A Bootstrap Test for the Comparison of Nonlinear Time Series - with Application to Interest Rate Modelling

Holger Dette† & Rafael Weißbach‡*

†Lehrstuhl für Stochastik, Fakultät für Mathematik, Ruhr-Universität Bochum, Bochum, Germany

‡Institut für Wirtschafts- und Sozialstatistik, Fachbereich Statistik, Universität Dortmund, Dortmund, Germany

June 20, 2006

Abstract

We study the drift of stationary diffusion processes in a time series analysis of the autoregression function. A marked empirical process measures the difference between the nonparametric regression functions of two time series. We bootstrap the distribution of a Kolmogorov-Smirnov-type test statistic for two hypotheses: Equality of regression functions and shifted regression functions. Neither markovian behavior nor Brownian motion error of the processes are assumed. A detailed simulation study finds the size of the new test near the nominal level and a good power for a variety of parametric models. The two-sample result serves to test for mean reversion of the diffusion drift in several examples. The interest rates Euribor, Libor as well as T-Bond yields do not show that stylized feature often modelled for interest rates.

1 Introduction

Modelling the development of interest rates is of ongoing interest in finance. Interest rate risk is of great concern in risk management, and derivative products are a common way to hedge risk in the capital market. Rational prices of interest rate derivatives depend upon the underlying model. Triggered by the work of Black and Scholes (1973), stochastic differential equations defining diffusion process models still underpin the valuation of financial derivatives written on many references nowadays. In the seventies lack of computational power enforced a closed form solution. To that end, Black and Scholes proposed the geometric Brownian motion as a model for the underlying.

The volatility of the process is a predominant pricing factor for non-linear products, like options, and much attention has been and still is payed to it. The second risk factor is the diffusion's drift.

*JEL classifications. C32, C52, E43. AMS subject classification 62G08.

Keywords. Nonparametric autoregressive time series, comparision of conditional expectations, wild bootstrap, local linear estimation, interest rate, mean reversion, Ornstein-Uhlenbeck, Cox-Ingersoll-Ross.

On the one hand, drift is sometimes considered negligible as a risk factor, e.g. in the calculation of regulatory capital for market risk according to the Basel regulations [see McNeil, Frey and Embrechts (2005), pg 38] and especially for interest rates the macroeconomic theory does not guarantee a drift as is the case for stock prices. On the other hand, drift is important for pricing and many interest rate models devote effort to it [see e.g. Vasicek (1977), Cox, Ingersoll and Ross (1985), Hull and White (1980), Brace, Gatarek and Musiela (1997)]. The statistical questions that arise are the following: May drift be tested for existence, or for equality to a pre-specified value, or for equality between different interest rates? We will restrict ourselves to the last case because the first two have been discussed extensively in the literature [see e.g. Hjellvik, Yao and Tjøstheim (1998) or Dette and Spreckelsen (2004) and the references in these papers].

Valuation nowadays can be based on Monte Carlo simulation so that closed-from valuation is not mandatory anymore. The specification of the right model for the reference asset has taken its place. Many models can be simulated e.g. with the Euler scheme [see Fan (2005)]. The geometric Brownian motion, as a starting point, models a constant drift. Additionally it models independence of increments on the log-scale, implying that risk, e.g. the Value-at-Risk, increases at a "square-root of time" order with respect to the risk horizon. That may not be crucial in Value-at-Risk calculations of market risk, as those have one and ten days horizons. However, it affects the pricing of long-range interest rate derivatives, like Swaps that can last twenty years. Another area where the long range behavior becomes important is credit risk management: The "credit equivalent" of a derivative product is a long range higher quantile of its mark-to-model distribution. For interest rate derivatives it depends on the respective interest rate model.

As a consequence, in the interest rate market, risk-reducing *mean reversion* of the drift has been introduced, e.g. with homoscedastic error in the Ornstein-Uhlenbeck process [Vasicek (1977)] and heteroscedastic by Cox, Ingersoll and Ross (1985). A generalization is non-linear mean reversion [Aït-Sahalia (1996)]. Mean reversion is not solely a model aspect for interest rates. Vlaar and Palm (1993) model mean reversion of exchange rate and find evidence for their model with a χ^2-goodness-of-fit test. An early economic explanation of potential reversion of prices is arbitrage. However, e.g. for the Standard & Poor's 500 Index basis changes Miller, Muthuswamy and Whaley (1994) argue that corresponding negative autocorrelations are statistical illusion. Bonomo and Garcia (1994) propose to refrain for modelling mean reversion for asset returns, as done before. Our non-parametric test for the curve comparison surprisingly proves usefulness to seek for empirical evidence of mean reversion. We demonstrate the deed in the field of interest rate modelling.

Before that we must be specific about the model we assume. We follow the stationarity model (5) of Fan (2005), where drift (and volatility) are unspecified functions of the underlying and need nonparametric estimation. In contrast to that model with Brownian motion error we account for the well accepted deed that returns have non-gaussian error. Extensions include modelling of jump events with Lévy-processes [Barndorff-Nielson and Shepard (2001)] and the work of Lando (1998) where Cox processes model credit risk. Our bootstrap test does however assume existence of of error-mean and variance. We do not consider the class of time-dependent models where the drift is allowed to depend on time [see e.g. Hull and White (1990) or the popular model by Brace, Gatarek and Musiela (1997)]. Our technical analysis refrains from considering covariates, even though we are aware of long-term predictors like exchange rates having macroeconomic foundation as well as empirical evidence [see e.g. Hoffmann and MacDonald (2006)].

Inference about the drift of time-continuous continuous state processes is in practice based on time series data. The geometric Brownian motion implies a markov process [Karatzas and Shreve (1991), Theorem 5.4.10]. It may be tested whether a diffusion is markovian and e.g. for a financial time-continuous discrete-state process Lando and Skødeberg (2002) finds non-markovian behavior using a Cox regression. As our test is based on a time series model, we are able to relieve the markovian assumption and allow the process to depend on the former realizations up to a specific lag.

Nonlinear autoregressive models (NLAR) are a broad family and nonparametric models [see Jones (1978)] allow for an arbitrary form of the regression function. The problem of nonparametric estimation of the autoregression function has been considered frequently in the statistical literature, see Masry (1996) as a reference. However two sample inference, e.g. to test the equality of two regression functions, has received little attention although linked closely to the framework where observations are independent. Here the literature is rich [see e.g. King, Hart and Wehrly (1991), Delgado and González Manteiga (2001), Kulasekera (1995), Cabus (1998), Munk and Dette (1998), Neumeyer and Dette (2003) among many others]. We adopt a Kolmogorov-Smirnov-type test of Cabus (1998) to autoregressive dependent data with a marked empirical process as test statistic. Along with the test for equality we develop a test for the hypothesis that the regression functions are shifted.

A Monte Carlo simulation finds that size of the new test is sufficiently accurate for many drift shapes. The power is good for several parametric alternative hypotheses, including linear mean reversion. These results prove validity for homoscedastic and heteroscedastic errors and are especially true for the sample sizes of our applications. One example compares two very closely related interest rates, the inter-bank offer rates Libor (in British Pound) and Euribor (in Euro). The second example compares two loosely connected rates, the Euribor and the US treasury bond yield. The example reveals the regression functions of the Libor and Euribor to be equal. As a consequence we do not find mean reversion. The curves for the Euribor differ from the T-bond yield curve, however, not significantly.

2 Nonparametric autoregressive models

Nonparametric autoregressive models were introduced by Jones (1978) and for the particular case of comparing two (or more) financial time series we consider the heteroscedastic nonparametric conditional heteroscedastic autoregressive models

$$X_t = m_1(X_{t-1}, \dots, X_{t-p}) + \sigma_1(X_{t-1}, \dots, X_{t-p})\varepsilon_t; \qquad t = 1, \dots, n_1 \tag{2.1}$$

$$Y_t = m_2(Y_{t-1}, \dots, Y_{t-p}) + \sigma_2(Y_{t-1}, \dots, Y_{t-p})\eta_t; \qquad t = 1, \dots, n_2 \tag{2.2}$$

where $p \in \mathbb{N}$ is fixed and the random variables ε_t and η_t are assumed to be i.i.d. with mean 0 and variance 1. Throughout this paper we assume that $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are independent and strictly stationary processes. The problem of estimating the regression functions m_1 and m_2 (and σ_1^2, σ_2^2) nonparametrically has found considerable interest in the literature [see e.g. Robinson (1983), Tjøstheim (1994), Masry (1996) among many others]. Note that the models (2.1) or (2.2) are only useful when p is small. For moderately large p the nonparametric form is usually difficult to estimate because of the "curse of dimensionality".

The problem of comparing the two regression curves m_1 and m_2 corresponding to nonparametric conditional heteroscedastic autoregressive models has not been studied so far. On the other hand the comparison of curves has been recently investigated in the context of samples with independent observations e.g. by Neumeyer and Dette (2003).
In the present paper we are interested in the performance of the test proposed by Cabus (1998) in the context of nonparametric conditional heteroscedastic autoregressive models. The results of Neumann and Kreiss (1998) indicate that (under a suitable assumption of ergodicity) many of the asymptotic properties for the independent case can be transferred to autoregressive models. Therefore the main object of this paper is the investigation of the finite sample properties of Cabus' (1998) test in the autoregressive setup. For the sake of brevity we concentrate on one testing procedure but it is notable that similar results can be obtained for the procedures proposed by Kulasekera (1995), Munk and Dette (1998), Dette and Neumeyer (2001) and Neumeyer and Dette (2003) for a comparison of two regression curves based on samples with independent observations. We are interested in the hypothesis of equal regression curves

$$H_0^{=} : m_1 = m_2 \quad \text{versus} \quad H_1^{\neq} : m_1 \neq m_2 \tag{2.3}$$

and in the hypothesis of a shift between the two regression curves, that is

$$H_0^c : m_1 = m_2 + c \quad \text{for some} \quad c \in \mathbb{R}; \quad H_1^c : m_1 \neq m_2 + c \ \forall\, c \in \mathbb{R}. \tag{2.4}$$

In the context of two samples with independent observations Cabus (1998) adapted a proposal of Zheng (1996) to the problem of testing the hypothesis (2.3) which can easily be transferred to the problem of comparing curves corresponding to two autoregressive models. Therefore we propose the marked empirical process

$$\begin{aligned} (2.5)\qquad & C_{n_1,n_2}(x_1,\dots,x_p) = \\ & \frac{1}{n_1 n_2 h_1 \dots h_p} \sum_{s=p+1}^{n_1} \sum_{t=p+1}^{n_2} (X_s - Y_t) \prod_{j=1}^{p} K\Big(\frac{X_{s-j} - Y_{t-j}}{h_j}\Big) I\{X_{s-j} \leq x_j\} I\{Y_{t-j} \leq x_j\} \end{aligned}$$

as basic tool for the construction of a test for the hypothesis (2.3). In (2.5) the quantities n_1, n_2 denote the sample size of the two samples in (2.1) and (2.2), $K : \mathbb{R} \to \mathbb{R}$ is a symmetric kernel (integrating to 1) and $h_1, \dots, h_p$ are bandwidths converging to 0 with increasing sample size. The order of bandwidth convergence is tied to the consistency of the estimate.
Note that under appropriate assumptions [see e.g. Fan and Yao (2000)] it follows that

$$\begin{aligned} (2.6)\quad & E\Big[(X_s - Y_t) \prod_{j=1}^{p} \frac{1}{h_j} K\Big(\frac{X_{s-j} - Y_{t-j}}{h_j}\Big) I\{X_{s-j} \leq x_j, Y_{t-j} \leq x_j\}\Big] \\ & \approx \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} \Big\{ m_1(u_1,\dots,u_p) - m_2(u_1,\dots,u_p) \Big\} f_1(u_1,\dots,u_p) f_2(u_1,\dots,u_p) du_1 \dots du_p \end{aligned}$$

where f_1 and f_2 denote the density of $(X_{s-1},\dots,X_{s-p})$ and $(Y_{s-1},\dots,Y_{s-p})$, respectivley (for this motivation we assume that the processes $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are strictly stationary). Obviously, the right hand side of (2.6) vanishes for all $x_1,\dots,x_p \in \mathbb{R}$ if and only if the null hypothesis (2.3) is valid and therefore it is reasonable to reject this hypothesis for large values of the Kolmogorov Smirnov statistic

$$T_{n_1,n_2} = \sup_{x_1,\dots,x_p} |C_{n_1,n_2}(x_1,\dots,x_p)| \tag{2.7}$$

In the following Section 3 we will investigate the finite sample properties of a bootstrap test based on the statistic T_{n_1,n_2}.

The problem of testing the hypothesis (2.4) of a shift between the two regression curves can be reduced to the problem of comparing two regression curves. For this purpose we define $\hat{m}_1$ and $\hat{m}_2$ as the local linear estimate for the regression function based on the sample $X_1,\dots,X_{n_1}$ and $Y_1,\dots,Y_{n_2}$, respectively [see e.g. Wand and Jones (1995)] and note that for any $x_0 = (x_0^{(1)},\dots,x_0^{(p)})^T$ with $f_1(x_0)f_2(x_0) > 0$ it follows approximately

$$\bar{X}_t = X_t - \hat{m}_1(x_0) \approx m_1(X_{t-1},\dots,X_{t-p}) - m_1(x_0) + \sigma_1(X_{t-1},\dots,X_{t-p})\varepsilon_t, \tag{2.8}$$

$$\bar{Y}_t = Y_t - \hat{m}_2(x_0) \approx m_2(Y_{t-1},\dots,X_{t-p}) - m_2(x_0) + \sigma_2(Y_{t-1}.\dots,Y_{t-p})\eta_t. \tag{2.9}$$

Obviously, the hypothesis (2.4) is satisfied if and only if

$$\bar{m}_1(x_1,\dots,x_p) := m_1(x_1,\dots,x_p) - m_1(x_0) = \bar{m}_2(x_1,\dots,x_p) := m_2(x_1,\dots,x_p) - m_2(x_0)$$

and consequently we propose

$$\bar{C}_{n_1,n_2}(x_1,\dots,x_p) = \frac{1}{n_1 n_2 h_1 \dots h_p} \sum_{s=p+1}^{n_1} \sum_{t=p+1}^{n_2} (\bar{X}_s - \bar{Y}_t) \prod_{j=1}^{p} K\Big(\frac{X_{s-j} - Y_{s-j}}{h_j}\Big) I\{X_{s-j} \le x_j\} I\{Y_{s-j} \le y_j\} \tag{2.10}$$

as basic process for testing the hypothesis (2.4) of a shift between the regression functions m_1 and m_2. This hypothesis is rejected for large values of the Kolmogorov-Smirnov statistic

$$\bar{T}_{n_1,n_2} = \sup_{x_1,\dots,x_p} |\bar{C}_{n_1,n_2}(x_1,\dots,x_p)| \tag{2.11}$$

and the finite sample properties of a bootstrap test based on the statistic $\bar{T}_{n_1,n_2}$ for the hypothesis (2.4) will be investigated in Section 4.

3 Comparing curves with a wild bootstrap procedure

In this section we investigate the finite sample properties of the test which rejects the hypothesis of equal regression curves (2.3) for large values of the Kolmogorov Smirnov statistic T_{n_1,n_2} defined in (2.7). The asymptotic properties of a standardized version of the stochastic process $C_{n_1,n_2}(\cdot)$ defined in (2.5) have been studied by Cabus (1998) in the case of two samples with independent observations, who proved weak convergence to a Gaussian process with a covariance kernel depending on certain features of the data generating process. Using similar arguments as given in Hjellvik, Yao and Tjøstheim (1998) and Dette and Spreckelsen (2004) similar results could be derived in the case of strictly stationary processes under appropriate mixing conditions. However,

even in the independent case, the limiting distribution of the normalized Kolmogorov Smirnov statistic is so complicated such that it cannot be directly used for the calculation of critical values [see e.g. Cabus (1998)] and for this reason we propose to use the wild bootstrap procedure for this purpose [see e.g. Wu (1986)].
To be precise let $\hat{m}_g$ denote the local linear estimate from the total sample $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ with bandwidth g, define nonparametric residuals by

$$\hat{\varepsilon}_t = X_t - \hat{m}_g(X_{t-1}, \ldots, X_{t-p}); \quad t = p+1, \ldots, n_1, \tag{3.1}$$

$$\hat{\eta}_t = Y_t - \hat{m}_g(Y_{t-1}, \ldots, Y_{t-p}); \quad t = p+1, \ldots, n_2, \tag{3.2}$$

and bootstrap residuals by

$$\varepsilon_t^* = \hat{\varepsilon}_t V_{1t}; \quad \eta_t^* = \hat{\eta}_t V_{2t} \tag{3.3}$$

where $V_{11}, \ldots, V_{1n_1}, V_{2n_1}, \ldots, V_{2n_2}$ are i.i.d. random variables with masses $(\sqrt{5}+1)/2\sqrt{5}$ and $(\sqrt{5}-1)/2\sqrt{5}$ at the points $(1-\sqrt{5})/2$ and $(1+\sqrt{5})/2$ (note that this distribution satisfies $E[V_{ij}] = 0, E[V_{ij}^2] = E[V_{ij}^3] = 1$)). Following Neumann and Kreiss (1998) we generate bootstrap observations from

$$X_t^* = \hat{m}_g(X_{t-1}, \ldots, X_{t-p}) + \varepsilon_t^*; \quad t = p+1, \ldots, n_1, \tag{3.4}$$

$$Y_t^* = \hat{m}_g(Y_{t-1}, \ldots, Y_{t-p}) + \eta_t^*; \quad t = p+1, \ldots, n_2, \tag{3.5}$$

and calculate the stochastic process

$$C_{n_1,n_2}^*(x_1, \ldots, x_p) = \frac{1}{n_2 n_2 h_1 \ldots h_p} \sum_{s=p+1}^{n_1} \sum_{t=p+1}^{n_2} (X_s^* - Y_t^*) \prod_{j=1}^{p} K\Big(\frac{X_{s-j} - Y_{t-j}}{h_j}\Big) I\{X_{s-j} \le x_j\} I\{Y_{t-j} \le x_j\}$$

to obtain the bootstrap analogue of the Kolmogorov Smirnov statistic T_{n_1,n_2}, i.e.

$$T_{n_1,n_2}^* = \sup_{x_1,\ldots,x_p} |C_{n_1,n_2}^*(x_1, \ldots, x_p)|.$$

For $\alpha \in (0,1)$ let $k_{n_1,n_2,1-\alpha}^*$ denote the $(1-\alpha)$ quantile corresponding to the bootstrap distribution, i.e.

$$\mathbb{P}(T_{n_1,n_2}^* \ge k_{n_1,n_2,1-\alpha}^* \mid \mathcal{Y}_{n_1,n_2}) = \alpha\,, \qquad i = 1, 2.$$

where

$$\mathcal{Y}_{n_1,n_2} := \Big\{X_s, Y_t \mid s = 1, \ldots, n_1, t = 1, \ldots n_2\Big\}$$

denotes the total sample. The null hypothesis (2.3) of equal regression curves is rejected whenever

$$T_{n_1,n_1} > k_{n_1,n_2,1-\alpha}^*. \tag{3.6}$$

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	(i)	0.034	0.061	0.117	0.021	0.053	0.121	0.022	0.050	0.110
	(ii)	0.034	0.058	0.109	0.037	0.061	0.113	0.021	0.048	0.103
	(iii)	0.024	0.056	0.108	0.020	0.047	0.095	0.023	0.047	0.112
50	(i)	0.037	0.062	0.112	0.037	0.061	0.115	0.026	0.046	0.108
	(ii)	0.027	0.048	0.101	0.028	0.057	0.112	0.025	0.048	0.107
	(iii)	0.022	0.050	0.113	0.024	0.048	0.100	0.023	0.046	0.093
100	(i)	0.021	0.048	0.108	0.020	0.055	0.109	0.030	0.048	0.108
	(ii)	0.029	0.056	0.106	0.031	0.057	0.108	0.028	0.051	0.100
	(iii)	0.022	0.050	0.113	0.021	0.045	0.094	0.027	0.050	0.098

Table 3.1. *Simulated rejection probabilities of the bootstrap test (3.6) for the regression models (i) - (iii) (corresponding to the null hypothesis) under homoscedasticity. The errors in the two samples are normally distributed with mean* 0 *and variances* $\sigma_1^2 = \sigma_2^2 = 0.25$.

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	(i)	0.032	0.064	0.113	0.038	0.062	0.121	0.029	0.050	0.104
	(ii)	0.023	0.059	0.109	0.027	0.055	0.108	0.021	0.042	0.096
	(iii)	0.014	0.032	0.084	0.017	0.041	0.089	0.023	0.046	0.104
50	(i)	0.034	0.061	0.120	0.026	0.046	0.107	0.023	0.055	0.112
	(ii)	0.026	0.048	0.105	0.032	0.050	0.109	0.026	0.051	0.103
	(iii)	0.021	0.041	0.089	0.021	0.052	0.102	0.020	0.041	0.089
100	(i)	0.026	0.058	0.109	0.031	0.057	0.110	0.033	0.059	0.108
	(ii)	0.031	0.054	0.106	0.022	0.047	0.099	0.024	0.053	0.098
	(iii)	0.021	0.041	0.092	0.019	0.042	0.093	0.023	0.047	0.096

Table 3.2. *Simulated rejection probabilities of the bootstrap test (3.6) for the regression models (i) - (iii) (corresponding to the null hypothesis) under heteroscedasticity. The errors* ε_t *and* η_t *in model (2.1) and (2.2) are normally distributed and the variance functions are given by* $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$.

In the simulation study we concentrate on the case of a first order autoregression, i.e. $p = 1$ [see also Neumann and Kreiss (1998)]. The bandwidth h_1 was chosen according to the simple rule of thumb

$$h_1 = \left\{ \frac{n_1\hat{\sigma}_2^2 + n_2\hat{\sigma}_1^2}{(n_1 + n_2)^2} \right\}^{1/5}, \tag{3.7}$$

where $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are estimates of the integrated variance $\int_0^1 \sigma_1^2(x) f_1(x)dx$ and $\int_0^1 \sigma_2^2(x) f_2(x)dx$, respectively [see Fan and Yao (2000), p. 375 or Rice (1984)]. The bandwidth for the calculation of the bootstrap residuals was chosen slightly larger, i.e. $g = h_1^{5/6}$.

In all cases we simulated data according to the models (2.1) and (2.2) where the errors are standard normally distributed and different variance functions are considered. $B = 200$ bootstrap replications were used for the calculation of the critical values $k^*_{n_1,n_2,1-\alpha}$ and the rejection probabilities were calculated from 1000 simulation runs. Our first results in Table 3.1 show the simulated rejection probabilities for three models corresponding to the null hypothesis of equal regression curves with homoscedastic errors, that is

$$\begin{array}{ll} \text{(i)} & m_1(x) = m_2(x) = 1 + 0.1x \\ \text{(ii)} & m_1(x) = m_2(x) = e^{-0.1x} \\ \text{(iii)} & m_1(x) = m_2(x) = \sin(2\pi x) \end{array}$$

The variances are given by $\sigma_1^2 = \sigma_2^2 = 0.25$. In all cases we observe a rather accurate approximation of the nominal level, even for very small sample sizes as $n_1 = n_2 = 25$. Other results, which are not displayed here for the sake of brevity show a similar picture and we conclude that the proposed bootstrap test is very reliable with respect to the approximation of the nominal level. In order to investigate how heteroscedasticity affects these results we display in Table 3.2 the rejection probabilities for the same regression functions, where $\sigma_1(x) = 0.1x$ and $\sigma_2(x) = 0.1\sqrt{x}$. Again the level is approximated rather precisely and no substantial differences can be observed between the homoscedastic and heteroscedastic case.
Secondly, we investigate the wild bootstrap test under several alternatives. For this purpose we consider six models

$$\begin{array}{lll} \text{(iv)} & m_1(x) = 1 + 0.1x; & m_2(x) = 1 + 0.9x \\ \text{(v)} & m_1(x) = e^{-0.1x}; & m_2(x) = e^{-0.1x} + x \\ \text{(vi)} & m_1(x) = \sin(2\pi x); & m_2(x) = \sin(2\pi x) + x \\ \text{(vii)} & m_1(x) = 1 + 0.1x; & m_2(x) = 2 + 0.1x \\ \text{(viii)} & m_1(x) = e^{-x}; & m_2(x) = 1 + e^{-x} \\ \text{(ix)} & m_1(x) = \sin(2\pi x); & m_2(x) = \sin(2\pi x) + 1 \end{array}$$

with standard normal distributed errors. In Table 3.3 we show the rejection probabilities of the test in the case of homoscedasticity, that is $\sigma_1 = \sigma_2 = 0.5$, while Table 3.4 corresponds to the case of heteroscedasticity considered in Table 3.2, that is $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$. We observe that in both cases all alternatives are rejected with reasonable probabilities. In the homoscedastic case the results for models (vii) - (ix) indicate that the test is very sensitive with respect to shifts between the two regression functions. In these examples the 5% rejection probabilities are at least 90 % and in nearly all cases larger than 99%. In the examples (v) and (vi) the difference between the regression functions is $m_1(x) - m_2(x) = x$, while it is $0.8x$ in the case (iv). In theses cases the rejection probabilities vary between 65% and 85% depending on the sample sizes. The situation in the heteroscedastic case is again very similar.

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	(iv)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(v)	0.632	0.696	0.809	0.650	0.712	0.836	0.589	0.679	0.811
	(vi)	0.537	0.628	0.744	0.479	0.570	0.703	0.473	0.584	0.715
	(vii)	0.762	0.906	0.982	0.922	0.977	0.994	0.956	0.990	0.998
	(viii)	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(ix)	0.986	0.994	0.998	0.999	1.000	1.000	1.000	1.000	1.000
50	(iv)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(v)	0.666	0.731	0.842	0.644	0.726	0.839	0.656	0.739	0.848
	(vi)	0.552	0.632	0.734	0.574	0.648	0.773	0.589	0.691	0.802
	(vii)	0.906	0.966	0.998	0.996	1.000	1.000	1.000	1.000	1.000
	(viii)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(ix)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	(iv)	1.000	1.000	1.000	1.000	.1.000	.1.000	1.000	1.000	1.000
	(v)	0.652	0.717	0.821	0.670	0.750	0.852	0.686	0.763	0.847
	(vi)	0.564	0.652	0.761	0.578	0.662	0.796	0.664	0.738	0.833
	(vii)	1.000	1.000	1.000	0.997	0.999	1.000	1.000	1.000	1.000
	(viii)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(ix)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.3. *Simulated rejection probabilities of the bootstrap test (3.6) for the regression models (iv) - (ix) (corresponding to the alternative). The errors* ε_t *and* η_t *in model (2.1) and (2.2) are standard normally distributed and the variance functions are given by* $\sigma_1^2(x) = \sigma_2^2(x) = 0.25$ *(corresponding to the homoscedastic case).*

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	(iv)	0.885	0.907	0.922	0.911	0.927	0.958	0.961	0.971	0.982
	(v)	0.472	0.501	0.578	0.489	0.553	0.612	0.517	0.571	0.687
	(vi)	0.502	0.618	0.722	0.573	0.679	0.794	0.587	0.711	0.783
	(vii)	0.439	0.699	0.875	0.792	0.847	0.951	0.840	0.928	0.979
	(viii)	0.801	0.902	0.968	0.990	0.995	1.000	1.000	1.000	1.000
	(ix)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	(iv)	0.901	0.937	0.955	0.931	0.950	0.972	0.972	0.981	0.990
	(v)	0.498	0.521	0.602	0.513	0.554	0.634	0.543	0.597	0.712
	(vi)	0.544	0.631	0.731	0.618	0.697	0.799	0.682	0.771	0.865
	(vii)	0.627	0.808	0.838	0.901	0.961	0.992	0.982	0.996	1.000
	(viii)	0.815	0.929	0.987	0.995	0.998	1.000	1.000	1.000	1.000
	(ix)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	(iv)	0.921	0.961	0.978	0.951	0.971	0.989	0.999	0.999	0.999
	(v)	0.521	0.578	0.683	0.541	0.599	0.702	0.579	0.618	0.753
	(vi)	0.561	0.631	0.740	0.584	0.665	0.786	0.708	0.779	0.858
	(vii)	0.706	0.866	0.961	0.958	0.988	0.995	0.999	1.000	1.000
	(viii)	0.842	0.941	0.992	0.998	1.000	1.000	1.000	1.000	1.000
	(ix)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.4. *Simulated rejection probabilities of the bootstrap test (3.6) for the regression models (iv) - (ix) (corresponding to the alternative). The errors ε_t and η_t in model (2.1) and (2.2) are normally distributed and the variance functions are given by $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$.*

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	[i]	0.027	0.048	0.105	0.029	0.057	0.105	0.023	0.057	0.109
	[ii]	0.022	0.045	0.107	0.020	0.043	0.094	0.021	0.046	0.096
	[iii]	0.040	0.078	0.141	0.041	0.074	0.129	0.037	0.068	0.118
50	[i]	0.022	0.057	0.109	0.019	0.047	0.098	0.020	0.041	0.088
	[ii]	0.021	0.060	0.111	0.026	0.050	0.099	0.021	0.051	0.108
	[iii]	0.042	0.073	0.136	0.033	0.065	0.117	0.023	0.060	0.112
100	[i]	0.025	0.056	0.106	0.025	0.054	0.108	0.028	0.059	0.110
	[ii]	0.024	0.043	0.096	0.025	0.049	0.107	0.022	0.048	0.108
	[iii]	0.041	0.071	0.121	0.034	0.068	0.114	0.028	0.058	0.109

Table 4.1. *Simulated rejection probabilities of the wild bootstrap test for the hypothesis (2.4) in the regression models [i] - [iii] in (4.1) (corresponding to the null hypothesis of a vertical shift). The errors are homoscedastic and normally distributed with mean 0 and variances $\sigma_1^2 = \sigma_2^2 = 0.25$.*

4 Testing for a vertical shift between the regression curves

In the present section we study the finite sample performance of a wild bootstrap test for the hypothesis of a vertical shift (2.4) which is based on the statistic $\bar{T}_{n_1,n_2}$ defined in (2.11). Again we restrict ourselves to the case $p = 1$ and note that the statistic $\bar{T}_{n_1,n_2}$ is based on the "data" $(\bar{X}_t, X_{t-1})$ and $(\bar{Y}_t, Y_{t-1})$ where $\bar{X}_t$ and $\bar{Y}_t$ are given by (2.8) and (2.9), respectively. Note that the definition of this sample requires the specification of a point x_0 such that the regression functions m_1 and m_2 can be estimated at x_0 with $f_1(x_0)f_2(x_0) > 0$, i.e. $x_0 \in supp(f_1) \cap supp(f_2)$. Because in practice it is difficult to identify the support of the densities of the stationary distributions of the time series $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$, we choose

$$x_0 = \frac{n_1 X. + n_2 Y.}{n_1 + n_2}$$

for this purpose, where $X.$ and $Y.$ denote the sample means. We used the re-scaled data (2.8) and (2.9) in the wild bootstrap test described in Section 2 to obtain a test for a vertical shift between the regression curves m_1 and m_2. In Table 4.1 we display the simulated rejection probabilities under the null hypothesis $H_0^c : m_1 = m_2 + c$ for some $c \in \mathbb{R}$. In particular we consider the 3 models

$$\begin{array}{lll} & [i] & m_1(x) = 0.1x; \quad m_2(x) = 1 + 0.1x \\ (4.1) & [ii] & m_1(x) = \exp(-0.1x); \quad m_2(x) = -0.5 + \exp(-0.1x) \\ & [iii] & m_1(x) = \sin(2\pi x); \quad m_2(x) = 1 + \sin(2\pi x) \end{array}$$

where the errors are homoscedastic and normally distributed with variances $\sigma_1^2 = \sigma_2^2 = 0.25$. We observe a very precise approximation of the nominal level in the models [i] and [ii], even if the sample sizes are $n_1 = n_2 = 25$. On the other hand in the periodic model [iii] the approximation of the nominal level is less accurate for smaller sample sizes, but rather reliable in the case $n_1, n_2 \geq 50$.

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	[i]	0.021	0.055	0.114	0.025	0.050	0.101	0.022	0.049	0.109
	[ii]	0.023	0.059	0.112	0.027	0.055	0.108	0.021	0.043	0.094
	[iii]	0.015	0.040	0.083	0.017	0.041	0.086	0.020	0.046	0.106
50	[i]	0.021	0.046	0.105	0.026	0.053	0.104	0.027	0.051	0.112
	[ii]	0.020	0.047	0.100	0.032	0.050	0.111	0.032	0.058	0.108
	[iii]	0.029	0.068	0.115	0.019	0.040	0.090	0.021	0.052	0.109
100	[i]	0.031	0.059	0.110	0.028	0.055	0.107	0.025	0.054	0.112
	[ii]	0.030	0.061	0.112	0.022	0.054	0.096	0.022	0.045	0.093
	[iii]	0.021	0.058	0.113	0.027	0.059	0.121	0.032	0.059	0.109

Table 4.2. *Simulated rejection probabilities of the wild bootstrap test for the hypothesis (2.4) in the regression models [i] - [iii] in (4.1) (corresponding to the null hypothesis of a vertical shift) under heteroscedasticity. The errors ε_t and η_t in model (2.1) and (2.2) are normally distributed and the variance functions are given by $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$.*

In Table 4.2 we show the situation corresponding to the heteroscedastic case, where the variance functions are given by $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$ [see also Table 3.2 for the simulated rejection probabilities of the test for the hypothesis (2.3) in the case of heteroscedastic errors]. Again we do not observe substantial differences in the approximation of the nominal level between the homoscedastic and heteroscedastic case.
In the remaining part of this section we study the behaviour of the wild bootstrap test for the hypothesis (2.4) under some alternatives, that is

$$
\begin{array}{lll}
\text{[iv]} & m_1(x) = 1 + 0.1x; & m_2(x) = 1 + 0.1x + 1/x \\
\text{[v]} & m_1(x) = 1 + 0.1x; & m_2(x) = 0.1x + \sin(2\pi x) \\
\text{[vi]} & m_1(x) = -0.5 + \exp(-0.1x); & m_2(x) = \exp(-0.1x) + 0.6x \\
\text{[vii]} & m_1(x) = \sin(2\pi x); & m_2(x) = 1 + \sin(2\pi x) + \cos(2\pi x) \\
\text{[viii]} & m_1(x) = \sin(2\pi x); & m_2(x) = \sin(2\pi x) + \sin(\pi x) \\
\text{[ix]} & m_1(x) = e^{-0.1x}; & m_2(x) = e^{-0.1x} + \sin(2\pi x)
\end{array}
$$

In Table 4.3 we display the rejection probabilities in the case where the errors are again standard normal distributed with variances $\sigma_1^2 = \sigma_2^2 = 0.25$. We observe that the power is usually lower as in the case of testing for equality and that it depends sensitively on the alternative under consideration. For example it is difficult for the bootstrap test to detect the differences $m_1(x) - m_2(x) = c + 1/x$ or $m_1(x) - m_2(x) = ax + c$ for some $a \in \mathbb{R}$ [see example [iv] and [vi] in Table 4.3], while the test has more power for alternatives of trigonometric type (see example [v], [vii] - [ix]). The situation for the heteroscedastic case is depicted in Table 4.4 and shows a similar picture. The alternative [iv] yields the lowest rejection probabilities, followed by the alternative [vi]. Again differences of trigonometric type can be detected easier (see the examples [v], [vii], [viii] and [ix]).

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	[iv]	0.048	0.092	0.196	0.058	0.133	0.225	0.095	0.168	0.281
	[v]	0.218	0.333	0.442	0.439	0.553	0.664	0.553	0.647	0.724
	[vi]	0.035	0.091	0.182	0.053	0.095	0.198	0.081	0.134	0.243
	[vii]	0.118	0.177	0.270	0.173	0.257	0.346	0.295	0.403	0.520
	[viii]	0.298	0.410	0.520	0.350	0.431	0.520	0.385	0.543	0.628
	[ix]	0.161	0.201	0.313	0.348	0.442	0.567	0.425	0.522	0.630
50	[iv]	0.057	0.121	0.211	0.092	0.151	0.265	0.119	0.215	0.341
	[v]	0.245	0.352	0.458	0.514	0.607	0.694	0.700	0.760	0.818
	[vi]	0.058	0.097	0.188	0.072	0.130	0.228	0.091	0.146	0.277
	[vii]	0.135	0.201	0.293	0.283	0.346	0.449	0.509	0.604	0.686
	[viii]	0.371	0.465	0.556	0.649	0.706	0.774	0.776	0.819	0.851
	[ix]	0.170	0.253	0.359	0.456	0.562	0.664	0.568	0.641	0.717
100	[iv]	0.072	0.151	0.231	0.137	0.225	0.331	0.171	0.287	0.401
	[v]	0.265	0.381	0.479	0.529	0.627	0.719	0.775	0.823	0.874
	[vi]	0.072	0.129	0.221	0.087	0.167	0.278	0.121	0.202	0.341
	[vii]	0.148	0.221	0.317	0.304	0.382	0.497	0.634	0.699	0.752
	[viii]	0.435	0.513	0.583	0.683	0.736	0.794	0.869	0.904	0.927
	[ix]	0.179	0.274	0.386	0.491	0.636	0.712	0.652	0.741	0.823

Table 4.3. *Simulated rejection probabilities of the wild bootstrap test for the hypothesis (2.4) in the regression models [iv] - [ix] (corresponding to the alternative). The errors ε_t and η_t in model (2.1) and (2.2) are standard normally distributed and the variance functions are given by $\sigma_1^2(x) = \sigma_2^2(x) = 0.25$ (corresponding to the homoscedastic case).*

n_1	n_2	25			50			100		
	α	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
25	[iv]	0.070	0.148	0.267	0.108	0.178	0.276	0.123	0.198	0.284
	[v]	0.170	0.278	0.408	0.370	0.481	0.608	0.359	0.570	0.688
	[vi]	0.080	0.156	0.272	0.089	0.167	0.289	0.118	0.209	0.361
	[vii]	0.114	0.176	0.269	0.211	0.299	0.402	0.198	0.321	0.471
	[viii]	0.187	0.256	0.348	0.382	0.447	0.519	0.571	0.763	0.801
	[ix]	0.232	0.333	0.457	0.429	0.335	0.636	0.481	0.591	0.687
50	[iv]	0.079	0.171	0.293	0.129	0.212	0.332	0.167	0.246	0.376
	[v]	0.232	0.319	0.444	0.469	0.578	0.672	0.625	0.712	0.805
	[vi]	0.107	0.182	0.310	0.146	0.238	0.384	0.228	0.350	0.502
	[vii]	0.163	0.221	0.299	0.369	0.438	0.519	0.625	0.792	0.748
	[viii]	0.219	0.295	0.379	0.413	0.489	0.562	0.773	0.823	0.868
	[ix]	0.275	0.373	0.499	0.499	0.588	0.674	0.656	0.743	0.826
100	[iv]	0.085	0.199	0.317	0.148	0.241	0.352	0.198	0.304	0.432
	[v]	0.281	0.415	0.507	0.551	0.655	0.756	0.741	0.815	0.865
	[vi]	0.178	0.213	0.357	0.188	0.288	0.425	0.274	0.403	0.558
	[vii]	0.198	0.257	0.381	0.406	0.464	0.527	0.741	0.782	0.825
	[viii]	0.244	0.321	0.413	0.459	0.528	0.592	0.902	0.920	0.944
	[ix]	0.294	0.421	0.587	0.557	0.648	0.735	0.739	0.808	0.857

Table 4.4. *Simulated rejection probabilities of the wild bootstrap test for the hypothesis (2.4) in the regression models [iv] - [ix] (corresponding to the alternative) under heteroscedasticity. The errors ε_t and η_t in model (2.1) and (2.2) are normally distributed and the variance functions are given by $\sigma_1(x) = 0.1x, \sigma_2(x) = 0.1\sqrt{x}$.*

5 Data example

In the present section we illustrate the application of the new tests in an analysis of interest rates. Important interest rates are 3-month (London) inter-bank offer rates like the Libor (for British Pound quoted debt) and the Euribor (for Euro quoted debt). Figure 5.1 depicts daily quotes for the period of 12/30/1998 till 10/24/2005. Note that the curves show the interest rates as a function of time, not as a function of tenor.

Inter-bank rates are the reference for many interest rate derivatives. Especially the Swap market has become very liquid due to its standardized documentation [see International Swaps and Derivatives Association Inc., New York (2002)]. Swap prices are even used for policy making of central banks. The European Central Bank bases macroeconomic analysis of market expectations on the Euribor-Swap rates [Durré (2006)].

It is clearly interesting to assess whether Euribor and Libor behave similar in terms of their expected development over successive periods, say days. Financial derivatives are to a large extent used to exploit arbitrage opportunities, i.e. to earn money risk-free. The idea of a technical time

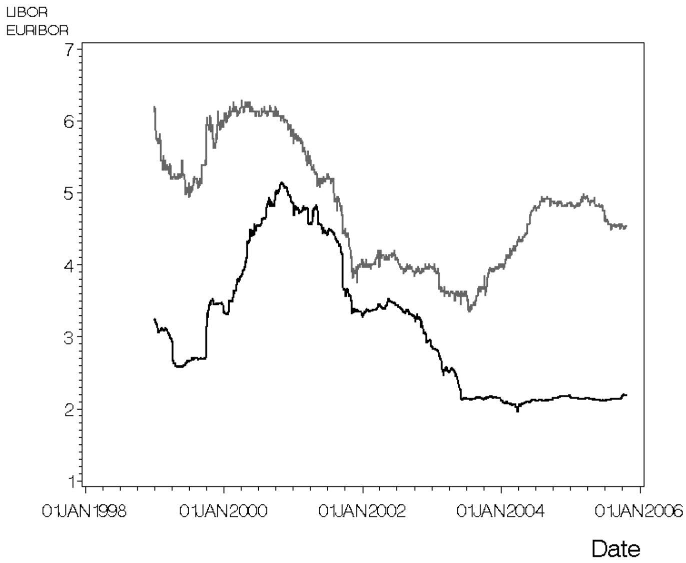

Figure 5.1: *Daily 3-month Euribor (black line) and 3-month Libor (grey line) for 12/30/1998 till 10/24/2005.*

series analysis is now the following: Suppose an arbitrageur knows tomorrow's expected price of a financial product to be different from today's. If tomorrow's price is larger, she buys the product now and sells it tomorrow. If the price is smaller she sells the product today - so-called short-selling - and buys it tomorrow. It appears to be a one-sample problem, however, buying essentially means exchanging cash with the product, two goods are involved. In general, arbitrage is possible using several goods [see e.g. Neftci (2000)].
Hedging with cash is rare, we consider the case of two similar products and their prices. Exchanging interest rates in different currency is easy: One can "swap" the one into the other with a cross-currency Swap [see e.g. Miron and Swannell (1992)]. Using interest rates, i.e. prices for debt, has serval points. On the one hand, we have long times series (around 1,800 days), so that our findings for the size and power of the tests for moderate sample sizes (25 to 100 observations) suggest that virtually any deviation from the null hypothesis will be detected. We may consider the test as equivalence test [see e.g. Munk and Weißbach (1999)] and prove the null hypothesis. On the other hand, the highly competitive and liquid market of interest rate derivatives suggests arbitrage opportunities are already continuously exploited. This means for $p = 1$, the regression curves should both be equal to bisecting lines, i.e.

$$m_1(x) = m_2(x) = x \ . \tag{5.1}$$

Libor and Euribor should be martingales. We will see which effect dominates.
Being more specific we could assume a certain model, e.g. a diffusion model like the Ornstein-Uhlenbeck process (or Vasicek model [see Vasicek (1977)]). For the Euribor a stochastic differential

equation (SDE) characterizes the behavior:

$$dX_t = \kappa_1(\alpha_1 - X_t)dt + \sigma_1 dW_t^1. \tag{5.2}$$

Here, W^1 denotes a standard Brownian motion. The component α_1 in the drift is the long-run mean, the "gravity" center of the mean reversion. Drift in the SDE is linked to the regression function in definition (2.1). For a one day step, i.e. $dt = 1$, we have

$$E(X_t|X_{t-1}) = X_{t-1} + E(dX_t|X_{t-1}) = X_{t-1} + \kappa_1(\alpha_1 - X_{t-1}).$$

The regression curve writes as $m_1(x) = a_1 x + b_1$ with $a_1 = 1 - \kappa_1$ and $b_1 = \kappa_1\alpha_1$. A similar time series model holds for the Libor as an example for Y_t in (2.1) with σ_2, W^2, κ_2, α_2, a_2 and b_2. Figure 5.1 shows that the Libor is larger than the Euribor, so that clearly we can assume $\alpha_2 > \alpha_1$. The regression curves $m_1(x)$ and $m_2(x)$ are now equal (and equal to the bisecting line) if and only if the reversion "speed" vanishes, i.e. if $\kappa_1 = \kappa_2 = 0$. Under the parametric assumption of an Ornstein-Uhlenbeck process our test for equality of regressors is a globally consistent test for the hypothesis

$$H_0: \ \kappa_1 = \kappa_2 = 0.$$

Cox, Ingersoll and Ross (1985) assume a heteroscedastic extension of (5.2). Instead of the a constant volatility σ_1, a time-homogeneous factor $\sigma_1\sqrt{X_t}$ models the stylized fact of higher volatility during high interest rate regimes. Our formulation (2.1) accommodates for arbitrary time-homogeneous heteroscedasticity. Again in the CIR model, testing for equal regression curves corresponds to testing for $\kappa_1 = \kappa_2 = 0$. Note that the linear regression function with one regressor is an example studied in the simulation study of Sections 3.
In Figure 5.2 we display the regression $E[X_t \mid X_{t-1}] = m_1(X_{t-1})$ and $E[Y_t \mid Y_{t-1}] = m_2(Y_{t-1})$, based on a local linear fit, where the bandwidth was chosen by least squares cross validation [see Fan and Yao (2000)]. Note that we have assumed that the two curves can be represented by the models (2.1) and (2.2) with $p = 1$. We observe that both regression functions are close to the bisecting line. We adjust the descriptive result for noise by testing the (first) equality in (5.1) with our test for the hypothesis of equal regression curves in (2.3). The bootstrap procedure yields the p-value 0.822. Given the sample size, this gives evidence for the equality of the two regression curves. Comparing Euribor and Libor gives strong evidence for both rates to be free of mean reversion.
We have seen that beyond their different location - rates in British pound are uniformly larger than rates in Euro - the two rates are very similar. What if we compare one of them to an interest rate that has a different tenor (10 years instead of 3 months), another currency (US dollar instead of Euro), a different originator (the United States instead of large banks) and a different data recording frequency (monthly instead of daily).
For this purpose we consider a comparison of the 3-month Euribor (with monthly geometric averages) with the yields of the 10-year treasury bill, the T-bond. We have collected monthly averages of the yields in 10-year US T-bonds (quoted in USD). The data comprises the 74 observations between January 1999 and February 2005. Figure 5.3 on the left displays the rates, again stationarity is plausible. The difference between the two curves is much more pronounced than between

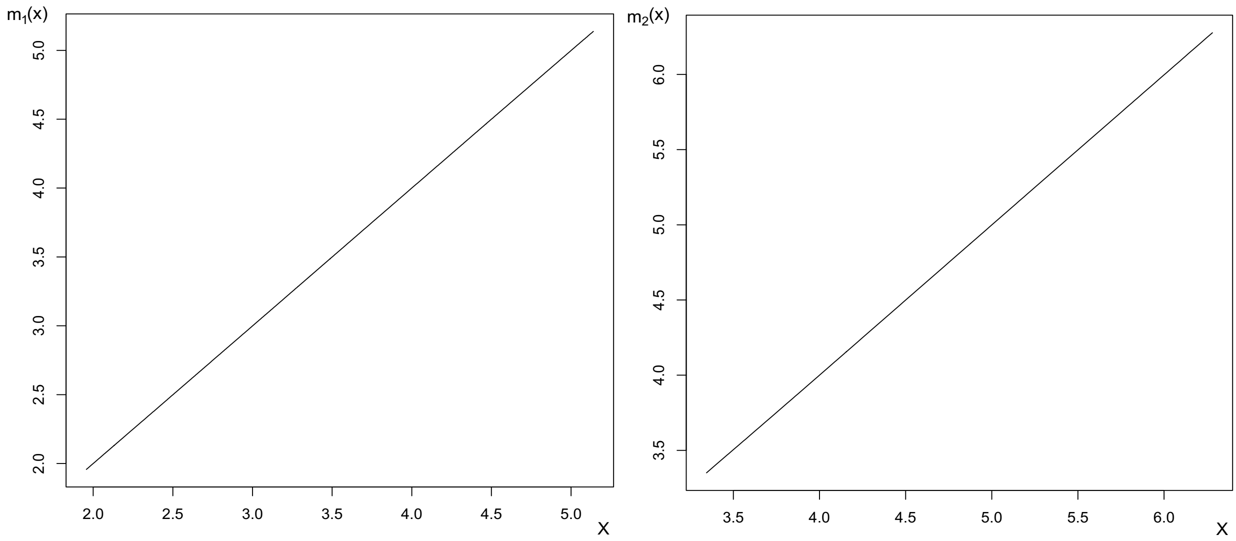

Figure 5.2: *The local linear estimates of the regression functions $m_1(x) = E[X_t|X_{t-1}]$ for 3-month Euribor (left) and $m_2(x) = E[Y_t|Y_{t-1}]$ for 3-month Libor (right).*

Euribor and Libor. Figure 5.3 on the right displays again the local linear estimates for the regression functions. The slopes of the two curves do not appear to be totally equal. However, the result of the test for equality is a p-value of 0.628. (The test for a shift yields a larger p-value.) Considering the moderate number of observations, the large p-value can not be interpreted as lack of arbitrage opportunities, but rather that large deviations from the null hypothesis do not become evident. Substantial arbitrage is not possible.

6 Conclusion

In the construction of tests to compare autoregression functions we concentrate on broad assumptions. The shape of regression function, the volatility function and the dependence structure are to a large extent unspecified. Motivation for our tests is the frequent occurrence of time series data in econometrics and especially in finance. Marked empirical processes are typical building blocks in time series analysis and guided by the quest for globally consistent tests we decide to use a Kolmogorov-Smirnov-tpye statistic. Given the ambitious practical target we sacrifice the tractability of asymptotic theory and bootstrap the null distribution with the additional gain of good results in size and power. Aspects of bandwidth selection are kept scarce because we experienced those to have little impact. Application to interest rate models demonstrates that the test can serve to detect e.g. mean reversion, a stylized feature we have not been able to prove for the three sample time series we analyzed.

Acknowledgements. The authors are grateful to Melanie Birke for computational assistance and to Isolde Gottschlich who typed numerous versions of this paper with considerable technical

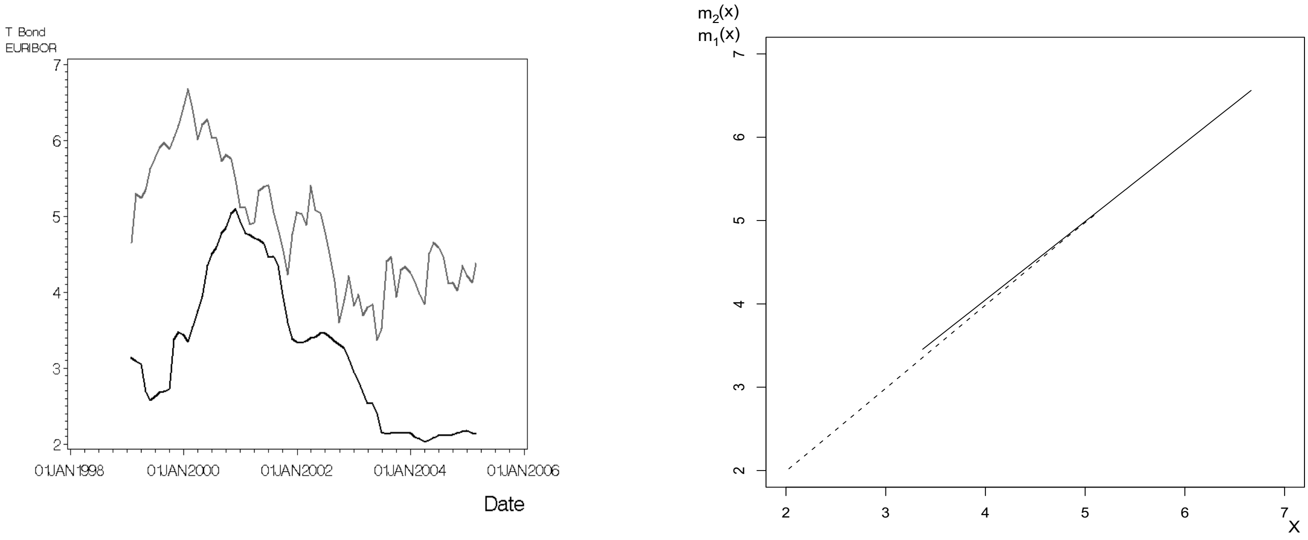

Figure 5.3: *Monthly 3-month Euribor (black line) and 10-year US treasury bond yield (grey line) for 01/29/1999 till 02/28/2005 (left) and regression functions* $m_1(x) = E[X_t|X_{t-1}]$ *for 3-month Euribor (dashed line) and* $m_2(x) = E[Y_t|Y_{t-1}]$ *for 10-year US-treasury debt (solid line) (right).*

expertise. The work of the authors was supported by the Deutsche Forschungsgemeinschaft (SFB 475, ‘Reduction of complexity in multivariate data structures’).

References

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

O.E. Barndorff-Nielson, N. Shepard (2001). Non-gaussian Ornstein-Uhlenbeck models and some of their uses in financial economics (with discussion). J. R. Stat. Soc. Ser. B Stat. Methodol. 63, 167-241.

F. Black, M. Scholes (1973). The pricing of options and corporate liabilities. J. Political Econom. 81, 637-645.

M. Bonomo, R. Garcia (1994). Can a well-fitted equilibrium asset-pricing model produce mean reversion?, J. Appl. Econom. 9, 19-29.

A. Brace, D. Gatarek, M. Musiela (1997). The market model of interest rate dynamics. Math. Financ. 7, 27-155.

P. Cabus (1998). Un test de type Kolmogorov-Smirnov dans le cadre de comparaison de fonctions de regression. (A Kolmogorov-Smirnov test for checking the equality of two regression curves). (French) C. R. Acad. Sci., Paris, Sr. I, Math. 327, 939-942 .

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

M.A. Delgado, W. González Manteiga (2001). Significance testing in nonparametric regression based on the bootstrap. Ann. Statist. 29, 1469-1507.

H. Dette, N. Neumeyer (2001). Nonparametric analysis of covariance. Ann. Statist. 29, 1361-1400.

H. Dette, I. Spreckelsen (2004). Some comments on specification tests in nonparametric absolutely regular processes. J. Time Ser. Anal. 25, 159-172.

A. Durré (2006). The liquidity premium in the money market: A comparison of the german mark period and the euro area. Germ. Econ. R. 7(2), 163–187.

J. Fan, Q. Yao (2000). Nonlinear time series. Nonparametric and parametric methods. Springer Series in Statistics. New York

J. Fan (2005). A selective overview of nonparametric methods in financial econometrics. Stat. Sci. 20, 317-337.

Hjellvik, V., Yao, Q., Tjøstheim, D. (1998). Linearity testing using local polynomial approximation. J. Stat. Plann. Inference 68, No.2, 295-321.

M. Hoffmann, R. MacDonald (2006). A re-examination of the link between real exchange rates and real interest rate differentials. CESifo Working Paper Series, No. 894.

J. Hull, A. White (1990). Pricing interest-rate-derivative securities. Rev. Financ. Stud. 18, 37-84.

International Swaps and Derivatives Association Inc., New York (2002). 2002 ISDA Master Agreement, Technical report.

D.A. Jones (1978). Nonlinear regressive processes. Proc. Roy. Soc. London, Ser. A, 360, 71-95.

E.C. King, J.D. Hart, T.E. Wehrly (1991). Testing the equality of regression curves using linear smoothers. Statist. Probab. Lett. 12, 239-247.

I. Karatzas, S.S. Shreve (1991). Brownian motion and stochastic calculus, 2nd ed. Springer, New York.

K.B. Kulasekera (1995). Comparison of regression curves using quasi residuals. J. Amer. Stat. Assoc. 90, 1085-1093.

D. Lando and T. Skødeberg (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations. J. Bank. Financ. 26, 423-444.

E. Masry (1996). Multivariate local polynomial regression for time series: uniform strong consistency and rates. J. Time Ser. Anal. 17, 571-599.

A.J. McNeil, R. Frey, P. Embrechts (2005). Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, Princeton.

M.H. Miller, J. Muthuswamy, R.E. Whaley (1994). Mean Reversion of Standard & Poor's 500 Index Basis Changes: Arbitrage-induced of Statistical Illusion?, J. Financ., 48, 479-513.

P. Miron, P. Swannell (1992). Pricing and Hedging Swaps. Euromoney Books, New York.

A. Munk, H. Dette (1998). Nonparametric comparison of several regression functions: exact and asymptotic theory. Ann. Statist. 26, 2339-2368.

A. Munk, R. Weißbach. $1-\alpha$ Equivariant Confidence Rules for Convex Alternatives are $\frac{\alpha}{2}$-Level Tests. J. Amer. Stat. Assoc. 94, 1311-1319.

S.N. Neftci (2000). An Introduction to the Mathematics of Financial Derivatives, Academic Press, London.

M. Neumann, J.P. Kreiss (1998). Regression-type inference in nonparametric autoregression. Ann. Statist. 26, 1570-1613.

N. Neumeyer, H. Dette (2003). Nonparametric comparison of regression curves: an empirical process approach. Ann. Statist. 31, 880 - 920.

J. Rice (1984). Bandwidth choice for nonparametric regression. Ann. Statist. 12, 1215-1230.

P.M. Robinson (1983). Nonparametric estimators for time series. J. Time Sr. Anal. 4, 185-207.

D. Tjøstheim (1994). Non-linear time series: a selective review. Scand. J. Statist. 21, 97-130.

O.A. Vasicek (1977). An equilibrium characterization of the term structure. J. Financ. Econ. 5, 177-188.

P.J.G. Vlaar, F.C. Palm (1993). The message in weekly exchange rates in the European Monetary System: mean reversion, conditional heteroscedasticity, and jumps, J. Bus. Econ. Stat. 11, 351-360.

M.P. Wand, M.C. Jones (1995). Kernel Smoothing. Chapman and Hall, London.

C.F.Y. Wu (1986). Jacknife, bootstrap and other resampling methods in regression analysis. Ann. Statist. 14, 1261-1350.