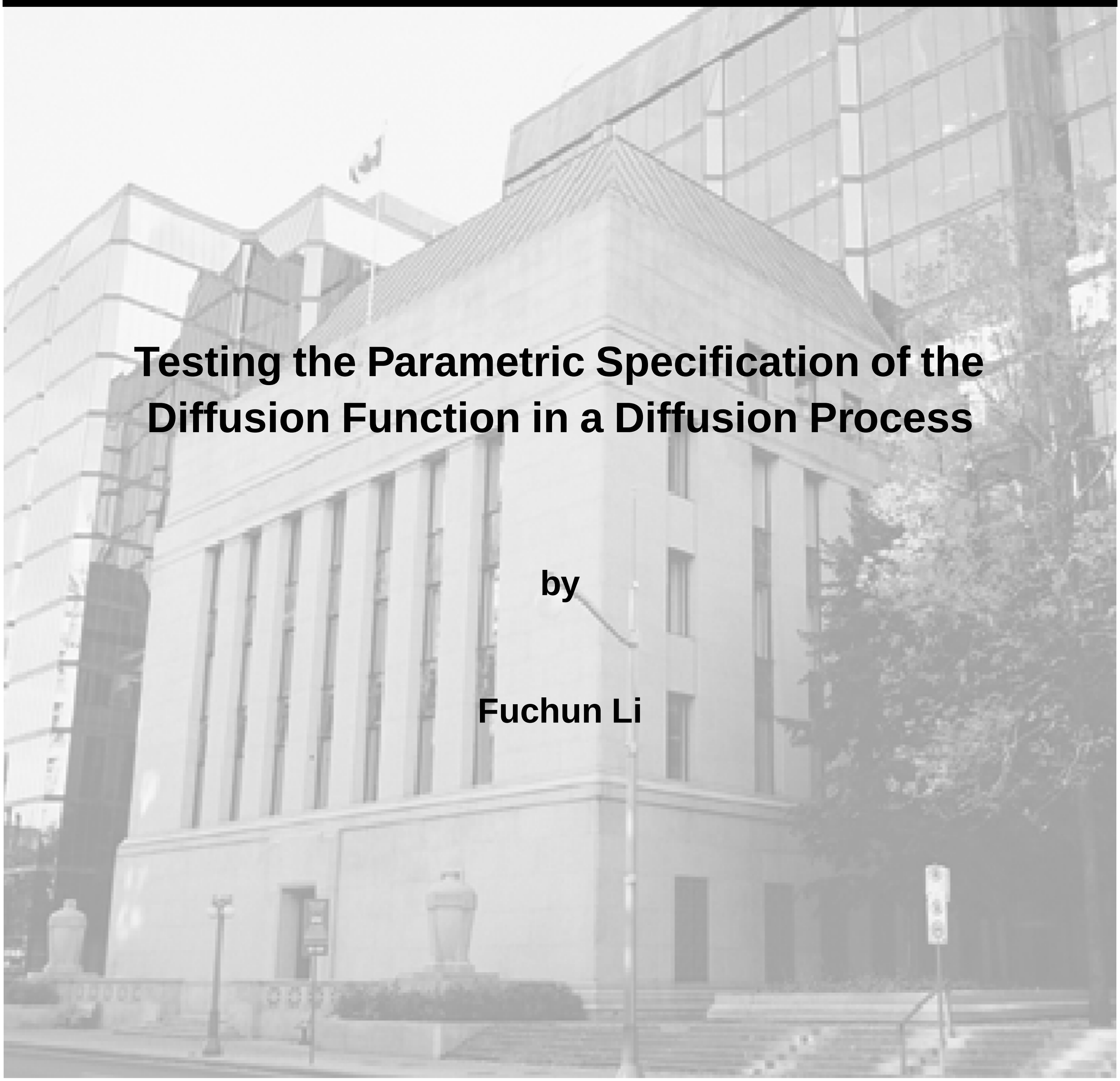
Bank of Canada Banque du Canada

Working Paper 2005-35 / Document de travail 2005-35



Testing the Parametric Specification of the Diffusion Function in a Diffusion Process

by

Fuchun Li

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2005-35

November 2005

Testing the Parametric Specification of the Diffusion Function in a Diffusion Process

by

Fuchun Li

Monetary and Financial Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
fuchunli@bankofcanada.ca

Contents

Acknowledgements iv
Abstract/Résumé v

1. Introduction 1

2. The Hypothesis and Test Statistic 4

3. Assumptions and the Limiting Distribution of the Test Statistic 8

4. Monte Carlo Simulations 11

5. Conclusion 17

References 19

Tables 22

Appendix 25

Acknowledgements

The author is grateful to Yacine Aït-Sahalia, John Knight, Oliver Linton, Greg Tkacz, Jun Yang, and three anonymous referees for helpful comments and suggestions. He also thanks seminar participants at the Bank of Canada, the 2004 Semiparametrics Conference in Rio de Janeiro, and the 2005 Econometric Study Group in Bristol. Any errors or omissions are those of the author.

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 that 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.

JEL classification: C12, C14
Bank classification: Econometric and statistical methods; Interest rates

Résumé

L'auteur propose un nouveau test convergent pour vérifier la validité de la spécification paramétrique de la fonction de diffusion d'un processus où la forme fonctionnelle de la dérive n'est soumise à aucune contrainte. Les données sont tirées par hypothèse d'un échantillon discret constitué sur un intervalle de temps qui peut être fixe ou infini. L'auteur montre que la statistique du test admet pour loi asymptotique la loi normale si l'hypothèse nulle que les paramètres de la fonction de diffusion sont spécifiés correctement est vraie. Il fait appel à des simulations de Monte-Carlo pour analyser la performance du test en échantillon fini. Le niveau et la puissance du test se révèlent satisfaisants.

Classification JEL : C12, C14
Classification de la Banque : Méthodes économétriques et statistiques; Taux d'intérêt

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 modelling 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) propose a minimum chi-square specification test for continuous-time models using the efficient method of moments. Aït-Sahalia (1996b) proposes two specification tests by comparing the model-implied parametric density with the same density estimated nonparametrically. Diebold, Gunther, and Tay (1998), Thompson (2001), and Hong and Li (2005) propose 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) propose 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) propose a Kolmogorov-type conditional distribution test. Because the limiting distribution of their test statistic is nuisance parameters free, Corradi and Swanson (2004) use a nonparametric bootstrap procedure to construct the critical values.

The null hypothesis of all the above-noted 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 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 advantageous to be able to develop a reliable test that can detect whether the model misspecification comes from the drift function or the diffusion function. Note that transition density-based tests can be used to test the specification of the diffusion function (drift function) only by 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 (Wong 1964).

These limitations of the above-mentioned tests and the recent developments in nonparametric estimation techniques of a continuous-time model prompt us to use nonparametric estimation techniques to develop 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 (1999) provide a first step in this direction. With knowledge of the functional form of the drift function not being required, they propose a specification test for the diffusion function based on discrete sampling observations.

As Corradi and White (1999) point out, however, their test can be used to test a parametric diffusion function at only a given point, and the time span of observations is fixed. Their test cannot be used to detect diffusion function misspecifications over a continuous range of the state variable.

Using discrete observations, I propose a new test for the functional specification of the diffusion function without placing any restriction on the functional form for the drift function. The test can be used to test the parametric specification of the diffusion function over a time interval that 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. The test can be applied to a wide variety of continuous-time models in economics and finance. For example, in the finance literature, when applied to eurodollar interest rates, Hong and Li's test (2005) rejects a wide range of popular interest rate models, including the linear specifications of the diffusion function in Vasicek (1977) and Cox, Ingersoll, and Ross (1985), and the nonlinear specifications in Chan et al. (1992), Aït-Sahalia (1996b), and Ahn and Gao (1999). However, Hong and Li's tests (2005) cannot indicate whether the rejection is due to the misspecification from the diffusion function, which is the critical component in the specification of a continuous-time model of the spot interest rate (Durham 2003). Taking advantage of our test, we can apply it to further explore whether there is statistically significant evidence in favour of any potential choice among these competing specifications of the diffusion function, or whether none of them is appropriate and an alternative specification is needed.

The rest of this paper is organized as follows. In section 2, I state the hypothesis of interest and introduce the test statistic. In section 3, I discuss the asymptotic properties of the test. In section 4, I use Monte Carlo simulations to examine the test's size and power performance. Section 5 concludes. Proofs are provided in the appendix.

2. The Hypothesis and Test Statistic

The model I 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 $\{w_t : t \geq 0\}$ 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 $\{w_t : t \geq 0\}$.

I assume that the process $\{w_t : t \geq 0\}$ is observed at $t = t_1, t_2, \ldots, t_n$ in the time interval $[t_0, T]$, and 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 $\{x_t : t \geq 0\}$ 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.

I use the notation $x_{n,j}$ to express the observation on the process $\{w_t : t \geq 0\}$ 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. I 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,

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

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

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

The 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) are 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 Stoker (2001), and Li and Tkacz (2004).

The density function $\pi(x)$ is introduced in (2.4) to avoid the problem of trimming the small values of the random denominator in the nonparametric estimation of the diffusion function. The weighting function $a(x)$ is included in (2.4) to allow me to focus goodness-of-fit testing on particular ranges of the state variable. Specifically, I will assume that $a(x)$ is bounded with compact support $S \subset R$ (Assumption 8 in section 3). This assumption will help prevent 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 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 that cross the interval [0.05, 0.10] are used in the specification analysis.

Under the null hypothesis, H_0, $I(\theta_0, \sigma_0^2) = 0$, and under the alternative, H_1, $I(\theta, \sigma_0^2) > 0$ 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 $I(\theta_0, \sigma_0^2)$ can be estimated 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, it is necessary 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 (2.5) shows, 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; Bandi and Phillips 2003). In fact, Nicolau (2003) shows 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) requires that the sampling interval Δ_n be fixed in order to obtain asymptotic results.

Since my aim is to construct a test for the parametric specification of the diffusion function without a 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 the test statistic.

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

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

Corradi and White (1997) provide regularity conditions under which $\hat{\theta}_n$ is a quasi-maximum-likelihood estimator. These regularity conditions are given in the assumptions listed in section 3.

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,

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

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

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

The test statistic is a properly centred and scaled version of I_n,

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

where

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

3. Assumptions and the Limiting Distribution of the Test Statistic

I 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 is satisfied; i.e., for any compact subset $A \subset D$ there exists a positive constant C_A such that, for every $x, y \in C_A$,*

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

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

Assumption 3. *The global growth condition is satisfied; i.e., there exists a positive constant,* C_D *, such that, for every* $x \in D$ *,*

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

Assumption 4. *There exists a positive constant,* C_1*, such that, for every* $x, y \in D$*:*

$$\left|\sigma^2(x) - \sigma^2(y)\right| \le C_1|x-y| \, . \tag{3.3}$$

Assumption 5. $\text{Lim}_{|x| \to \infty}|\sigma(x)/[2\mu(x) - \sigma(x)\sigma'(x)]| < \infty, \text{Lim}_{|x| \to \infty}(\sigma(x)\pi(x) = 0).$

Assumption 6. *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 7. *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* $0 < C_2 \le \sigma_0(x, \theta_0) \le C_2^{-1}$*.* $P_{x_0}(\sigma_0^2(x_t, \theta) > 0) = 1$ *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 8. $a(x)$ *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 9. $K(.)$ *is a bounded and symmetric function about* 0, *with* $\int K(u)du = 1$, $\int \|u\| K(u)du < \infty$, *and* $\int uK(u)du = 0$.

Assumptions 1 and 2 ensure the existence and uniqueness of a strong solution to the stochastic differential equation (2.1). Assumptions 3 and 4 are used to establish some important moment inequalities (for example, Theorem 2.2 in Friedman 1975, 127) that are needed to derive asymptotic results. Assumption 5 is taken from Hansen and Scheinkman (1995, 801). Aït-Sahalia (1996a, 552) proves that, under Assumption 5, 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 a geometric decay rate. Without this assumption, the central limit theorem for second-order degenerate U-statistics of absolutely regular processes can fail. Corradi and White (1999, Theorem 3.2) use Assumptions 6 and 7 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 8 requires $a(x)$ to be bounded with compact support. As stated earlier, without this assumption I cannot prove uniform convergence of the nonparametric estimations of the marginal density and diffusion function on S. In practice, $a(x)$ can be taken as the indicator function of a compact set related to the empirical question of interest. Assumption 9 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 that Assumptions 1-9 hold and that* $h_n = O(n^{-1/\gamma})$, *where* $1.5 < \gamma < 4.5$. *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,* *(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 nonstochastic sequence* $B_n = o(nh_n^{1/2})$.

Proof: *See the appendix.*

4. Monte Carlo Simulations

In this section, I examine the finite-sample performance of the test using Monte Carlo simulations. As stated in section 1, assuming that the drift function is specified correctly and that the closed-form expression of the model-implied transition density is available, transition density-based tests can also be used for the specification of the diffusion function in a diffusion process. Hong and Li (2005) conduct a simulation study to examine the size and power of their tests. For comparison, I adopt simulation designs that are similar to Hong and Li's (2005).

To examine the test's size performance, I simulate data from the Vasicek (1977), Cox, Ingersoll, and Ross (1985) (CIR hereafter), and drift-misspecified CIR (DMCIR hereafter) models, respectively.

Vasicek's model is:

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

CIR's model is:

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

The DMCIR model is:

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

For Vasicek's model, the null hypothesis is that the diffusion function is a constant; i.e., H_0: $\sigma^2(x) = $ 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 (2005) test can be used to test the null hypothesis by testing whether the data are generated from a normal transition density (Hong and Li 2005, 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.

As with Hong and Li (2005) and Pritsker (1998), to examine the impact of the level of persistence on the size performance of our test, I consider both low and high levels of persistent dependence and adopt their parameter values. The parameter values for low and high levels of persistent dependence are, respectively, $(\beta, \alpha, \sigma^2) = (0.85837,$

$0.089102, 0.002185)$ and $(\beta, \alpha, \sigma^2) = (0.214592, 0.089102, 0.000546)$.

To examine the test's size performance when the drift function is misspecified, I consider two cases. For Case 1, the data are assumed to be from CIR's model, but they are generated from the DMCIR model. For Case 2, the data are assumed to be from the DMCIR model, but they are generated from CIR's model. For both cases, I 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 CIR's model and the DMCIR model, with the drift functions being misspecified, the diffusion functions are correctly specified. The parameter values of CIR's model are taken as $(\beta, \alpha, \sigma^2) = (0.89218, 0.090495, 0.032742)$, which are from Hong and Li (2005), whereas the parameter values of the DMCIR model are taken 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 $\hat{\sigma}_n^2 = \sum_{t=1}^{n-1} \frac{(x_{n,t+1} - x_{n,t})^2}{Tx_{n,t}}$.

Since Vasicek's and CIR's models have closed-form transition density and marginal density functions (Pritsker 1998, 456; Hong and Li 2005, 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 a sampling interval of $\Delta_n = T/n$. For the DMCIR model (4.3), because its transition density has no closed form, data are simulated using Milstein's scheme (see (4.7)). Throughout the experiment, I generate 500 realizations of a random sample $\{x_{n,j}\}_{j=1}^{n}$ for sample sizes $n = 250, 500, 1000, 2500$, respectively. I 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 test's power performance, I consider two cases. For Case 1, the null hypothesis stipulates that the data are generated by a model with a constant diffusion function; i.e., H_0: $\sigma^2(x) =$ constant. However, I simulate data from three different models: CIR's model, Chan et al.'s (1992) (CKLS hereafter) model, and Aït-Sahalia's (1996b) nonlinear drift model. The same parameter values as in Hong and Li (2005) are again adopted. If I impose the assumption that the drift function is correctly specified as $\beta(\alpha - x)$, Hong and Li's (2005) test can be used to test the null hypothesis by testing whether the data are from a normal transition density.

The 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 value $(\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \sigma^2, \rho) = (0.00107, -0.0517, 0.877, -4.604, 0.64754, 1.50)$.

For Case 2, the null hypothesis stipulates that the data are generated by a model with the diffusion function $\sigma^2 x$. However, the 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:

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

with the parameter values $(\beta, \alpha, \sigma^2) = (0.89218, 0.090495, 0.032742)$ used in CIR's model. Note that the process (4.6) has a nonlinear diffusion function and the same drift function as in CIR's model. Particularly, the linear diffusion function in CIR's model is tangential to the diffusion function in the MCKLS model at point $x = 0.090495$. This design helps in evaluating the test's power for testing curvature.

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

Throughout this experiment, I use the standard normal kernel. The bandwidth parameter h_n is chosen according to $h_n = c\sigma_x n^{-1/\gamma}$, where σ_x is the standard deviation of observations. I choose $\gamma = 2.1, 3.5$. The choice of h_n satisfies the conditions of Theorem 1. To check the sensitivity of the test with respect to the choice of bandwidth h_n, I 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 $z_{0.1} = 1.28$.

Table 1 reports the estimated size of the test. Four general conclusions can be drawn from the table. First, the test has satisfactory size performance at all three levels for sample sizes as small as $n = 250$. In contrast, it is clear that, under the same simulation setting, Hong and Li's tests show strong overrejections under the 1 per cent level (Hong and Li 2005, Figures 1 and 2), and the size of their tests is about 2.1 per cent on average, even if the sample size increases to 5500. Second, the impact of the level of the dependent persistence on the size of the test is minimal, which suggests that the test achieves robustness to the persistent dependence. This result can be explained by the fact that the test statistic is independent of the specification of the drift function, which determines the level of the persistent dependence. Third, the test still exhibits a satisfactory size performance even if the drift function is misspecified. In contrast, the Monte Carlo simulation shows that, under the null hypothesis that the data are generated from CIR's model, the power of Hong and Li's (2005) tests for rejecting the DMCIR model is about 59 per cent, even if the sample size is increased to 2500 across lag orders from 1 to 20. In other words, Hong and Li's (2005) tests strongly reject the correct null hypothesis of $\sigma^2(x) = \sigma^2 x$. It is obvious that the rejection arises because of the misspecification of the transition density. Fourth, note that the estimated size of the test is quite stable over different choices of bandwidth, particularly for large samples.

Table 2 reports the estimated power of the 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 reports the estimated power of the 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, Aït-Sahalia's (1996b) nonlinear drift model, and the MCKLS model, respectively.

The simulation results of the test's power performance lead to three conclusions. First, Tables 2 and 3 indicate that the test detects the misspecifications of the diffusion functions quite well in both Vasicek's and CIR's models against their respective alternatives. For a given alternative, the test's power always increases rapidly with respect to the sample size, in line with the test's consistency property. By comparison, the power of Hong and Li's (2005) tests in detecting Vasicek's model against CIR's model is about 50 per cent when n increases to 2500, which is noticeably worse than against the CKLS model and Aït-Sahalia's nonlinear drift model (Hong and Li 2005, Figure 3). However, under the same simulation setting, the power of the test is above 90 per cent. Second, the test has good power in detecting CIR's model against the MCKLS model when the sample size increases to 2500. This test, however, has a lower power in detecting CIR's model against the MCKLS model than the CKLS model and 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 CIR's model than in the two other models. Third—even though the test's power is already quite stable over different choices of h_n for large samples (n = 1000, n = 2500)—the higher the value of the bandwidth, h_n (i.e., the higher the value of c), the higher the test's power. This result can be explained by the fact that the test statistic diverges to $+\infty$ at the rate of $nh_n^{1/2}$ under the alternative. Hence, a higher h_n (in a 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 it would oversmooth 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 it could result in an inaccurate kernel estimation. Specifically, an h_n that is too small tends to make the test less powerful. Since the test is based on high-frequency data, the large sample sizes available should make the choice of h_n less crucial than the moderate sample size. How to choose the bandwidth optimally, so that the test's power is maximized and the size is kept under control, is left for future research.

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

5. Conclusion

In this paper, I 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 a comparison of the kernel estimate of the true unknown diffusion function with the parametric specification of the diffusion function. It is shown to have the standard normal distribution under the null hypothesis. 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 evaluate the performance of a variety of specifications for the diffusion function in the spot interest rate process (Durham 2003) will be considered for future work. It would also be useful to develop a test for the parametric specification of the drift function in a diffusion process.

References

Ahn, D., and B. Gao. 1999. "A Parametric Nonlinear Model of Term Structure Dynamics." *Review of Financial Studies* 12: 721—62.

Aït-Sahalia, Y. 1996a. "Nonparametric Pricing of Interest Rate Derivatives Securities." *Econometrica* 64: 527—60.

———1996b. "Testing Continuous-Time Models of the Spot Interest Rate." *Review of Financial Studies* 2: 385—426.

Aït-Sahalia, Y., P.J. Bickel, 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—95.

Bandi, F.M. and P.C.B. Phillips. 2003. "Fully Nonparametric Estimation of Scalar Diffusion Models." *Econometrica* 71: 241—83.

Bickel, P.J. and M. Rosenblatt. 1973. "On Some Global Measures of the Deviations of Density Function Estimates." *Annals of Statistics* 1: 1071—95.

Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders (CKLS). 1992. "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate." *Journal of Finance* 47: 1209—27.

Corradi, V. and H. White. 1997. "Specification Tests for the Variance of a Diffusion." *Journal of Time Series Analysis* 20: 253—70.

Corradi, V. and N.R. Swanson. 2004. "Bootstrap Conditional Distribution Tests under Dynamic Misspecification." *Forthcoming in Journal of Econometrics.*

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—83.

Durham, G.B. 2003. "Likelihood-Based Specification Analysis of Continuous-Time Models of the Short-Term Interest Rate." *Journal of Financial Economics* 70: 436—87.

Fan, Y. 1994. "Testing the Goodness-of Fit of a Parametric Density Function by Kernel Method." *Econometric Theory* 10: 316—56.

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—71.

Friedman, A. 1975. "Stochastic Differential Equations and Applications." Volume 1. New York: Academic Press.

Gallant, A.R. and G. Tauchen. 1996. "Which Moments to Match?" *Econometric Theory* 12: 657-81.

Györfi, L., W. Härdle, P. Sarda, and P. Vieu 1989. "Nonparametric Curve Estimation from Time Series." In *Lecture Notes in Statistics,* edited by P. Bickel et al. Berlin: Springer-Verlag.

Hall, P. 1984. "Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators." *Journal of Multivariate Analysis* 14: 1—16.

Hansen, L.P. and J.A. Scheinkman. 1995. "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes." *Econometrica* 63: 767—804.

Hong, Y. and H. Li. 2005. "Nonparametric Specification Testing for Continuous-Time Models with Application to Spot Interest Rates." Forthcoming in *Review of Financial Studies.*

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—45.

Kristensen, D. 2004. "Estimation in Two Classes of Semiparametric Diffusion Models." Department of Economics, University of Wisconsin-Madison. Working Paper.

Li, F. and G. Tkacz. 2004. "A Consistent Bootstrap Test for Conditional Density Functions with Time-Series Data." *Forthcoming in Journal of Econometrics.*

Nicolau, J. 2003. "Bias Reduction in Nonparametric Diffusion Coefficient Estimation." *Econometric Theory* 19: 754—77.

Pritsker, M. 1998. "Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models." *Review of Financial Studies* 11: 449-87.

Thompson, T. 2001. "Specification Tests for Continuous-Time Models." Department of Economics, Harvard University. Working Paper.

Vasicek, O. 1977. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* 5: 177—88.

Wong, E. 1964. "The Construction of a Class of Stationary Markov Processes." In *Stochastic Processes in Mathematical Physics and Engineering,* edited by R. Bellman. 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: 237-52.

Table 1: Percentage of 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's 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's 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: Data are assumed to be from CIR model but in fact are 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: Data are assumed to be from DMCIR model but in fact are 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 of 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 et al.'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 of 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 et al.'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	0.010	0.010	0.012	0.014	0.016	0.018	0.026	0.044	0.104
500	0.052	0.064	0.114	0.062	0.090	0.126	0.092	0.098	0.142
1000	0.154	0.254	0.330	0.256	0.356	0.420	0.328	0.386	0.472
2500	0.468	0.630	0.662	0.618	0.706	0.740	0.670	0.734	0.802
5500	0.818	0.854	0.886	0.834	0.862	0.890	0.850	0.884	0.904

Appendix: Proofs

Let $\{X_{n,l}\} \equiv \{X_{n,l}: l \le n;\ n \ge 1\}$ be a triangular array of random variables, and $M_{s,t}^{n}$ denote the sigma algebra generated by $(X_{n,s}, \dots, X_{n,t})$ for $s \le t$.

Definition. (i). *$\{X_{n,l}\}$ is said to be a strictly stationary triangular array of random variables if for positive integers s_1, s_2, and k, $\{X_{n,s_1}, X_{n,s_1+1}, \dots, X_{n,s_1+k}\}$ and $\{X_{n,s_2}, X_{n,s_2+1}, \dots, X_{n,s_2+k}\}$ have the same joint distribution, where $s_1 + k \le n$, $s_2 + k \le n$.*
(ii). *Let $\{X_{n,l}\}$ be a strictly stationary triangular array of random variables and $\beta_{n\tau} = sup_{s+\tau \le n} E[sup_{A \in M_{s+\tau,n}^{n}}\{|P(A|M_{1,s}^{n}(n)) - P(A)|\}]$. Then, $\{X_{n,l}\}$ is said to satisfy an absolute regularity condition with the mixing coefficient $\beta_{n\tau}$ if $\beta_{n\tau} \to 0$ as $\tau \to \infty$.*

Lemma 1. *Let $\xi_{n,1}, \dots, \xi_{n,n}$ be random vectors taking values in R^d satisfying an absolute regularity condition with the mixing coefficient $\beta_{n\tau}$. Let $a(y_i, \dots, y_k)$ be a Borel measurable function such that, for some $\delta > 0$,*

$$\overline{M}_n = max\left\{E[|a(\xi_{n,i_1}, \dots, \xi_{n,i_k})|^{1+\delta}], E\left[E_{\xi_{n,i_1}, \dots, \xi_{n,i_j}}[|a(\xi_{n,i_1}, \dots, \xi_{n,i_k})|^{1+\delta}]\right]\right\}$$

exists. Then,

$$\left|E[a(\xi_{n,i_1}, \dots, \xi_{n,i_k})] - E[E_{\xi_{n,i_1}, \dots, \xi_{n,i_j}}[a(\xi_{n,i_1}, \dots, \xi_{n,i_k})]]\right| \le 4\overline{M}_n^{1/(1+\delta)} \beta_{nm}^{\delta/(1+\delta)}$$

where,

$$\begin{aligned} & E_{\xi_{n,i_1}, \dots, \xi_{n,i_j}}[|a(\xi_{n,i_1}, \dots, \xi_{n,i_k})|^{1+\delta}] \\ & = \int |a(y_{i_1}, \dots, y_{i_j}, \xi_{n,i_{j+1}}, \dots, \xi_{n,i_k})|^{1+\delta} dF_n^{i_1, \dots, i_j}(y_{i_1}, \dots, y_{i_j}) \end{aligned},$$

$m = i_{j+1} - i_j$, $F_n^{i_1, \dots, i_j}(y_{i_1}, \dots, y_{i_j})$ is the distribution function of random vectors $\xi_{n,i_1}, \dots, \xi_{n,i_j}$, and $i_1 < i_2 < \dots < i_k$.

Proof: Lemma 1 follows straightforwardly from Lemma 1 in Yoshihara (1976).

For simplicity, h_n and $K\left(\frac{x_{n,i}-x}{h_n}\right)$ are expressed by h and $K_i(x)$, respectively. E^j is used to express the conditional expectation with respect to the σ-*field* generated by $\{x_u: u \le t_0 + j\Delta_n\}$. The symbol C denotes a generic big enough positive constant. The notation $A \sim B$ means that A has an order no larger than that of B. Also, let $B(.,.,.)$ be a Borel measurable function, and $F_n(x, \bar{x})$ be the joint distribution function for $(x_{n,i}, x_{n,j})$, where $i \ne j$. I denote: $E_i[B(x_{n,i}, x_{n,j}, x_{n,k})] \equiv \int B(x, x_{n,j}, x_{n,k})dF(x)$,

$E_k[B(x_{n,i}, x_{n,j}, x_{n,k})] = \int B(x_{n,i}, x_{n,j}, x)dF(x)$, and

$E_{i,j}[B(x_{n,i}, x_{n,j}, x_{n,k})] \equiv \int\int B(x, \bar{x}, x_{n,k})dF_n(x, \bar{x})$.

Lemma 2. *Suppose that Assumptions 1-4 and 6 hold and* $E(x_{t'}^{2l}) < +\infty$ *for some positive integer* l, *and* $t' \in [t_0, T)$. *Then, for* $t' > t_0 + j\Delta_n$ *:*

$$(a) \quad E^j(x_{t'} - x_{n,j})^{2l} \le D_n(1 + x_{n,j}^{2l})(t' - t_0 - j\Delta_n)^l, \tag{A.1}$$

where $D_n = 2^{2(2l-1)}C_D^{2l}\{(t' - t_0 - j\Delta_n)^l + [l(2l-1)]^l\}$.

(b) *Let* $r_n(Z_{n,j}, x, \hat{\theta}_n) = \frac{1}{nh\Delta_n}K_j(x)[(x_{n,j+1} - x_{n,j})^2 - \sigma_0^2(x, \hat{\theta}_n)\Delta_n]$,

where $Z_{n,j} = (x_{n,j}, x_{n,j+1})$. *Then, under the null hypothesis, for* $x \in S$*:*

$$E[r_n(Z_{n,j}, x, \hat{\theta}_n)] = O(n^{-1}\Delta_n^{1/2}) + O(n^{-1}h^2) + O(n^{-3/2}h^{1/\xi - 1}), \tag{A.2}$$

$$E\int[r_n(Z_{n,j}, x, \hat{\theta}_n)]^2 a(x)dF(x) = \frac{2}{n^2h}\int\sigma^4(x)\pi(x)a(x)dF(x)\int k^2(x)dx + O(\Delta_n^{1/2}(n^2h)^{-1} + n^{-2}), \tag{A.3}$$

where $\xi > 1$.

Proof of (a) of Lemma 2: (A.1) directly follows Theorem 2.2 of Friedman (1975), simply by replacing the unconditional expectation with the conditional expectation.

Proof of (b) of Lemma 2: For simplicity, $\int_{t_0 + j\Delta_n}^{t_0 + (j+1)\Delta_n} g(u)du$ is denoted by $\int_{\Delta_n} g(u)du$, where $g(u)$ is any integrable function. An application of Itô's formula yields:

$$E(r_n(Z_{n,j}, x, \hat{\theta}_n)) = \frac{1}{n\Delta_n h}E\left\{K_j(x)E^j\left[\int_{\Delta_n}(x_u - x_{n,j})\mu(x_u)du + \int_{\Delta_n}(\sigma^2(x_u) - \sigma^2(x_{n,j}))du\right]\right\}$$

$$+ \frac{1}{n\Delta_n h}E\left\{K_j(x)\int_{\Delta_n}(\sigma^2(x_{n,j}) - \sigma^2(x))du\right\}$$

$$- \frac{1}{nh}E[K_j(x)(\sigma_0^2(x, \hat{\theta}_n) - \sigma^2(x))]$$

$$\equiv A_{n1} + A_{n2} - A_{n3}.$$

I will prove (b) of Lemma 2 by showing $A_{n1} = O(\Delta_n^{1/2}n^{-1})$. $A_{n2} = O(n^{-1}h^2)$, and $A_{n3} = O(n^{-3/2}h^{1/\xi - 1})$. From Assumptions 3 and 4, (A.1), and Schwarz's inequality, it is straightforward to have that $A_{n1} = O(\Delta_n^{1/2}n^{-1})$. For A_{n2}, by assumption 4, I have,

$$A_{n2} = \frac{1}{nh}\int K\left(\frac{u - x}{h}\right)[\sigma^2(u) - \sigma^2(x)]\pi(u)du = n^{-1}\int K(v)[\sigma^2(x + hv) - \sigma^2(x)]\pi(x + hv)dv$$

$= O(n^{-1}h\int K(v)v\pi(x + hv)dv) = O(n^{-1}h^2)$, where I use the Taylor extension $\pi(x + hv) = \pi(x) + \pi'(x)hv + \frac{1}{2}\pi''(\vartheta)h^2v^2$, and $\vartheta \in (x, x + hv)$. For A_{n3}, under H_0, from Assumption 6 and Holder 's inequality, there exists $\xi > 1$ such that,

$$|A_{n3}| \le C(nh)^{-1}[E(K_j(x))^{\xi}]^{1/\xi}[E|\hat{\theta}_n - \theta_0|^{\eta}]^{1/\eta} = O(n^{-3/2}h^{1/\xi - 1}),$$

where $\eta = (1 - \xi^{-1})^{-1}$. Hence (A.2) holds. I show (A.3).

$$E\int(r_n(Z_{n,j}, x, \hat{\theta}_n))^2 a(x)dF(x)$$

$$= (nh\Delta_n)^{-2}E\left\{\int K_j^2(x)[(x_{n,j+1} - x_{n,j})^2 - \sigma^2(x, \theta_0)\Delta_n]^2 a(x)dF(x)\right\}$$

$$+ (nh)^{-2} E\left\{\int K_j^2(x) v^2(x, \theta_0, \hat{\theta}_n)\ a(x) dF(x)\right\}$$

$$+ (nh)^{-2}\Delta_n^{-1} E\left\{\int K_j^2(x)[(x_{n,\,j+1} - x_{n,\,j})^2 - \sigma^2(x, \theta_0)\Delta_n]\ v(x, \theta_0, \hat{\theta}_n) a(x) dF(x)\right\}$$

$$\equiv A_{n3}^1 + A_{n3}^2 + A_{n3}^3,$$

where $v(x, \theta_0, \hat{\theta}_n) = \sigma_0^2(x, \hat{\theta}_n) - \sigma^2(x, \theta_0)$.

Using (a) of Lemma 2, the mean value theorem (to $\sigma_0^2(x, \hat{\theta}_n) - \sigma^2(x, \theta_0)$), $\hat{\theta}_n - \theta_0 = O_p(n^{-1/2})$, and changing the variable by $\frac{x_{n,\,j} - x}{h} = v$, it is straightforward to have $A_{n3}^2 = O((n^3 h)^{-1})$ and $A_{n3}^3 = O(\Delta_n^{1/2}(n^{5/2}h)^{-1}) + O(n^{-5/2})$. I next consider A_{n3}^1.

Applying Itô's formula to $(x_{n,\,j+1} - x_{n,\,j})^4$ and $(x_{n,\,j+1} - x_{n,\,j})^2$:

$$A_{n3}^1 = (nh\Delta_n)^{-2}\left\{\int E\left[K_j^2(x)\int_{\Delta_n} 4E^j(x_u - x_{n,\,j})^3 \mu(x_u) du\right] a(x) dF(x)\right.$$

$$-4\Delta_n \int E\left[K_j^2(x)\int_{\Delta_n} E^j(x_u - x_{n,\,j})\mu(x_u) du\right]\sigma^2(x) a(x) dF(x)$$

$$+ 6\int E\left[K_j^2(x)\int_{\Delta_n} E^j(x_u - x_{n,\,j})^2(\sigma^2(x_u) - \sigma^2(x)) du\right] a(x) dF(x)$$

$$+ 12\int E\left[K_j^2(x)\int_{\Delta_n}\int_{t_0+\Delta_n}^{u} E^j(x_s - x_{n,\,j})\mu(x_s) ds du\right]\sigma^2(x) a(x) dF(x)$$

$$+ 6\int E\left[K_j^2(x)\int_{\Delta_n}\int_{t_0+\Delta_n}^{u} E^j(\sigma^2(x_s) - \sigma^2(x)) ds du\right]\sigma^2(x) a(x) dF(x)$$

$$-2\int E\left[K_j^2(x)\Delta_n\int_{t_0+\Delta_n}^{u} E^j(\sigma^2(x_u) - \sigma^2(x)) du\right]\sigma^2(x) a(x) dF(x)$$

$$\left.+ 2\Delta_n^2 \int E[K_j^2(x)\sigma^4(x)] a(x) dF(x)]\right\}$$

$$\equiv A_{n3}^{11} + A_{n3}^{12} + A_{n3}^{13} + A_{n3}^{14} + A_{n3}^{15} + A_{n3}^{16} + A_{n3}^{17}\ .$$

By using (a) of Lemma 2, it can be shown that:

$A_{n3}^{11} = O(\Delta_n^{1/2}(n^2h)^{-1})$, $A_{n3}^{12} = O(\Delta_n^{1/2}(n^2h)^{-1})$, $A_{n3}^{13} = O(\Delta_n^{1/2}(n^2h)^{-1} + n^{-2})$,

$A_{n3}^{14} = O(\Delta_n^{1/2}(n^2h)^{-1})$, $A_{n3}^{15} = O(\Delta_n^{1/2}(n^2h)^{-1} + n^{-2})$, $A_{n3}^{16} = O(\Delta_n^{1/2}(n^2h)^{-1} + n^{-2})$,

$$A_{n3}^{17} = 2(n^2h)^{-1}\int[\pi(x+hv)\sigma^4(x)a(x)dF(x)]\int K^2(v)dv = 2(n^2h)^{-1}\int[\pi(x)\sigma^4(x)a(x)dF(x)]$$

$$\times\int K^2(v)dv + O(n^{-2}) .$$

To summarize the above, I have shown that $max\left\{A_{n3}^2, A_{n3}^3, A_{n3}^{11}, A_{n3}^{12}, A_{n3}^{14}\right\} =$ $O(\Delta_n^{1/2}(n^2h^{-1}))$, $max\{A_{n3}^{13}, A_{n3}^{15}, A_{n3}^{16}\} = O(\Delta_n^{1/2}(n^2h)^{-1} + n^{-2})$, and

$A_{n3}^{17} = 2(n^2h)^{-1}\int\pi(x)\sigma^4(x)a(x)dF(x)\int K^2(v)dv + O(n^{-2})$. These results imply that (A.3) holds.

Lemma 3. *Under Assumptions 1-9 and the null hypothesis,* I_n *can be written as:*

$$\mathrm{I}_n = \bar{I}_n + o_p((nh^{1/2})^{-1}), \tag{A.4}$$

where $$\bar{I}_n = \int[(\hat{\sigma}_n^2(x) - \sigma_0^2(x, \hat{\theta}_n))\hat{\pi}(x)]^2 a(x)dF(x) .$$

Lemma 3 indicates that the only difference between $\bar{I}_n$ and I_n is that the latter is average over the empirical conditional distribution function, instead of F. Lemma 3 shows that this difference is inconsequential for the asymptotic distribution of the test statistic.

Proof of Lemma 3: I need to prove that***:***

$$\varepsilon_n \equiv \int[(\hat{\sigma}_n^2(x) - \sigma_0^2(x, \hat{\theta}_n))\hat{\pi}(x)]^2 a(x)d(\hat{F}(x) - F(x))$$

$$= \sum_{i,j=1}^{n} \int(nh\Delta_n)^{-2}K_i(x)K_j(x)[(x_{n,i+1} - x_{n,i})^2 - \sigma_0^2(x, \hat{\theta}_n)\Delta_n]$$

$$\times[(x_{n,j+1}-x_{n,j})^2-\sigma_0^2(x,\hat{\theta}_n)\Delta_n]a(x)d(\hat{F}(x)-F(x))$$

$$= o_p((nh^{1/2})^{-1}). \qquad \text{(A.5)}$$

First, under Assumption 5, the observed data sequence $\{x_{n,t}\}$ is absolutely regular with a geometric decay rate; i.e., $\{x_{n,t}\}$ satisfies the absolute regular condition with mixing coefficient $\beta_{nm} = O(\lambda^m)$ (Aït-Sahalia 1996a, 552), where λ is a positive constant determined by the integral operator of the diffusion process (2.1). Let $m = [b\log n]$ and $\kappa = -\log\lambda > 0$, where b is a sufficiently large positive constant. Then, $\beta_{nm} = O(\lambda^m) = O(\lambda^{-b\kappa\log_\lambda n}) = O(n^{-b\kappa})$.

Recalling that $\sigma_0^2(x,\hat{\theta}_n)-\sigma_0^2(x,\theta_0) = v(x,\theta_0,\hat{\theta}_n)$, ε_n can be written as follows:

$$\varepsilon_n = n^{-3}(h\Delta_n)^{-2}\sum_{i,j,k=1}^{n}\{K_i(x_{n,k})K_j(x_{n,k})[(x_{n,i+1}-x_{n,i})^2-\sigma_0^2(x_{n,k},\theta_0)\Delta_n]$$

$$\times[(x_{n,j+1}-x_{n,j})^2-\sigma_0^2(x_{n,k},\theta_0)\Delta_n]a(x_{n,k})$$

$$-\int K_i(x)K_j(x)[(x_{n,i+1}-x_{n,i})^2-\sigma_0^2(x,\theta_0)\Delta_n][(x_{n,j+1}-x_{n,j})^2-\sigma_0^2(x,\theta_0)\Delta_n]a(x)dF(x)\}$$

$$-2n^{-3}h^{-2}\Delta_n^{-1}\sum_{i,j,k=1}^{n}\{K_i(x_{n,k})K_j(x_{n,k})[(x_{n,i+1}-x_{n,i})^2-\sigma_0^2(x_{n,k},\theta_0)\Delta_n]$$

$$\times v(x_{n,k},\theta_0,\hat{\theta}_n)a(x_{n,k})$$

$$-\int K_i(x)K_j(x)[(x_{n,i+1}-x_{n,i})^2-\sigma_0^2(x,\theta_0)\Delta_n]v(x,\theta_0,\hat{\theta}_n)a(x)dF(x)\}$$

$$+n^{-1}(nh)^{-2}\sum_{i,j,k=1}^{n}\{K_i(x_{n,k})K_j(x_{n,k})v^2(x_{n,k},\theta_0,\hat{\theta}_n)-\int K_i(x)K_j(x)v^2(x,\theta_0,\hat{\theta}_n)a(x)dF(x)\}$$

$$\equiv \varepsilon_{n1}-2\varepsilon_{n2}+\varepsilon_{n3}\ . \qquad \text{(A.6)}$$

I shall show that $\varepsilon_{ni} = o_p((nh^{1/2})^{-1})$ for $i = 1, 2, 3$. I will first show that $\varepsilon_{n1} = o_p((nh^{1/2})^{-1})$. Let

$$W_{i,j,k}(x_{n,k}) = \{K_i(x_{n,k})K_j(x_{n,k})[(x_{n,i+1}-x_{n,i})^2 - \sigma_0^2(x_{n,k},\theta_0)\Delta_n]$$

$$\times[(x_{n,j+1}-x_{n,j})^2 - \sigma_0^2(x_{n,k},\theta_0)\Delta_n]a(x_{n,k})$$

$$-\int K_i(x)K_j(x)[(x_{n,i+1}-x_{n,i})^2 - \sigma_0^2(x,\theta_0)\Delta_n]$$

$$\times[(x_{n,j+1}-x_{n,j})^2 - \sigma_0^2(x,\theta_0)\Delta_n]a(x)dF(x)\} \ ,$$

then,

$$\begin{aligned}\varepsilon_{n1} &= n^{-3}(h\Delta_n)^{-2}\sum_{i,j,k=1}^{n} W_{i,j,k}(x_{n,k}) \\ &= n^{-3}(h\Delta_n)^{-2}\sum_{i=1}^{n} W_{i,i,i}(x_{n,i}) + n^{-3}(h\Delta_n)^{-2}\sum_{i\neq k} W_{i,i,k}(x_{n,k}) \\ &\quad + n^{-3}(h\Delta_n)^{-2}\sum_{i\neq j,k} W_{i,j,k}(x_{n,k}) \\ &\equiv \varepsilon_{n1}^1 + \varepsilon_{n1}^2 + \varepsilon_{n1}^3 . \end{aligned} \quad \text{(A.7)}$$

To prove that $\varepsilon_{n1} = o_p((nh^{1/2})^{-1})$, I need to prove that $\varepsilon_{n1}^i = o_p((nh^{1/2})^{-1})$ for $i = 1, 2, 3$. For ε_{n1}^1, because $(x_{n,i+1}-x_{n,i})^{2l} = O_p(\Delta_n^l)$ for $l = 1, 2$ by (A.1) in Lemma 2, I immediately obtain:

$$\varepsilon_{n1}^1 = O_p((nh\Delta_n)^{-2}\Delta_n^2) = o_p((nh^{1/2})^{-1}). \quad \text{(A.8)}$$

For ε_{n1}^2, I denote $(\varepsilon_{n1}^2)^2$ as:

$$\begin{aligned}(\varepsilon_{n1}^2)^2 &= n^{-2}(nh\Delta_n)^{-4}\sum_{i\neq k} W_{i,i,k}\sum_{i'\neq k'} W_{i',i',k'} \\ &\equiv (\varepsilon_{n11}^2)^2 + (\varepsilon_{n12}^2)^2 + (\varepsilon_{n13}^2)^2 \ ,\end{aligned}$$

where $(\varepsilon_{n11}^2)^2$, $(\varepsilon_{n12}^2)^2$, and $(\varepsilon_{n13}^2)^2$, respectively, denote the cases where the summation indices satisfy $min\{|k-k'|, |k-i|, |k-i'|\} > m+1$, either $i = i', k = k'$ or $i = k', k = i'$ and all remaining cases. For $(\varepsilon_{n11}^2)^2$, using $E_k(W_{i,i,k}(x_{n,k})) = 0$, and Lemma 1:

$$E(n^2 h(\varepsilon_{n11}^2)^2) \leq 0 + Ch^{-3}(n\Delta_n)^{-4}(h\Delta_n^4)n^4\beta_{nm}^{1/2} = O(h^{-2}\beta_{nm}^{1/2}) = o(1). \quad \text{(A.9)}$$

For $(\varepsilon_{n12}^2)^2$:

$$E(n^2 h(\varepsilon_{n12}^2)^2) \leq Ch^{-3}(n\Delta_n)^{-4}(n^2\Delta_n^4) = O((nh^{3/2})^{-2}) = o(1). \quad \text{(A.10)}$$

For $(\varepsilon_{n13}^2)^2$, using Holder inequality:

$$E(n^2 h(\varepsilon_{n13}^2)^2) \leq C(nh^3)^{-1} h^{4/3} m = O(m(nh^{1.5})^{-1}) = o(1). \quad \text{(A.11)}$$

From (A.9), (A.10), and (A.11):

$$\varepsilon_{n1}^2 = o_p((nh^{1/2})^{-1}). \quad \text{(A.12)}$$

To evaluate ε_{n1}^3, I consider the second moment of $(nh^{1/2}\varepsilon_{n1}^3)^2$:

$$E(nh^{1/2}\varepsilon_{n1}^3)^2 = h^{-3}(n\Delta_n)^{-4} \sum_{i \neq j, k} \sum_{i' \neq j', k'} E[W_{i,j,k} W_{i',j',k'}]. \quad \text{(A.13)}$$

I consider four different cases: (a) for any two summation indices l and L from k, i, j, k', i', j', $|l - L| > m + 1$ for all $L \neq l$; (b) there exist exactly four different summation indices such that, for any index l from these four indices, $|l - L| > m + 1$ for all $L \neq l$; (c) there exist exactly three different summation indices such that, for any index l from these three indices, $|l - L| > m + 1$ for all $L \neq l$; and (d) all the other remaining cases. I will use EA_s to denote these cases $(s = a, b, c, d)$. For case (a), using $E_k W_{i,j,k} = 0$ or $E_{k'} W_{i',j',k'} = 0$ and Lemma 1:

$$EA_a \leq 0 + C(h^{-3})(n\Delta_n)^{-4} n^6 (h\Delta_n^4)\beta_{nm}^{1/2} = Cn^2 h^{-2}\beta_{nm}^{1/2} = O(n^4\beta_{nm}^{1/2}) = o(1). \quad \text{(A.14)}$$

For case (b), it is necessary to consider only the case $|k - k'| \leq m + 1$, since otherwise I will have k or k' is at least m periods away from any other indices and, by

Lemma 1, I know it is bounded by $Cn^6\beta_{nm}^{1/2} = o(1)$. With $|k - k'| \le m + 1$, l must be at least $m + 1$ periods away from any other indices for $l = i, j, i', j'$. Hence, repeating the application of Lemma 1 yields:

$$EA_b \le h^{-3}(n\Delta_n)^{-4} \sum_{i \ne j, k} \sum_{i' \ne j', k'} E[E_j E_i(W_{i,j,k}) E_{j'} E_{i'}(W_{i',j',k'})] + Cn^6\beta_{nm}^{1/2} .$$

$$\le Ch^{-3}(n\Delta_n)^{-4}(\Delta_n^4 h^8 + \Delta_n^6 h^4) m n^5 + Cn^6\beta_{nm}^{1/2} = o(1) . \qquad \text{(A.15)}$$

For case (c), it is necessary to consider only $|k - k'| \le m + 1, |k - l| \le m + 1$ for exactly one $l \in \{i, j, i', j'\}$, since otherwise it will be bounded by $Cn^6\beta_{nm}^{1/2}$ by Lemma 1. By symmetry it is necessary to consider only $l = i$. Repeating the application of Lemma 1 yields:

$$EA_{(c)} \le h^{-3}(n\Delta_n)^{-4} \sum_{i \ne j, k} \sum_{i' \ne j', k'} E[E_j(W_{i,j,k}) E_{j'} E_{i'}(W_{i',j',k'})] + Cn^6\beta_{nm}^{1/2}$$

$$\le Ch^{-3}(n\Delta_n)^{-4}(\Delta_n^2 h^4 + \Delta_n^3 h^2)(\Delta_n^2 h^2 + \Delta_n^{5/2} h) n^4 m^2 + Cn^6\beta_{nm}^{1/2}$$

$$= C(h^3 + h^2\Delta_n^{1/2} + \Delta_n h + \Delta_n^{3/2}) m^2 + Cn^6\beta_{nm}^{1/2} = o(1) . \qquad \text{(A.16)}$$

For case (d), for any three different l's, $|k - l| \le m + 1$ for all $k \ne l$; case (d) has, at most, $m^3 n^3$ terms. Hence, using Lemma 1:

$$EA_{(d)} \le Ch^{-3}(n\Delta_n)^{-4} m^3 n^3 h^2 \Delta_n^4 + Cn^6\beta_{nm}^{1/2} = Cm^3(nh)^{-1} = o(1) . \qquad \text{(A.17)}$$

From (A.13)-(A.17):

$$\varepsilon_{n1}^3 = o_p((nh^{1/2})^{-1}) . \qquad \text{(A.18)}$$

Finally, from (A.8), (A.12), and (A.18), $\varepsilon_{n1} = o_p((nh^{1/2})^{-1})$.

To prove that $\varepsilon_{n2} = o_p((nh^{1/2})^{-1})$, expanding $\sigma_0(x, \hat{\theta}_n)$ around θ_0 yields:

$$\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0) = \nabla_\theta'\sigma_0^2(x, \theta_0)(\hat{\theta}_n - \theta_0) + \frac{1}{2}(\hat{\theta}_n - \theta_0)'\nabla_\theta^2\sigma_0^2(x, \bar{\theta}_n)(\hat{\theta}_n - \theta_0) . \quad (A.19)$$

Using (A.19), ε_{n2} can be expressed as:

$$\varepsilon_{n2} = -(n\Delta_n)^{-1}(nh)^{-2}\sum_{i, j, k = 1}^{n}\{K_i(x_{n, k})K_j(x_{n, k})[(x_{n, i+1} - x_{n, i})^2 - \sigma_0^2(x_{n, k}, \theta_0)\Delta_n]$$

$$\times \nabla_\theta'\sigma_0^2(x_{n, j}, \theta_0)a(x_{n, k})$$

$$-\int K_i(x)K_j(x)[(x_{n, i+1} - x_{n, i})^2 - \sigma_0^2(x, \theta_0)\Delta_n]\nabla'_\theta\sigma_0^2(x, \theta_0)a(x)dF(x)\}(\hat{\theta}_n - \theta_0)$$

$$+ (\hat{\theta}_n - \theta_0)'(n\Delta_n)^{-1}(nh)^{-2}\sum_{i, j, k = 1}^{n}\{K_i(x_{n, k})K_j(x_{n, k})[(x_{n, i+1} - x_{n, i})^2 - \sigma_0^2(x_{n, k}, \theta_0)\Delta_n]$$

$$\times \nabla_\theta^2\sigma_0^2(x_{n, j}, \bar{\theta}_n)a(x_{n, k}) - \int K_i(x)K_j(x)[(x_{n, i+1} - x_{n, i})^2 - \sigma_0^2(x, \theta_0)\Delta_n]\nabla_\theta^2\sigma_0^2(x, \bar{\theta}_n)$$

$$\times a(x)dF(x)\}(\hat{\theta}_n - \theta_0)$$

$$= \varepsilon_{n2}^1(\hat{\theta}_n - \theta_0) + (\hat{\theta}_n - \theta_0)'\varepsilon_{n2}^2(\hat{\theta}_n - \theta_0) . \quad (A.20)$$

Using the same approach I used to prove that $\varepsilon_{n1} = o_p((nh^{1/2})^{-1})$, I can prove that $\varepsilon_{n2}^1 = o_p((nh^{1/2})^{-1})$. By $(x_{n, i+1} - x_{n, i})^2 = O_p(\Delta_n)$ and Assumption 5, it follows that $\varepsilon_{n2}^2 = O_p(1)$. Hence, $\varepsilon_{n2} = o_p((nh^{1/2})^{-1})$.

Finally, for ε_{n3}:

$$\varepsilon_{n3} = n^{-1}(nh)^{-2}\sum_{i, k = 1}^{n}\{K_i(x_{n, k})K_i(x_{n, k})v^2(x_{n, k}, \theta_0, \hat{\theta}_n) - \int K_i(x)K_i(x)v^2(x, \theta_0, \hat{\theta}_n)a(x)dF(x)\}$$

$$+ n^{-1}(nh)^{-2}\sum_{i \neq j, k}^{n}\{K_i(x_{n, k})K_j(x_{n, k})v^2(x_{n, k}, \theta_0, \hat{\theta}_n) - \int K_i(x)K_j(x)v^2(x, \theta_0, \hat{\theta}_n)a(x)dF(x)\}$$

$$\equiv \varepsilon_{n3}^1 + \varepsilon_{n3}^2 . \quad (A.21)$$

It is easy to show that $\varepsilon_{n3}^1 = O_p(n^{-2}h^{-1}) = o_p(1)$ and $\varepsilon_{n3}^2 = O_p(n^{-1}h^{-1}) = o_p(1)$. Hence, $\varepsilon_{n3} = o_p(nh^{1/2})$.

A central limit theorem is required for degenerate U-statistics of the triangular arrays of random variables $\{X_{n,l}\}$, $l \le n$, and $n \ge 1$, which is used to derive the asymptotic distribution of the test statistic proposed in this paper:

$$U_n = \sum\sum_{1 \le s < t \le n} H_n(X_{n,t}, X_{n,s}),$$

where $H_n(.,.)$ depends on n and satisfies $\int H_n(x, y) dF_n(x) = 0$ for all y, and $F_n(.)$ is the marginal distribution function of $X_{n,t}$. For every n, let $\{\tilde{X}_{n,l}\}_{l=1}^{n}$ be an independent, identically distributed (i.i.d.) sequence that has the same marginal distribution as $\{X_{n,l}\}$. Following the same approach as in Fan and Li (1999), I define $\sigma_n^2 \equiv E[H_n^2(\tilde{X}_{n,1}, \tilde{X}_{n,2})]$. In Lemma 4, Assumptions (A1)-(A3) in Fan and Li (1999) are said to be satisfied by $\{X_{n,t}\}$ if the conditions in Assumptions (A1)-(A3) in Fan and Li (1999) calculated by every row of $\{X_{n,t}\}$ are satisfied.

Lemma 4. *Let* $\{X_{n,t}\}$ *be strictly stationary and satisfy the absolutely regular condition with mixing coefficient* β_{nm}. *If Assumptions (A1)-(A3) in Fan and Li (1999) are satisfied by* $\{X_{n,t}\}$ *, then:* $\frac{\sqrt{2}U_n}{n\sigma_n} \to N[0, 1]$ *in distribution as* $n \to \infty$.

Proof of Lemma 4: The proof follows the same way as the proof of Theorem 2.1 (Fan and Li 1999), and thus is omitted from here.

Proof of part (a) of Theorem 1: From Lemma 3, I will complete the proof of part (a) of Theorem 1 by showing***:***

(i) $\bar{J}_n \equiv (nh^{1/2})\left[\bar{I}_n - \frac{2}{n^2 h}\sum_{t=1}^{n}(\hat{\sigma}_n^2(x_{n,t}))^2 a(x_{n,t})\hat{\pi}(x_{n,t})\int k^2(u)du\right] \to N[0, v^2]$ in distribution as $n \to \infty$.

(ii) $\hat{v}_n^2 \to v^2$ in probability as $n \to \infty$.

Proof of (i): $\bar{I}_n$ can be rewritten as:

$$\bar{I}_n = \int\left\{\sum_{j=1}^{n-1} \frac{1}{nh\Delta_n} K_j(x)[(x_{n,j+1} - x_{n,j})^2 - \sigma_0^2(x, \hat{\theta}_n)\Delta_n]\right\}^2 a(x)dF(x) \quad .$$

Let $\bar{r}_n(z_{n,j}, x, \theta) = r_n(z_{n,j}, x, \theta) - E(r_n(z_{n,j}, x, \theta))$, where $z_{n,j} = (x_{n,j}, x_{n,j+1})$. I decompose $\bar{I}_n$ according to

$$\begin{aligned} \bar{I}_n = {} & 2\sum_{1 \le j < k \le n} \int \bar{r}_n(z_{n,j}, x, \hat{\theta}_n)\bar{r}_n(z_{n,k}, x, \hat{\theta}_n)a(x)dF(x) \\ & + \sum_{j=1}^{n} \int r_n^2(z_{n,j}, x, \hat{\theta}_n)a(x)dF(x) \\ & + 2(n-1)\sum_{j=1}^{n} \int \bar{r}_n(z_{n,j}, x, \hat{\theta}_n)E(r_n(z_{n,j}, x, \hat{\theta}_n))a(x)dF(x) \\ & + n(n-1)\int [E(r_n(z_{n,1}, x, \hat{\theta}_n))]^2 a(x)dF(x) \\ \equiv {} & \bar{I}_{n1} + \bar{I}_{n2} + \bar{I}_{n3} + \bar{I}_{n4} \quad . \end{aligned} \tag{A.22}$$

I will show under the assumptions that $\bar{I}_{n1}$ is asymptotically normal in distribution, $\bar{I}_{n3}$ and $\bar{I}_{n4}$ are asymptotically negligible in probability, and $\bar{I}_{n2}$ gives a bias term. First, I prove that $nh^{1/2}(\bar{I}_{n2} - E\bar{I}_{n2}) \to 0$ in probability under the null hypothesis:

$$\begin{aligned} Var[\bar{I}_{n2}] = {} & \sum_{j=1}^{n} Var[\int r_n^2(z_{n,j}, x, \hat{\theta}_n)a(x)dF(x)] \\ & + 2\sum_{1 \le j < k \le n} \{E[\int r_n^2(z_{n,j}, x, \hat{\theta}_n)a(x)dF(x) \times \int r_n^2(z_{n,k}, x, \hat{\theta}_n)a(x)dF(x)] \\ & - \int E[r_n^2(z_{n,j}, x, \hat{\theta}_n)]a(x)dF(x) \times \int E[r_n^2(z_{n,k}, x, \hat{\theta}_n)]a(x)dF(x)\} \\ \equiv {} & \bar{I}_{n2}^1 + \bar{I}_{n2}^2 \quad . \end{aligned} \tag{A.23}$$

Changing variables by $(x_{n,j} - x)/h = u$ and using (a) of Lemma 2:

$$n^{-1}\bar{I}_{n2}^{1} \le E[\int r_n^2(z_{n,j}, x, \hat{\theta}_n)a(x)dF(x)]^2$$

$$= E\{\int (nh\Delta_n)^{-2} K_j^2(x)[(x_{n,j+1} - x_{n,j})^4 - 2\sigma_0^2(x, \hat{\theta}_n)\Delta_n(x_{n,j+1} - x_{n,j})^2$$

$$+ \sigma_0^4(x, \hat{\theta}_n)\Delta_n^2]a(x)dF(x)\}^2$$

$$\le C((n\Delta_n)^4 h^2)^{-1} E\left\{[\int K^2(u)a(x_{n,j} + hu)\pi(x_{n,j} + hu)du]^2 (x_{n,j+1} - x_{n,j})^8\right\}$$

$$+ C(n^4h^2)^{-1} E[\int K^2(u)\sigma_0^8(x_{n,j} + hu, \hat{\theta}_n)a(x_{n,j} + hu)\pi(x_{n,j} + hu)du]^2$$

$$+ C(n^4h^2)^{-1} E\left\{[\int K^2(u)\sigma_0^4(x_{n,j} + hu, \hat{\theta}_n)a(x_{n,j} + hu)\pi(x_{n,j} + hu)du]^2 (x_{n,j+1} - x_{n,j})^4\right\}$$

$$= O((n^4h^2)^{-1}). \tag{A.24}$$

For $\bar{I}_{n2}^{2}$, I consider two different cases: (a) $min\{|j-k|\} > m+1$ and (b) $min\{|j-k|\} \le m+1$. I will use EB_a and EB_b to denote cases (a) and (b), respectively. By Schwarz's inequality and (A.24), uniformly for j and k:

$$E[\int r_n^2(Z_{n,j}, x, \hat{\theta}_n)a(x)dF(x) \times \int r_n^2(Z_{n,k}, y, \hat{\theta}_n)a(y)dF(y)]^2 = O((n^4h^2)^{-1}) \ . \tag{A.25}$$

Hence, by Lemma 1:

$$EB_a \le 0 + C(nh)^{-2}\beta_{nm}^{1/2}, \tag{A.26}$$

$$EB_b \le Cmn^{-3}h^{-2} \ . \tag{A.27}$$

By (A.3), (A.25), (A.26), (A.27), and Chebyshev's inequality, it follows that:

$$nh^{1/2}\left[\bar{I}_{n2} - \frac{2}{nh}\int \sigma^4(x)\pi(x)a(x)dF(x)\int k^2(x)dx\right] = o_p(1) \ , \tag{A.28}$$

which characterizes the asymptotic bias term in the test statistic. Using Theorem 3.3.2 and Remark 3.3.4 and Remark 3.3.5 in Györfit et al. (1989, Section III.3), or Theorem 1(b) of

Andrews (1995), I obtain the following results:

$$sup_{x \in S}|\hat{\pi}(x) - \pi(x)| = O_p(h^2 + n^{-1/2}h^{-1/2}(\ln(n))^{1/2}) , \quad (A.29)$$

$$sup_{x \in S}|\hat{\sigma}^2(x) - \sigma^2(x)| = O_p(h^2 + n^{-1/2}h^{-1/2}\ln(n)) . \quad (A.30)$$

The proofs of (A.29) and (A.30) are available from the author upon request. By the central limit theorem it is easy to verify that,

$$\frac{2}{nh}\int\sigma^4(x)\pi(x)a(x)dF(x) = \frac{2}{n^2h}\sum_{i=1}^{n}\sigma^4(x_{n,i})\pi(x_{n,i})a(x_{n,i}) + O_p(n^{-3/2}h^{-1}). \quad (A.31)$$

From (A.28)-(A.31), I obtain the sample analogue of (A.28):

$$nh^{1/2}\left[\bar{I}_{n2} - \frac{2}{n^2h}\sum_{i=1}^{n}(\hat{\sigma}_n^2(x_{n,i}))^2 a(x_{n,i})\hat{\pi}(x_{n,i})\int k^2(u)du\right] = o_p(1) . \quad (A.32)$$

To prove that $nh^{1/2}\bar{I}_{n3} = o_p(1)$, I evaluate $E(nh^{1/2}\bar{I}_{n3})^2$. Using Lemma 2 and choosing $1 < \xi < 4/3$ in (A.2):

$$E(nh^{1/2}\bar{I}_{n3})^2 = n^2(n-1)^2h\sum_{j,k=1}^{n}\iint E[r_n(z_{n,j}, x, \hat{\theta}_n)r_n(z_{n,k}, y, \hat{\theta}_n)]E[r_n(z_{n,1}, x, \hat{\theta}_n)]$$
$$\times E[r_n(z_{n,1}, y, \hat{\theta}_n)]a(x)a(y)dF(x)dF(y) - n^4(n-1)^2h\left\{\int[Er_n(z_{n,1}, x, \hat{\theta}_n)]^2a(x)dF(x)\right\}^2$$
$$= o(1). \quad (A.33)$$

From $E(\bar{I}_{n3}) = 0$, (A.33), and Chebyshev’s inequality, it follows that $nh^{1/2}\bar{I}_{n3} = o_p(1)$. For $\bar{I}_{n4}$, choosing $1 < \xi < 4/3$ from (A.2):

$$nh^{1/2}\bar{I}_{n4} = n^2(n-1)h^{1/2}\int[E(r_n(z_{n,1}, x, \hat{\theta}_n))]^2a(x)dF(x)$$
$$= O(Th^{1/2}) + O(nh^{4.5}) + O(h^{2/\xi - 3/2}) = o(1) \cdot \quad (A.34)$$

Next, I express $\bar{I}_{n1}$ in U-statistic form as:

$$\bar{I}_{n1} = 2 \sum_{1 \le j < k \le n} \int \bar{r}_n(z_{n,j}, x, \theta_0)\bar{r}_n(z_{n,k}, x, \theta_0)a(x)dF(x)$$

$$+ (nh)^{-2} \sum_{j \ne k} \int \{K_j(x)[\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0)] - E[K_j(x)(\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0))]\}$$

$$\times \{K_k(x)[\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0)] - E[K_k(x)(\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0))]\}a(x)dF(x)$$

$$-2 \sum_{j \ne k} \int \bar{r}_n(z_{n,j}, x, \theta_0)\{K_k(x)(\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0)) - E[K_k(x)(\sigma_0^2(x, \hat{\theta}_n) - \sigma_0^2(x, \theta_0))]\}$$

$$\times a(x)dF(x)$$

$$\equiv \bar{I}_{n11} + \bar{I}_{n12} - 2\bar{I}_{n13}. \tag{A.35}$$

Using Holder inequality, $\bar{I}_{n12} = o_p((nh^{1/2})^{-1})$ and $\bar{I}_{n13} = o_p((nh^{1/2})^{-1})$. I will use Lemma 4 to prove that $\bar{I}_{n11}$ is asymptotically normal in distribution. I next verify that Assumptions (A1)-(A3) in Fan and Li (1999) are satisfied under Assumptions 1-9.

Let $H(z_{n,j}, z_{n,k}) \equiv \int \bar{r}_n(z_{n,j}, x, \theta_0)\bar{r}_n(z_{n,k}, x, \theta_0)a(x)dF(x)$ and $\{\tilde{z}_{n,j}\}_{j=1}^n$ be an i.i.d. sequence having the same marginal distribution as $\{z_{n,j}\}_{j=1}^n$. Then:

$$\sigma_n^2 \equiv E[H^2(\tilde{z}_{n,1}, \tilde{z}_{n,2})] = \int E[\bar{r}_n(\tilde{z}_{n,1}, x, \theta_0)\bar{r}_n(\tilde{z}_{n,1}, y, \theta_0)]$$

$$\times E[\bar{r}_n(\tilde{z}_{n,2}, x, \theta_0)\bar{r}_n(\tilde{z}_{n,2}, y, \theta_0)]a(x)a(y)dF(x)dF(y)$$

$$= \int\int E[r_n(\tilde{z}_{n,1}, x, \theta_0)r_n(\tilde{z}_{n,1}, y, \theta_0)]$$

$$\times E[r_n(\tilde{z}_{n,2}, x, \theta_0)r_n(\tilde{z}_{n,2}, y, \theta_0)]a(x)a(y)dF(x)dF(y)$$

$$-2\int\int E[r_n(\tilde{z}_{n,1}, x, \theta_0)r_n(\tilde{z}_{n,1}, y, \theta_0)]$$

$$\times E[r_n(\tilde{z}_{n,2}, x, \theta_0)]E[r_n(\tilde{z}_{n,2}, y, \theta_0)]a(x)a(y)dF(x)dF(y)$$

$$+ \int\int E[r_n(\tilde{z}_{n,1}, x, \theta_0)]E[r_n(\tilde{z}_{n,1}, y, \theta_0)]$$

$$\times E[r_n(\tilde{z}_{n,2}, x, \theta_0)]E[r_n(\tilde{z}_{n,2}, y, \theta_0)]a(x)a(y)dF(x)dF(y).$$

To evaluate σ_n^2, I first evaluate $E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)]$:

$$E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)] = (nh\Delta_n)^{-2} E\left\{K_1(x)K_1(y)E_1[(x_{n,2} - x_{n,1})^2 - \sigma^2(x)\Delta_n]^2\right\}$$

$$+ \sigma^2(x)\Delta_n(nh\Delta_n)^{-2} E\left\{K_1(x)K_1(y)E_1[(x_{n,2} - x_{n,1})^2 - \sigma^2(x)\Delta_n]\right\}$$

$$-\sigma^2(y)\Delta_n(nh\Delta_n)^{-2} E\left\{K_1(x)K_1(y)E_1[(x_{n,2} - x_{n,1})^2 - \sigma^2(x)\Delta_n]\right\}.$$

Applying Itô's formula to $(x_{n,2} - x_{n1})^2$ and $(x_{n,2} - x_{n1})^4$, and using similar arguments as in the proof of Lemma 2, I obtain:

$$E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)] = 2\sigma^4(x)(n^2h)^{-1} \int K(u) K\left(\frac{x-y}{h} + u\right) \pi(x + hu) du$$

$$+ \int K(u) K\left(\frac{x-y}{h} + u\right) \pi(x + hu) du \times O(n^{-2} + \Delta_n^{1/2}(n^2h)^{-1})$$

$$= 2\sigma^4(x)(n^2h)^{-1} \int K(u) K\left(\frac{x-y}{h} + u\right) \pi(x + hu) du$$

$$+ \int K(u) K\left(\frac{x-y}{h} + u\right) \pi(x + hu) du \times o(n^2h)^{-1}, \qquad \text{(A.36)}$$

$$E[r_n(\tilde{z}_{n,1}, x, \theta_0)] E[r_n(\tilde{z}_{n,1}, y, \theta_0)] = O(n^{-2}\Delta_n) + O(n^{-2}h^2) + O(n^{-2}h\Delta_n^{1/2}) \quad \text{.(A.37)}$$

From (A.36) and (A.37), it follows that:

$$E[r_n(\tilde{z}_{n,1}, x, \theta_0)] E[r_n(\tilde{z}_{n,1}, y, \theta_0)] = o(E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)]).$$

Hence, I get:

$$\sigma_n^2 = \iint E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)] E[r_n(\tilde{z}_{n,2}, x, \theta_0) r_n(\tilde{z}_{n,2}, y, \theta_0)] a(x) a(y) dF(x) dF(y)$$

$$+ o(\iint E[r_n(\tilde{z}_{n,1}, x, \theta_0) r_n(\tilde{z}_{n,1}, y, \theta_0)] E[r_n(\tilde{z}_{n,2}, x, \theta_0) r_n(\tilde{z}_{n,2}, y, \theta_0)] a(x) a(y) dF(x) dF(y))$$

$$= \frac{4}{n^4 h^2} \iint \sigma^4(x) \sigma^4(y) \iint K(u) K(v) K\left(\frac{x-y}{h} + u\right) K\left(\frac{x-y}{h} + v\right) \pi(x + hu) \pi(y + hv) du dv$$

$$\times a(x)a(y)\pi(x)\pi(y)dxdy + o(n^{-4}h^{-1})$$

$$= 4(n^4h)^{-1}\int\sigma^8(y)\pi^4(y)a^2(y)dy\int\int\int K(u)K(v)K(w+u)K(w+v)dudvdw + o((n^4h)^{-1})$$

$$= 4(n^4h)^{-1}\int\sigma^8(y)\pi^4(y)a^2(y)dy\int[\int K(u)K(w+u)du]^2dw + o((n^4h)^{-1}). \quad \text{(A.38)}$$

To conserve space, the detailed proof of the rest of the conditions of the CLT in Lemma 4 is not incorporated here, but it is available from the author.

$$\mu_{n4} \equiv E[H^4(\tilde{z}_{n,1}, \tilde{z}_{n,2})] = O((n^8h^3)^{-1}) \ . \quad \text{(A.39)}$$

$$\gamma_{n11} \equiv max_{t \neq s, t' \neq s'}E[H(z_{n,t}, z_{n,s})H(z_{n,t'}, z_{n,s'})]\gamma_{n11} = O(n^{-4}h^{4/\eta - 4}) \quad \text{(A.40)}$$

where ζ is slightly larger than 2 and $1 < \eta = (1-\zeta^{-1})^{-1} < 4/3$.

$$\gamma_{n22} \equiv max_{t \neq s, t' \neq s'}E[H(z_{n,t}, z_{n,s})^2H(z_{n,t'}, z_{n,s'})^2] = O(n^{-8}h^{-6+4/\eta}), \quad \text{(A.41)}$$

where $1 < \eta \leq 4/3$.

$$\gamma_{n13} \equiv max_{t \neq s, t' \neq s'}E[H(z_{n,t}, z_{n,s})^1H(z_{n,t'}, z_{n,s'})^3] = O(n^{-8}h^{-6+4/\eta}), \quad \text{(A.42)}$$

where $1 < \eta \leq 4/3$.

$$\tilde{\gamma}_{n14} \equiv max_{t \neq s}\int\{E[H(z, z_{n,t})H(z, z_{n,s})]\}^2dF_n(z) = (O(n^{-8}h^{-5+5/\eta})) \quad \text{(A.43)}$$

where $1 < \eta < 5/3$.

$$\tilde{\gamma}_{n22} \equiv E[H^2(\tilde{z}_{n,1}, \tilde{z}_{n,2})H^2(\tilde{z}_{n,1}, \tilde{z}_{n,3})] = O(n^{-8}h^{-4+2/\eta}) \quad \text{(A.44)}$$

where $1 < \eta < 2$. To summarize the above, I have shown that $\sigma_n^2 = O(n^{-4}h^{-1})$, $\mu_{n4} = O(n^{-8}h^{-3})$, $\gamma_n \equiv max\{\gamma_{n11}, \bar{\gamma}_{n22}, \bar{\gamma}_{n14}\} = O(n^{-4}h^{4/\eta-4})$, and $v_n \equiv max\{\gamma_{n22}, \gamma_{n13}\}$

$= O(n^{-8}h^{-6+4/\eta})$. These results, together with the assumption $h = O(n^{-1/\gamma})$ and $1.5 < \gamma < 4.5$, imply (A1) (i)-(iii) in Fan and Li (1999).

$$G_n(z_t, z_s) \equiv EH(z_{n,1}, z_t)H(z_{n,1}, z_s) \quad ,$$

$$\sigma_{nG}^2 \equiv E[G^2(z_{n,t}, z_{n,t})] = O(n^{-8}h^{-5+3/\eta}) \qquad \text{(A.45)}$$

where $1 < \eta \le 3/2$.

$$\mu_{nG2} \equiv max_{t \ne s}\int G^2(z_{n,t}, z_{n,s})dQ(z_{n,t}, z_{n,s}) = O(n^{-8}h^{-4+3/\eta}) \qquad \text{(A.46)}$$

where $1 < \eta < 3/2$.

$$\gamma_{nG11} \equiv max\{max_{s \ne s' \ne s''}\left|E[G(z_{n,s}, z_{n,s})G(z_{n,s'}, z_{n,s''})]\right| = O(n^{-8}h^{-4+3/\eta}) \qquad \text{(A.47)}$$

Thus, (A2)(i)-(iii) in Fan and Li (1999) are satisfied.

Finally, it is easy to show that M_n in Fan and Li (1999) is bounded by some positive constant. From $\beta_{nm} = O(n^{-b\kappa})$, $m^2n^2\beta_{nm}^{1/2}/\sigma_n^4 = o(1)$, provided I choose b sufficiently large. Thus, (A3) (i)-(ii) in Fan and Li (1999) are all satisfied.

Proof of (ii): From Assumption 8, (A.29) and (A.30), we have:

$$inf_{x \in S}\pi(x) \equiv C_0 > 0 \qquad \text{(A.48)}$$

$$sup_{x \in S}\left|(\hat{\sigma}^2(x)\hat{\pi}(x))^4 - (\sigma^2(x)\pi(x))^4\right| = o_p(1). \qquad \text{(A.49)}$$

From (A.48) and (A.49) it follows that:

$$\left|\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(\hat{\sigma}^2(x_{n,i})\hat{\pi}(x_{n,i}))^4}{\hat{\pi}(x_{n,i})}a^2(x_{n,i}) - \frac{(\sigma^2(x_{n,i})\pi(x_{n,i}))^4}{\pi(x_{n,i})}a^2(x_{n,i})\right)\right|$$
$$\le sup_{x \in S}\left|\frac{(\hat{\sigma}^2(x)\hat{\pi}(x))^4 - (\sigma^2(x)\pi(x))^4}{\pi(x)}a(x)\right| + sup_{x \in S}\left|\frac{(\hat{\sigma}^2(x)\hat{\pi}(x))^4(\hat{\pi}(x) - \pi(x))}{\pi(x)\hat{\pi}(x)}a(x)\right|$$

$= o_p(1)$, which implies,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{\sigma}^2(x_{n,i})\hat{\pi}(x_{n,i}))^4}{\hat{\pi}(x_{n,i})}a^2(x_{n,i}) = \frac{1}{n}\sum_{i=1}^{n}\frac{(\sigma^2(x_{n,i})\pi(x_{n,i}))^4}{\pi(x_{n,i})}a^2(x_{n,i}) + o_p(1)$$

$$= \int\sigma^8(x)\pi^4(x)a^2(x)dx + o_p(1)$$

by the law of large numbers. Thus, $v_n^2 = v^2 + o_p(1)$.

Proof of part (b) of Theorem 1:

Given Assumptions 1-9, from Theorem 3.2 in Corradi and White (1997), $\hat{\theta}_n$ is a $\sqrt{n}$-consistent estimator of some $\theta^* \in \Theta$. I can show by using the similar arguments as those in the proof of part (a) of the theorem that, under H_1,

$$I_n = \int[(\sigma^2(x) - \sigma_0^2(x, \theta^*))\pi(x)]^2 a(x)dF(x) + o_p(1),$$

where $\int[(\sigma^2(x) - \sigma_0^2(x, \theta^*))\pi(x)]^2 a(x)dF(x) > 0$, and $plim_{n\to\infty} v_n^2 = v^2$. Hence, $J_n = O_p(nh^{1/2})$.

Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue*.**

2005

2005-34	The Exchange Rate and Canadian Inflation Targeting	C. Ragan
2005-33	Does Financial Structure Matter for the Information Content of Financial Indicators?	R. Djoudad, J. Selody, and C. Wilkins
2005-32	Degree of Internationalization and Performance: An Analysis of Canadian Banks	W. Hejazi and E. Santor
2005-31	Forecasting Canadian GDP: Region-Specific versus Countrywide Information	F. Demers and D. Dupuis
2005-30	Intertemporal Substitution in Macroeconomics: Evidence from a Two-Dimensional Labour Supply Model with Money	A. Dib and L. Phaneuf
2005-29	Has Exchange Rate Pass-Through Really Declined in Canada?	H. Bouakez and N. Rebei
2005-28	Inflation and Relative Price Dispersion in Canada: An Empirical Assessment	A. Binette and S. Martel
2005-27	Inflation Dynamics and the New Keynesian Phillips Curve: An Identification-Robust Econometric Analysis	J.-M. Dufour, L. Khalaf, and M. Kichian
2005-26	Uninsured Idiosyncratic Production Risk with Borrowing Constraints	F. Covas
2005-25	The Impact of Unanticipated Defaults in Canada's Large Value Transfer System	D. McVanel
2005-24	A Search Model of Venture Capital, Entrepreneurship, and Unemployment	R. Boadway, O. Secrieru, and M. Vigneault
2005-23	Pocket Banks and Out-of-Pocket Losses: Links between Corruption and Contagion	R.H. Solomon
2005-22	The Effects of the Exchange Rate on Investment: Evidence from Canadian Manufacturing Industries	T. Harchaoui, F. Tarkhani, and T. Yuen
2005-21	The Effectiveness of Official Foreign Exchange Intervention in a Small Open Economy: The Case of the Canadian Dollar	R. Fatum and M.R. King
2005-20	La fonction de production et les données canadiennes	P. Perrier

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :

Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
E-mail: publications@bankofcanada.ca
Web site: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca