

Kleinow, Torsten

**Working Paper**

## Testing the diffusion coefficient

SFB 373 Discussion Paper, No. 2002,38

**Provided in Cooperation with:**

Collaborative Research Center 373: Quantification and Simulation of Economic Processes,  
Humboldt University Berlin

*Suggested Citation:* Kleinow, Torsten (2002) : Testing the diffusion coefficient, SFB 373 Discussion Paper, No. 2002,38, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin,  
<https://nbn-resolving.de/urn:nbn:de:kobv:11-10048964>

This Version is available at:

<https://hdl.handle.net/10419/65295>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Testing the Diffusion Coefficient

Torsten Kleinow

Humboldt-Universität zu Berlin, Institut für Statistik und Ökonometrie, Spandauer Strasse  
1, 10178 Berlin, Germany, E-mail: kleinow@wiwi.hu-berlin.de

This version: April 16, 2002

---

**Abstract:**

In mathematical finance diffusion models are widely used and a variety of different parametric models for the drift and diffusion coefficient coexist in the literature. Since derivative prices depend on the particular parametric model of the diffusion coefficient function of the underlying, a misspecification of this function leads to misspecified option prices. We develop two tests about a parametric form of the diffusion coefficient. The finite sample properties of the tests are investigated in a simulation study and the tests are applied to the 7-day Eurodollar rate, the German stock market index DAX and five German stocks. For all observed processes, we find in the empirical analysis that our tests reject all tested parametric models. We conclude that affine diffusion processes might not be appropriate to model the evolution of financial time series and that a successful model for a financial market should incorporate the history of the observed processes or additional sources of randomness like stochastic volatility models.

*JEL classification:* C12, C14, C22, C52

*Keywords:* Diffusion, Continuous-time financial models, Nonparametric methods, Kernel smoothing, Goodness of fit test, spot rate models, interest rate, stock market index, Empirical Likelihood

---

## 1 Introduction

In mathematical finance diffusion processes are widely used to model the dynamics of stock prices, interest rates or other processes observed in the market. In the last decades many parametric models have been proposed to capture the dynamics of these processes. Since the prices of derivative securities and the calculation of risk measures depend on the particular choice of the model, each model yields different prices and risk measures.

In our approach we assume that the price process of an underlying is a one-dimensional diffusion process  $\{X(t), t \in [0, T]\}$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0, T]})$ . We assume that  $X$  is a strong solution of the stochastic differential equation

$$dX(t) = m\{X(t)\}dt + \sigma\{X(t)\}dW(t) \quad t > 0 \quad (1.1)$$

where  $m$  and  $\sigma$  are smooth functions, such that a unique strong solution of (1.1) exists and  $\{W(t), t \in [0, T]\}$  is a standard Brownian Motion adapted to the filtration  $\{\mathcal{F}_t\}_{t \in [0, T]}$ . Furthermore we assume that  $\sigma^2$  has continuous derivatives up to the second order.

Of particular interest in mathematical finance is the pricing of contingent claims with a payoff function that depends on the evolution of the underlying process  $X$  during the time period  $[0, T]$ . Since the derivative prices are calculated under the risk neutral equivalent martingale measure, the drift  $m$  does not influence the prices, Karatzas & Shreve (1998) and Musiela & Rutkowski (1991). The parameter of interest is the diffusion function  $\sigma$  that captures the volatility of the underlying. In financial application it is often assumed that the diffusion coefficient  $\sigma$  belongs to a set of parametric functions, i.e. there exists an unknown parameter  $\theta_0 \in \Theta$  such that  $\sigma(x) = \sigma(\theta_0, x)$ . A misspecification of the diffusion coefficient's parametric form could lead to misspecified derivative prices.

In the field of interest rate modelling, the specification of  $m$  is also important. However, a first step in evaluating a particular parametric interest rate model could be a test of the parametric form of its diffusion coefficient.

For the above reasons, we propose a procedure that tests a parametric form of  $\sigma^2$  using  $Tn$  discrete observations of  $X$ , where  $n$  is the number of observations per unit of time. The hypotheses of the test are

$$\begin{aligned} H_0 & : \exists \theta_0 \in \Theta : \text{for every } t \in [0, T] : \sigma^2\{X(t)\} = \sigma^2\{\theta_0, X(t)\} && \text{P-a.s.} \\ H_1 & : \forall \theta \in \Theta : \text{for every } t \in [0, T] : \\ & \quad |\sigma^2\{X(t)\} - \sigma^2\{\theta, X(t)\}| \geq c_n \Delta_n(X(t)) && \text{P-a.s.} \end{aligned}$$

where  $\Delta_n$ , the local shift in the alternative, is a sequence of bounded functions and  $c_n$  is the order of difference between  $H_0$  and  $H_1$ . This choice of the alternative ensures that the power of the proposed test depends on the number of observations  $n$ .

A related problem is already discussed in the literature by Aït-Sahalia (1996b) and Hong & Li (2002). Their aim is to test the complete dynamics of the diffusion in (1.1) which is determined by the diffusion coefficient  $\sigma$  as well as the drift  $m$ . The test by Aït-Sahalia (1996b) is based on the comparison of the parametric marginal density with its nonparametric estimator. Since the dependency structure of the observations is not reflected in the marginal distribution, this test can not distinguish between processes with the same marginal density but different dependency structures. Hong & Li (2002) extend the results by applying the Markov property of  $X$ . Their test compares the nonparametric estimator of the transition density with the parametrically estimated transition density, as implied by the null hypothesis.

As mentioned above, in many financial applications the parameter of interest is the diffusion coefficient function  $\sigma$  of  $X$ . On the other hand the estimation of the drift function  $m$  as well as the estimation of parameters of  $m$  is a difficult task, even if a long history of data is given, Merton (1980) and Aït-Sahalia (1996a). For these reasons we introduce a test about  $\sigma^2$  that does not incorporate the drift  $m$  and is thus robust to its misspecification.

The proposed test statistic is based on the comparison of a parametric estimator  $\sigma^2(\hat{\theta}, x)$  and a nonparametric estimator  $S_t^{(n)}(x)$  of  $\sigma^2(x)$ . We apply an estimator  $S_t^{(n)}(x)$  that is proposed by Florens-Zmirou (1993) and is a particular choice of the more general estimator proposed by Jacod (2000).

We assume that the diffusion process  $X$  is observed at equidistant time

points

$$0, 1/n, \dots, ([Tn] - 1)/n, [Tn]/n$$

where  $[a]$  denotes the integer part of  $a$ . We denote by  $X_i$  the value of  $X$  at  $i/n$ , i.e.  $X_i := X(i/n)$  for  $i = 0, \dots, [Tn]$ .

For the remainder of the paper we make the following assumptions

**(A1)** The process  $X$  given as the solution of (1.1) is stationary and  $\alpha$ -mixing, i.e.

$$\alpha(u) := \sup_{\mathcal{A} \in \mathcal{F}_t; \mathcal{B} \in \mathcal{F}_{t+u}^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)| \leq a\rho^u$$

for some  $a > 0$  and  $\rho \in [0, 1)$ . Here  $\mathcal{F}_t^\infty$  denotes the  $\sigma$ -algebra generated by  $\{(X_u), u \geq t\}$ . For an introduction into  $\alpha$ -mixing processes see Bosq (1998) or Billingsley (1968). As shown by Genon-Catalot, Jeantheau & Larédo (2000) this assumption implies that the both  $\{X_i, i = 0, \dots, [nt]\}$  and  $\{(X_{i+1} - X_i)^2, i = 0, \dots, [nt] - 1\}$  are  $\alpha$ -mixing.

**(A2)** The following holds for  $\sigma^2$ :

$$|\sigma^2(\theta, x) - \sigma^2(\theta_0, x)| \leq D(x)\|\theta - \theta_0\| \quad \forall x \in I_X$$

where  $D(x)$  is a constant depending on  $x$  and the set  $I_X$  is defined by

$$I_X := \{x \mid f(x) \geq \varepsilon > 0\}$$

with an arbitrary number  $\varepsilon$ , where  $f(x)$  denotes the marginal density of  $X$ .

**(A3)**  $\hat{\theta}$  is a square root consistent parametric estimator of  $\theta$  within the family of the parametric model, i.e.  $\|\hat{\theta} - \theta\| = \mathcal{O}_p((nt)^{-1/2})$ .

The particular choice of  $\hat{\theta}$  does not influence the test as long as assumption (A3) holds. However, there exists a large literature about parameter estimation in diffusion models, see among others Bibby & Sørensen (1996).

The paper is organized as follows. In Section 2 we introduce two testing procedures that are based on the comparison between parametric and nonparametric estimation of  $\sigma^2$ . In Section 3 we analyze the finite sample properties of both tests by a simulation study. Finally we apply one proposed test to the 7-day Eurodollar rate, the German stock market index DAX and five German stocks.

## 2 Testing procedure

As mentioned in the introduction our test is based on the comparison of a parametric estimates and a nonparametric estimates of  $\sigma^2$ . In a first step we discuss nonparametric estimators of  $\sigma^2$ .

## 2.1 Nonparametric estimation of the Diffusion coefficient

Florens-Zmirou (1993), Jiang & Knight (1997) and in a more general framework Jacod (2000) introduce a nonparametric estimator for  $\sigma^2$  that is based on the local time of  $X$ .

**DEFINITION 2.1** *For a diffusion  $X$  we define*

- *occupation measure  $\nu_t$ :  $\nu_t(B, \omega) = \int_0^t \mathbf{1}_B\{X(u, \omega)\} du$*
- *Local Time:  $L_t(\cdot, \omega) = \frac{d\nu_t}{d\lambda}$  for  $P - a.e. \omega \in \Omega$*

This definition is given in Bosq (1998). Heuristically speaking, the occupation measure  $\nu_t(B)$  measures the time that the process  $X$  spends in the set  $B$  up to time  $t$  and the local time  $L_t(x)$  measures the time that  $X$  spends in a neighborhood of  $x$ .

Note, that a different definition of the semimartingale local time is given by Karatzas & Shreve (1991). They define the semimartingale local time  $\Lambda_t(x)$  as a random field such that for every Borel-measurable function  $k : \mathbb{R} \rightarrow [0, \infty)$  the following equation holds among others:

$$\int_0^t k\{X(u, \omega)\} \sigma^2(X(u, \omega)) du = 2 \int_{-\infty}^{\infty} k(a) \Lambda_t(a, \omega) da \quad (2.1)$$

for  $P - a.e. \omega \in \Omega$ . The relationship between  $L$  and  $\Lambda$  is

$$\sigma^2(x) = \frac{2\Lambda_t(x)}{L_t(x)} = \frac{2\Lambda_t(x)}{tf(x)}$$

where  $f(x)$  denotes the marginal density of  $X$ .

For the nonparametric estimator based on the discrete sample of  $X$  we introduce the notation

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right)$$

where  $K$  is a Lipschitz continuous kernel with support  $[-1; 1]$ , i.e. a probability density function on  $[-1, 1]$  and  $h > 0$  is a bandwidth parameter. From the definition of  $L$  it follows

$$L_t(x) = \lim_{h \rightarrow 0} \int_0^t K_h(X(u) - x) du. \quad (2.2)$$

A nonparametric estimator  $L_t^{(n)}(x)$  of  $L_t(x)$  based on the observation  $X_0, \dots, X_{[tn]}$  is then given by an approximation of the integral in (2.2), i.e.

$$L_t^{(n)}(x) := \frac{1}{n} \sum_{i=0}^{[tn]} K_{h_n}(X_i - x) \quad (2.3)$$

where  $h_n$  is bandwidth tending to 0 for  $n \rightarrow \infty$ . Florens-Zmirou (1993) proves that  $L_t^{(n)}(x)$  converges in the  $L^2$  sense to  $L_t(x)$ , i.e. if  $nh_n^4 \rightarrow 0$  then

$$E \left[ \{L_t^{(n)}(x) - L_t(x)\}^2 \right] \rightarrow 0. \quad (2.4)$$

Using equation (2.1) we get

$$2\Lambda_t(x) = \lim_{h \rightarrow 0} \int_0^t K_h(X(u) - x) d\langle X \rangle_u$$

where  $\langle X \rangle_t$  denotes the quadratic variation of  $X$  up to time  $t$ . We define our estimator of  $\Lambda(x)$  by approximating the integral,

$$2\Lambda_t^{(n)}(x) := \sum_{i=0}^{[tn]-1} K_{h_n}(X_i - x) \{X_{i+1} - X_i\}^2. \quad (2.5)$$

Combining (2.3) and (2.5) yields a nonparametric estimator  $S_t^{(n)}(x)$  for  $\sigma^2(x)$ ,

$$S_t^{(n)}(x) = \frac{n \sum_{i=0}^{[tn]-1} K\left(\frac{X_{i+1}-x}{h_n}\right) \{X_{i+1} - X_i\}^2}{\sum_{i=0}^{[tn]} K\left(\frac{X_i-x}{h_n}\right)} \quad (2.6)$$

which was first given in Florens-Zmirou (1993).

Other estimators of  $\sigma^2$  are discussed in the literature. Kutoyants (1998) propose a discrete approximation of the Tanaka-Meyer formula,

$$\Lambda_t(x) = (X(t) - x)^- - (X(0) - x)^- + \int_0^t \mathbf{1}_{(-\infty, x]}(X(u)) dX(u)$$

while Aït-Sahalia (1996a) applies the Kolmogorov forward equation for ergodic diffusions to get an estimator for  $\Lambda$ ,

$$E[\Lambda_t(x)] = \int_{-\infty}^x m(a) f(a) da$$

where  $f$  denotes the marginal density of  $X$ . Note, that the last equation directly follows from the Tanaka-Meyer formula. Stanton (1997) proposes a nonparametric method that separately estimates the drift  $m$  and the diffusion coefficient  $\sigma$ .  $S_t^{(n)}(x)$  coincides with one of the estimators given there. As we are interested in a smooth estimator for  $\sigma^2$  and do not want to use any information about the drift coefficient  $m$ , we will apply the estimator in (2.6).

Another approach to estimate the diffusion coefficient is based on time discretization. From Itô's formula we get with an appropriate function  $\mu$  depending on  $m$  and  $\sigma$

$$\begin{aligned} \{X_{i+1} - X_i\}^2 &= \int_{i/n}^{(i+1)/n} \sigma^2\{X(u)\} du \\ &\quad + \int_{i/n}^{(i+1)/n} \mu\{X(u)\} dW(u) + \mathcal{O}(n^{-2}) \\ &\approx \sigma^2(X_i) \frac{1}{n} + \mu(X_i) \sqrt{\frac{1}{n}} w_i \end{aligned} \quad (2.7)$$

where  $w_i \sim N(0, 1)$  for all  $i = 0, \dots, [Tn] - 1$ . This approach is applied by Hoffmann (1999) to develop an adaptive nonparametric estimation procedure for  $\sigma$  using wavelets. In Härdle, Kleinow, Korostelev, Logeay & Platen (2001)

a discrete time approximation similar to (2.7) is applied to nonparametrically estimate the drift  $m$  and the diffusion coefficient  $\sigma$  in (1.1) with a Nadaraya-Watson estimator. Note, that a Nadaraya-Watson estimator for  $\sigma^2(x)$  applied in (2.7) coincides with  $S_t^{(n)}(x)$ .

To achieve a test statistic for a particular set of points  $x_1, \dots, x_k$  we will now investigate the joint asymptotic distribution of  $S_t^{(n)}(x_1), \dots, S_t^{(n)}(x_k)$ .

**PROPOSITION 2.1** *Given  $k$  points  $x_1, \dots, x_k$  and under the assumption  $nh_n^3 \rightarrow 0$  the random vector*

$$Z_t^{(n)} = (Z_t^{(n)}(x_1), \dots, Z_t^{(n)}(x_k))^\top$$

with

$$Z_t^{(n)}(x_l) = \sqrt{nhL_t^{(n)}(x_l)} \left( \frac{S_t^{(n)}(x_l)}{\sigma^2(x_l)} - 1 \right) \quad l = 1, \dots, k \quad (2.8)$$

converges in distribution to a random vector  $Z$  where  $Z$  has a joint standard normal distribution.

**PROOF:**

We proof the result only for  $t = 1$ . From Theorem 1 in Florens-Zmirou (1993) and Theorem 1 in Jiang & Knight (1997) we know that  $Z_1^{(n)}(x_l)$  converges in distribution to  $Z_l \sim N(0, 1)$  for every  $l = 1, \dots, k$ . We introduce the notation

$$\begin{aligned} a &= \min\{|x_{l_1} - x_{l_2}|; l_1, l_2 = 1, \dots, k; l_1 \neq l_2\} \\ n_0 &= \min\{n|a > 2h_n\}. \end{aligned} \quad (2.9)$$

Following Florens-Zmirou (1993) we define

$$m_{i+1}(x_l) := \sqrt{\frac{n}{h_n}} K \left( \frac{X_i - x_l}{h_n} \right) \left[ \{X_{i+1} - X_i\}^2 - \frac{\sigma^2(x_l)}{n} \right]$$

$$M_t^{(n)}(x_l) = \sum_{i=0}^{[nt]-1} m_{i+1}(x_l)$$

and get

$$Z_1^{(n)}(x_l) = \frac{M_1^{(n)}(x_l)}{\sigma^2(x_l) \sqrt{L_1^{(n)}(x_l)}}.$$

From (2.4) we have that

$$\tilde{Z}_n(x_l) := \frac{M_1^{(n)}(x_l)}{\sigma^2(x_l) \sqrt{L_1(x_l)}}.$$

also converges in distribution to a standard normal random variable.

For  $k$  arbitrary numbers  $u_l$  we define

$$C_l := \frac{u_l}{\sigma^2(x_l) \sqrt{L_t(x_l)}}$$

and

$$\tilde{Z}_n^{(u)} = \sum_{l=1}^k u_l \tilde{Z}_n(x_l) = \sum_{i=0}^{[nt]-1} \sum_{l=1}^k C_l m_{i+1}(x_l)$$

Then we have from Lemma 2 in Florens-Zmirou (1993) for every  $n > n_0$  and with the notation  $\mathbb{E}^{i,n}[\cdot] = \mathbb{E}[\cdot | \mathcal{F}_{i/n}]$  that

$$\begin{aligned} \sum_{i=0}^{[nt]-1} \mathbb{E}^{i,n} \left[ \left( \sum_{l=1}^k C_l m_{i+1}(x_l) \right)^2 \right] &= \sum_{l=1}^k C_l^2 \sum_{i=0}^{[nt]-1} \mathbb{E}^{i,n} [m_{i+1}^2(x_l)] \\ &= \sum_{l=1}^k C_l^2 \sigma^4(x_l) L_t(x_l) = \sum_{l=1}^k u_l^2 \end{aligned}$$

and

$$\sum_{i=0}^{[nt]-1} \mathbb{E}^{i,n} \left| \sum_{l=1}^k C_l m_{i+1}(x_l) \right|^3 \leq \sum_{l=1}^k |C_l|^3 \sum_{i=0}^{[nt]-1} \mathbb{E}^{i,n} |m_{i+1}(x_l)|^3.$$

Applying proposition 5 in Florens-Zmirou (1993) we have that

$$\tilde{Z}_n^{(u)} \longrightarrow \tilde{Z}^{(u)}$$

in distribution, where  $\tilde{Z}^{(u)}$  is normal distributed with zero expectation and variance  $\sum_{l=1}^k u_l^2$ . Since  $L_t^{(n)}(x_l)$  converges in the  $L_2$  sense to  $L_t(x_l)$  the same follows for

$$\sum_{l=1}^k u_l Z_1^{(n)}(x_l).$$

The convergence in distribution of  $Z_t^{(n)}$  to a joint normal distribution follows from applying the Cramér-Wold device, Billingsley (1968).  $\square$

We remark that Proposition 2.1 is also valid if we replace  $\sigma^2(x)$  in (2.8) by its smoothed version

$$\tilde{\sigma}^2(x) := \frac{\sum_{i=0}^{[nt]} K_{h_n}(X_i - x) \sigma^2(X_i)}{\sum_{i=0}^{[nt]} K_{h_n}(X_i - x)}. \quad (2.10)$$

It is well known that for fixed  $n$  and  $h_n$ ,  $S_t^{(n)}(x)$  is a biased estimator of  $\sigma^2(x)$ . Thus we will not compare it directly with  $\sigma^2(\hat{\theta}, x)$  but with  $\tilde{\sigma}^2(\hat{\theta}, x)$ .

## 2.2 Goodness of Fit Test

For an arbitrary point  $x \in I_X$  we introduce the test statistic

$$\begin{aligned} T_t^{(n)}(x) &= \sqrt{nh_n L_t^{(n)}(x_l)} \left( \frac{S_t^{(n)}(x)}{\tilde{\sigma}^2(\hat{\theta}, x)} - 1 \right) \\ &= \sqrt{nh_n L_t^{(n)}(x_l)} \left( \frac{S_t^{(n)}(x)}{\tilde{\sigma}^2(x)} - 1 \right) + R_t^{(n)}(x) \end{aligned} \quad (2.11)$$



From Proposition 2.1 we know that  $T_t^{(n)}(x)$  converges in distribution to  $Z + R$ , where  $Z$  is standard normal distributed and  $R$  is the limit of

$$R_t^{(n)}(x) = \sqrt{nh_n L_t^{(n)}(x)} \frac{\tilde{\sigma}^2(x) - \tilde{\sigma}^2(\hat{\theta}, x)}{\tilde{\sigma}^2(\hat{\theta}, x)} \frac{S_t^{(n)}(x)}{\tilde{\sigma}^2(x)}$$

for  $n \rightarrow \infty$ . Proposition 1 and 3 in Florens-Zmirou (1993) imply that  $S_t^{(n)}(x)/\tilde{\sigma}^2(x)$  converges to 1 in the  $L^2$  sense if  $nh^4$  tends to 0. Under  $H_0$  assumption (A2) and (A3) imply

$$\frac{\tilde{\sigma}^2(x) - \tilde{\sigma}^2(\hat{\theta}, x)}{\tilde{\sigma}^2(\hat{\theta}, x)} = \mathcal{O}_p\left(\frac{1}{\sqrt{nt}}\right)$$

and it follows with Proposition 1 in Florens-Zmirou (1993) and  $L_t(x) = tf(x)$  that

$$R_t^{(n)}(x) = \sqrt{f(x)} \mathcal{O}_p(\sqrt{h_n}) \{1 + \mathcal{O}_p(nh_n^4)\} = \mathcal{O}_p(\sqrt{h_n}).$$

To study the properties of the test statistic  $T_t^{(n)}(x)$  under the alternative  $H_1$ , we make the following assumption about  $c_n$  and  $\Delta_n$ .

**(A4)**  $\Delta_n(x)$  is bounded with respect to  $n$  and  $x$ . and  $c_n = 1/\sqrt{nth_n}$ .

With assumption (A4) we have under the alternative

$$\frac{\tilde{\sigma}^2(x) - \tilde{\sigma}^2(\hat{\theta}, x)}{\tilde{\sigma}^2(\hat{\theta}, x)} \rightarrow \frac{\sigma^2(x) - \sigma^2(\hat{\theta}, x)}{\sigma^2(\hat{\theta}, x)} \rightarrow \frac{c_n \Delta_n(x)}{\sigma^2(\theta_0, x)}$$

and thus  $R_t^{(n)}(x) = \Delta_n(x)/\sigma^2(\theta_0, x)$

To get a global Goodness-of-Fit test we choose  $k$  arbitrary points  $x_1, \dots, x_k \in I_X$  and built the test statistic

$$T_t^{(n)} = \sum_{l=1}^k \{T_t^{(n)}(x_l)\}^2. \quad (2.12)$$

We now study the asymptotic distribution of  $T_t^{(n)}$  under the null hypothesis.

**PROPOSITION 2.2** *If  $H_0$  holds and  $nh_n^3$  tends to zero, we have for every  $k$  and every set of points  $x_1, \dots, x_k \in I_X$  with  $x_i \neq x_j$  for  $i \neq j$  that  $T_t^{(n)}$  converges in distribution to a  $\chi^2$ -distributed random variable with  $k$  degrees of freedom.*

**PROOF:**

With (2.9) we have for every  $n > n_0$  and for every  $i \neq j$  that

$$\text{Cov}\{T_t^{(n)}(x_i), T_t^{(n)}(x_j)\} = 0$$

and from Proposition 2.1 it follows that  $T_t^{(n)}$  is asymptotically  $\chi^2$ -distributed with  $k$  degrees of freedom.  $\square$

With a similar proof we obtain that  $T_t^{(n)}$  converges under the alternative to a non-central  $\chi^2$ -distributed random variable with  $k$  degrees of freedom and non-centrality parameter  $\sum_{l=1}^k \Delta_n^2(x_l)/\sigma^4(\theta_0, x_l)$ .

We remark that the proposed test statistic is asymptotically equivalent to the  $L_2$  distance between  $S_t^{(n)}$  and  $\tilde{\sigma}^2(\hat{\theta}, \cdot)$ . In a nonparametric regression context Härdle & Mammen (1993) propose a  $L_2$  test statistic

$$T_{HM} := nh_n^{\frac{1}{2}} \int \{S_t^{(n)}(x) - \tilde{\sigma}^2(\hat{\theta}, x)\}^2 \pi(x) dx$$

with a certain weight function  $\pi(x)$ . For a fixed bandwidth  $h_n$  and with  $k_n = 1/(2h_n)$ ,  $x_l = h_n + 2h_n(l-1)$  for  $l = 1, \dots, k_n$  we get that

$$\frac{1}{k_n} T_t^{(n)} = \frac{1}{k_n} nh_n \sum_{l=1}^{k_n} \frac{L_t^{(n)}(x_l)}{\tilde{\sigma}^4(\hat{\theta}, x)} \{S_t^{(n)}(x_l) - \tilde{\sigma}^2(\hat{\theta}, x_l)\}^2$$

is the Riemann approximation of  $\sqrt{h_n} T_{HM}$  with the weight function  $\pi(x) = L_t^{(n)}(x)/\tilde{\sigma}^4(\hat{\theta}, x)$ .

In the simulation study in Section 3 we find that a test based on  $T_t^{(n)}$  is too conservative for all considered models. The reason could be that  $\pi(x)$  does not reflect features of the empirical distribution of  $S_t^{(n)}(x_l) - \tilde{\sigma}^2(\hat{\theta}, x_l)$ . One way to improve the test is to change the weighting function  $\pi(x)$ . The approximation in (2.7) suggest to modify the test statistic in a way that captures the heteroscedasticity of the error terms  $\mu(X_i)\sqrt{1/n}w_i$ . For this reason we will now propose a test statistic based on the empirical likelihood concept.

### 2.3 Empirical Likelihood

The main advantage of Empirical Likelihood methods is their ability to studentize internally and to correct test statistics and confidence intervals for empirically properties of the data. This is the reason, why we introduce a test about  $\sigma$  based on the EL methodology. For a detailed discussion of EL tests and confidence bands we refer to Owen (2001). In a time series context a EL test about a parametric model of the drift of a time series is proposed by Chen, Härdle & Kleinow (2001) and in a diffusion context by Chen, Härdle & Kleinow (2002). We will follow the results of Chen et al. (2001) without giving the proofs. We remark that all proofs are also valid in the context of this paper. However, for the convenience of the reader we show the main principles of the EL as applied to the present situation.

For the sake of simplicity we study the test of  $\sigma$  based on the observations up to time  $t = 1$ . The general case for  $t \in [0, T]$  follows directly.

With the notation

$$\eta_i^{(x)}\{s\} := K\left(\frac{X_i - x}{h_n}\right) \left[ n \{X_{i+1} - X_i\}^2 - s(x) \right] \quad i = 0, \dots, n-1$$

for a positive function  $s$  with support  $I_X$ , we get from the definition of  $S_1^{(n)}(x)$  for any  $x \in I_X$

$$S_1^{(n)}(x) - \tilde{\sigma}^2(\hat{\theta}, x) = \frac{1}{nh_n L_1^{(n)}(x)} \sum_{i=0}^{n-1} \eta_i^{(x)}\{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}$$

and might rewrite  $T_1^{(n)}(x)$  in the following way

$$T_1^{(n)}(x) = \frac{1}{\tilde{\sigma}^2(\hat{\theta}, x) \sqrt{L_1^{(n)}(x)}} \sqrt{\frac{n}{h_n}} \sum_{i=0}^{n-1} \frac{1}{n} \eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\} \quad (2.13)$$

The first part of (2.13) is a factor to standardize the variance of  $T_1^{(n)}(x)$ . The second part is a mean over  $\eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}$  that gives equal weight  $1/n$  to every  $i$ .

To introduce the EL concept we now replace  $T_1^{(n)}(x)$  by a similar statistic which gives different weights to each  $i$ .

$$\tilde{T}_1^{(n)}(x) = \frac{1}{\tilde{\sigma}^2(\hat{\theta}, x) \sqrt{L_1^{(n)}(x)}} \sqrt{\frac{n}{h_n}} \sum_{i=0}^{n-1} p_i \eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\} \quad (2.14)$$

with  $\sum_{i=0}^{n-1} p_i = 1$ . For a fixed point  $x$  we follow Chen et al. (2001) to derive an EL test statistic.

The empirical likelihood  $\mathcal{L}$  for  $s(x)$  is defined by

$$\mathcal{L}\{s(x)\} := \max \prod_{i=0}^{n-1} p_i(x) \quad (2.15)$$

subject to

$$\sum_{i=0}^{n-1} p_i(x) = 1 \quad \text{and} \quad \sum_{i=0}^{n-1} p_i(x) \eta_i^{(x)} \{s\} = 0. \quad (2.16)$$

The second condition reflects that under the null hypothesis  $E[\eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}]$  converges to 0 for  $n \rightarrow \infty$  and  $h_n \rightarrow 0$ . The test is based on the EL ratio  $\mathcal{L}\{\tilde{\sigma}^2(\hat{\theta}, x)\} / \mathcal{L}\{S_1^{(n)}(x)\}$ , which should be close to 1 if the null hypothesis is true. To formalize this idea and to derive a test statistic we study the properties of  $\mathcal{L}\{s(x)\}$ .

Following Owen (2001), we find the maximum of  $\mathcal{L}\{s(x)\}$  by introducing Lagrange multipliers and maximizing the Lagrangian function

$$\mathcal{H}(p, \lambda_1, \lambda_2) = \sum_{i=0}^{n-1} \log p_i(x) - \lambda_1 \sum_{i=0}^{n-1} p_i(x) \eta_i^{(x)} \{s\} - \lambda_2 \left\{ \sum_{i=0}^{n-1} p_i(x) - 1 \right\}$$

where  $\lambda_1$  and  $\lambda_2$  depend on  $x$ . The first order conditions are the equations in (2.16) and

$$\frac{\partial \mathcal{H}(p, \lambda_1, \lambda_2)}{\partial p_i(x)} = \frac{1}{p_i(x)} - \lambda_1 \eta_i^{(x)} \{s\} - \lambda_2 = 0$$

for all  $i = 0, \dots, n-1$ . With the normalization  $\lambda_2 = n$  and  $\lambda(x) = \lambda_1 / \lambda_2$  we obtain the optimal weights

$$p_i(x) = n^{-1} \left[ 1 + \lambda(x) \eta_i^{(x)} \{s\} \right]^{-1} \quad (2.17)$$

where  $\lambda(x)$  is the root of

$$\sum_{i=0}^{n-1} \frac{\eta_i^{(x)}\{s\}}{1 + \lambda(x)\eta_i^{(x)}\{s\}} = 0. \quad (2.18)$$

The maximum empirical likelihood is achieved at  $p_i(x) = n^{-1}$  corresponding to the nonparametric curve estimate  $s(x) = S_1^{(n)}(x)$ . For a parameter estimate  $\hat{\theta}$  we get the maximum empirical likelihood for the smoothed parametric model  $\mathcal{L}\{\tilde{\sigma}^2(\hat{\theta}, x)\}$ . The log-EL ratio is

$$\ell\{\tilde{\sigma}^2(\hat{\theta}, x)\} := -2 \log \frac{\mathcal{L}\{\tilde{\sigma}^2(\hat{\theta}, x)\}}{\mathcal{L}\{S_1^{(n)}(x)\}} = -2 \log[\mathcal{L}\{\tilde{\sigma}^2(\hat{\theta}, x)\}n^n].$$

To study properties of the empirical likelihood based test statistic we need to evaluate  $\ell\{\tilde{\sigma}^2(\hat{\theta}, x)\}$  at an arbitrary  $x$  first, which requires the following lemma about the Lagrange multipliers  $\lambda(x)$  that is proved in Chen et al. (2001).

**LEMMA 2.1** *For  $s(x) = \tilde{\sigma}^2(\hat{\theta}, x)$  and under the assumptions (A1) - (A4) and the additional assumption  $E\{\exp(a_0|n(X_{i+1} - X_i)^2 - \sigma^2(X_1)|)\} < \infty$  for some  $a_0 > 0$  we have*

$$\sup_{x \in I_X} |\lambda(x)| = \mathcal{O}_p\{(nh_n)^{-1/2} \log(n)\}.$$

where we use the notation  $A_n = \mathcal{O}_p(B_n)$  iff  $\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P[|A_n/B_n| > \varepsilon] = 0$ .

An application of the power series expansion of  $1/(1 - \cdot)$  applied to (2.18) and Lemma 2.1 yields

$$\sum_{i=0}^{n-1} \eta_i^{(x)}\{\tilde{\sigma}^2(\hat{\theta}, \cdot)\} \left[ \sum_{j=0}^{\infty} (-\lambda(x))^j (\eta_i^{(x)}\{\tilde{\sigma}^2(\hat{\theta}, \cdot)\})^j \right] = 0. \quad (2.19)$$

With the notation

$$\bar{U}_j(x) = \frac{1}{nh_n} \sum_{i=0}^{n-1} [\eta_i^{(x)}\{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}]^j$$

we have from Lemma 2.1 and (2.19)

$$\lambda(x) = \bar{U}_2^{-1}(x)\bar{U}_1(x) + \tilde{\mathcal{O}}_p\{(nh_n)^{-1} \log^2(n)\}. \quad (2.20)$$

From (2.17), Lemma 2.1 and the Taylor expansion of  $\log(1 + \cdot)$  we get

$$\begin{aligned} \ell\{\tilde{\sigma}^2(\hat{\theta}, x)\} &= -2 \log[\mathcal{L}\{\tilde{\sigma}^2(\hat{\theta}, x)\}n^n] \\ &= 2 \sum_{i=0}^{n-1} \log[1 + \lambda(x)\eta_i^{(x)}\{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}] \\ &= 2nh_n\lambda(x)\bar{U}_1 - nh_n\lambda^2(x)\bar{U}_2 + \tilde{\mathcal{O}}_p\{(nh_n)^{-1/2} \log^3(n)\} \end{aligned} \quad (2.21)$$

Inserting (2.20) in (2.21) yields

$$\ell\{\tilde{\sigma}^2(\hat{\theta}, x)\} = nh_n\bar{U}_2^{-1}(x)\bar{U}_1^2(x) + \tilde{\mathcal{O}}_p\{(nh_n)^{-1/2} \log^3(n)\}$$

and with the definition of  $U_1$  and  $U_2$  we approximate  $\ell\{\tilde{\sigma}^2(\hat{\theta}, x)\}$  by

$$\ell\{\tilde{\sigma}^2(\hat{\theta}, x)\} = \frac{\sum_{i=0}^{n-1} \eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}}{\sum_{i=0}^n (\eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\})^2}$$

and for the general case  $t \in [0, T]$  we have

$$\ell_t\{\tilde{\sigma}^2(\hat{\theta}, x)\} = \frac{\sum_{i=0}^{[tn]-1} \eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\}}{\sum_{i=0}^{[tn]-1} (\eta_i^{(x)} \{\tilde{\sigma}^2(\hat{\theta}, \cdot)\})^2}. \quad (2.22)$$

For  $k$  points  $x_1, \dots, x_k$  we define the global EL Goodness of Fit test statistic  $\mathcal{T}_1^{(n)}$  as in Chen et al. (2001),

$$\mathcal{T}_1^{(n)} := \sum_{l=1}^k \ell^2\{\tilde{\sigma}^2(\hat{\theta}, x_l)\}$$

and for  $t \in [0, T]$

$$\mathcal{T}_t^{(n)} := \sum_{l=1}^k \ell_t^2\{\tilde{\sigma}^2(\hat{\theta}, x_l)\}. \quad (2.23)$$

As in Chen et al. (2001) we can show that the asymptotic distribution of  $\mathcal{T}_t^{(n)}$  under the null hypothesis is again a  $\chi^2$ -distribution with  $k$  degrees of freedom and that  $1/k \mathcal{T}_t^{(n)}$  is asymptotically equivalent to a  $L_2$  distance between  $S_t^{(n)}$  and  $\tilde{\sigma}^2(\hat{\theta}, x)$ . This means that both test statistics,  $T_t^{(n)}$  and  $\mathcal{T}_t^{(n)}$ , are asymptotically equivalent. However, the simulation study shows that the ability of the EL test statistic to internally use features of the empirical distribution of  $S_t^{(n)} - \tilde{\sigma}^2(\hat{\theta}, x)$  results in a smaller empirical level and thus produces more reliable results.

## 2.4 Extension to Time-inhomogeneous Diffusion Coefficients

To extend the proposed methodology to time-inhomogeneous coefficients, we now assume that the diffusion process  $X$  is given as the solution of

$$dX(t) = m\{X(t), t\}dt + \sigma\{X(t), t\}dW(t) \quad t > 0$$

and we replace our null hypothesis by

$$\begin{aligned} H'_0 &: \exists \theta_0 \in \Theta : \text{for every } t \in [0, T] : \\ &\quad \sigma^2\{X(t), t\} = \sigma^2\{\theta_0, X(t), t\} \quad \text{P-a.s.} \end{aligned}$$

Furthermore we replace assumption (A2) by

(A2')

$$|\sigma^2(\theta, x, t) - \sigma^2(\theta_0, x, t)| \leq D(x, t)\|\theta - \theta_0\| \quad \forall x \in I_X, \forall t \in [0, T]$$

where  $D(x, t)$  is a constant depending on  $x$  and  $t$  and the set  $I_X$  is defined as in (A2).

Applying Itô's formula to  $g(x, t) := \int_0^x 1/\sigma(\hat{\theta}, z, t)dz$ , Karatzas & Shreve (1991), we get for  $Y(t) := g(X(t), t)$

$$dY(t) = m_Y\{X(t), t\}dt + \frac{\sigma(X(t), t)}{\sigma(\hat{\theta}, X(t), t)}dW(t) \quad t > 0$$

where  $m_Y(x, t)$  is given by

$$m_Y(x, t) = \frac{\partial}{\partial t}g(x, t) + \frac{\partial}{\partial x}g(x, t)m(x, t) + 0.5\frac{\partial}{\partial x^2}g(x, t)\sigma^2(x, \hat{\theta}, s)$$

By replacing  $x$  by  $g^{-1}(y)$  in the last equation, we get from the assumptions (A2') and (A3) under the null hypothesis a diffusion  $Y$  with constant diffusion coefficient equal to  $1 + \mathcal{O}_p((nt)^{-1/2})$ , for which 1 is a square root consistent estimator. Since the proposed tests do not depend on the drift, and the diffusion coefficient of  $Y$  is asymptotically independent of  $t$ , we are now in the situation described above.

### 3 Finite Sample Properties

We investigate the finite sample properties of the two proposed tests by simulating various models and applying the test to simulated data. The simulated process  $X$  follows the general stochastic differential equation

$$dX(t) = m\{\theta, X(t)\}dt + \sigma\{\theta, X(t)\}dW(t) \quad t > 0 \quad (3.1)$$

where  $\theta$  is a parameter vector. To get discrete observations of  $X$  we use a Milstein scheme

$$X(t + \delta) = X(t) + m(\theta, X(t))\delta + \sigma(\theta, X(t))\sqrt{\delta}\varepsilon(t) + \frac{1}{2}\sigma^2(\theta, X(t))\delta\{\varepsilon(t)^2 - 1\}$$

for  $\delta > 0$  and a sequence of independent standard normal random variables  $\varepsilon(t)$ , Kloeden & Platen (1999).

In the empirical analysis we test parametric models for the diffusion coefficient of the spot rate. For this reason we will investigate the fixed sample properties of the tests applied to these models. A summary of the investigated models is given in Table 1. The parameters are chosen according to the esti-

Name	$\sigma(x)$	$\theta_3$
constant (VC)	$\theta_3$	0.013
square root (CIR)	$\theta_3\sqrt{x}$	0.066
Chan, Karolyi, Longstaff, Sanders (CKLS)	$\theta_3x^{1.5}$	1.2

Table 1: Diffusion coefficient models used in the simulation study

mated values in Ahn & Gao (1999). They estimate the parameters of different

models applied to the one month US treasury bill rate. Note that they did not estimate the parameters for all combination of drift and diffusion coefficients that we use here. However, the parameters in our simulation study generate trajectories that are positive and in the range of about 0.02 - 0.2 and thus might be a good choice to simulate interest rate processes.

To study the influence of the unknown drift function  $m(x)$  on our test results, we use two different drift functions in our simulation, a linear mean reverting drift and a non linear function proposed by Ahn & Gao (1999). Both functions are given in Table 2 together with the parameter values used for the simulation.

Name	$m(x)$	$\theta_1$	$\theta_2$
Linear mean reverting model (LMR)	$\theta_1(\theta_2 - x)$	0.13	0.08
Ahn, Gao model (AG)	$\theta_1(\theta_2 - x)x$	3.4	0.08

Table 2: Drift models used in the simulation study

Ahn & Gao (1999) also estimate a model introduced by Duffie & Kan (1996), where the diffusion coefficient is given by  $\sigma(x) = \sqrt{\theta_3 + \theta_4 x}$ . Since this model is not consistent for values of  $X(t)$  smaller than  $-\theta_3/\theta_4$  we will not use it in our simulation study.

For the simulation we use every combination of the given diffusion coefficient and drift function, i.e. we simulate paths from 6 models. For every model we simulate 1000 paths of length  $nt = 1000$ ,  $nt = 3000$  and  $nt = 5000$ . For  $nt = 1000, 3000$  we simulate 10 observations each day, but sample the process daily. For  $nt = 5000$  we simulate 20 observations per day and sample the data daily. Since the parameter values given in Tables 1 and 2 are annual values, we choose  $n = 250$  (250 trading days per year) and  $t = 4, 12, 20$  years. The parameter estimates are obtained from the quadratic variation.

Both test statistics,  $T_t^{(n)}$  and  $\mathcal{T}_t^{(n)}$ , depend on the choice of the degrees of freedom  $k$ , on the bandwidth  $h$  and on the points  $x_1, \dots, x_k$ . For given degrees of freedom  $k$  we choose

$$h = 1/(2k) \quad \text{and} \quad x_l = h + 2h(l - 1) \quad (3.2)$$

for  $l = 1, \dots, k$ . This choice guarantees that the random variables  $T_t^{(n)}(x_l)$  and  $\ell\{\hat{\sigma}^2(\hat{\theta}, x_l)\}$  are uncorrelated.

For  $nt = 1000$  the empirical levels of both tests are shown in Figure 1. The results indicate that the empirical level of the EL test statistic  $\mathcal{T}_t^{(n)}$  is close to the nominal level only for degrees of freedom between about 4 and 6 and the test based on  $T_t^{(n)}$  is too conservative even for small degrees of freedom. This statement holds independently of the model that is tested. The nonlinearity of the drift seems to have almost no impact on the empirical level of the test.

Figure 2 shows the empirical level for the test about the CKLS diffusion coefficient when the length of the simulated paths is 3000 (left column) and 5000 (right column), respectively. For the simulation we used a nonlinear drift (the AG model). As we expected the empirical level is closer to the nominal level when the sample size is increasing. For the LMR drift and the other two diffusion functions we get similar pictures.

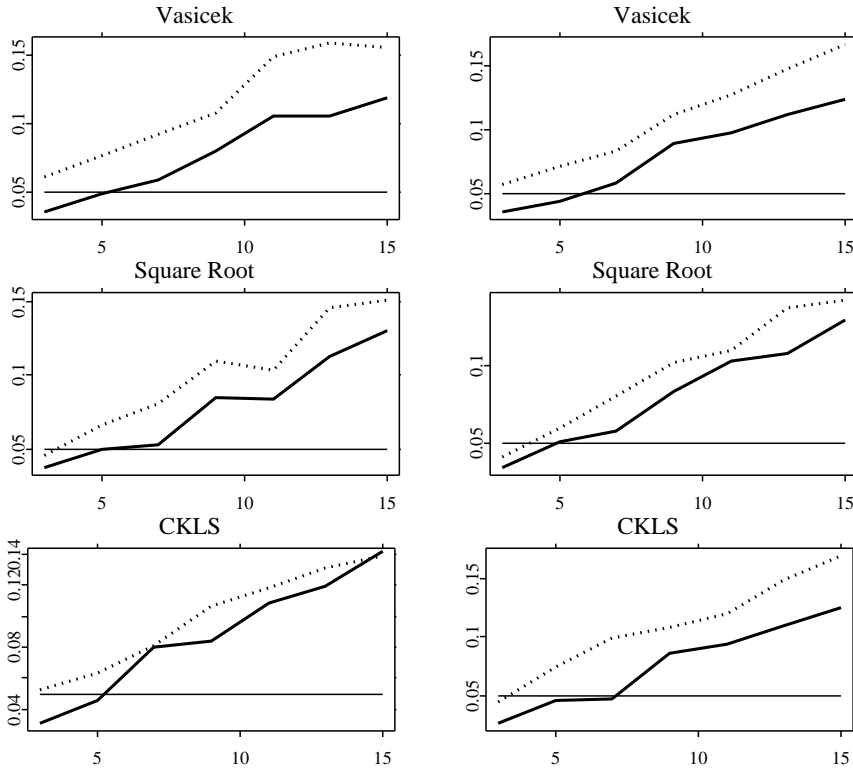


Figure 1: Empirical level of  $T_t^{(n)}$  and  $\mathcal{T}_t^{(n)}$  for different models and path length  $nt = 1000$ . On the vertical axis the empirical level is displayed and the horizontal axis shows the degrees of freedom ( $k$ ). The solid line is the level of  $\mathcal{T}_t^{(n)}$ , the dotted line is the level of  $T_t^{(n)}$  and the thin vertical line is the nominal level of 0.05.

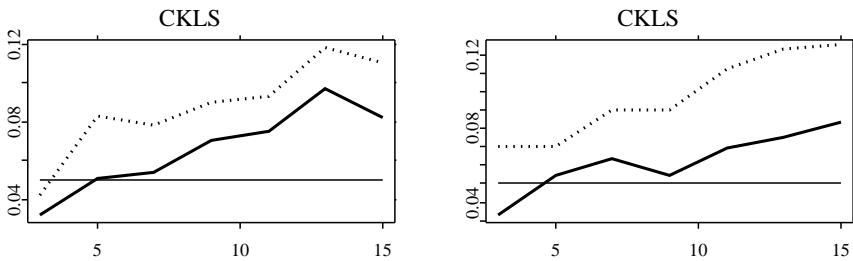


Figure 2: Empirical level of  $T_t^{(n)}$  and  $\mathcal{T}_t^{(n)}$  for the CKLS model with AG drift and path lengths  $nt = 3000$  (left) and  $nt = 3000$  (right). On the vertical axis the empirical level is displayed and the horizontal axis shows the degrees of freedom ( $k$ ). The solid line is the level of  $\mathcal{T}_t^{(n)}$ , the dotted line is the level of  $T_t^{(n)}$  and the thin vertical line is the nominal level of 0.05.



The simulations show that the performance of the test strongly depends on the choice of  $k$ , the degrees of freedom of the asymptotic  $\chi^2$ -distribution. If  $k$  is too large, the approximation of  $T_t^{(n)}(x)$  and  $\ell_t\{\tilde{\sigma}^2(\hat{\theta}, x)\}$  by normally distributed random variables fails and thus the test statistics  $T_t^{(n)}$  and  $\mathcal{T}_t^{(n)}$  are not  $\chi^2$ -distributed. In addition we see from Figure 1 that the empirical level of the test increases with  $k$ . The reason seems to be clear, the larger  $k$  the smaller is  $h_n$ . It is a well known feature of nonparametric estimators that the variance of the estimator is decreasing in  $h_n$ . Thus a larger  $k$ , i.e. smaller  $h_n$ , yields a larger variance of  $T_t^{(n)}(x)$  and  $\ell_t\{\tilde{\sigma}^2(\hat{\theta}, x_l)\}$  and thus a larger expectation of the test statistics. For the 6 simulated models we report the estimated variance and mean of the test statistics in Table 3. It also appears from Figure 1 and Table 3 that the internal studentization of the EL test statistics reduces the variance of  $\mathcal{T}_t^{(n)}$  and thus the empirical level of the test is closer to the nominal level than the empirical level of the  $T_t^{(n)}$  test.

		Vasicek		Square Root		CKLS	
		mean	Var	mean	Var	mean	Var
EL test statistic $\mathcal{T}_t^{(n)}$							
LMR	3	2.42	5.32	2.31	4.89	2.37	5.38
	7	6.87	16.84	6.75	16.25	7.08	19.06
	11	11.60	35.04	11.74	31.12	12.18	34.40
AG	3	2.43	5.49	2.34	5.11	2.25	4.88
	7	6.89	16.51	6.91	16.66	6.85	17.18
	11	11.81	34.30	12.05	37.71	11.93	34.31
test statistic $T_t^{(n)}$							
LMR	3	2.82	7.54	2.61	6.75	2.82	7.44
	11	12.70	41.56	12.26	37.10	12.42	37.98
AG	3	2.83	7.45	2.63	6.72	2.66	6.71
	11	12.67	37.16	12.43	39.56	12.51	41.07

Table 3: Mean and variance of the two test statistics estimated from a sample of 1000 paths with length  $nt = 1000$ .

On the other hand, the comparison of the parametric function  $\tilde{\sigma}^2(\hat{\theta}, \cdot)$  and  $S_t^{(n)}$  is done only at  $k$  points. This means that the smaller  $k$  the less function values are used for the test decision. One way to solve this trade off is to use overlapping intervals for the calculation of the smoother. But in this approach we lost the asymptotic independence of  $T_t^{(n)}(x_l)$  and thus  $T_t^{(n)}$  is not asymptotically  $\chi^2$ -distributed. A similar argument holds for  $\mathcal{T}_t^{(n)}$ .

One possible solution to solve the problem of small sample sizes and to make the test more reliable in such situations is the use of a bootstrap approximation of the asymptotic distribution. Using the bootstrap methodology we could construct the test statistics from small overlapping intervals  $(x_l - h_n, x_l + h_n)$ . One possible bootstrap approach that could be applied in this situation is the local bootstrap method introduced by Paparoditis & Politis (2000). It captures the dependency structure of the data. However, the application of bootstrap is beyond the scope of this paper.

To investigate the power of the EL test we simulate 1000 paths of the Vasicek model with linear drift ( $nt = 1000$ ) and test the three diffusion coefficient models given in Table 1 with this data. The result is shown in Figure 3. It appears from that figure that the power of the test for the square root model is smaller than that of the CKLS model. However, the difference of the empirical rejection level between the (true) Vasicek model and the square root model is significant. This means that the proposed test is able to distinguish these two models. An inclusion of the Ahn, Gao drift does not change the result in principle. Since the test based on  $T_t^{(n)}$  does not hold its nominal level, we will not use it in our empirical analysis and we do not investigate its power.

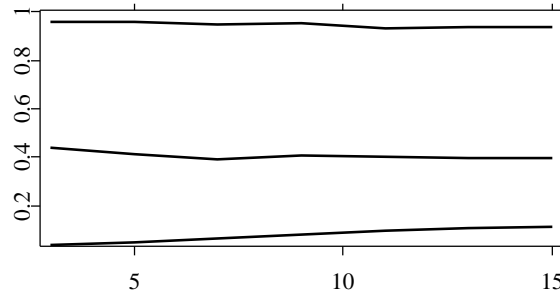


Figure 3: Empirical power of the EL test  $\mathcal{T}_t^{(n)}$  the upper line corresponds to the CKLS model and the middle one to the square root model. The lower line represents the empirical level of the Vasicek model. The paths are simulated from the Vasicek model ( $nt = 1000$ , 1000 trajectories)

### 3.1 Empirical Analysis

We now apply the proposed EL test to the observations of the 7-day Eurodollar rate (interest rate), the DAX stock market index and five German stocks. As already mentioned above, the  $T_t^{(n)}$  test statistic is not applied to the empirical data, since it does not produce reliable results when applied to finite samples.

We start with the analysis of the 7-day Eurodollar rate. The data we use are daily observations of the spot rate from 1975/01/02 to 2002/02/18. This are 7078 observations.

All models in Table 1 are tested. The parameters estimated from  $n = 250$  trading days per year are given in Table 4 along with the values of the EL test statistic. As can be seen, the EL test rejects all models, independently of the chosen degrees of freedom of their asymptotic  $\chi^2$ -distribution. This result indicates that the deviations of the estimated parametric diffusion functions from the nonparametrically estimated one can not be explained by random fluctuations. Since the drift is not used in the construction of the test, an inclusion of a specific drift function will not change the result. In particular, we have seen in the simulation study that the test is robust at least against quadratic drift functions. This means that despite the importance of the drift

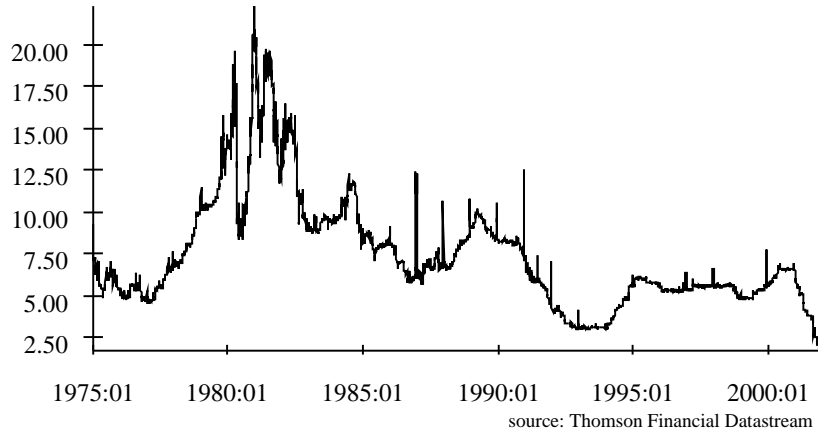


Figure 4: The 7-day Eurodollar rate

function of interest rate models for the valuation of continent claims, all spot rate models that use one of the tested diffusion coefficients can be rejected. This result coincides with the empirical findings by Hong & Li (2002).

$k$	value of $\mathcal{T}_t^{(n)}$			$\hat{\theta}$
	3	7	11	
0.05 critical values	7.815	14.067	19.675	
$\sigma(\theta, x) = \theta$	68.906	274.721	550.362	0.04260
$\sigma(\theta, x) = \theta\sqrt{x}$	19.481	56.268	111.100	0.12554
$\sigma(\theta, x) = \theta x^{1.5}$	89.528	260.461	195.044	1.77895

Table 4: Values of the EL test statistic and estimated parameters for the 7-day Eurodollar rate.

The EL test is also applied to the German stock market index DAX and to the German stocks Allianz, Bayer, Deutsche Bank, RWE and VW. The data we use are daily observations of the assets from 01.07.1991 to 19.02.2002. These are 2778 observations. We apply the test not to the original data but to the log prices,  $X(t) = \log P(t)$ , where  $P(t)$  is the observed price of the asset at time  $t$ . The results of the EL test are given in Table 5.

As for the interest rate, all supposed models are rejected by the test, except the CKLS model, which is not rejected for the VW stock price process when  $k = 5$ .

The empirical results indicate that affine diffusion processes might not be appropriate to model financial time series, like interest rates or stock prices. A number of alternative models is proposed in the literature. Hobson & Rogers (1998) propose a complete model, i.e. without an additional source of randomness. They model price processes as the solution of a stochastic delay differential

$k$		value of $\mathcal{T}_t^{(n)}$			$\hat{\theta}$
		3	7	11	
	0.05 critical values	7.815	14.067	19.675	
DAX	$\sigma(\theta, x) = \theta$	95.495	274.544	276.595	0.181
	$\sigma(\theta, x) = \theta\sqrt{x}$	78.983	332.413	367.068	0.073
	$\sigma(\theta, x) = \theta x^{1.5}$	33.779	239.898	253.275	0.009
Allianz	$\sigma(\theta, x) = \theta$	71.274	260.952	266.548	0.259
	$\sigma(\theta, x) = \theta\sqrt{x}$	59.163	335.127	336.712	0.126
	$\sigma(\theta, x) = \theta x^{1.5}$	21.158	210.860	220.641	0.024
Bayer	$\sigma(\theta, x) = \theta$	119.001	135.356	248.024	0.221
	$\sigma(\theta, x) = \theta\sqrt{x}$	145.471	130.443	255.342	0.138
	$\sigma(\theta, x) = \theta x^{1.5}$	18.203	55.027	117.207	0.041
Deutsche Bank	$\sigma(\theta, x) = \theta$	164.887	101.440	455.881	0.243
	$\sigma(\theta, x) = \theta\sqrt{x}$	232.280	105.342	569.002	0.145
	$\sigma(\theta, x) = \theta x^{1.5}$	122.952	55.802	421.737	0.036
RWE	$\sigma(\theta, x) = \theta$	120.130	172.589	289.700	0.229
	$\sigma(\theta, x) = \theta\sqrt{x}$	113.045	187.370	314.577	0.136
	$\sigma(\theta, x) = \theta x^{1.5}$	37.562	103.120	174.242	0.038
VW	$\sigma(\theta, x) = \theta$	43.655	199.624	196.352	0.290
	$\sigma(\theta, x) = \theta\sqrt{x}$	20.487	187.114	183.105	0.164
	$\sigma(\theta, x) = \theta x^{1.5}$	3.817	117.371	137.169	0.047

Table 5: Values of the EL test statistic and estimated parameters for the DAX and the five German stocks.

equation, where the diffusion and drift coefficients depend on the whole history of the process. Stochastic volatility models, where the diffusion coefficient depends on an additional non observable volatility process are another way to capture the dynamics observed in the market, Hofmann, Platen & Schweizer (1992). As these models yield incomplete markets, derivative prices are not unique.

## References

- Ahn, D.-H. & Gao, B. (1999). A parametric nonlinear model of term structure dynamics, *The Review of Financial Studies* **12**(4): 721–762.
- Aït-Sahalia, Y. (1996a). Nonparametric pricing of interest rate derivative securities., *Econometrica* **64**(3): 527–560.
- Aït-Sahalia, Y. (1996b). Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies* **9**: 385–426.
- Bibby, B. M. & Sørensen, M. (1996). On estimation of discretely observed diffusions, a review, *Theory of Stochastic Processes* **2**(18): 49–56.
- Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.
- Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes*, Vol. 110 of *Lecture Notes in Statistics*, Springer-Verlag, Heidelberg.

- Chen, S. X., Härdle, W. & Kleinow, T. (2001). An empirical likelihood goodness-of-fit test for time series, Discussion paper 1/2001, Sonderforschungsbereich 373, Humboldt-Universität zu Berlin.
- Chen, S. X., Härdle, W. & Kleinow, T. (2002). An empirical likelihood goodness-of-fit test for diffusions, in W. Härdle, T. Kleinow & G. Stahl (eds), *Applied Quantitative Finance*, Springer Verlag.
- Duffie, D. & Kan, R. (1996). A yield-factor model of interest rates., *Math. Finance* **6**(4): 379–406.
- Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations., *J. Appl. Probab.* **30**(4): 790–804.
- Genon-Catalot, V., Jeantheau, T. & Larédo, C. (2000). Stochastic volatility models as hidden markov models and statistical applications, *Bernoulli* **6**(6).
- Härdle, W., Kleinow, T., Korostelev, A., Logeay, C. & Platen, E. (2001). Semiparametric diffusion estimation and application to a stock market index, Discussion Paper 24/2001, Sonderforschungsbereich 373, Humboldt-Universität zu Berlin.
- Härdle, W. & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits, *Ann. Statist.* **21**: 1926–1947.
- Hobson, D. G. & Rogers, L. (1998). Complete models with stochastic volatility., *Math. Finance* **8**(1): 27–48.
- Hoffmann, M. (1999). Adaptive estimation in diffusion processes., *Stochastic Processes Appl.* **79**(1): 135–163.
- Hofmann, N., Platen, E. & Schweizer, M. (1992). Option pricing under incompleteness and stochastic volatility., *Math. Finance* **2**(3): 153–187.
- Hong, Y. & Li, H. (2002). Nonparametric specification testing for continuous-time models with application to spot interest rates, Working paper, Cornell University.
- Jacod, J. (2000). Non-parametric kernel estimation of the coefficient of diffusion., *Scand. J. Stat.* **27**(1): 83–96.
- Jiang, G. & Knight, J. (1997). A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model, *Econometric Theory* **13**(5).
- Karatzas, I. & Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, Springer Verlag New York.
- Karatzas, I. & Shreve, S. E. (1998). *Methods of Mathematical Finance*, Vol. 39 of *Applications of Mathematics, Stochastic Modelling and Applied Probability*, Springer Verlag New York.
- Kloeden, P. E. & Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*, Vol. 23 of *Applications of Mathematics*, Springer Verlag Berlin Heidelberg.

- Kutoyants, Y. (1998). Efficient density estimation for ergodic diffusion processes., *Stat. Inference Stoch. Process.* **1**(2): 131–155.
- Merton, R. C. (1980). On estimating the expected return on the market, *Journal of Financial Economics* **8**: 141–183.
- Musiela, M. & Rutkowski, M. (1991). *Martingale methods in financial modelling*, Springer Verlag New York.
- Owen, A. B. (2001). *Empirical Likelihood*, Vol. 92 of *Monographs on Statistics and Applied Probability*, Chapman & Hall/CRC.
- Paparoditis, E. & Politis, D. N. (2000). The local bootstrap for kernel estimators under general dependence conditions., *Ann. Inst. Stat. Math.* **52**(1): 139–159.
- Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* **52**: 1973–2002.