TESTING THE PARAMETRIC SPECIFICATION OF THE DIFFUSION FUNCTION IN A DIFFUSION PROCESS

Fuchun Li[1]

Bank of Canada

Abstract

A new consistent test is proposed for the parametric specification of the diffusion function in a diffusion process without any restrictions on the functional form of the drift function. The data are assumed to be sampled discretely in a time interval which can be fixed or lengthened to infinity. The test statistic is shown to follow an asymptotic normal distribution under the null hypothesis that the parametric diffusion function is correctly specified. Monte Carlo simulations are conducted to examine the finite sample performance of the test, revealing that the test has good size and power.

Keywords: Consistent Test; Diffusion Process; Absolutely Regular Process; Degenerate U-statistics; Kernel estimation.

JEL Classification Codes: C12, C14.

---

1. The author is grateful to Yacine Aït-Sahalia, John Knight, Oliver Linton (the co-editor), Greg Tkacz, Jun Yang and three anonymous referees for helpful comments and suggestions. He also thanks seminar participants at the Bank of Canada, and the 2004 Semiparametrics Conference in Rio de Janeiro. The views expressed in this paper are those of the author. No responsibility for them should be attributed to the Bank of Canada. Any errors or omissions are those of the author. Address correspondence to: Fuchun Li, Department of Monetary and Financial Analysis, Bank of Canada, 234 Wellington, Ottawa Canada K1A 0G9. Tel. 613-782-7392, E-mail: fuchunli@bank-banque-canada.ca.

1. Introduction

In economics and finance, continuous-time models have been widely used to study the dynamics of underlying state variables, such as asset prices, exchange rates, or spot interest rates. The basic modeling approach in this literature is to assume that the underlying state variables follow a stochastic differential equation.

In the parametric specification of a stochastic differential equation, it is assumed that the functional forms of the drift and diffusion functions are known, apart from a finite number of unknown parameters. Given the parametric specification of a stochastic differential equation, researchers have proposed many different methods to estimate the unknown parameters and to derive the statistical inferences from the discrete observations.

The validity of these estimation and inference procedures, however, is conditional on the hypothesis that the continuous-time model described by a stochastic differential equation is correctly specified. Unfortunately, economic theory typically does not suggest functional forms for the continuous-time model. Model misspecification may lead to misleading conclusions in inference and hypothesis testing. This motivates the development of model specification tests for continuous-time models.

Gallant and Tauchen (1996) proposed a minimum chi-square specification test for continuous-time models using the efficient method of moments. Aït-Sahalia (1996b) proposed two specification tests by comparing the model-implied parametric density to the same density estimated nonparametrically. Diebold, Gunther and Tay (1998), Thompson (2001), and Hong and Li (2002) proposed transition density-based specification tests based on the fact that the probability integral transform of the model-implied transition density would be distributed as an independent and identical uniform distribution under the correct model specification. Li and Tkacz (2004) proposed a parametric bootstrap procedure to approximate the finite sample distribution of a goodness-of-fit test statistic of a parametric transition density. Corradi and Swanson (2004) proposed a Kolmogorov type conditional distribution test. As the limiting distribution of their test statistic is nuisance parameters free, Corradi and Swanson (2004) used a nonparametric bootstrap procedure to construct the critical values.

It is noted that the null hypothesis of all above-mentioned tests is that both the drift and diffusion functions are specified correctly. Under such a null hypothesis, while these tests can detect a wide range of model misspecifications, they cannot reveal possible sources of model misspecifications. However, for the specification analysis of the continuous-time model, when a misspecified model is rejected, one would like to explore the possible reasons for the rejection. Specifically, is the rejection due to the misspecification from the drift function or the diffusion function? When economic theory provides little guidance about the specification of the drift and diffusion functions, it is desirable to be able to develop a reliable test which can detect whether the model misspecification comes from the drift function or the diffusion function. It should be mentioned that transition density-based tests can be used to test the specification of the diffusion function (drift function) only through presupposing both the correct specification of the drift function (diffusion function) and the availability of the closed-form expression of the model-implied transition density. Unfortunately, even if the drift function (diffusion function) is specified correctly, the closed-form expression of a transition density still cannot be available for most continuous-time models. In fact, Wong (1964) showed that a stationary diffusion process has a closed-form expression of its transition density only if it has a linear specification for the drift function and a quadratic specification for the diffusion function. As a consequence, these transition density-based tests are only applicable to very rare continuous-time models.

The limitations of the above-mentioned tests and the recent developments in nonparametric estimation techniques of a continuous-time model prompt us to make use of nonparametric estimation techniques in developing tests for a parametric form of a continuous-time model by directly testing the specifications of its drift and diffusion functions, without relying on the model-implied density function or model-implied moment condition. Corradi and White (1997) provided a first step in this direct. With knowledge of the functional form of the drift function being not required, Corradi and White (1997) proposed a specification test for the diffusion function based on discrete sampling observations.

However, as pointed out by Corradi and White (1997), their test can only be used to test a parametric diffusion function at a given point and the time span of observations is fixed. Thus, their test cannot be used to detect diffusion function misspecifications over a continuous range of the state variable.

Based on discrete observations, we propose a new test of the functional specification of the diffusion function without any restriction on the functional form for the drift function. Our test can be used to test the specification of the parametric diffusion function over a time interval which can be fixed or lengthened to infinity. Using theories of degenerate U-statistics, the test statistic is shown to be asymptotically distributed standard normal under the null hypothesis, while diverging to infinity if the parametric specification is misspecified over a significant range.

This paper is organized as follows. In Section 2 we state the hypothesis of interest and introduce our test statistic. In Section 3 the asymptotic properties of the test are discussed. The size and power performance of the test are examined by Monte Carlo study in section 4. Section 5 concludes. The proofs are provided in Appendix.

2. The Hypothesis and Test Statistic

The model we consider is the following autonomous stochastic differential equation,

$$dx_t = \mu(x_t)dt + \sigma(x_t)dw_t, \tag{2.1}$$

with initial condition x_{t_0}, where x_t is the state variable, and $\{w_t\text{:t} \geq 0\}$ is a standard Brownian motion process. The functions $\mu(\cdot)$ and $\sigma^2(\cdot)$ are, respectively, the drift function and the diffusion function of the process $\{x_t\text{:t} \geq 0\}$.

We assume that the process $\{x_t\text{:t} \geq 0\}$ is observed at $\{t = t_1, t_2, \ldots, t_n\}$ in the time interval $[t_0, T]$. Further assume that the observations are equispaced. Then,

$\{x_t = x_{t_0+\Delta_n}, x_{t_0+2\Delta_n}, \ldots, x_{t_0+n\Delta_n}\}$ are n observations on the process $\{x_t\colon t \geq 0\}$ at dates $\{t_1 = t_0 + \Delta_n, t_2 = t_0 + 2\Delta_n, \ldots, t_n = t_0 + n\Delta_n\}$, where $\Delta_n = (T - t_0)/n$ is the sampling interval.

We use the notation $x_{n,j}$ to express the observation on the process $\{x_t\colon t \geq 0\}$ at $t = t_0 + j\Delta_n$, i.e., $x_{n,j} \equiv x_{t_0+j\Delta_n}$, where $j = 1, 2, \ldots, n$ and $n \geq 1$.

A parametric family of the specification of the diffusion function $\sigma^2(\cdot)$ is

$\{\sigma_0^2(x, \theta) : \theta \in \Theta\}$ with Θ being a subset of R^d. We want to justify the use of a parametric specification of the diffusion function without knowledge of the functional form of the drift function. Thus, the null hypothesis to be tested is that the parametric specification of $\sigma^2(.)$ is correct,

$$H_0\colon \qquad \sigma^2(x) = \sigma_0^2(x, \theta_0) \quad \text{almost everywhere for some } \theta_0 \in \Theta. \tag{2.2}$$

The alternative hypothesis is $\sigma^2(x) \neq \sigma_0^2(x, \theta)$ for all $\theta \in \Theta$ over a significant range, that is,

$$H_1\colon \qquad \sigma^2(x) \neq \sigma_0^2(x, \theta) \quad \text{on a subset S with positive measure for any } \theta \in \Theta. \tag{2.3}$$

Our testing approach is based on the squared error goodness-of-fit function between $\sigma^2(x)$ and $\sigma_0^2(x, \theta)$,

$$\begin{aligned} I(\theta, \sigma_0^2) &\equiv E\left\{[(\sigma^2(x_t) - \sigma_0^2(x_t, \theta))\pi(x_t)]^2 a(x_t)\right\} \\ &= \int [(\sigma^2(x) - \sigma_0^2(x, \theta))\pi(x)]^2 a(x) dF(x), \end{aligned} \tag{2.4}$$

where $\pi(x)$ and $F(x)$ are respectively the unknown density function and cumulative distribution function of x_t. Distance measures similar to (2.4) were used as a basis for testing the model specifications of either a parametric density function or a general regression function by, for example, Bickel and Rosenblatt (1973), Hall (1984), Fan (1994), Aït-Sahalia, Bickel and Stock (2001), and Li and Tkacz (2004) etc.

The introduction of the density function $\pi(x)$ in (2.4) is to avoid the problem of trimming the small values of the random denominator in nonparametric estimation of the diffusion function. The inclusion of the weighting function $a(x)$ in (2.4) allows us to focus goodness-of-fit testing on particular ranges of the state variable. Specifically, we will assume $a(x)$ to be bounded with compact support $S \subset R$ (Assumption 8 in Section 3). This assumption will help us to avoid technical problems in proving uniform convergence of the nonparametric estimations of the marginal density and diffusion function on S. By choosing an appropriate $a(x)$, the specification test can be tailored to the empirical question of interest. In practice, $a(x)$ can typically be chosen as the indicator function of a compact set related to the empirical question of interest. For example, to infer the behaviour of the short-term interest rate within a range of levels, say, $[0.05, 0.10]$, only the paths of the state variable which cross the interval [0.05, 0.10] are used in the specification analysis.

Under the null hypothesis H_0, we have $I(\theta_0, \sigma_0^2) = 0$, and under the alternative H_1, we have $I(\theta, \sigma_0^2) > 0$ for any $\theta \in \Theta$. Hence, the measure $I(\theta, \sigma_0^2)$ can be used as an indicator for the misspecification of the diffusion function $\sigma^2(\cdot)$.

If $\sigma^2(x)$, θ_0 and $\pi(x)$ were available, then we could estimate $I(\theta_0, \sigma_0^2)$ by its sample analogue $\frac{1}{n}\sum_{i=1}^{n}[(\sigma^2(x_{n,i}) - \sigma_0^2(x_{n,i}, \theta))\pi(x_i)]^2 a(x_{n,i})$. To get a feasible test statistic, we need to estimate $\sigma^2(x)$, θ_0, and $\pi(x)$.

Under both H_0 and H_1, the true, unknown $\sigma^2(x)$ can be estimated by the kernel method which is proposed by Jiang and Knight (1997) and Bandi and Phillips (2003),

$$\hat{\sigma}_n^2(x) = \frac{\sum_{i=1}^{n-1} K\left(\frac{x_{n,i} - x}{h_n}\right)[x_{n,i+1} - x_{n,i}]^2}{\Delta_n \sum_{i=1}^{n-1} K\left(\frac{x_{n,i} - x}{h_n}\right)}, \tag{2.5}$$

where $K(.)$ is a kernel function, and h_n is a sequence of bandwidth parameters.

As can be seen from (2.5), the nonparametric estimator $\hat{\sigma}_n^2(x)$ is built without imposing any restrictions on the functional form of the drift function. The derivation of the asymptotic distribution of the nonparametric estimator $\hat{\sigma}_n^2(x)$ depends crucially on the assumption $\Delta_n = (T - t_0)/n \to 0$ as $n \to \infty$ (Jiang and Knight (1997) and Bandi and Phillips (2003)). In fact, Nicolau (2003) showed that without the assumption $\Delta_n = (T - t_0)/n \to 0$ as $n \to \infty$, the nonparametric estimator (2.5) is not consistent. In contrast, in a semiparametric model with the drift function specified parametrically, the semiparametric diffusion function estimator proposed by Aït-Sahalia (1996a) and Kristensen (2004) required the sampling interval Δ_n be fixed in order to obtain asymptotic results.

Since the objective of this paper is to construct a test for the specification of the diffusion function without any functional form specification of the drift function, the nonparametric estimation procedure (2.5) based on $\Delta_n = (T - t_0)/n \to 0$ as $n \to \infty$ is used to construct our test statistic.

The estimator of θ_0, $\hat{\theta}_n$, is defined as follows,

$$\hat{\theta}_n = \arg min_{\theta \in \Theta} \sum_{i=1}^{n-1} [\log \sigma_0^2(x_{n,i}, \theta) + (\sigma_0^2(x_{n,i}, \theta)\Delta_n)^{-1}(x_{n,i+1} - x_{n,i})^2] \ . \qquad (2.6)$$

Corradi and White (1997) provided regularity conditions under which $\hat{\theta}_n$ is a quasi-maximum likelihood estimator. These regularity conditions will be given in the Assumptions listed in the next section.

The parametric function $\sigma_0^2(x, \theta_0)$ is estimated by $\sigma_0^2(x, \hat{\theta}_n)$. The unknown density function of x_t, $\pi(x)$, can be consistently estimated by the kernel estimator,

$$\hat{\pi}(x) = \frac{1}{nh_n} \sum_{i=1}^{n-1} K\left(\frac{x_{n,i} - x}{h_n}\right) \ . \qquad (2.7)$$

Let $\hat{F}(x)$ be the empirical cumulative distribution estimator of $F(x)$. Inserting these estimates above into the definition of $I(\theta_0, \sigma_0^2)$, given by (2.4), yields the following estimator of $I(\theta_0, \sigma_0^2)$,

$$\begin{aligned} I_n &= \int [(\hat{\sigma}_n^2(x) - \sigma_0^2(x, \hat{\theta}_n))\hat{\pi}(x)]^2 a(x) d\hat{F}(x) \\ &= \frac{1}{n} \sum_{i=1}^{n} [(\hat{\sigma}_n^2(x_{n,i}) - \sigma_0^2(x_{n,i}, \hat{\theta}_n))\hat{\pi}(x_{n,i})]^2 a(x_{n,i}) \ . \qquad (2.8) \end{aligned}$$

Our test statistics is a properly centered and scaled version of I_n,

$$J_n \equiv (nh_n^{1/2})\left[I_n - \frac{2}{n^2 h_n} \sum_{i=1}^{n} (\hat{\sigma}_n^2(x_{n,i}))^2 a(x_{n,i}) \hat{\pi}(x_{n,i}) \int k^2(u) du\right] / \hat{v}_n \ , \qquad (2.9)$$

where

$$\hat{v}_n^2 = \frac{8}{n}\sum_{i=1}^{n}(\hat{\sigma}^2(x_{n,i}))^4\hat{\pi}^3(x_{n,i})a^2(x_{n,i})\int[\int K(u)K(w+u)du]^2 dw. \tag{2.10}$$

3. Assumptions and The Limiting Distribution of The Test Statistic

We specify assumptions for the functions $\mu(.)$, $\sigma(.)$, and the parametric family of $\{\sigma_0^2(x,\theta) : \theta \in \Theta\}$, under which the asymptotic validity of this test statistic J_n can be established.

Assumption 1. *Let* $D = (l, r)$ *be an open interval with* $-\infty \le l < r \le \infty$. $\mu(.)$ *and* $\sigma(.)$ *are twice continuously differentiable on* D *, and Lipschitz continuity and growth conditions are satisfied, i.e., for any compact subset* $A \subset D$ *there exists a positive constant* C_A *such that for every* $x, y \in C_A$,

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \le C_A|x - y| \ . \tag{3.1}$$

$$\mu^2(x) + \sigma^2(x) \le C_A(1 + x^2) \ . \tag{3.2}$$

Assumption 2. $\sigma^2(x) > 0$ for any $x \in D$.

Assumption 3. *We define:*

$$S(\alpha) = \int_c^{\alpha} \exp\left\{\int_c^y \left[-\frac{2\mu(x)}{\sigma^2(x)}\right]dx\right\}dy. \tag{3.3}$$

where c *is a generic fixed number belonging to* D. *We require* $S(\alpha)$ *to satisfy* $\lim_{\alpha \to l} S(\alpha) = -\infty$, *and* $\lim_{\alpha \to u} S(\alpha) = \infty$.

Assumption 4. $\lim_{|x| \to \infty} |\sigma(x)/[2\mu(x) - \sigma(x)\sigma'(x)]| < \infty$, $\lim_{|x| \to \infty} \sigma(x)\pi(x) = 0$.

Assumption 5. *The parametric space* Θ *is compact. For any* $\theta \in \Theta$, *the given function* $\sigma_0(x, \theta)$ *satisfies Assumptions 1-5, and* $\partial\sigma_0^2(x, \theta)/\partial\theta$, $\partial^2\sigma_0^2(x, \theta)/\partial\theta\partial\theta'$, $\partial\sigma_0^2(x, \theta)/\partial x$, $\partial^2\sigma_0^2(x, \theta)/\partial x\partial\theta$ *exist and are continuous on* $R \times \Theta$.

Assumption 6. *For almost all* $(x, \theta) \in R \times \Theta$, $\sigma_0^2(x, \theta_1) \neq \sigma_0^2(x, \theta_2)$ *if* $\theta_1 \neq \theta_2$. *For at least finitely many* x, *there exists a constant* C_2 *such that* $0 < C_2 \leq \sigma_0(x, \theta_0) \leq C_2^{-1}$. $P_{x_0}(\sigma_0^2(x_t, \theta) > 0) = 1$ *for* $(t, \theta) \in [t_0, T] \times \Theta$, *where* P_{x_0} *denotes the probability measure generated by the initial value* x_{t_0}.

Assumption 7. $a(x)$ *is a given Borel measurable function and bounded with compact support* $S \subset D$. $\pi(x)$ *and its derivative are continuous and bounded on* D, *and* $\pi(x)$ *is bounded away from zero on the compact support* S *of* $a(x)$. *There exists* $\alpha > 0$, *such that* $\int \exp(\alpha x^2)\pi(x)dx < \infty$.

Assumption 8. $K(.)$ *is a bounded and symmetric function about* 0, *with* $\int K(\mathrm{u})du = 1$, $\int \|u\| K(u)du < \infty$, *and* $\int uK(u)du = 0$.

Assumptions 1 is sufficient for pathwise uniqueness of the solution to the stochastic differential equation (2.1). Assumption 1 and Assumption 2 ensure the existence and uniqueness of a strong solution up to an explosion time. Assumptions 3 is used to guarantee that the exit time from D is infinite (Karatzas and Shreve (1991, Proposition 5.5.22, p.345). Assumption 4 is from Hansen and Scheinkman (1995, p. 801). Aït-Sahalia (1996a, p. 552) proved that under Assumption 4 the various classical mixing properties of the discrete observations from the stochastic differential equation (2.1) are satisfied. In particular, the observation process is absolutely regular with geometric decay rate. Without this assumption, the Central Limit Theorem for second order degenerate U-statistics of absolutely regular processes can fail. Assumptions 5 and 6 were used by Corradi and White (1997, Theorem 3.2., p. 261) to ensure that $\hat{\theta}_n$ is a $\sqrt{n}$ – consistent estimator of θ_0 under the null hypothesis, whereas under the alternative $\hat{\theta}_n$ is a $\sqrt{n}$ – consistent estimator of some θ^*, where $\theta^* \in \Theta$. Assumption 7 requires $a(x)$ to be bounded with compact support. As we mentioned earlier, without this assumption, we can not prove uniform convergence of the nonparametric estimations of the marginal density and diffusion function on S. In practice, $a(x)$ can typically be taken as the indicator function of a compact set related to the empirical question of interest. Assumption 8 is a standard regularity condition imposed on a kernel function.

The asymptotic null distribution and consistency of J_n is provided in the following theorem.

Theorem 1: *Suppose Assumptions 1-8 hold and* $h_n = O(n^{-1/\gamma})$*, where* $2 < \gamma < 2.5$*. If* T *is either fixed or* $T \to \infty$, $Th_n^{1/2} \to 0$*, and* $\Delta_n \to 0$ *as* $n \to \infty$*, then we have,*

(a) under H_0, $J_n \to N(0,1)$ *in distribution as* $n \to \infty$*, and* $\hat{v}_n^2$ *is a consistent estimator of* v^2*, where* $v^2 = 8\int \sigma^8(x)\pi^4(x)a^2(x)dx \int [\int K(u)K(w+u)du]^2 dw$.

(b) under H_1, $Pr(J_n \geq B_n) \to 1$*, for any non-stochastic sequence* $B_n = o(nh_n^{1/2})$.

Proof: *See Appendix.*

4. Monte Carlo Study

In this section, we examine the finite sample performance of our test through a Monte Carlo experiment. As we mentioned in introduction section, under the assumption that the drift function is specified correctly and the closed-form expression of the model-implied transition density is available, transition density-based tests can also be used to test the specification of the diffusion function in a diffusion process. Hong and Li (2002) have conducted a simulation study of the size and power of their tests. For comparison, we adopt similar simulation designs as in Hong and Li (2002).

To examine the size performance of our test, we simulate data from the Vasicek (1977), Cox, Ingersoll and Ross (1985, CIR), and drift-misspecified CIR (DMCIR) models, respectively.

The Vasicek model is,

$$dx_t = \beta(\alpha - x_t)dt + \sigma dw_t \,. \tag{4.1}$$

The CIR model is,

$$dx_t = \beta(\alpha - x_t)dt + \sigma\sqrt{x_t}dw_t. \tag{4.2}$$

The DMCIR is designed as,

$$dx_t = (\alpha_{-1}x_t^{-1} + \alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2)dt + \sigma\sqrt{x_t}dw_t. \tag{4.3}$$

For the Vasicek model, the null hypothesis is that the diffusion function is a constant, i.e. H_0: $\sigma^2(x) = \text{Constant}$. Under the null hypothesis, the estimator of $\theta = \sigma^2$ is given by $\hat{\sigma}_n^2 = \sum_{t=1}^{n-1} (x_{n,t+1} - x_{n,t})^2 / T$. Under the assumption that the drift function is correctly specified as $\beta(\alpha - x)$, Hong and Li's (2002) test can be used to test the null hypothesis by testing that the data is generated from a normal transition density (Hong and Li (2002), page 21). In the Vasicek model, the parameter β determines the persistence of the process. The smaller β is, the higher the level of persistence in the process, and consequently, the slower the convergence to the long-run mean α.

Like Hong and Li (2002) and Pritsker (1998), in order to look at the impact of the level of persistence on the size performance of our test, we consider both low and high levels of persistent dependence and adopt the same parameter values used in Hong and Li (2002) and Pritsker (1998). The parameter values for low and high levels of persistent dependence are respectively set as $(\beta, \alpha, \sigma^2) = (0.85837, 0.089102, 0.002185)$ and $(\beta, \alpha, \sigma^2) = (0.214592, 0.089102, 0.000546)$.

To examine the size performance of our test when the drift function is misspecified, we consider two cases. For Case 1, the data is assumed to be from the CIR model, but in fact it is generated from the DMCIR model. For Case 2, the data is assumed to be from the DMCIR model but in fact it is generated from the CIR model. For both cases, we test the null hypothesis that the diffusion function is $\sigma^2 x$, i.e. H_0: $\sigma^2(x) = \sigma^2 x$. Obviously, in both the CIR and DMCIR models, with the drift functions being misspecified, the diffusion functions are correctly specified. The parameter values of the CIR model are taken as $(\beta, \alpha, \sigma^2) = (0.89218, 0.090495, 0.032742)$, which are from Hong and Li (2002), while the parameter values of the DMCIR model are designed as

$(\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \sigma^2) = (0.00107, -0.0517, 0.877, -4.604, 0.032742)$. Under the null hypothesis, the estimator of the parameter $\theta = \sigma^2$ is given by $\hat{\sigma}_n^2 = \sum_{t=1}^{n-1} \frac{(x_{n,t+1} - x_{n,t})^2}{T x_{n,t}}$.

Since the Vasicek and CIR models have closed-form transition density and marginal density functions (Pritsker, 1998, p.456, and Hong and Li, 2002, p.22), the simulated sample path can be constructed by their transition densities. The initial values are drawn from their marginal densities. The discrete observations of sample size n are generated over a time period $[0, T]$ with sampling interval $\Delta_n = T/n$. For the DMCIR model (4.3), because its transition density has no closed form, we simulate data using the Milstein scheme (see (4.7)). Throughout the experiment, we generate 500 realizations of a random sample $\{x_{n,j}\}_{j=1}^{n}$ for sample sizes $n = 250, 500, 1000, 2500$ respectively. We discard the first 500 observations to eliminate any start-up effects. T is set to 1 and 5 to consider the impact of the sample interval on the test performance.

To study the power performance, we consider two cases. For Case1, the null hypothesis stipulates the data to be generated by a model with a constant diffusion function, i.e. H_0: $\sigma^2(x) = \text{Constant}$. However, in fact we simulate data from three different models, namely, the CIR model, the Chan, Karolyi, Longstaff and Sanders' (1992, CKLS) model, and the Aït-Sahalia's (1996b) nonlinear drift model. The same parameter values as in Hong and Li (2002) are again taken. If we impose the assumption that the drift function is correctly specified as $\beta(\alpha - x)$, Hong and Li's (2002) test can also used to test the null hypothesis by testing that the data is from a normal transition density.

CKLS model is,

$$dx_t = \beta(\alpha - x_t)dt + \sigma x_t^{\rho} dw_t \quad , \tag{4.4}$$

with parameter values $(\alpha, \beta, \sigma^2, \rho) = (0.0808, 0.0972, 0.52186, 1.46)$.

Aït-Sahalia's nonlinear drift model (1996b) is,

$$dx_t = (\alpha_{-1}x_t^{-1} + \alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2)dt + \sigma x_t^{\rho} dw_t, \tag{4.5}$$

with parameter values $(\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \sigma^2, \rho) = (0.00107, -0.0517, 0.877, -4.604, 0.64754$

$, 1.50)$.

For Case 2, the null hypothesis stipulates that the data are generated by a model with the diffusion function $\sigma^2 x$. However, in fact data are simulated from three different models, the CKLS model (4.4), the nonlinear drift model (4.5), and a modified CKLS (MCKLS) model. The MCKLS model is designed as,

$$dx_t = \beta(\alpha - x_t)dt + \sigma/(\sqrt{3})(x_t + 2/3)^{1.5} dw_t. \tag{4.6}$$

with parameter values $(\beta, \alpha, \sigma^2) = (0.89218, 0.090495, 0.032742)$ used in CIR model. It is should be noted that the process (4.6) has a nonlinear diffusion function and the same drift function as in the CIR model. Particularly, the linear diffusion function in the CIR model is tangential to the diffusion function in the MCKLS model at point $x = 1/3$. This design helps us to evaluate the power of our test for testing curvature.

For the CKLS model (4.4), the Aït-Sahalia's nonlinear drift model (1996b) (4.5) and the MCKLS model (4.6), since their transition densities have no closed forms, we simulate data using the Milstein scheme,

$$x_{t+\Delta_n} = x_t + \mu(x_t)\Delta_n + \sigma(x_t)\sqrt{\Delta_n}\varepsilon_t + \frac{1}{2}\sigma^2(x_t)\Delta_n(\varepsilon_t^2 - 1) \quad , \tag{4.7}$$

where ε_t is a standard normal distribution. The initial value is set to equal the average interest rate level of the data set in Aït-Sahalia (1999b).

Throughout this experiment, we use the standard normal kernel. The bandwidth parameter h_n is chosen according to $h_n = c\sigma_x n^{-1/2.1}$, where σ_x is the standard deviation of observations. The above choice of h_n satisfies the conditions of Theorem 1. To check the sensitivity of our test with respect to the choice of bandwidth h_n, we change h_n through different values of c: $c = 0.5, 1, 1.5$. The function $a(x)$ is the indicator function of the interval $S = \{x|x \in [0.002, 2]\}$. The critical value z_α is from the standard normal distribution, i.e., $z_{0.01} = 2.33$, $z_{0.05} = 1.645$ and $z_{0.1} = 1.28$.

The estimated size of our test are reported in Table 1. Four general conclusions can be drawn from Table 1. First, we observe that our test has satisfactory size performance at all three levels for sample sizes as small as $n = 250$. In contrast, it is clearly noted that under the same simulation setting, Hong and Li's (2002) tests show strong over rejections under 1% level (Hong and Li (2002), Figure 1 and Figure 2), and the size of their tests is about 2.1% in average even if sample size increases to 5500. Second, we notice that the impact of the level of the dependent persistence on the size of our test is minimal. This suggests that our test achieves robustness to the persistent dependence. This result can be explained by the fact that our test statistic is independent on the specification of the drift function, which determines the level of the persistent dependence. Third, this test still exhibits a satisfactory size performance even if the drift function is misspecified. In contrast, our Monte Carlo simulation shows that under the null hypothesis that the data are generated from the CIR model, Hong and Li's (2002) tests reject the DMCIR model about 59% even if the sample size is increased to 2500 across lag orders from 1 to 20. In other words, Hong and Li's (2002) tests strongly reject the correct null hypothesis $\sigma^2(x) = \sigma^2 x$. It is obvious that the rejection arises because of the misspecification of the transition density. Finally, it should be noted that the estimated size of our test is quite stable over different choice of the bandwidth, which is particularly true for large samples.

Table 2 gives the estimated power of our test when the null hypothesis is that the diffusion function is a constant, but in fact the data are generated from the CIR model, the CKLS model, and Aït-Sahalia's (1996b) nonlinear drift model, respectively. Table 3 gives the estimated power of our test when the null hypothesis is that the diffusion function is $\sigma^2 x$, but in fact the data are generated from the CKLS model, the Aït-Sahalia's (1996b) nonlinear drift model, and the MCKLS's model, respectively.

The simulation results of the power performance of our test can be summarised as following three conclusions. First, both Table 2 and Table 3 indicate that the power of the test performs quite well in detecting the misspecifications of the diffusion functions in both the Vasicek and CIR models against their respective alternatives. For a given alternative, the power of our test always increases rapidly with respect to the sample size, which is in line with the consistency property of our test. By comparison, we note that the power of Hong and Li's (2002) tests in detecting the vasicek model against the CIR model is about 50% when n increases to 2500, which is noticeably worse than against the CKLS model and Aït-Sahalia nonlinear drift model (Hong and Li (2002), Figure 3). However, under the same simulation setting the power of our test is above 90%. Second, it is noted that our test has a good power in detecting the CIR model against the MCKLS model when the sample size increases to 2500. However, we also observe that our test has a lower power in detecting the CIR model against the MCKLS model than the two other models, namely, the CKLS model, and the Aït-Sahalia's (1996b) nonlinear drift model. This result can be explained by the fact that the diffusion function in the MCKLS model is closer to the diffusion function in the CIR model than the two other models. Third, even though the power of our test is already quite stable over different choices of h_n for large samples (n=1000, n=2500), nevertheless, it is still noted that the higher the value of the bandwidth h_n (i.e., the higher the value of c), the higher the power of our test. This result can be explained by the fact that the test statistic diverges to $+\infty$ at the rate of $nh_n^{1/2}$ under alternative. Hence, a higher h_n (in certain range) will lead to a more powerful test against some fixed alternatives (in finite samples). This result does not mean that one should use a very large value of h_n in practice, because a very large value of h_n will over smooth the data, and hence obliterate any deviation of the data from the null data-generating process. Of course, one should not use a very small value of h_n, because a too small h_n may result in an inaccurate kernel estimation. Specifically, a too small h_n tends to make the test less powerful. Since our test is based on the high frequency data, the large sample sizes available should make the choice of h_n be not as crucial as the moderate sample size. How to choose the bandwidth optimally in the sense that the power of the tests are maximized and at the same time to keep the size under control is an important future research.

Simulation results for $T = 5$ is not presented, but are available from author. They are qualitatively similar to those for $T = 1$.

5. Conclusion

In this paper, we propose a consistent test for the parametric specification of the diffusion function in a diffusion process without any restrictions on the functional form of the drift function. The test is based on comparing kernel estimate of the true unknown diffusion function to the parametric specification of the diffusion function. The test is shown to have the standard normal distribution under the null. The Monte Carlo simulation results suggest that the overall performance of the test is satisfactory.

Extensions to multi-dimensional diffusion processes (including unobservable state variables) and applications to derivative pricing will be considered in future work. In addition, another research topic that deserves effort is to develop a test for the parametric specification of the drift function in a diffusion process.

Table 1: Percentage rejections of the true H_0

	c=0.5			c=1			c=1.5		
n	1%	5%	10%	1%	5%	10%	1%	5%	10%
Vasicek Model with Low Level of Persistent Dependence									
250	0.006	0.026	0.040	0.010	0.032	0.044	0.012	0.028	0.054
500	0.018	0.046	0.062	0.010	0.036	0.076	0.020	0.046	0.072
1000	0.020	0.058	0.086	0.014	0.062	0.090	0.012	0.054	0.082
2500	0.010	0.054	0.088	0.012	0.056	0.096	0.008	0.052	0.090
Vasicek Model with High Level of Persistent Dependence									
250	0.010	0.022	0.044	0.008	0.028	0.046	0.022	0.048	0.072
500	0.012	0.046	0.064	0.020	0.046	0.084	0.020	0.052	0.074
1000	0.008	0.044	0.068	0.020	0.038	0.090	0.020	0.048	0.076
2500	0.008	0.046	0.076	0.010	0.048	0.092	0.012	0.050	0.086
Case 1: The data is assumed to be from CIR model but in fact is generated from DMCIR model									
250	0.020	0.030	0.050	0.022	0.046	0.074	0.004	0.024	0.044
500	0.016	0.056	0.108	0.014	0.044	0.072	0.022	0.060	0.090
1000	0.014	0.052	0.066	0.012	0.054	0.104	0.020	0.058	0.094
2500	0.012	0.052	0.078	0.010	0.048	0.102	0.014	0.054	0.098
Case 2: The data is assumed to be from DMCIR model but in fact is generated from CIR model									
250	0.014	0.040	0.064	0.014	0.042	0.068	0.020	0.042	0.068
500	0.016	0.050	0.076	0.020	0.050	0.074	0.028	0.056	0.082
1000	0.016	0.054	0.082	0.020	0.048	0.072	0.020	0.052	0.084
2500	0.014	0.052	0.090	0.012	0.052	0.078	0.012	0.050	0.096

Table 2: Percentage rejections of the false H_0

		c=0.5			c=1			c=1.5	
n	1%	5%	10%	1%	5%	10%	1%	5%	10%
Cox, Ingersoll and Ross's Model (CIR)									
250	0.114	0.166	0.202	0.146	0.216	0.280	0.196	0.282	0.344
500	0.230	0.324	0.400	0.308	0.416	0.474	0.362	0.458	0.578
1000	0.534	0.632	0.696	0.596	0.692	0.748	0.668	0.738	0.792
2500	0.904	0.922	0.950	0.942	0.968	0.980	0.944	0.960	0.974
Chan, Karolyi, Longstaff and Sanders's Model (CKLS)									
250	0.262	0.360	0.426	0.270	0.370	0.452	0.328	0.424	0.496
500	0.752	0.860	0.908	0.822	0.892	0.926	0.900	0.946	0.960
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Aït-Sahalia's Nonlinear Drift Model									
250	0.276	0.362	0.438	0.342	0.432	0.482	0.424	0.518	0.582
500	0.660	0.714	0.752	0.718	0.798	0.832	0.736	0.802	0.836
1000	0.904	0.938	0.952	0.936	0.956	0.968	0.952	0.974	0.978
2500	0.994	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: Percentage rejections of the false H_0

		c=0.5			c=1			c=1.5	
n	1%	5%	10%	1%	5%	10%	1%	5%	10%
				Chan, Karolyi, Longstaff and Sanders's Model (CKLS)					
250	0.176	0.260	0.314	0.250	0.360	0.434	0.316	0.374	0.430
500	0.394	0.490	0.550	0.446	0.552	0.608	0.528	0.624	0.668
1000	0.604	0.680	0.718	0.676	0.728	0.782	0.696	0.776	0.818
2500	0.842	0.892	0.912	0.878	0.904	0.922	0.904	0.938	0.956
				Aït-Sahalia's Nonlinear Drift Model					
250	0.058	0.110	0.148	0.060	0.102	0.156	0.064	0.124	0.162
500	0.228	0.330	0.402	0.252	0.350	0.420	0.302	0.396	0.458
1000	0.700	0.824	0.880	0.784	0.858	0.890	0.824	0.875	0.908
2500	0.974	0.990	0.996	0.992	0.992	0.996	1.000	1.000	1.000
				Modified CKLS (MCKLS) model					
250	0.012	0.024	0.060	0.014	0.020	0.072	0.034	0.102	0.114
500	0.126	0.156	0.202	0.164	0.200	0.282	0.188	0.250	0.312
1000	0.398	0.438	0.462	0.414	0.452	0.500	0.422	0.492	0.514
2500	0.602	0.654	0.706	0.688	0.716	0.798	0.740	0.802	0.824

REFERENCES

Aït-Sahalia, Y. (1996a) Nonparametric pricing of interest rate derivatives securities, *Econometrica 64, 527-560.*

Aït-Sahalia, Y. (1996b) Testing continuous-time models of the spot interest rate, *Review of Financial Studies 2, 385-426.*

Aït-Sahalia, Y., J.B. Peter, and T.M.Stoker. (2001) Goodness-of-Fit tests for regression using kernel methods, *Journal of Econometrics 105, 363-412.*

Andrews, D.W.K. (1995)Nonparametric kernel estimation for semiparametric models, *Econometric Theory 11, 560-595.*

Bandi, F.M., and Phillips, P.C.B. (2003) Fully nonparametric estimation of scalar diffusion models, *Econometrica 71, 241-283.*

Bickel, P.J. and M. Rosenblatt (1973) On some global measures of the deviations of density function estimates, *Annals of Statistics 1, 1071-1095.*

Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders. (1992) An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance 47, 1209-1227.*

Corradi, V., and H. White (1997) Specification tests for the variance of a diffusion, *Journal of Time Series Analysis 20, 253-270.*

Corradi, V., N.R.Swanson (2004) Bootstrap conditional distribution tests under dynamic misspecification, *Journal of Econometrics, forthcoming.*

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

Diebold, F.X., T.Gunther and A.Tay (1998) Evaluating Density Forecasts with Applications to finance and management, *International Economic Review, 39, 863-883.*

Fan, Y., 1994, Testing the goodness-of-fit of a parametric density function by kernel method, *Econometric Theory 10, 316-356.*

Fan, Y. and Q. Li (1999) Central limit theorem for degenerate U-statistic of absolute regular processes with application to model specification testing, *Journal of Nonparametric Statistics 10, 245-271.*

Friedman, A. (1975) Stochastic Differential Equations and applications, *Volume 1, Academic Press.*

Gallant, A.R. and G. Tauchen. (1996) Which moments to match?, *Econometric Theory 12, 657-681.*

Gyorfi,L., Härdle, W., Sarda, P. and Vieu, P. (1989) Nonparametric Curve Estimation from Time Series, *Lecture Notes in Statistics, Vol.60, Springer-Verlag, Berlin.*

Hansen, L.P., and J.A. Scheinkman (1995) Back to the future: Generating moment implications for continuous time Markov processes, *Econometrica 63, 767-804.*

Hall, P., (1984) Central limit theorem for integrated square error of multivariate nonparametric density estimators, *Journal of Multivariate Analysis 14, 1-16.*

Hong, Y., and H. Li (2002) Nonparametric specification testing for continuous-time models with application to spot interest rates, *Review of Financial Studies, forthcoming.*

Jiang, G.J. and J.L. Knight (1997) A nonparametric approach to the estimation of diffusion process, with an application to a short-term interest rate model, *Econometric Theory 13, 615-645.*

Karatzas, I. and S.E. Shreve (1991) Brownian motion and stochastic calculus, *Springer.*

Kristensen, D. (2004) Estimation in two classes of semiparametric diffusion models, *Department of Economics, University of Wisconsin-Madison.*

Li, F. and G. Tkacz (2004) A consistent bootstrap test for conditional density functions with time-series data. *Journal of Econometrics, forthcoming.*

Nicolau, J. (2003) Bias reduction in nonparametric diffusion coefficient estimation, *Econometric Theory 19, 754-777.*

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

Thompson, T. (2001) Specification tests for continuous-time models, *Working Paper, Department of Economics, Harvard University.*

Vasicek, O. (1977) An equilibrium characterization of the term structure, *Journal of Financial Economics 5, 177-188.*

Wong, E. (1964) The construction of a class of stationary Markov processes, in: *Stochastic Processes in Mathematical physics and engineering, proceedings of symposia in applied mathematics, Vol. XVI.*

Yoshihara, K. (1976) Limiting behavior of U-statistic for stationary, absolutely regular processes, *Z. Wahrscheinlichkeitstheorie verw. Gebiete 35, 237-252.*