2 is convex, which suggests that the conditional variance of V_1^2 increases faster in V_1^2 than the square root diffusion process dictates.

In the right panels of Figure 7, the slope of the drift at high levels of V_{18}^2 is steeper than that at low levels, suggesting that mean reverting is stronger when V_{18}^2 is high. For quite a wide range in the medium level, V_{18}^2 does not mean revert at all and behaves like

⁷The skewness of the distribution of V_1^2 is 1.62. Using $c = 1$ in the bandwidth selection, the main pattern in the figures is preserved. However, the figures are rough at high levels of V_1^2 due to the relatively sparse data.

a random walk. The unconditional mean of V_{18}^2 is 0.0444, and the drift of V_{18}^2 is close to zero at that value. In comparison with V_1^2 , V_{18}^2 mean reverts much more slowly. The squared diffusion of V_{18}^2 is also convex, similar to that of V_1^2 , suggesting that the square root diffusion process does not capture the dynamics of V_{18}^2 well.

Using a parametric approach, Jones (2003) finds that the stochastic volatility models in the constant elasticity of variance (CEV) class are more consistent with the underlying asset and options data. He reaches the conclusion without allowing jumps in the underlying asset price. Since adding jumps may affect the estimation of the volatility process, his conclusion may change when jumps are allowed. Jones (2003) also finds that the parameter representing the speed of mean-reversion in the volatility process is insignificant and concludes that the observed mean-reversion is generated by the nonlinear conditional variance of the volatility process. By contrast, our approach is non-parametric and our setting allows jumps. Therefore, our findings about the nonlinearity of both the drift and the squared diffusion are more general than those of Jones (2003).

Figure 7 Here

For the bivariate case, the drift of V_1^2 is a function of V_1^2 and V_{18}^2 . Its estimate is presented conditional on V_{18}^2 at low, medium and high levels in the left panels of Figure 8. Conditional on low and medium levels of V_{18}^2 , V_1^2 shows a very weak tendency to reverting to mean. Conditional on the high level of V_{18}^2 , the mean reversion is stronger. The results suggest that the dynamic of V_1^2 is determined jointly by the lagged value of itself and the lagged value of V_{18}^2 . The different slopes of the drift of V_1^2 at different levels of V_{18}^2 suggests that the drift of V_1^2 cannot be simply an affine function of V_1^2 and V_{18}^2 . The drift of V_{18}^2 as a function of V_1^2 and V_{18}^2 is shown in the right panels of Figure 8. The magnitude of the drift of V_{18}^2 is smaller than that of V_1^2 , suggesting that V_{18}^2 is more persistent. The mean reversion is stronger conditional on the low level of V_1^2 than on the medium level of V_1^2 . The strongest mean reversion of V_{18}^2 also occurs conditional on the high level of V_1^2 .

Figure 8 Here

The estimation result of the bivariate squared diffusion is shown in Figure 9. In the left panels, the estimated squared diffusion of V_1^2 is shown as a function of V_1^2 conditional on the low, medium and high levels of V_{18}^2 . The right panels are for the estimated squared diffusion of V_{18}^2 . We note the difference in scales for different panels. For both V_1^2 and V_{18}^2 , the squared diffusion is a convex function of their levels, especially conditional on the medium and high levels of the variance-swap price of the other maturity. V_1^2 has a higher variability than V_{18}^2 has.

Figure 9 Here

The estimation result of the conditional covariance is shown in Figure 10. In the left panels, the estimated conditional covariance between V_1^2 and V_{18}^2 is shown as a function of V_1^2 for the low, medium and high levels of V_{18}^2 . The right panels show the estimated conditional covariance as a function of V_{18}^2 for the different levels of V_1^2 . The figures for the conditional covariance exhibit a similar pattern as those for the squared diffusion: the conditional covariance is also a convex function of the level of the variance-swap prices. This is again inconsistent with the affine jump-diffusion models in which the conditional covariance is an affine function of the state variables.

Figure 10 Here

In summary, the drifts of V_1^2 and V_{18}^2 do not appear to be affine, especially for V_1^2 . When the variance-swap prices are low, their mean reversion is slower than what the affine jump-diffusion models suggest. In a wide range, the variance-swap prices behave like random walks. The squared diffusion is also not consistent with the square root process. If the diffusion is specified as a power function of its level, the power should be greater than 0.5. The bivariate analysis of the conditional moments shows that the dynamics of V_1^2 and V_{18}^2 depend on each another.

We now provide a formal test of the linearity of the drift, squared diffusion, and conditional covariance of V_1^2 and V_{18}^2 . For each of five functions involved, we consider several hypothesized functional forms that are linear in at least one component of (V_1^2, V_{18}^2) . We implement the same bootstrap test as in testing the number of state variables. The results are listed in Table 7.

Table 7 here

The first column of Table 7 lists the functional forms in various null hypotheses where g indicates a unrestricted, general nonparametric function. For example, for the first univariate linear function $a + bV_1^2$ as the drift of V_1^2 , a and b are estimated using OLS in $V_{1,t+1}^2 - V_{1,t}^2 = a + bV_{1,t}^2 + \eta_{1,t+1}$, and the alternative model is unspecified and estimated nonparametrically. We calculate the $R^2 = 1 - \frac{\text{MSE}(H_1)}{\text{MSE}(H_0)}$, where $\text{MSE}(H_0)$ and $\text{MSE}(H_1)$ are the mean squared error for the models under null and alternative hypotheses, respectively. We then construct a bootstrap sample with 100 replications and compare the R^2 from the original sample with those from the bootstrap samples to calculate its p-value. We set $c = 2$ in the bandwidth for all the linearity tests.⁸ For the first univariate linear drift function of V_1^2 , the p-value is 0.00.

The function in null hypotheses involving a nonparametric term is estimated as follows. Consider $\mu_1(V_1^2, V_{18}^2) = g(V_1^2) + bV_{18}^2$ as an example. The regression is

$$V_{1,t+1}^2 - V_{1,t}^2 = g(V_{1,t}^2) + bV_{18,t}^2 + \eta_{1,t+1}, \quad (20)$$

where g is unspecified. Taking the expectation of (20) conditional on $V_{1,t}^2$ gives

$$E(V_{1,t+1}^2 - V_{1,t}^2 | V_{1,t}^2) = g(V_{1,t}^2) + bE(V_{18,t}^2 | V_{1,t}^2). \quad (21)$$

Subtracting (21) from (20) yields

$$V_{1,t+1}^2 - V_{1,t}^2 - E(V_{1,t+1}^2 - V_{1,t}^2 | V_{1,t}^2) = b[V_{18,t}^2 - E(V_{18,t}^2 | V_{1,t}^2)] + \eta_{1,t+1}. \quad (22)$$

⁸The results only change marginally for the range of c from 1 to 2.5.

$E(V_{1,t+1}^2 - V_{1,t}^2 | V_{1,t}^2)$ and $E(V_{18,t}^2 | V_{1,t}^2)$ are estimated nonparametrically as in (15)-(16). Given the estimates of the conditional expectations, b is estimated by regressing the residual $V_{1,t+1}^2 - V_{1,t}^2 - E(V_{1,t+1}^2 - V_{1,t}^2 | V_{1,t}^2)$ on the residual $V_{18,t}^2 - E(V_{18,t}^2 | V_{1,t}^2)$ using OLS. With the estimated \hat{b} , $g(V_{1,t}^2)$ can be estimated by regressing $V_{1,t+1}^2 - V_{1,t}^2 - \hat{b}V_{18,t}^2$ on $V_{1,t}^2$ nonparametrically.

To test the linearity of the squared diffusion and the conditional covariance, we estimate the drift nonparametrically first as in (15)- (16). We then use the residuals to construct the squared diffusion and the conditional covariance to test the specific linearity hypothesis against the nonparametric alternative as in (17)-(19).

The results of the formal tests are in line with the impression from Figures 7-10. We find strong nonlinearity in the drift of V_1^2 , evidenced by the very low p-values of the univariate linear and bilinear models. However, the linearity of the drift of V_{18}^2 cannot be rejected. The squared diffusion of V_1^2 and V_{18}^2 both show strong nonlinearity. The partially linear model allowing for the nonparametric component of $g(V_1^2)$ captures the nonlinearity in the drift and the squared diffusion of V_1^2 , and the model with the nonparametric component of $g(V_{18}^2)$ cannot be rejected for the squared diffusion of V_{18}^2 . The conditional covariance between V_1^2 and V_{18}^2 does not exhibit strong nonlinearity.

The findings on the functional form of the drift of the variance-swap prices have implications on the discussion about the number of state variables in the literature. In Bates (2000), the residuals from his most general two-factor stochastic volatility/jump-diffusion model, SVJD2, are found to have substantial and persistent common shock components. The model also cannot explain the fact that the volatility smile is equally pronounced for short- and long-maturity options Christoffersen, Heston and Jacobs (2006) find that their two-factor stochastic volatility model encounters some difficulty in capturing both moneyness and term structure effects. These authors suggest that a richer model is needed either with more state variables or with alternative functional specifications. In both papers, the stochastic volatility process is affine. Our empirical results suggest that the

inadequacy of their two-factor models is due to the mis-specification of the functional form, rather than the number of state variables.

It should be noted that our analysis is based on the transformed state variables, rather than on the original state variables. The dynamic features of the variance-swap prices may not be the same as those of the true state variables. However, if the true state variables follow the process as the example in Section 1, which generalizes the affine jump-diffusion models in Duffie et al. (2000), the variance-swap prices should be affine in the true state variables and, therefore, inherit all the features of the true state variables. In this sense, our result is useful in suggesting that the functional form of the drift and diffusion terms of the affine jump-diffusion models, which have been popular in theoretical analysis, are responsible for the rejections of two-factor models in the literature.

5. Robustness Checks

5.1. Alternative Factor Construction

To check the robustness of the results on the number of state variables in the options pricing, we use an alternative method to construct the proxies for the state variables. We use the weekly average implied volatility, AIV, for six different maturity ranges. The six maturity ranges are chosen similarly as for constructing V_{τ_j} s.⁹ The bandwidth selection is chosen as before with $c = 1$. The MSE is used to measure the goodness of fit for models M'_1 , M'_2 and M'_3 , where M'_1 is the model with AIV_1 as the proxy for one state variable, and M'_2 is the model with AIV_1 and AIV_{18} as the proxies for the two state variables, and M'_3 is a three state-variable model with AIV_6 , AIV_9 or AIV_{12} as the additional factor. The improvement from adding a factor is measured by the R^2 defined similarly as in (10). The result is reported in Table 8. When AIV_1 is added to the model, the

⁹The implied volatility from options with less than 39 days to maturity is used to construct the shortest maturity factor AIV_1 , and options with maturities between 39 days and 69 days are used to constructed AIV_2 , between 69 days and 189 days for AIV_6 , between 189 days and 279 days for AIV_9 , between 269 days and 369 days for AIV_{12} , and greater or equal to 350 days for AIV_{18} .

MSE is reduced significantly to 0.4704×10^{-3} with a R^2 of 0.7958. For the model with AIV_1 and AIV_{18} as the proxies for the state variables, the MSE is further reduced to 0.2635×10^{-3} and the R^2 is 0.4398. The additional factor only marginally reduces the MSE. The bootstrap test indicates that the first two state variable proxies, AIV_1 and AIV_{18} , are statistically significant. Comparing the results with those reported in Table 4, we find that given the same bandwidth selection criteria, V_{τ_j} s fit the implied volatility better than AIVs, suggesting that V_{τ_j} s contain less noise and capture the state variables better. Nevertheless, the results using AIVs as the proxy for state variables confirm our conclusion that two state variables are needed to explain the cross-sectional options prices.

Table 8 Here

5.2. Robustness of Nonlinear Drift

We also use the approach in Chapman and Pearson (2000) to check the robustness of the nonlinearity of the drift in the variance-swap prices. In the literature on the short-term interest rate, Ait-Sahalia (1996) and Stanton (1997) use the nonparametric estimation method to conclude that the drift of the short rate contains important nonlinearity. However, Chapman and Pearson (2000) use a weighted least squares estimator to find that the nonlinearity of the short rate drift is not a robust stylized fact. They argue that the source of finite sample bias in the kernel regression estimator is from a truncation of the upper limit of any finite sample. The truncation bias from the finite sample also applies to this study. We extend their approach to test the nonlinearity of the variance-swap prices' drift for the bivariate process. From the theoretical point of view, the short rate process and the volatility process are modeled in a similar fashion. For example, the CIR model for the short rate and the Heston model for volatility are both a square-root diffusion process. Empirically, they exhibit persistence and mean reverting properties. As far as we know, there is no nonlinear drift parametric volatility model in the literature. We therefore borrow the parametric model of the short rate in Ait-Sahalia (1996) and

Chapman and Pearson (2000) and extend it to the bivariate case.

The drift and the squared diffusion are specified as

$$\begin{aligned} \mu_{\tau_j}(V_1^2, V_{18}^2) &= \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_1^2 + \alpha_{2,\tau_j} V_{18}^2 + \alpha_{3,\tau_j} V_1^4 + \alpha_{4,\tau_j} V_{18}^4 \\ &\quad + \alpha_{5,\tau_j} (V_1^2 V_{18}^2) + \alpha_{6,\tau_j} (1/V_1^2) + \alpha_{7,\tau_j} (1/V_{18}^2), \end{aligned} \quad (23)$$

$$\sigma_{\tau_j}^2(V_{\tau_j}^2) = \beta_{\tau_j} V_{\tau_j}^{2\gamma_{\tau_j}}, \quad (24)$$

where μ_{τ_j} is the drift of $V_{\tau_j}^2$ and $\sigma_{\tau_j}^2$ is the squared diffusion of $V_{\tau_j}^2$ for $\tau_j = 1, 18$. The parameter values can be determined using linear regression. In order to reduce multicollinearity problems, we orthogonalize the explanatory variables as in Chapman and Pearson (2000). V_{18}^2 is regressed on a constant and V_1^2 , and the regression residual is the orthogonalized V_{18}^2 , denoted as \widetilde{V}_{18}^2 . Likewise, V_1^4 is regressed on a constant, V_1^2 and \widetilde{V}_{18}^2 and the regression residual is the orthogonalized V_1^4 , denoted as \widetilde{V}_1^4 . Other variables, \widetilde{V}_{18}^4 , $(\widetilde{V}_1^2 V_{18}^2)$, $(1/\widetilde{V}_1^2)$ and $(1/\widetilde{V}_{18}^2)$ are defined similarly. We use OLS to estimate the following regression,

$$\begin{aligned} V_{\tau_j,t+1}^2 - V_{\tau_j,t}^2 &= \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{1,t}^2 + \alpha_{2,\tau_j} \widetilde{V}_{18,t}^2 + \alpha_{3,\tau_j} \widetilde{V}_{1,t}^4 + \alpha_{4,\tau_j} \widetilde{V}_{18,t}^4 \\ &\quad + \alpha_{5,\tau_j} (V_{1,t}^2 \widetilde{V}_{18,t}^2) + \alpha_{6,\tau_j} (1/\widetilde{V}_{1,t}^2) + \alpha_{7,\tau_j} (1/\widetilde{V}_{18,t}^2) + \eta_{\tau_j,t+1}, \end{aligned} \quad (25)$$

for $\tau_j = 1, 18$. Given the fitted values for the drift, the diffusion parameters are estimated from a log linear regression using the squared residuals from the first regression as

$$\ln(\hat{\eta}_{\tau_j,t+1}^2) = \ln(\beta_{\tau_j}) + \gamma_{\tau_j} \ln(V_{\tau_j,t}^2) + \xi_{\tau_j,t+1}, \quad (26)$$

where $\hat{\eta}_{\tau_j,t+1}$ is the residual from the drift function regression for $\tau_j = 1, 18$. Given the fitted values of the residuals, the drift parameters are re-estimated by weighted least squares. The iterative procedure is applied until the parameter estimate converges. The parameter estimates and t-statistics adjusted for heteroscedasticity and 24 lags of autocorrelation using Newey and West (1987) for the drift and the diffusion are reported in Table 9. For V_1^2 , α_3 , α_6 and α_7 are statistically significant, which suggests that the nonlinearity of the drift of V_1^2 is robust. The evidence of the nonlinearity of the drift in V_{18}^2 is not robust as

indicated by the insignificant coefficient estimates of the nonlinear terms. However, this may be due to the poor finite sample performance of the covariance matrix estimator of Newey and West (1987) or the specific functional form of the drift. The estimated γ for both of V_1^2 and V_{18}^2 is greater than 1, suggesting that the data are inconsistent with the square root volatility process.

Table 9 Here

For completeness, a univariate model is estimated as well and reported in Table 10. The result indicates that the drift of V_1^2 is nonlinear and the squared diffusion of V_1^2 and V_{18}^2 are disproportionately large at high variance-swap prices. Overall, the results of the robustness check are consistent with the nonparametric linearity tests in the previous section.

Table 10 Here

6. Conclusion

In this paper, we provide evidence for the number of state variables needed in pricing options on the S&P 500 index. We avoid the mis-specification of the processes of the underlying security price and state variables using a nonparametric approach. We construct model-free variance-swap prices with different maturities as proxies for unobserved state variables. The results indicate that two variance-swap prices, with a short maturity and a long maturity, capture the dynamics of the state variables in the options prices. The model with the two state variables is able to eliminate most of the persistent pricing errors. We further explore the functional forms of the drift and the diffusion terms of the state variables, and find that the drift is nonlinear and the diffusion term is inconsistent with the square root diffusion process. Our results suggest that extending the volatility

process to higher dimensions is of little help. Rather, improving upon the specification of the two-factor volatility process is a promising direction in modeling options prices.

References

- Ait-Sahalia, Y., 1996, Testing continuous-time models of the spot interest rate, *Review of Financial Studies*, 9, 385-426.
- Ait-Sahalia, Y., P. J. Bickel and T. M. Stoker, 2001, Goodness-of-fit tests for kernel regression with an application to option implied volatilities, *Journal of Econometrics*, 105, 363-412.
- Ait-Sahalia, Y. and A. W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance*, 53, 499-547.
- Bakshi, G., C. Cao and Z. Chen, 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. and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy*, 81, 637-659.
- Bondarenko, O., 2007, Nonparametric test of affine option models, Working paper.
- Carr, P., and D. Madan, 1998, Towards a theory of volatility trading, in *Risk Book on Volatility*, ed. by R. Jarrow. Risk, New York, pp. 417-427.
- Carr, P. and L. Wu, 2008, Variance risk premiums, *Review of Financial Studies*, forthcoming.
- Chapman, D. A. and N. D. Pearson, 2000, Is the short rate drift actually nonlinear? *Journal of Finance*, 55, 355-388.
- Chernov, M., A. R. Gallant, E. Ghysels and G. Tauchen, 2003, Alternative models for stock price dynamics, *Journal of Econometrics*, 116, 225-257.

- Chicago Board Options Exchange, 2003, VIX-CBOE volatility index,
<http://www.cboe.com/micro/vix/vixwhite.pdf>.
- Christoffersen, P., S. Heston and K. Jacobs, 2006, The shape and term structure of the index option smirk: why multifactor stochastic volatility models work so well, Working paper.
- Christoffersen, P., K. Jacobs and Y. Wang, 2006, Option valuation with long-run and short-run volatility components, Working paper.
- Demeterfi, K., E. Derman, M. Kamal, and J. Zou, 1999, A guide to volatility and variance swaps, *Journal of Derivatives*, 6, 9-32.
- Duan, C. 1995, The GARCH option pricing model, *Mathematical Finance*, 5, 351-362.
- Duan, J. and C. Yeh, 2007, Jump and volatility risk premiums implied by VIX, working paper, University of Toronto.
- Duffie, D, J. Pan and K. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68, 1343-1376.
- Egloff, D., M. Leippold and L. Wu, 2007, Variance risk dynamics, variance risk premia, and optimal variance swap investments, Working paper.
- Fan, Y. and Q. Li, 1996, Consistent model specification tests: omitted variables and semiparametric functional forms, *Econometrica*, 64, 865-890.
- 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 and S. Nandi, 2000, A closed-form GARCH option valuation model, *Review of Financial Studies*, 13, 585-625.
- Hull, J. C and A. White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance*, 42, 281-300.

- Jiang, G., and Y. Tian, 2005, Model-free implied volatility and its information content, *Review of Financial Studies*, 18, 1305-1342.
- Jones, C. S., 2003, The dynamics of stochastic volatility: evidence from underlying and option markets, *Journal of Econometrics*, 116, 181-224.
- Lavergne, P. and Q. Vuong, 2000, Nonparametric significance testing, *Econometric Theory*, 16, 576-601.
- Li, Q. and S. Wang, 1998, A simple consistent bootstrap test for a parametric regression function, *Journal of Econometrics*, 1998, 145-165.
- Li, Q. and J. S. Racine, 2007, *Nonparametric econometrics*, Princeton University Press, New Jersey.
- Newey, W. K. and K. D. West, 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.
- Stanton, R., 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance*, 52, 1973-2002.
- Stein, E. M. and J. C. Stein, 1991, Stock price distributions with stochastic volatility: an analytic approach, *Review of Financial Studies*, 4, 727-752.
- Wiggins, J. B., 1987, Options values under stochastic volatilities, *Journal of Financial Economics*, 19, 351-372.

Table 1**Summary statistics of implied volatility from the S&P 500 index out-of-the-money options**

The implied volatilities of the S&P 500 index out-of-the-money options are sorted in moneyness and maturity groups. We use weekly data for the period of 1996-2005. T denotes time-to-maturity in terms of months and K/F denotes moneyness. $K/F \leq 1.00$ is for the out-of-the-money put options, and $K/F > 1.00$ is for the out-of-the-money call options. Panel A shows the number of option contracts in each moneyness and maturity group, and Panel B shows the average implied volatility in each moneyness and maturity group.

A. Number of contracts							
	T						
	$1m$	$2m$	$3 - 6m$	$7 - 9m$	$10 - 12m$	$> 12m$	Total
$K/F \leq 0.94$	5433	5258	11895	6821	5880	11160	46447
$0.94 < K/F \leq 0.97$	2516	1377	1909	814	677	1189	8482
$0.97 < K/F \leq 1.00$	2927	1636	2079	800	651	1138	9231
$1.00 < K/F \leq 1.03$	2877	1638	2053	796	709	1083	9156
$1.03 < K/F \leq 1.06$	2156	1288	1763	797	645	982	7631
$1.06 < K/F$	1672	2149	5888	4307	3816	6795	24627
Total	17581	13346	25587	14335	12378	22347	105574

B. Average implied volatility							
	T						
	$1m$	$2m$	$3 - 6m$	$7 - 9m$	$10 - 12m$	$> 12m$	Total
$K/F \leq 0.94$	0.3062	0.2932	0.2844	0.2702	0.2607	0.2556	0.2759
$0.94 < K/F \leq 0.97$	0.2113	0.2077	0.2073	0.2093	0.2101	0.2095	0.2093
$0.97 < K/F \leq 1.00$	0.1859	0.1883	0.1962	0.2020	0.2039	0.2016	0.1933
$1.00 < K/F \leq 1.03$	0.1680	0.1733	0.1833	0.1939	0.1947	0.1957	0.1800
$1.03 < K/F \leq 1.06$	0.1672	0.1625	0.1769	0.1863	0.1902	0.1921	0.1758
$1.06 < K/F$	0.2199	0.1904	0.1809	0.1765	0.1758	0.1792	0.1823
Total	0.2358	0.2282	0.2322	0.2269	0.2226	0.2234	0.2288

Table 2**Principal components analysis of implied volatility**

The principal components analysis is employed on the weekly average implied volatility in each moneyness and maturity bin as specified in Table 1. VAR denotes the variance explained (multiplied by 100) by all the 36 principal components. PVAR denotes the proportion of the variance explained by each principal component out of the total variance.

	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	8.0560	0.4846	0.1269	0.0607	0.0337	0.0260
PVAR	0.9066	0.0545	0.0143	0.0068	0.0038	0.0029
	P7	P8	P9	P10	P11	P12
VAR($\times 10^{-2}$)	0.0219	0.0203	0.0094	0.0090	0.0072	0.0056
PVAR	0.0025	0.0023	0.0011	0.0010	0.0008	0.0006
	P13	P14	P15	P16	P17	P18
VAR($\times 10^{-2}$)	0.0041	0.0032	0.0023	0.0022	0.0019	0.0015
PVAR	0.0005	0.0004	0.0003	0.0002	0.0002	0.0002
	P19	P20	P21	P22	P23	P24
VAR($\times 10^{-2}$)	0.0012	0.0012	0.0010	0.0007	0.0007	0.0006
PVAR	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	P25	P26	P27	P28	P29	P30
VAR($\times 10^{-2}$)	0.0006	0.0005	0.0005	0.0004	0.0003	0.0003
PVAR	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000
	P31	P32	P33	P34	P35	P36
VAR($\times 10^{-2}$)	0.0003	0.0002	0.0002	0.0002	0.0002	0.0001
PVAR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 3**Summary statistics of the V_τ s**

This table presents the mean, standard deviation, skewness, excess kurtosis, 5th percentile and 95th percentile of the V_τ s for different maturities from 1-month to 18-month in Panel A. Panel B shows the results for the principal components analysis of the V_τ s. VAR denotes the variance explained by each principal component, and PVAR denotes the proportion of the variance explained by each principal component out of the total variance. The loadings on each principal component are also shown.

A. Summary statistics						
	Mean	Std	Skew	Kurt	5P	95P
V_1	0.2153	0.0639	0.7659	0.6286	0.1240	0.3361
V_2	0.2115	0.0578	0.6728	0.5050	0.1278	0.3210
V_6	0.2135	0.0464	0.1979	-0.6726	0.1438	0.2908
V_9	0.2103	0.0429	0.2385	-0.6098	0.1472	0.2857
V_{12}	0.2107	0.0432	0.2482	-0.7033	0.1458	0.2894
V_{18}	0.2069	0.0402	0.1646	-0.7260	0.1458	0.2741

B. Principal components analysis						
Variance explained						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	1.2805	0.0899	0.0064	0.0047	0.0018	0.0013
PVAR	0.9247	0.0649	0.0046	0.0034	0.0014	0.0010

Loadings						
	P1	P2	P3	P4	P5	P6
V_1	0.5001	-0.5986	0.5231	-0.3363	0.0005	0.0679
V_2	0.4762	-0.3483	-0.4670	0.6155	0.2243	0.0671
V_6	0.4031	0.1374	-0.2604	-0.1986	-0.4209	-0.7308
V_9	0.3708	0.2866	-0.2684	-0.2407	-0.4422	0.6743
V_{12}	0.3487	0.4230	-0.0754	-0.3576	0.7513	-0.0366
V_{18}	0.3183	0.4902	0.6021	0.5316	-0.1115	-0.0270

Table 4
Testing the number of state variables

Column two of this table reports the mean squared errors (MSE) for different models. M_0 is the model with underlying as the only state variable; M_1 , M_2 and M_3 represent the models with one, two and three additional state variables, respectively. The numbers in the brackets indicate which variance-swap prices are used for the fitting. For example, $M_2(1, 18)$ is the model with two additional state variables, with V_1 and V_{18} as the proxies for the state variables. The third column indicates the restricted models, which are tested against the more general models. R^2 indicates the improvement of the more general models upon the restricted models, defined as $1 - \frac{\text{MSE}(M_{i+1})}{\text{MSE}(M_i)}$. The bootstrap p-value is reported in the last column.

Model	MSE($\times 10^{-3}$)	Restricted	R^2	p-value
M_0	2.3037			
$M_1(1)$	0.3894	M_0	0.8310	0.00
$M_2(1, 18)$	0.2070	$M_1(1)$	0.4683	0.00
$M_3(1, 6, 18)$	0.1945	$M_2(1, 18)$	0.0604	0.18
$M_3(1, 9, 18)$	0.2010	$M_2(1, 18)$	0.0290	0.30
$M_3(1, 12, 18)$	0.1986	$M_2(1, 18)$	0.0407	0.11

Table 5**Principal components analysis of residuals**

Principal components analysis is employed on the weekly average of the implied volatility fitted residuals in each moneyness and maturity bin as specified in Table 1. We use the following models to fit the implied volatility. M_0 is the model with underlying as the only state variable; M_1 , M_2 and M_3 represent the models with one, two and three additional state variables, respectively. The numbers in the bracket indicate which variance-swap prices are used for the fitting. For example, $M_2(1, 18)$ is the model with two additional state variables, with V_1 and V_{18} as the proxies for the state variables. VAR denotes the variance explained (multiplied by 100) by the first six principal components. PVAR denotes the proportion of variance explained by each component out of the total variance.

A. Model M_0						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	7.2993	0.3908	0.0917	0.0316	0.0251	0.0171
PVAR	0.9233	0.0494	0.0116	0.0040	0.0032	0.0022
B. Model $M_1(1)$						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	0.5888	0.1289	0.0391	0.0317	0.0188	0.0132
PVAR	0.6845	0.1499	0.0454	0.0369	0.0219	0.0153
C. Model $M_2(1, 18)$						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	0.1048	0.0825	0.0406	0.0349	0.0246	0.0175
PVAR	0.2985	0.2350	0.1155	0.0995	0.0702	0.0499
D. Model $M_3(1, 6, 18)$						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	0.1061	0.0569	0.0320	0.0294	0.0217	0.0090
PVAR	0.3629	0.1946	0.1094	0.1004	0.0741	0.0307
E. Model $M_3(1, 9, 18)$						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	0.1000	0.0577	0.0329	0.0296	0.0201	0.0086
PVAR	0.3489	0.2015	0.1149	0.1035	0.0702	0.0300
F. Model $M_3(1, 12, 18)$						
	P1	P2	P3	P4	P5	P6
VAR($\times 10^{-2}$)	0.1064	0.0569	0.0379	0.0280	0.0199	0.0090
PVAR	0.3601	0.1925	0.1283	0.0947	0.0675	0.0305

Table 6**Simulation result of the bootstrap test**

This table shows the number of rejections of the null hypothesis out of 100 tests using the bootstrap test and the simulated data. Panel A is for the case of two useful regressors and one useless regressor, i.e., the data are simulated from model M_0 . Panel B is for the case of three useful regressors and one useless regressor, i.e. the data are simulated from model $M_1(1)$. Panel C is for the case of four useful regressors and one useless regressor, i.e., the data are simulated from model $M_2(1, 18)$. The results are shown for different values of the constant, c , in the bandwidth. The results for different rejection rules are also shown. 95% indicates that the null is rejected if the R^2 calculated from the original sample is greater than the 95th percentile of the empirical distribution of R^2 from the bootstrap samples. 90% indicates that the threshold for rejection is 90th percentile. The results for the sample sizes of 500, 1000 and 2000 are shown.

A. M_0	$c = 0.85$		$c = 1$		$c = 1.2$	
	95%	90%	95%	90%	95%	90%
$N = 500$	10	13	5	7	7	7
$N = 1000$	3	6	8	8	8	9
$N = 2000$	7	10	8	9	5	9
B. $M_1(1)$	$c = 0.85$		$c = 1$		$c = 1.2$	
	95%	90%	95%	90%	95%	90%
$N = 500$	17	25	12	15	7	7
$N = 1000$	18	25	9	13	3	5
$N = 2000$	13	20	6	9	3	10
C. $M_2(1, 18)$	$c = 0.85$		$c = 1$		$c = 1.2$	
	95%	90%	95%	90%	95%	90%
$N = 500$	11	21	4	7	4	5
$N = 1000$	19	29	13	14	2	6
$N = 2000$	18	28	8	10	3	5

Table 7**Testing the linearity in the drift, squared diffusion, and conditional covariance**

This table reports the results of bootstrap tests of the linearity of the drift, squared diffusion, and conditional covariance of V_1^2 and V_{18}^2 . The first column indicates the functional form in the null hypothesis that is linear in at least one component of (V_1^2, V_{18}^2) . The alternative hypothesis is unspecified and estimated nonparametrically. We calculate the $R^2 = 1 - \frac{\text{MSE}(H_1)}{\text{MSE}(H_0)}$, where $\text{MSE}(H_0)$ and $\text{MSE}(H_1)$ are the mean squared error for the models under null and alternative hypotheses, respectively. The numbers in the table are the p-value of the R^2 from the original sample against the empirical distribution of the R^2 s from 100 bootstrap samples.

	$\mu_1(V_1^2, V_{18}^2)$	$\mu_{18}(V_1^2, V_{18}^2)$	$\sigma_1^2(V_1^2, V_{18}^2)$	$\sigma_{18}^2(V_1^2, V_{18}^2)$	$\sigma_{1,18}(V_1^2, V_{18}^2)$
$a + bV_1^2$	0.00		0.00		
$a + bV_{18}^2$		0.13		0.00	
$a + b_1V_1^2 + b_2V_{18}^2$	0.01	0.41	0.01	0.04	0.32
$g(V_1^2) + bV_{18}^2$	0.43	0.33	0.20	0.02	0.31
$bV_1^2 + g(V_{18}^2)$	0.01	0.42	0.01	0.12	0.32

Table 8**Testing the number of state variables using an alternative factor proxy**

The average implied volatility for different maturities is used to proxy the state variables. Column two of this table reports the mean squared errors (MSE) for different models. M'_0 is the model with underlying as the only state variable; M'_1 , M'_2 and M'_3 represent the models with one, two and three additional state variables, respectively. The numbers in the brackets indicate which state variables are used for the fitting. For example, $M'_2(1, 18)$ is the model with two additional state variables, with AIV_1 and AIV_{18} as the proxy for the state variables. The third column indicates the restricted models, which are tested against the more general models. R^2 indicates the improvement of the more general models upon the restricted models, defined as $1 - \frac{MSE(M'_{i+1})}{MSE(M'_i)}$. The bootstrap p-value is reported in the last column.

Model	MSE($\times 10^{-3}$)	Restricted	R^2	p-value
M'_0	2.3037			
$M'_1(1)$	0.4704	M'_0	0.7958	0.00
$M'_2(1, 18)$	0.2635	$M'_1(1)$	0.4398	0.00
$M'_3(1, 6, 18)$	0.2520	$M'_2(1, 18)$	0.0438	0.09
$M'_3(1, 9, 18)$	0.2542	$M'_2(1, 18)$	0.0352	0.19
$M'_3(1, 12, 18)$	0.2494	$M'_2(1, 18)$	0.0536	0.12

Table 9**Weighted least squares estimates of a nonlinear model (bivariate)**

The parameters are estimated for the nonlinear drift function

$$V_{\tau_j,t+1}^2 - V_{\tau_j,t}^2 = \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{1,t}^2 + \alpha_{2,\tau_j} \widetilde{V}_{18,t}^2 + \alpha_{3,\tau_j} \widetilde{V}_{1,t}^4 + \alpha_{4,\tau_j} \widetilde{V}_{18,t}^4 \\ + \alpha_{5,\tau_j} (V_{1,t}^2 \widetilde{V}_{18,t}^2) + \alpha_{6,\tau_j} (1/\widetilde{V}_{1,t}^2) + \alpha_{7,\tau_j} (1/\widetilde{V}_{18,t}^2) + \eta_{\tau_j,t+1}, \quad \tau_j = 1, 18.$$

\widetilde{V}_{18}^2 is the residual of V_{18}^2 regressed on a constant and V_1^2 . \widetilde{V}_1^4 is the residual of V_1^4 regressed on a constant, V_1^2 and \widetilde{V}_{18}^2 . Other variables, \widetilde{V}_{18}^4 , $(V_1^2 \widetilde{V}_{18}^2)$, $(1/\widetilde{V}_1^2)$ and $(1/\widetilde{V}_{18}^2)$ are defined similarly. Given the fitted values for the drift, the diffusion parameters are estimated from a log linear regression using the squared residuals from the first regression as

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

where $\widehat{\eta}_{\tau_j,t+1}$ is the residual from the drift function regression. Given the fitted values of the residuals, the drift parameters are re-estimated by weighted least squares. The iterative procedure is applied until the parameter estimate converges. The t-statistics adjusted for heteroscedasticity and 24 lags of autocorrelation using Newey and West (1987) are reported in parentheses.

A. Drift of V_1^2							
$\alpha_{0,1} \times 10^{-3}$	$\alpha_{1,1}$	$\alpha_{2,1}$	$\alpha_{3,1}$	$\alpha_{4,1}$	$\alpha_{5,1}$	$\alpha_{6,1} \times 10^{-5}$	$\alpha_{7,1} \times 10^{-5}$
1.2114	-0.0246	0.0248	-0.2787	-0.4547	-1.6522	1.8570	7.7464
(3.72)	(-2.89)	(1.97)	(-1.95)	(-1.16)	(-1.74)	(2.10)	(2.73)
B. Drift of V_{18}^2							
$\alpha_{0,18} \times 10^{-3}$	$\alpha_{1,18}$	$\alpha_{2,18}$	$\alpha_{3,18}$	$\alpha_{4,18}$	$\alpha_{5,18}$	$\alpha_{6,18} \times 10^{-5}$	$\alpha_{7,18} \times 10^{-5}$
0.0071	-0.0004	-0.0090	-0.0045	-0.1409	-0.4266	0.4588	1.2677
(0.10)	(-0.25)	(-2.66)	(-0.12)	(-1.29)	(-1.55)	(1.42)	(1.15)
C. Diffusion of V_1^2							
$\ln(\beta_1)$	$\gamma_1 - 1$						
-5.0455	1.2610						
(-15.12)	(12.48)						
D. Diffusion of V_{18}^2							
$\ln(\beta_{18})$	$\gamma_{18} - 1$						
-8.7467	0.9214						
(-16.66)	(5.74)						

Table 10**Weighted least squares estimates of a nonlinear model (univariate)**

The parameters are estimated for the nonlinear drift function

$$V_{\tau_j,t+1}^2 - V_{\tau_j,t}^2 = \alpha_{0,\tau_j} + \alpha_{1,\tau_j} V_{\tau_j,t}^2 + \alpha_{2,\tau_j} \widetilde{V}_{\tau_j,t}^4 + \alpha_{3,\tau_j} (1/\widetilde{V}_{\tau_j,t}^2) + \eta_{\tau_j,t+1}, \quad \tau_j = 1, 18.$$

$\widetilde{V}_{\tau_j,t}^4$ is the residual of $V_{\tau_j,t}^4$ regressed on a constant and $V_{\tau_j,t}^2$. $(1/\widetilde{V}_{\tau_j,t}^2)$ is the residual of $1/V_{\tau_j,t}^2$ regressed on a constant, $V_{\tau_j,t}^2$ and $\widetilde{V}_{\tau_j,t}^4$. Given the fitted values for the drift, the diffusion parameters are estimated from a log linear regression using the squared residuals from the first regression as

$$\ln(\widehat{\eta}_{\tau_j,t+1}^2) = \ln(\beta_{\tau_j}) + \gamma_{\tau_j} \ln(V_{\tau_j,t}^2) + \xi_{\tau_j,t+1}, \quad j = 1, 18,$$

where $\widehat{\eta}_{\tau_j,t+1}$ is the residual from the drift function regression. Given the fitted values of the residuals, the drift parameters are re-estimated by weighted least squares. The iterative procedure is applied until the parameter estimate converges. The t-statistics adjusted for heteroscedasticity and 24 lags of autocorrelation using Newey and West (1987) are reported in parentheses.

A. Drift of V_1^2			
$\alpha_{0,1} \times 10^{-3}$	$\alpha_{1,1}$	$\alpha_{2,1}$	$\alpha_{3,1} \times 10^{-5}$
1.3046	-0.0265	-0.4345	1.3962
(3.78)	(-2.93)	(-2.97)	(1.97)
B. Drift of V_{18}^2			
$\alpha_{0,18} \times 10^{-3}$	$\alpha_{1,18}$	$\alpha_{2,18}$	$\alpha_{3,18} \times 10^{-5}$
0.2059	-0.0049	-0.1317	1.8192
(2.51)	(-2.24)	(-1.36)	(1.63)
C. Diffusion of V_1^2			
$\ln(\beta_1)$	$\gamma_1 - 1$		
-5.0953	1.2458		
(-15.54)	(12.60)		
D. Diffusion of V_{18}^2			
$\ln(\beta_{18})$	$\gamma_{18} - 1$		
-8.8993	0.8738		
(-17.28)	(5.57)		

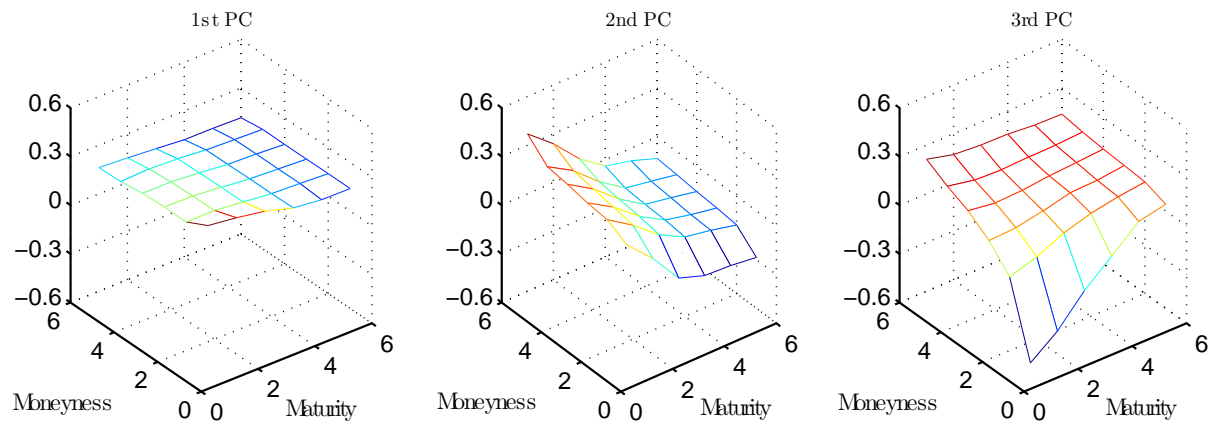


Figure 1. Loadings on the first three principal components

This figure shows the loadings on the first three principal components of the weekly average implied volatilities of thirty-six moneyness and maturity groups.

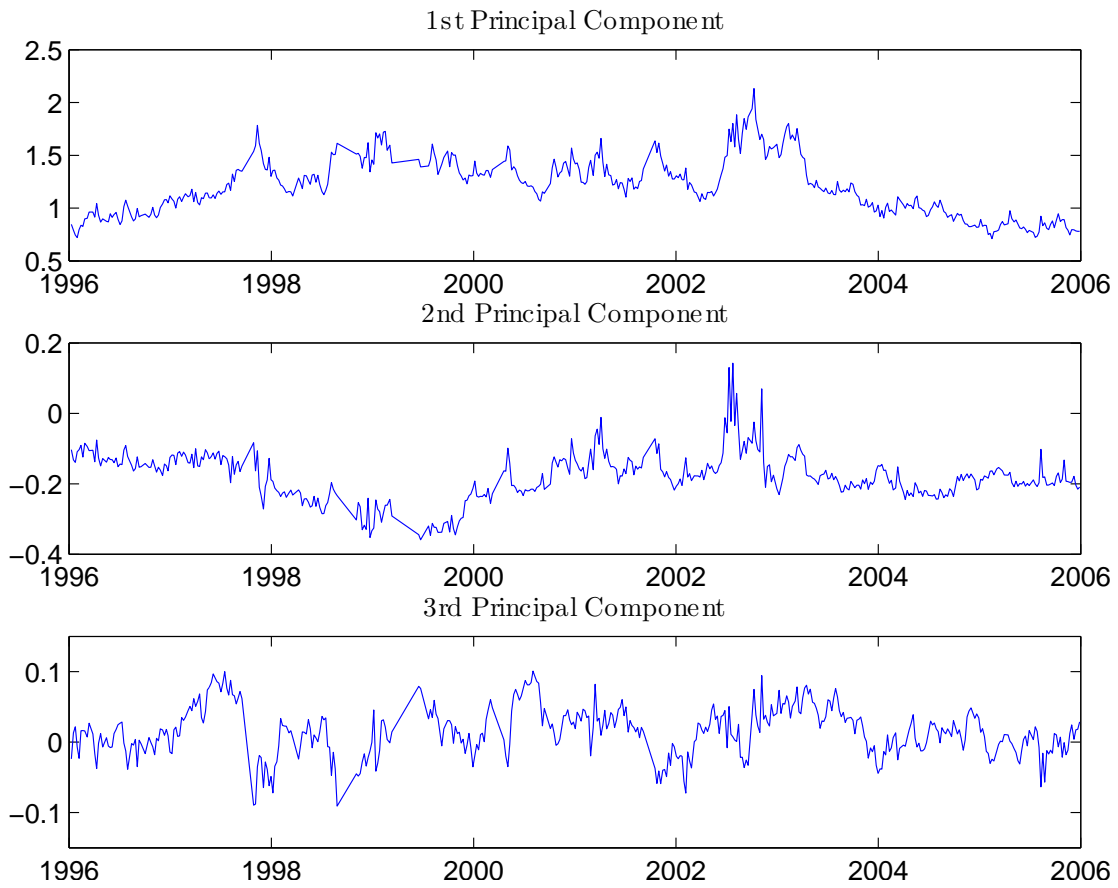


Figure 2. The principal components of the implied volatilities

This figure shows the time-series plot of the first three principal components of the weekly average implied volatility in thirty-six moneyness and maturity groups.

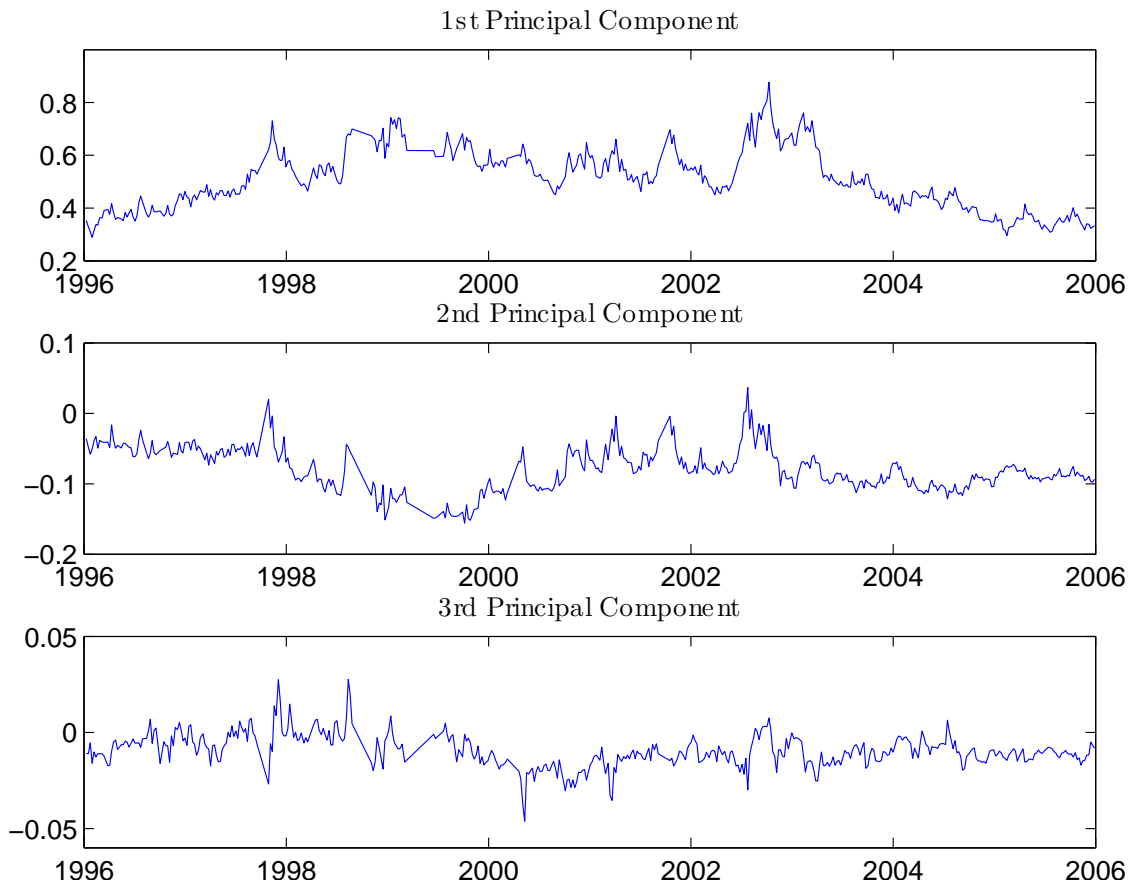


Figure 3. The principal components of the square root of variance-swap prices
This figure shows the time-series plot of the first three principal components of the six model-free square root variance-swap prices.

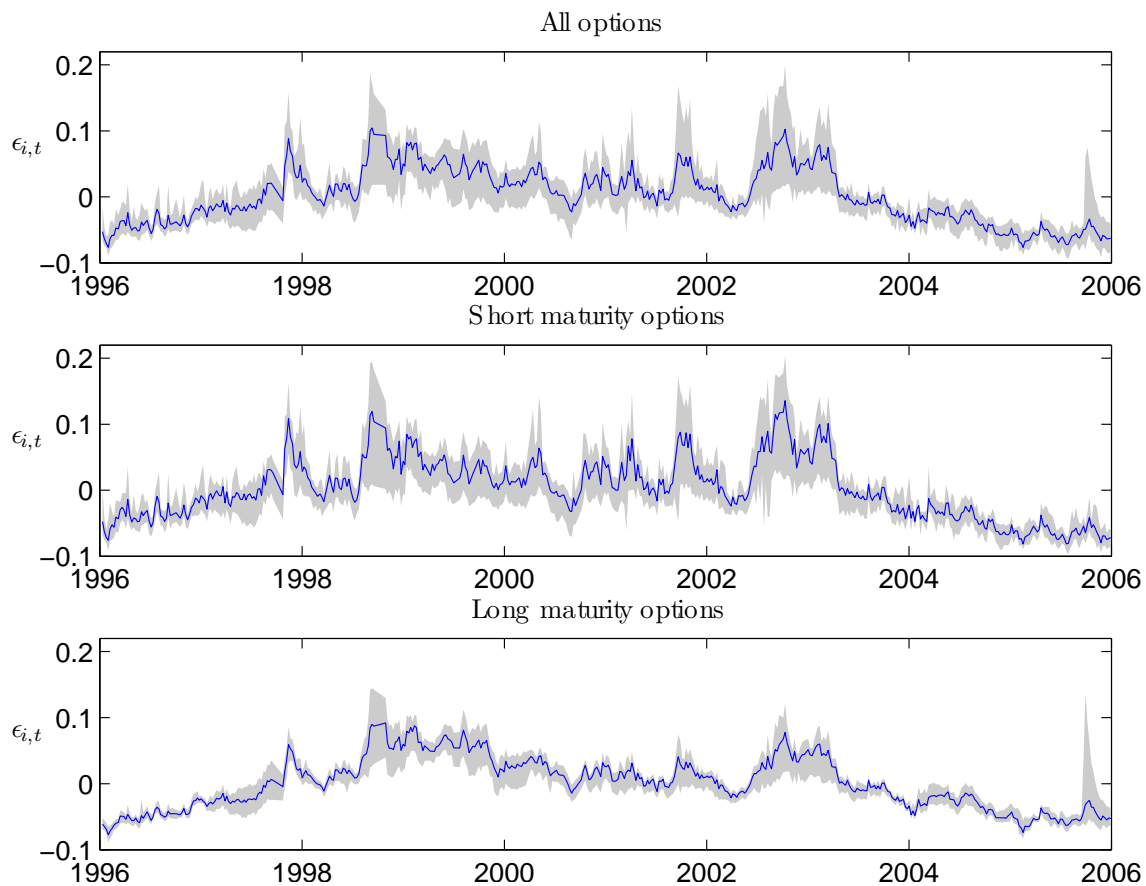


Figure 4. Residuals from model M_0

This figure shows the residuals from model M_0 . The solid line is the weekly average residual and the shaded area covers the 5 and 95 percent of the residual distribution. The upper panel is for the entire sample, the middle panel is for the options with maturities no greater than 180 days, and the lower panel is for the options with maturities longer than 180 days.

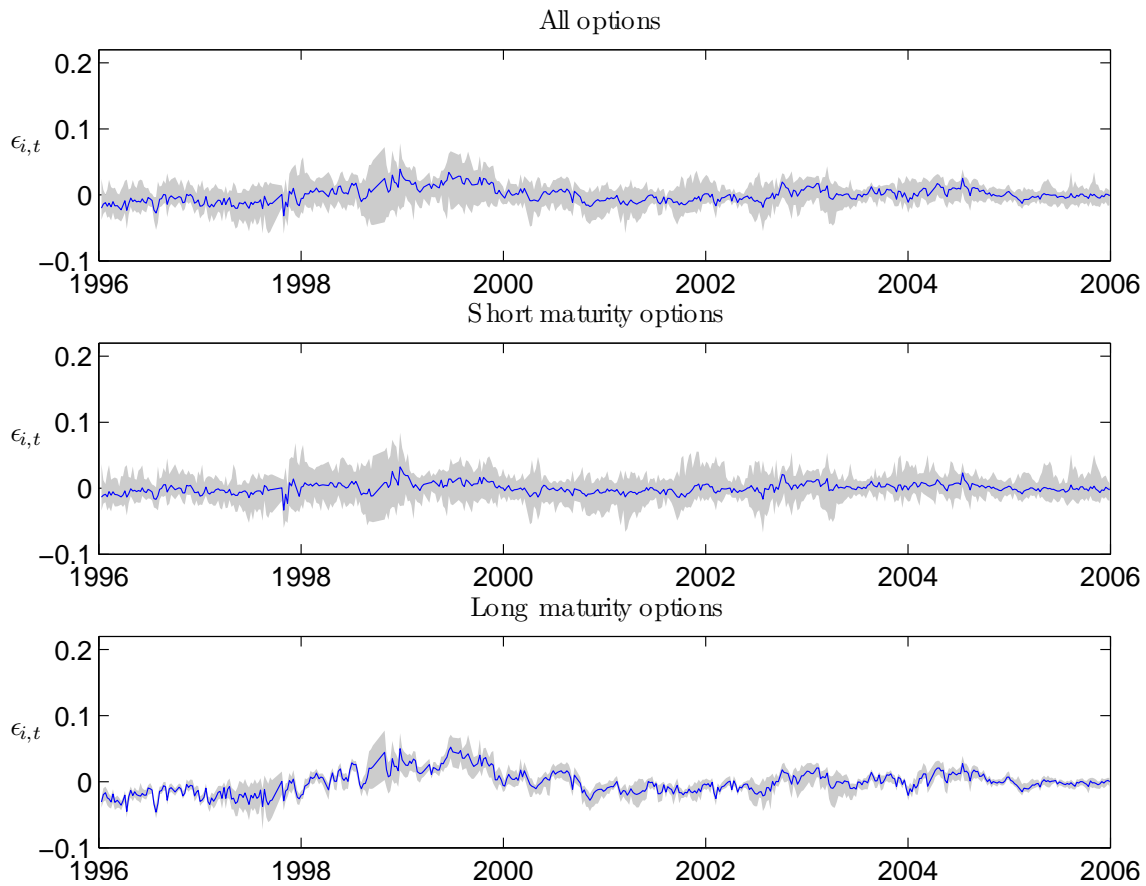


Figure 5. Residuals from model $M_1(1)$

This figure shows the residuals from model $M_1(1)$. The solid line is the weekly average residuals, and the shaded area covers the 5 and 95 percent of the residual distribution. The upper panel is for the entire sample, the middle panel is for the options with maturities no greater than 180 days, and the lower panel is for the options with maturities longer than 180 days.

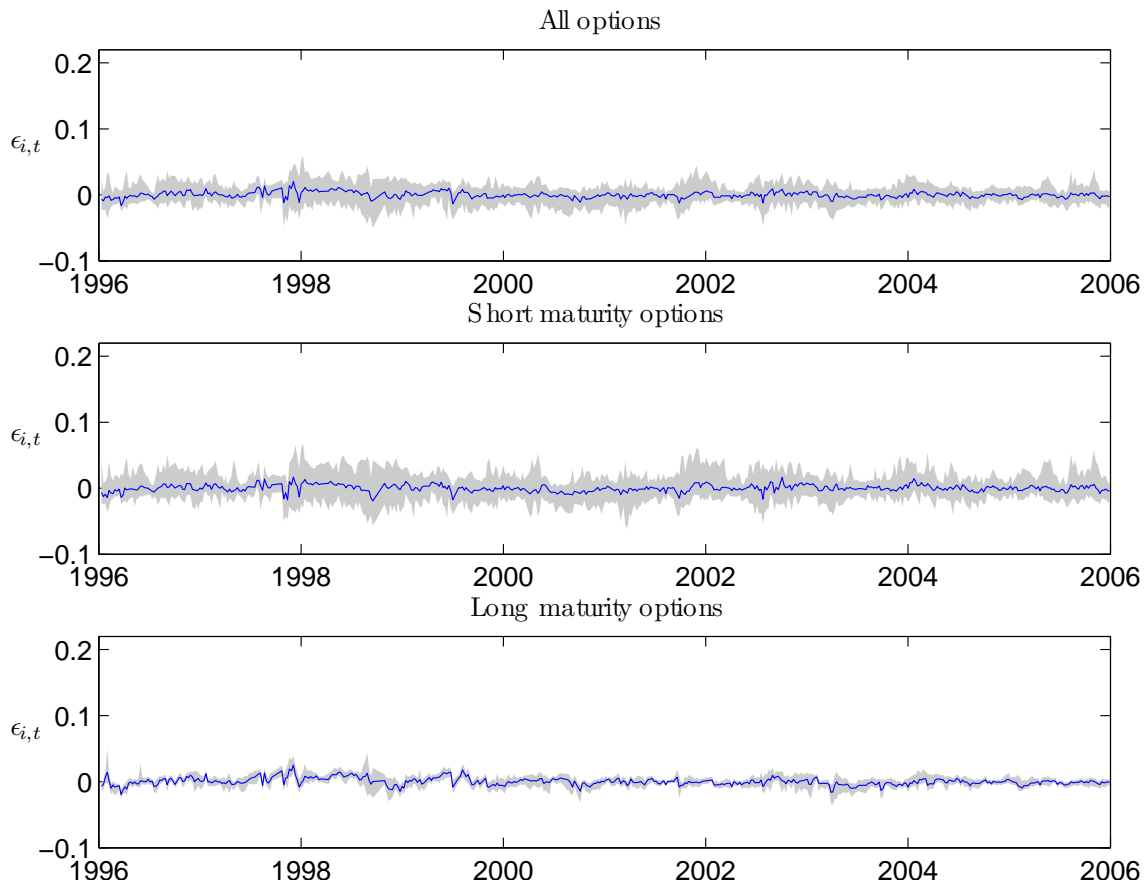


Figure 6. Residuals from model $M_2(1, 18)$

This figure shows the residuals from model $M_2(1, 18)$. The solid line is the weekly average residuals, and the shaded area covers the 5 and 95 percent of the residual distribution. The upper panel is for the entire sample, the middle panel is for the options with maturities no greater than 180 days, and the lower panel is for the options with maturities longer than 180 days.

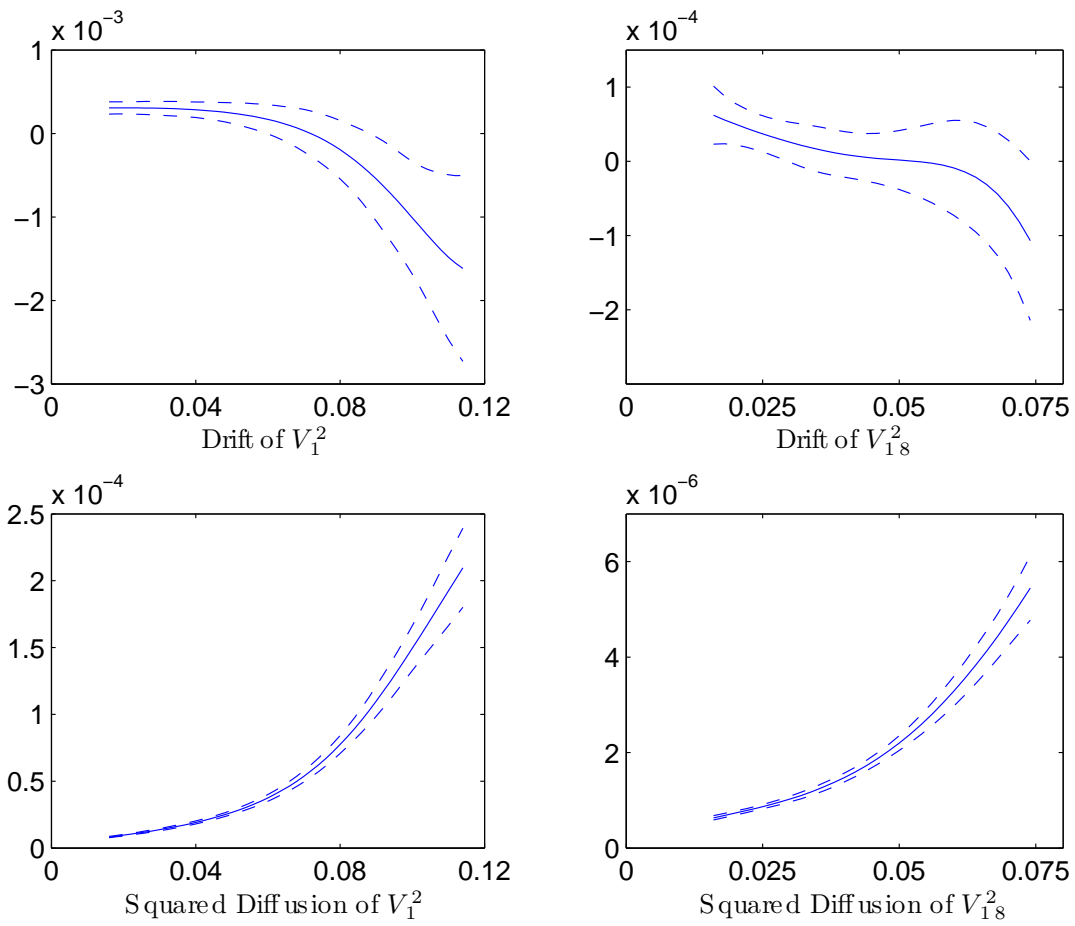


Figure 7. Estimated drift and squared diffusion (one-dimensional)

This figure shows the nonparametric estimation of the drift and squared diffusion for V_1^2 and V_{18}^2 . The solid line is the fitted curve. The dashed lines cover one standard deviation of the fitted value from both sides. The two panels on the left are for V_1^2 and the two panels on the right are for V_{18}^2 .

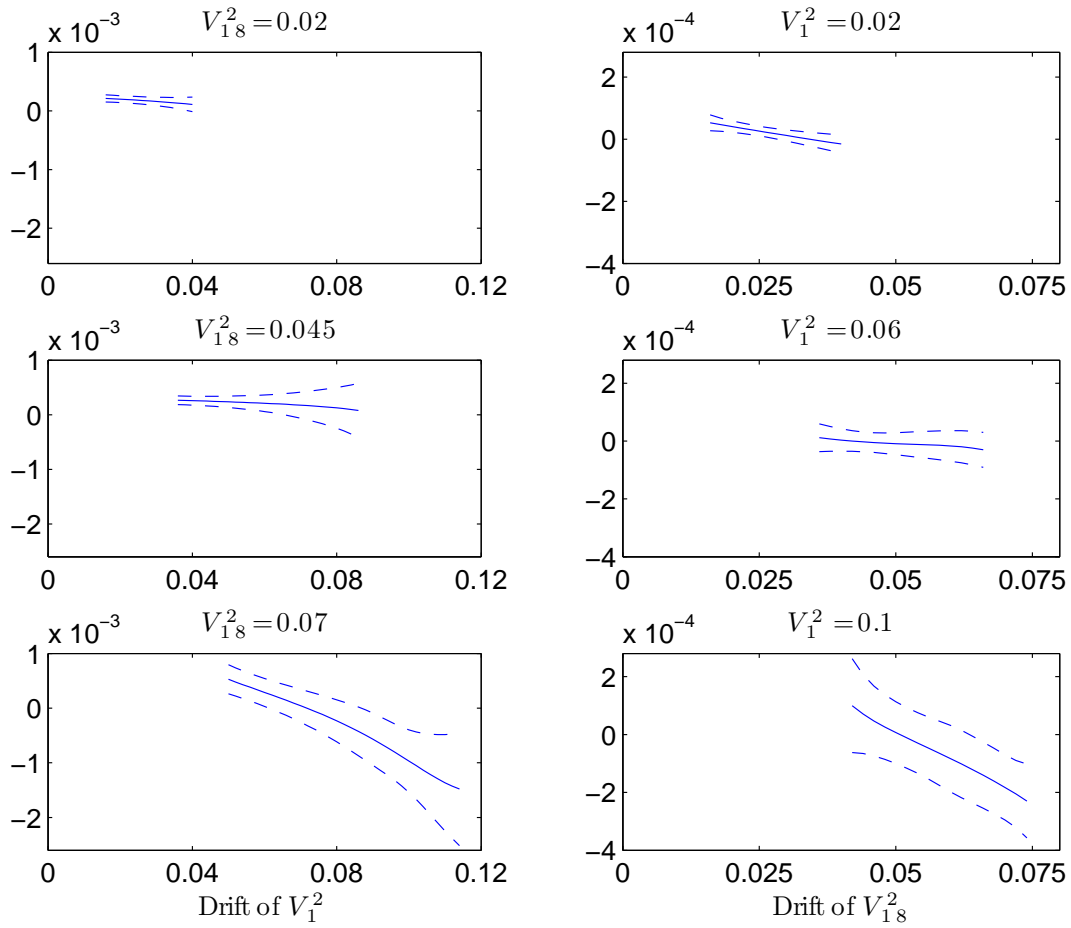


Figure 8. Estimated drift (two-dimensional)

The left panels of the figure are the nonparametric fit of the drift of V_1^2 as a function of V_1^2 and V_{18}^2 and the right panels are the drift of V_{18}^2 as a function of V_1^2 and V_{18}^2 . The solid line is the fitted value. The dashed lines cover one standard deviation from the fitted value on both sides.

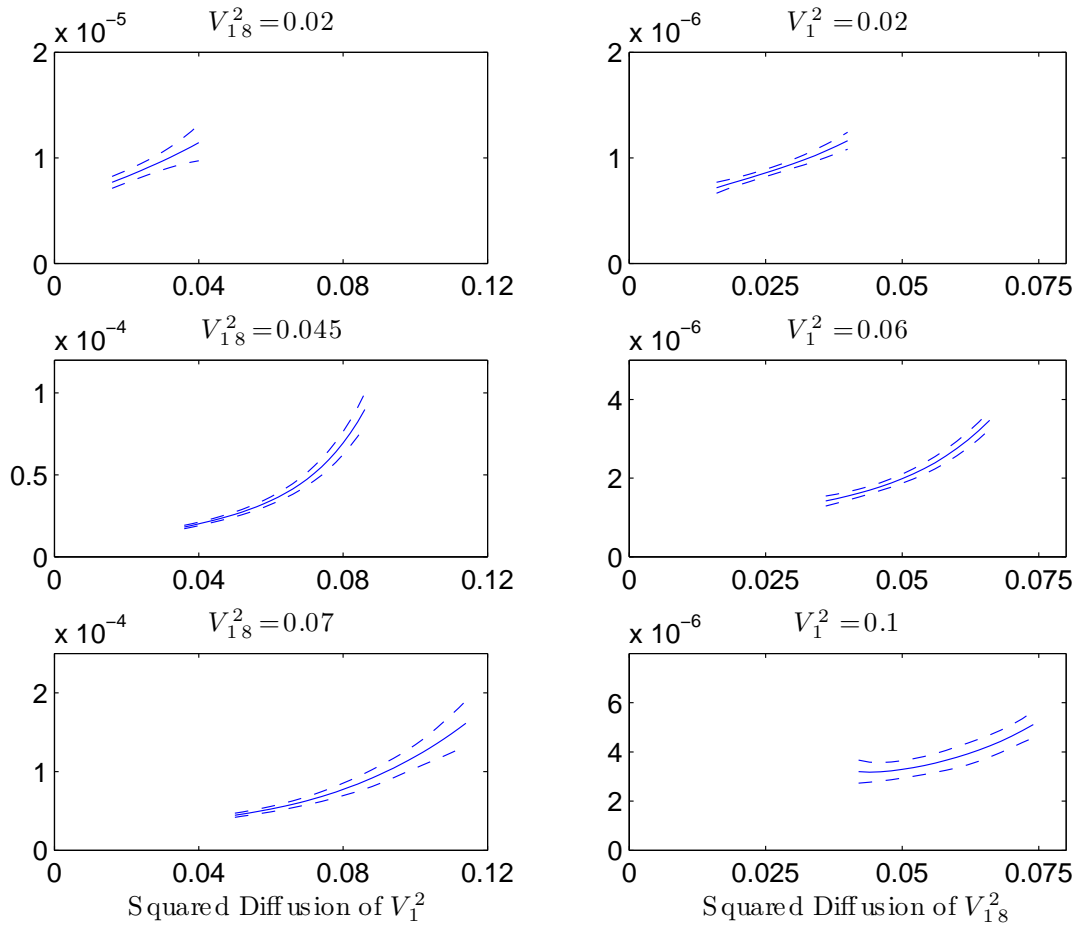


Figure 9. Estimated squared diffusion (two-dimensional)

The left panels of the figure are the nonparametric fit of the squared diffusion of V_1^2 as a function of V_1^2 and V_{18}^2 and the right panels are the squared diffusion of V_{18}^2 as a function of V_1^2 and V_{18}^2 . The solid line is the fitted value. The dashed lines cover one standard deviation from the fitted value on both sides.

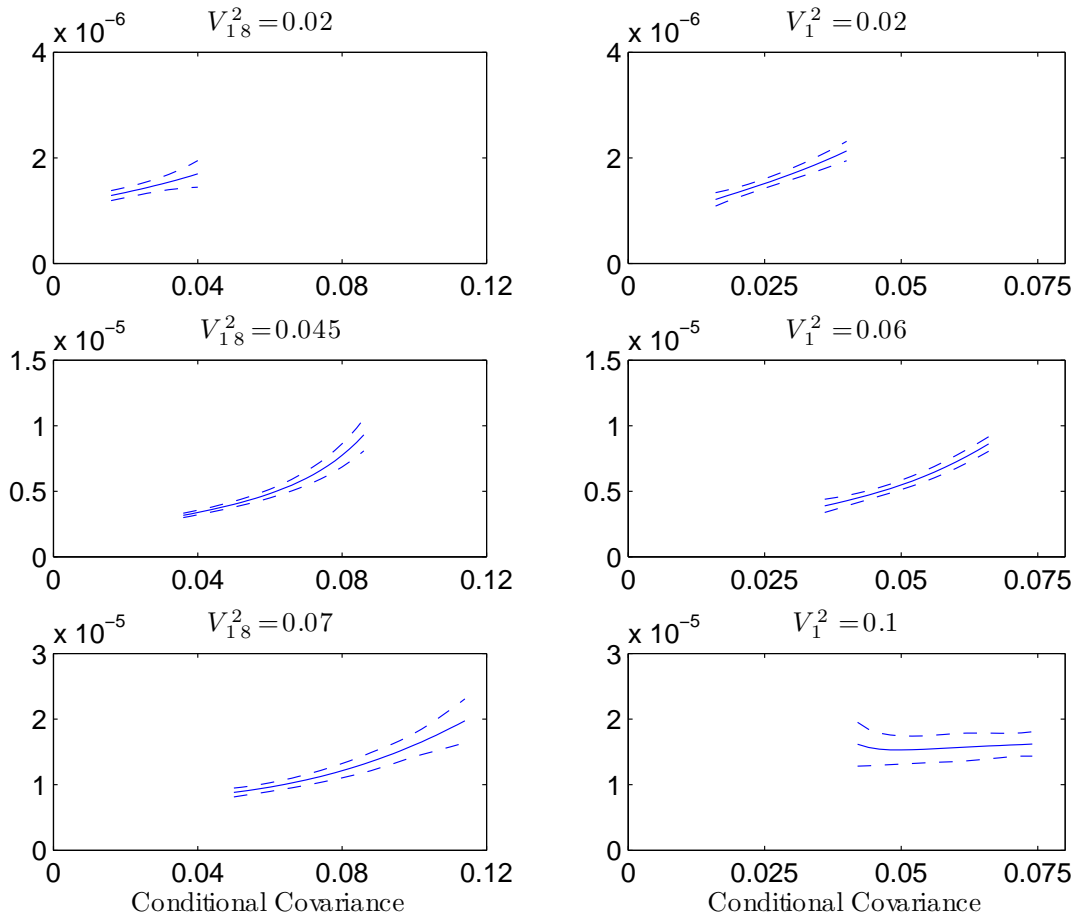


Figure 10. Estimated conditional covariance

This figure shows the nonparametric fit of the conditional covariance between V_1^2 and V_{18}^2 . The left panels show the conditional covariance as a function of V_1^2 for different levels of V_{18}^2 and the right panels show the conditional covariance as a function of V_{18}^2 for different levels of V_1^2 . The solid line is the fitted value. The dashed lines cover one standard deviation from the fitted value on both sides.