APPLICATIONES MATHEMATICAE
* (200*), pp. 1–24

D. Feldmann (Berlin)
W. Härdle (Berlin)
C. Hafner (Berlin)
M. Hoffmann (Paris)
O. Lepski (Marseille)
A. Tsybakov (Paris)

TESTING LINEARITY IN AN AR ERRORS-IN-VARIABLES MODEL WITH APPLICATION TO STOCHASTIC VOLATILITY

Abstract. Stochastic Volatility (SV) models are widely used in financial applications. To decide whether standard parametric restrictions are justified for a given data set, a statistical test is required. In this paper, we develop such a test of a linear hypothesis versus a general composite nonparametric alternative using the state space representation of the SV model as an errors-in-variables AR(1) model. The power of the test is analyzed. We provide a simulation study and apply the test to the HFDF96 data set. Our results confirm a linear AR(1) structure in log-volatility for the analyzed stock indices S&P500, Dow Jones Industrial Average and for the exchange rate DEM/USD.

1. Introduction. A good knowledge of path-dependent volatility structures is important for the analysis of high frequency data in finance (HFDF). Such knowledge enables multi-step forecasts of volatility, which can be used for derivative pricing, evaluation of risk exposure and prediction intervals for the mean. Potential applications of this knowledge are tests of economic or financial theories concerning the stock, bond and currency markets or

2000 *Mathematics Subject Classification*: 62G07, 62G10, 62M10.

Key words and phrases: autoregression with errors in variables, stochastic volatility, testing parametric versus nonparametric fit, minimax tests.

This research was financed by contributions from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 373 "Quantification and Simulation of Economic Processes".

studies of the link between short and long term interest rates. Another important set of applications concerns interventions on the markets based on portfolio choice, hedging portfolios, values at risk, the size and times of block trading.

Typically, the conditional volatility exhibits a strong dependence on past values of the observed process. In this context, autoregressive conditional heteroskedasticity (ARCH) models (Engle, 1982; Gouriéroux, 1997) and stochastic volatility (SV) models (Taylor, 1986) have been studied intensively (see Harvey, Ruiz and Shephard (1994), Shephard (1996)). Duan (1995) used an ARCH model for option pricing under time-varying volatility. Volatility models have consequences for the stationary distribution of the process, and thus, influence the calculation of tail indices and value at risk; see e.g. de Haan (1990) and de Vries (1994).

Starting with Taylor (1986), SV models are mostly specified as parametric AR(1)-type models. The question arises whether the parametric structure adequately describes the data. A similar question of appropriateness of simple parametric description was posed e.g. by Gouriéroux and Monfort (1992), Härdle and Tsybakov (1997), and Hafner (1998) in the context of ARCH models. In Härdle and Tsybakov (1997) and Härdle, Tsybakov and Yang (1998) nonparametric counterparts of ARCH, the CHARN (conditionally heteroskedastic autoregressive nonlinear) models are considered. Stylized facts of HFDF show that GARCH volatility models do not sufficiently capture the structure of HFDF. Thus, it is interesting to test parametric hypothesis versus nonparametric alternative in various volatility models (goodness-of-fit testing). Also, testing of purely nonparametric hypotheses (for example, the symmetry hypothesis for the volatility function) seems to be of interest. Such tests for nonparametric structures were recently developed by Leblanc and Lepski (1996) and by Gouriéroux, Monfort and Tenreiro (1995) in the time series context.

In this paper we consider nonparametric goodness-of-fit testing in the case of SV models. The discrete time SV model can be represented as an errors-in-variables autoregressive (AR) model. We propose the test which allows one to distinguish the linear parametric AR hypothesis from the set of nonparametric AR alternatives, and we analyze the power of the test. Next, we investigate its finite sample behavior by a simulation study. Finally, we apply it to HFDF96 data sets: the S&P500 and the Dow Jones stock price indices, as well as the DEM/USD exchange rate. Our findings support the hypothesis of a parametric volatility structure for all analyzed data sets.

2. State space representation of the SV model. Let S_t denote the underlying asset price at time t, $t = 1, \ldots, n$, and define returns h_t

as $h_t = \log(S_t/S_{t-1})$. The standard SV model as in Taylor (1986) can be written as

$$h_t = \exp(Y_t/2)\xi_t^*, \quad Y_t = \vartheta Y_{t-1} + \varepsilon_t,$$

where ξ_t^* and ε_t are i.i.d. random variables with $\mathbf{E}\xi_t^* = 0$ and $\mathbf{E}\varepsilon_t = 0$, where $\mathbf{E}$ denotes expectation. A discussion of this and related models is given by Harvey, Ruiz and Shephard (1994) and Shephard (1996). Let $\xi_t = 2\log(|\xi_t^*|) - \omega$ with $\omega = 2\mathbf{E}[\log(|\xi_t^*|)]$ and $Z_t = \log h_t^2$. Then we obtain the following linear state space model for the observables $Z_1, \ldots, Z_n$:

$$Z_t = \omega + Y_t + \xi_t, \tag{1}$$

$$Y_t = \vartheta Y_{t-1} + \varepsilon_t, \tag{2}$$

where $\mathbf{E}\xi_t = 0$.

We can write (2) as

$$Y_t = m(Y_{t-1}) + \varepsilon_t,$$

where $m(\cdot)$ is an unknown function. In general, there is no prior reason to assume that $m(\cdot)$ is a linear function. The shape of this function determines the type of impact of volatility on financial decision variables.

The aim of this paper is to propose a test of the composite hypothesis that the function $m(\cdot)$ is linear against a composite nonparametric alternative of rather general structure. In particular, no smoothness assumptions on $m(\cdot)$ are imposed under the alternative.

3. Main results. Let $Z_1, \ldots, Z_n$ be the observations obtained in the following model:

$$Z_t = Y_t + \xi_t, \tag{3}$$

$$Y_t = m(Y_{t-1}) + \varepsilon_t, \quad t = 1, \ldots, n, \tag{4}$$

where $\{\xi_t\}$ and $\{\varepsilon_t\}$ are i.i.d. zero mean random variables and $m(\cdot)$ is an unknown function. The values $\{Y_t\}$ are not observed.

The model (3) is simpler than (1), since here we put $\omega = 0$. This can be done without loss of generality if ω is known. An extension to the case of unknown ω appears in Section 4.

Equations (3)–(4) can be viewed as a nonparametric AR errors-in-variables model. In fact, by (3)–(4) the observations Z_t satisfy

$$Z_t = m(Z_{t-1} - \xi_{t-1}) + \varepsilon_t + \xi_t.$$

Our goal is to test the hypothesis that the function $m(\cdot)$ is linear, i.e.

$$\mathbf{H}_0: \quad m(x) = \vartheta x, \quad \vartheta \in [a, b], \tag{5}$$

where $0 \le a < b < 1$ are some known constants.

The problem of testing a linear hypothesis against a nonparametric alternative for the regression model was considered e.g. by Härdle and Mammen (1993), Härdle and Kneip (1999) and Spokoiny (1997). The case of single index regression was studied by Härdle, Sperlich and Spokoiny (1997). Here we consider an autoregressive model and we introduce a new way to describe the alternative adapted to settings with dependent data.

Assume the following.

A1. The sequences $\{\xi_t\}$ and $\{\varepsilon_t\}$ consist of i.i.d. random variables and these sequences are mutually independent. The value $Y_0 = y_0$ is fixed.

A2. $\mathbf{E}\varepsilon_1 = \mathbf{E}\xi_1 = 0$;
$\mathbf{E}\xi_1^2 = \eta^2$, $\mathbf{E}\xi_1^3 = 0$ and $\mathbf{E}\xi_1^4 = \mu$;
$\mathbf{E}\varepsilon_1^2 = \sigma^2$, $\mathbf{E}\varepsilon_1^3 = 0$ and $\mathbf{E}\varepsilon_1^4 = \nu$;
$\mathbf{E}|\xi_1|^{4+\delta} < \infty$, $\mathbf{E}|\varepsilon_1|^{4+\delta} < \infty$ with some $\delta > 0$.

A3. The random variable ε_1 has a density $p(\cdot)$ with respect to the Lebesgue measure, satisfying $p(x) > 0$ for all $x \in \mathbb{R}^1$.

Let us now introduce the test statistic T_n and the decision rule Δ_n. Set $N = [n/2]$, where $[\cdot]$ denotes the integer part. Suppose without loss of generality that $N \geq 4$. Consider the pilot statistic

$$\overline{\vartheta}_n = \frac{\sum_{t=3}^{N-1} Z_t Z_{t-2}}{\sum_{t=3}^{N-1} Z_{t-1} Z_{t-2}}, \tag{6}$$

which is the estimator of ϑ under the null hypothesis obtained by the instrumental variables method. It is easy to see that $\overline{\vartheta}_n$ is $\sqrt{n}$-consistent under $\mathbf{H}_0$.

Denote by $\widehat{\vartheta}_n$ the projection of $\overline{\vartheta}_n$ onto $[a, b]$,

$$\widehat{\vartheta}_n = \begin{cases} \overline{\vartheta}_n & \text{if } a \leq \overline{\vartheta}_n \leq b, \\ a & \text{if } \overline{\vartheta}_n < a, \\ b & \text{if } \overline{\vartheta}_n > b, \end{cases}$$

and let

$$M_n = M(\widehat{\vartheta}_n), \qquad B_n = B(\widehat{\vartheta}_n),$$

where

$$M(\vartheta) = \sigma^2 + \eta^2(1 + \vartheta^2),$$
$$B(\vartheta) = (\nu - \sigma^4) + \mu(1 + \vartheta^2)^2 - \eta^4(1 - \vartheta^2)^2 + 4\sigma^2\eta^2(1 + \vartheta^2).$$

Define the test statistic:

$$T_n = \frac{1}{\sqrt{N B_n}} \sum_{t=N+1}^{n} \{(Z_t - \widehat{\vartheta}_n Z_{t-1})^2 - M_n\}. \tag{7}$$

Fix some $0 < \alpha < 1$ and set

$$\Delta_n = \begin{cases} 0 & \text{if } T_n \leq t_\alpha, \\ 1 & \text{if } T_n > t_\alpha, \end{cases}$$

where t_α is the $(1 - \alpha)$-quantile of standard normal distribution.

The test accepts the hypothesis $\mathbf{H}_0$ if $\Delta_n = 0$ and rejects $\mathbf{H}_0$ if $\Delta_n = 1$.

Note that in the definition of the test we split the sample $Z_1, \ldots, Z_n$ into two parts. The first part $Z_1, \ldots, Z_{N-1}$ is used to find the preliminary estimator (6) and the second part $Z_N, \ldots, Z_n$ appears only in (7). This is done to make the proofs less technical. We believe that the result can be extended to the case where the entire sample is used both in (6) and (7). In simulations and in the real data example below we do not apply the splitting.

Let $\mathbf{P}_\vartheta$ be the probability measure generated by $(Z_1, \ldots, Z_n)$ satisfying (3)–(4) when the underlying $m(\cdot)$ has the form $m(x) = \vartheta x$.

THEOREM 1. *Assume* **A1** *and* **A2**. *Then*

$$\limsup_{n\to\infty} \sup_{\vartheta\in[a,b]} \mathbf{P}_\vartheta\{\Delta_n = 1\} \le \alpha.$$

Thus, the test based on the decision rule Δ_n is asymptotically of level α.

Now consider the power this test. We introduce a nonparametric set of alternatives and show that the probability to accept the hypothesis $\mathbf{H}_0$ for the case where the function $m(\cdot)$ belongs to this set (i.e. the second type error probability) is less than a given value β.

Let us define the set of alternatives. First, assume that the alternatives $m(\cdot)$ are such that Y_t does not explode as $t \to \infty$. This is guaranteed by the condition $m(\cdot) \in \mathcal{M}$, where $\mathcal{M} = \mathcal{M}(c, d)$ is the set of functions $m(\cdot)$ satisfying

$$|m(x)| \le c|x| + d, \qquad \forall x \in \mathbb{R}^1,$$

for some $c \in (0, 1),\ d > 0$.

Next we assume that the alternatives $m(\cdot)$ are bounded away from the set of linear functions at a certain distance. It would be natural to characterize the distance between a function $m(\cdot)$ and the hypothesis set (the set of linear functions) in the form

$$\zeta_n = \inf_{\vartheta\in[a,b]} \frac{1}{N} \sum_{t=N+1}^{n} (m(Y_{t-1}) - \vartheta Y_{t-1})^2. \tag{8}$$

However, this distance is random, which does not allow one to describe the set of alternatives in a relevant way. To avoid this inconvenience we replace ζ_n by its nonrandom analog:

$$d_n(m) = d_n(m(\cdot)) = \inf_{\vartheta\in[a,b]} \frac{1}{N} \sum_{t=N+1}^{n} \mathbf{E}_m (m(Y_{t-1}) - \vartheta Y_{t-1})^2, \tag{9}$$

where $\mathbf{E}_m$ is the expectation w.r.t. the probability measure $\mathbf{P}_m$ generated by the observations $(Z_1, \ldots, Z_n)$ satisfying (3)–(4) when the underlying autoregression function is $m(\cdot)$. The asymptotic equivalence of these two distance measures is justified by the following proposition.

PROPOSITION 1. *Assume* **A1**–**A3**. *Then*

$$\lim_{C\to\infty} \limsup_{n\to\infty} \sup_{m\in\mathcal{M}} \mathbf{P}_m\{\sqrt{n}\,|\zeta_n - d_n(m)| \geq C\} = 0. \tag{10}$$

Proposition 1 is a consequence of Lemma 2 proved below.

Now we complete the definition of the set of alternatives. For any $\lambda > 0$ and any $n \geq 1$ define

$$\mathcal{M}_n(\lambda) = \{m \in \mathcal{M} : d_n(m) \geq \lambda/\sqrt{n}\,\}.$$

Consider the set of alternatives

$$\mathbf{H}_n : \quad m \in \mathcal{M}_n(\lambda).$$

THEOREM 2. *Assume* **A1**–**A3**. *Then for any* $0 < \alpha, \beta < 1$ *there exists a constant* $\lambda(\alpha,\beta)$ *such that*

$$\limsup_{n\to\infty} \sup_{m\in\mathcal{M}_n(\lambda(\alpha,\beta))} \mathbf{P}_m\{\varDelta_n = 0\} \leq \beta.$$

This theorem shows that for $\lambda > 0$ large enough the proposed test attains the asymptotical power that is arbitrarily close to 1 uniformly on the set of nonparametric alternatives $\mathcal{M}_n(\lambda)$.

4. The case of unknown ω**.** Now we turn to the situation where the constant shift ω is not known. Note that we can rewrite the model (2) as

$$\begin{aligned} Z_t &= \omega(1-\vartheta) + \vartheta Z_{t-1} + \varepsilon_t + \xi_t - \vartheta\xi_{t-1} \\ &= \gamma + \vartheta Z_{t-1} + \nu_t(\vartheta), \end{aligned} \tag{11}$$

where $\gamma = \omega(1-\vartheta)$, and $\nu_t(\vartheta) = \varepsilon_t + \xi_t - \vartheta\xi_{t-1}$ with $\mathbf{E}\nu_t(\vartheta) = 0$. Thus, it is easy to see that the sample mean

$$\widehat{\omega} = \frac{1}{n}\sum_{t=1}^{n} Z_t$$

is a $\sqrt{n}$-consistent estimator for ω. However, in what follows we find it more convenient to work with estimators of γ rather than those of ω.

We define an iterative procedure to obtain estimates of γ and ϑ. This procedure will be used for the HFDF96 data set in Section 7. Here and in the numerical results below we do not apply the sample splitting that was necessary for the theory. Both the pilot statistic and the test statistic are computed from the entire sample $Z_1, \ldots, Z_n$.

Consider the centered observations $Z_t^* = Z_t - \widehat{\omega}$ and define the preliminary estimates for ϑ and γ:

$$\vartheta_n^{(1)} = \frac{\sum_{t=3}^n Z_t^* Z_{t-2}^*}{\sum_{t=3}^n Z_{t-1}^* Z_{t-2}^*}, \qquad \gamma^{(1)} = \frac{1}{n}\sum_{t=2}^{n}(Z_t - \vartheta_n^{(1)} Z_{t-1}).$$

The iterative procedure is suggested by the remark that (11) can be written as $\widetilde{z}_t = \vartheta Z_{t-1} + \nu_t(\vartheta)$ with $\widetilde{z}_t = Z_t - \gamma$, and $\nu_t(\vartheta)$ are zero mean random variables. Therefore, to estimate ϑ, one can iteratively regress $\widetilde{z}_t$ on Z_{t-1} adjusting at each step the γ values. At the ith step of iterations we compute

$$(12) \quad \widetilde{z}_t^{(i)} = Z_t - \gamma^{(i-1)}, \qquad \vartheta_n^{(i)} = \frac{\sum_{t=3}^n \widetilde{z}_t^{(i)} Z_{t-2}}{\sum_{t=3}^n Z_{t-1} Z_{t-2}}, \qquad \gamma^{(i)} = \widehat{\omega}(1 - \vartheta_n^{(i)}).$$

For n fixed and $i \to \infty$, $\vartheta_n^{(i)}$ converges to some limit $\overline{\vartheta}_{0n}$. Define $\widehat{\vartheta}_{0n}$ as the projection of this $\overline{\vartheta}_{0n}$ onto $[a, b]$ and replace the test statistic T_n by

$$(13) \qquad \widetilde{T}_n = \frac{1}{\sqrt{nB_n}} \sum_{t=2}^{n} \{(\widetilde{z}_t - \widehat{\vartheta}_{0n} Z_{t-1})^2 - M_n\}.$$

With this definition of a test statistic, one problem still remains: in practice we do not know the moments of the errors $\sigma^2, \eta^2, \mu, \nu$ that are needed to compute M_n and B_n. We return to this issue in Section 7 where a completely data-driven procedure is discussed.

5. Proofs

5.1. *Proof of Theorem 1.* Define

$$\nu_t(\vartheta) = \varepsilon_t + \xi_t - \vartheta \xi_{t-1}, \qquad \widehat{\nu}_t = \nu_t(\widehat{\vartheta}_n).$$

Note that under **A1** and **A2**,

$$(14) \quad \mathbf{E}\nu_t^2(\vartheta) = \sigma^2 + \eta^2(1 + \vartheta^2) = M(\vartheta),$$

$$(15) \quad \mathbf{E}\{(\nu_t^2(\vartheta) - M(\vartheta))^2 + 2(\nu_t^2(\vartheta) - M(\vartheta))(\nu_{t-1}^2(\vartheta) - M(\vartheta))\} = B(\vartheta),$$

$$(16) \quad \sup_{0 \le \vartheta < 1} \mathbf{E}|\nu_t(\vartheta)|^{4+\delta} < \infty.$$

The proof of Theorem 1 is based on the following lemma.

Lemma 1. *Assume* **A1** *and* **A2**. *Then*

$$(17) \qquad \limsup_{n \to \infty} \sup_{\vartheta \in [a,b]} \mathbf{E}_\vartheta (\sqrt{n}\, |\widehat{\vartheta}_n - \vartheta|)^{3/2} < \infty,$$

where $\mathbf{E}_\vartheta$ *denotes the expectation w.r.t.* $\mathbf{P}_\vartheta$.

Proof of Theorem 1. The summands in T_n are of the form

$$(18) \quad (Z_t - \widehat{\vartheta}_n Z_{t-1})^2 - (\sigma^2 + \eta^2(1 + \widehat{\vartheta}_n^2))$$
$$= (\vartheta - \widehat{\vartheta}_n)^2 Y_{t-1}^2 + [\widehat{\nu}_t^2 - (\sigma^2 + \eta^2(1 + \widehat{\vartheta}_n^2))] + 2(\vartheta - \widehat{\vartheta}_n) Y_{t-1} \widehat{\nu}_t.$$

Let $\mathcal{F}_k^m$ be the σ-algebra generated by $(Z_k, \ldots, Z_m)$. Note that Y_{t-1} is conditionally independent of $\widehat{\nu}_t$ given $\mathcal{F}_1^j$, $N - 1 \le j \le t - 2$, and thus,

$$(19) \qquad \mathbf{E}_\vartheta (Y_{t-1} \widehat{\nu}_t \,|\, \mathcal{F}_1^j) = 0$$

for $N-1 \le j \le t-2$. Note also that

$$B(\vartheta) \ge B^* = 4\sigma^2\eta^2, \qquad \forall \vartheta \in [0,1), \tag{20}$$

and

$$\sup_{\vartheta\in[a,b]} \sup_t \mathbf{E}_\vartheta Y_t^2 < \infty, \tag{21}$$

which is straightforward under **A1** and **A2**.

Since $\widehat{\vartheta}_n$ is independent of $Y_N, \ldots, Y_{n-1}$, we have $|\widehat{\vartheta}_n - \vartheta| \le 1$, and in view of (20) we get

$$\mathbf{E}_\vartheta\left[\frac{1}{\sqrt{NB_n}} \sum_{t=N+1}^{n} (\widehat{\vartheta}_n - \vartheta)^2 Y_{t-1}^2\right] \le \mathbf{E}_\vartheta|\widehat{\vartheta}_n - \vartheta|^{3/2}\mathbf{E}_\vartheta\left[\frac{1}{\sqrt{NB^*}} \sum_{t=N+1}^{n} Y_{t-1}^2\right].$$

This, together with (21) and Lemma 1, entails

$$\limsup_{n\to\infty} \sup_{\vartheta\in[a,b]} \mathbf{E}_\vartheta\left[\frac{1}{\sqrt{NB_n}} \sum_{t=N+1}^{n} (\widehat{\vartheta}_n - \vartheta)^2 Y_{t-1}^2\right] = 0. \tag{22}$$

Next, since $\widehat{\vartheta}_n$ is $\mathcal{F}_1^{N-1}$-measurable and in view of (19), (20) we find

$$\begin{aligned}
&\mathbf{E}_\vartheta\left|\frac{1}{\sqrt{NB_n}} \sum_{t=N+1}^{n} (\widehat{\vartheta}_n - \vartheta) Y_{t-1}\widehat{\nu}_t\right| \\
&\quad \le \mathbf{E}_\vartheta\left[\mathbf{E}_\vartheta\left(\frac{1}{\sqrt{NB^*}}\left|\sum_{t=N+1}^{n} Y_{t-1}\widehat{\nu}_t\right| \,\middle|\, \mathcal{F}_1^{N-1}\right)\mathbf{E}_\vartheta|\widehat{\vartheta}_n - \vartheta|\right] \\
&\quad \le \overline{C}\frac{1}{\sqrt{N}}\mathbf{E}_\vartheta|\widehat{\vartheta}_n - \vartheta|\left(\sup_{\vartheta\in[a,b]} \mathbf{E}_\vartheta\left(\sum_{t=N+1}^{n} Y_{t-1}^2\nu_t^2(\vartheta)\right)\right)^{1/2},
\end{aligned}$$

where $\overline{C}$ is a constant which does not depend on ϑ. This, together with Lemma 1, the fact that the random variable Y_{t-1} is independent of $\nu_t^2(\vartheta)$, (21) and **A1**–**A2**, yields

$$\limsup_{n\to\infty} \sup_{\vartheta\in[a,b]} \mathbf{E}_\vartheta\left|\frac{1}{\sqrt{NB_n}} \sum_{t=N+1}^{n} (\widehat{\vartheta}_n - \vartheta) Y_{t-1}\widehat{\nu}_t\right| = 0. \tag{23}$$

Hence, using (18), (22) and (23), we get

$$\begin{aligned}
\mathbf{P}_\vartheta(\Delta_n = 1) &= \mathbf{P}_\vartheta(T_n \ge t_\alpha) \\
&= \mathbf{P}_\vartheta\left\{\frac{1}{\sqrt{NB_n}} \sum_{t=N+1}^{n} (\widehat{\nu}_t^2 - (\sigma^2 + \eta^2(1 + \widehat{\vartheta}_n^2))) \ge t_\alpha\right\} + o(1)
\end{aligned} \tag{24}$$

as $n \to \infty$ uniformly in $\vartheta \in [a,b]$. The sequence $\{\widehat{\nu}_t\}$ for fixed $\widehat{\vartheta}_n = \vartheta_0 \in [a,b]$ is the sequence of 2-dependent random variables $\{\nu_t(\vartheta_0)\}$, satisfying (14), (15), (16) and such that (20) holds. For these variables we have the Berry–Esseen bound (Tikhomirov, 1980):

$$\left|\mathbf{P}\left\{\frac{1}{\sqrt{NB(\vartheta_0)}}\sum_{t=N+1}^{n}(\nu_t^2(\vartheta_0)-(\sigma^2+\eta^2(1+\vartheta_0^2)))<t_\alpha\right\}-\Phi(t_\alpha)\right| \le \frac{C_0(\log n)^{1+\delta/2}}{n^{\delta/4}}, \tag{25}$$

where C_0 is a constant independent of ϑ_0. Conditioning on $\widehat{\vartheta}_n$ in (24) and using (25) we arrive at the statement of the theorem.

5.2. *Proof of Theorem 2.* For all $0 \le \vartheta < 1$ and all $m \in \mathcal{M}$ define

$$d_n(m,\vartheta) = \frac{1}{N}\sum_{t=N+1}^{n}\mathbf{E}_m(m(Y_{t-1})-\vartheta Y_{t-1})^2,$$

$$\zeta_n(m,\vartheta) = \frac{1}{N}\sum_{t=N+1}^{n}(m(Y_{t-1})-\vartheta Y_{t-1})^2.$$

LEMMA 2. *Assume* **A1**–**A3**. *Then*

$$\lim_{C\to\infty}\limsup_{n\to\infty}\sup_{m\in\mathcal{M}}\mathbf{P}_m\{\sup_{\vartheta\in[a,b]}\sqrt{n}\,|\zeta_n(m,\vartheta)-d_n(m,\vartheta)|>C\}=0.$$

Write

$$T_n = \frac{A_n+W_n+R_n}{\sqrt{B_n}}, \tag{26}$$

where

$$A_n = \frac{1}{\sqrt{N}}\sum_{t=N+1}^{n}(m(Y_{t-1})-\widehat{\vartheta}_nY_{t-1})^2 = \sqrt{N}\,\zeta_n(m,\widehat{\vartheta}_n),$$

$$W_n = \frac{1}{\sqrt{N}}\sum_{t=N+1}^{n}[\widehat{\nu}_t^2-(\sigma^2+\eta^2(1+\widehat{\vartheta}_n^2))],$$

$$R_n = \frac{2}{\sqrt{N}}\sum_{t=N+1}^{n}\widehat{\nu}_t(m(Y_{t-1})-\widehat{\vartheta}_nY_{t-1}).$$

Furthermore, for all $\vartheta\in[a,b]$ set

$$\widetilde{d}_n(m,\vartheta) = \frac{1}{N}\sum_{t=N+1}^{n}\mathbf{E}_m[(m(Y_{t-1})-\vartheta Y_{t-1})^2\,|\,\mathcal{F}_1^{N-1}].$$

LEMMA 3. *Assume* **A1**–**A3**. *Then*

$$\lim_{C\to\infty}\limsup_{n\to\infty}\sup_{m\in\mathcal{M}}\mathbf{P}_m\{\sup_{\vartheta\in[a,b]}\sqrt{n}\,|\widetilde{d}_n(m,\vartheta)-d_n(m,\vartheta)|>C\}=0.$$

LEMMA 4. *Assume* **A1–A3**. *Then*

$$\lim_{n\to\infty}\sup_{m\in\mathcal{M}}\mathbf{P}_m\left\{|R_n|\ge\frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n);\ \widetilde{d}_n(m,\widehat{\vartheta}_n)>\frac{Q}{\sqrt{n}}\right\}=0$$

for all $Q>0$.

Proof of Theorem 2. Consider the random events

$$\Gamma_1=\left\{\widetilde{d}_n(m,\widehat{\vartheta}_n)\ge d_n(m,\widehat{\vartheta}_n)+\frac{Q_1}{\sqrt{n}}\right\};$$

$$\Gamma_2=\left\{\zeta_n(m,\widehat{\vartheta}_n)<d_n(m,\widehat{\vartheta}_n)-\frac{Q_2}{\sqrt{n}}\right\};$$

$$\Gamma_3=\left\{|R_n|\ge\frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n);\ \widetilde{d}_n(m,\widehat{\vartheta}_n)>\frac{Q_3}{\sqrt{n}}\right\},$$

where the positive constants Q_1, Q_2 and Q_3 are chosen so that for all $m\in\mathcal{M}$ and for all n large enough,

$$\mathbf{P}_m\{\Gamma_i\}\le\beta/8,\qquad i=1,2,3. \tag{27}$$

Such Q_1, Q_2 and Q_3 exist in view of Lemmas 2–4.

Note also that for all n and λ large enough,

$$\sup_{m\in\mathcal{M}_n(\lambda)}\mathbf{P}_m\{\Gamma_4\}\le\beta/8, \tag{28}$$

where

$$\Gamma_4=\{\widetilde{d}_n(m,\widehat{\vartheta}_n)\le Q_3/\sqrt{n}\}.$$

This follows from the definition of $\mathcal{M}_n(\lambda)$ and from Lemma 3.

Assume that λ is large enough to have (28) and $\lambda>Q_2$. Note that if $\lambda>Q_2$, we have in Γ_2: $d_n(m,\widehat{\vartheta}_n)-Q_2/\sqrt{n}>0$ for all $m\in\mathcal{M}$. Then, due to (20), (26), (27) and (28), we obtain

$$\mathbf{P}_m\{\Delta_n=0\}=\mathbf{P}_m\{T_n<t_\alpha\}=\mathbf{P}_m\left\{\frac{A_n+W_n+R_n}{\sqrt{B_n}}<t_\alpha\right\} \tag{29}$$

$$\le\mathbf{P}_m\left\{\frac{\sqrt{N}\,\zeta_n(m,\widehat{\vartheta}_n)+W_n}{\sqrt{B_n}}-\frac{1}{2}\sqrt{\frac{N}{B^*}}\,\widetilde{d}_n(m,\widehat{\vartheta}_n)<t_\alpha\right\}+\mathbf{P}_m\{\Gamma_3\}+\mathbf{P}_m\{\Gamma_4\}$$

$$\le\mathbf{P}_m\left\{\frac{W_n}{\sqrt{B_n}}+\frac{1}{2}\sqrt{\frac{N}{B^*}}\,d_n(m,\widehat{\vartheta}_n)-Q_4<t_\alpha\right\}+\sum_{i=1}^{4}\mathbf{P}_m\{\Gamma_i\}$$

$$\le\mathbf{P}_m\left\{\frac{W_n}{\sqrt{B_n}}<Q_5\right\}+\frac{\beta}{2},$$

where $Q_4=(Q_1/2+Q_2)/\sqrt{B^*}$ and $Q_5=-\frac{1}{2}\sqrt{\frac{N}{nB^*}}\lambda+Q_4+t_\alpha$. Let λ be so large that $Q_5<0$. Then, in view of (20), we get

$$\mathbf{P}_m\{W_n/\sqrt{B_n}<Q_5\}\le\mathbf{E}_m(\mathbf{P}_m\{|W_n|\ge\sqrt{B^*}|Q_5|\,|\,\mathcal{F}_1^{N-1}\}).$$

Applying here the Chebyshev inequality to the conditional probability and using the $\mathcal{F}_1^{N-1}$-measurability of $\widehat{\vartheta}_n$ and (14), (16), we infer that the last probability in (29) is less than $\beta/2$. This yields the result of the theorem.

5.3. *Proofs of Lemmas 1–4.* The following proposition will be used in the proofs.

PROPOSITION 2. *Let $\{Y_t\}$ be the Markov chain (4) starting at $t = t_0$ with $Y_0 = y_0$ where t_0 is an integer and $y_0 \in \mathbb{R}^1$. Assume* **A1**–**A3**. *Let the function $g : \mathbb{R}^1 \to \mathbb{R}^1$ satisfy*

$$|g(x)| \le g_0(1 + x^2), \qquad \forall x \in \mathbb{R}^1, \tag{30}$$

where g_0 is a finite constant. Then there exist finite constants $C_1 = C_1(c, d, p(\cdot), g_0, y_0)$ and $C_2 = C_2(c, d, p(\cdot), g_0)$ such that

$$\sup_{m\in\mathcal{M}} \sup_t \mathbf{E}_m |Y_t|^{4+\delta} \le C_1 \tag{31}$$

and

$$\sup_{m\in\mathcal{M}} \mathbf{P}_m \left\{ \left| \frac{1}{\operatorname{card} \tau_n} \sum_{t\in\tau_n} (g(Y_t) - \mathbf{E}_m g(Y_t)) \right| \ge \frac{v}{\sqrt{n}} \right\} \le \frac{C_2}{v^2} (|y_0|^{4+\delta} + 1) \tag{32}$$

for any integer $n \ge 1$, any $v > 0$ and any subset $\tau_n \subseteq \{t_0, \dots, n\}$ such that $\operatorname{card} \tau_n \ge [n/2]$.

The proof of Proposition 2 is given in the Appendix.

Proof of Lemma 1. It is straightforward to see that under $\mathbf{H}_0$, **A1** and **A2**,

$$\lim_{t\to\infty} \mathbf{E}_\vartheta Y_t^2 = \frac{\sigma^2}{1-\vartheta} = y_*, \qquad \sup_{\vartheta\in[a,b]} \sup_t \mathbf{E}_\vartheta Y_t^4 < \infty, \tag{33}$$

$$\sup_{\vartheta\in[a,b]} \sup_{N\ge 4} \mathbf{E}_\vartheta \left| \sqrt{n} \left(\frac{1}{N} \sum_{t=3}^{N-1} Y_{t-2}^2 - y_* \right) \right|^2 < \infty. \tag{34}$$

Next, under the null hypothesis we have, from (3) and (4),

$$Z_t = \vartheta Z_{t-1} + \nu_t(\vartheta), \qquad t = 2, \dots, n, \tag{35}$$

which, together with **A1** and **A2**, easily implies

$$\sup_{\vartheta\in[a,b]} \sup_t \mathbf{E}_\vartheta Z_t^2 < \infty. \tag{36}$$

Furthermore,

$$\begin{aligned} \sum_{t=3}^{N-1} Z_{t-1} Z_{t-2} &= \sum_{t=3}^{N-1} (\vartheta Y_{t-2} + \varepsilon_{t-1} + \xi_{t-1})(Y_{t-2} + \xi_{t-2}) \\ &= \vartheta \sum_{t=3}^{N-1} Y_{t-2}^2 + S_n, \end{aligned} \tag{37}$$

where

$$S_n = \sum_{t=3}^{N-1} [Y_{t-2}(\vartheta\xi_{t-2} + \varepsilon_{t-1} + \xi_{t-1}) + \xi_{t-2}(\varepsilon_{t-1} + \xi_{t-1})].$$

Using **A1**, **A2**, (33) and the fact that $\vartheta\xi_{t-2} + \varepsilon_{t-1} + \xi_{t-1}$ is independent of $\mathcal{F}_1^{t-2}$ we find

$$\sup_{\vartheta\in[a,b]} \mathbf{E}_\vartheta S_n^2 = O(n), \qquad n \to \infty. \tag{38}$$

Similarly, **A1**, **A2**, (36) and the fact that $\nu_t(\vartheta)$ is independent of $\mathcal{F}_1^{t-2}$ imply

$$\sup_{\vartheta\in[a,b]} \mathbf{E}_\vartheta \Big| \sum_{t=3}^{N-1} Z_{t-2}\nu_t(\vartheta)\Big|^2 = O(n), \qquad n \to \infty. \tag{39}$$

Now, (34), (37), (38) and Chebyshev's inequality yield

$$\begin{aligned} &\mathbf{P}_\vartheta\bigg(\frac{1}{N}\sum_{t=3}^{N-1} Z_{t-1}Z_{t-2} \le ay_*/2\bigg) \\ &\quad\le \mathbf{P}_\vartheta\bigg(\frac{S_n}{N} + \vartheta\bigg(\frac{1}{N}\sum_{t=3}^{N-1} Y_{t-2}^2 - y_*\bigg) \le -ay_*/2\bigg) \\ &\quad\le \mathbf{P}_\vartheta(S_n \ge Nay_*/4) + \mathbf{P}_\vartheta\bigg(\bigg|\frac{1}{N}\sum_{t=3}^{N-1} Y_{t-2}^2 - y_*\bigg| \ge ay_*/4\bigg) \le \frac{K_1}{n} \end{aligned} \tag{40}$$

for any $N \ge 4$ and $\vartheta \in [a,b]$, where K_1 is a constant which does not depend on ϑ.

Using (35) we get

$$\overline{\vartheta}_n = \vartheta + \frac{\sum_{t=3}^{N-1} Z_{t-2}\nu_t(\vartheta)}{\sum_{t=3}^{N-1} Z_{t-1}Z_{t-2}}.$$

This representation, together with (39), (40) and Chebyshev's inequality, gives

$$\begin{aligned} &\mathbf{P}_\vartheta(|\sqrt{n}\,(\widehat{\vartheta}_n - \vartheta)| \ge u) \le \mathbf{P}_\vartheta(|\sqrt{n}\,(\overline{\vartheta}_n - \vartheta)| \ge u) \\ &\le \mathbf{P}_\vartheta\bigg(\frac{1}{N}\sum_{t=3}^{N-1} Z_{t-1}Z_{t-2} \le ay_*/2\bigg) + \mathbf{P}_\vartheta\bigg(\frac{2}{ay_*}\bigg|\frac{1}{N}\sum_{t=3}^{N-1} Z_{t-2}\nu_t(\vartheta)\bigg| \ge \frac{u}{\sqrt{n}}\bigg) \\ &\le K_2\bigg(\frac{1}{n} + \frac{1}{u^2}\bigg) \end{aligned} \tag{41}$$

for any $u > 0$, any $N \ge 4$, any $\vartheta \in [a,b]$, and some $K_2 > 0$ independent of n, ϑ. Observing that $|\widehat{\vartheta}_n - \vartheta| \le 1$ and using (41) we get

$$\begin{aligned}
\mathbf{E}_\vartheta|\sqrt{n}\,(\widehat{\vartheta}_n - \vartheta)|^{3/2} &= \int_0^{n^{3/4}} \mathbf{P}_\vartheta(|\sqrt{n}\,(\widehat{\vartheta}_n - \vartheta)|^{3/2} \ge x)\,dx \\
&\le 1 + \int_1^{n^{3/4}} \mathbf{P}_\vartheta(|\sqrt{n}\,(\widehat{\vartheta}_n - \vartheta)| \ge x^{2/3})\,dx \\
&\le 1 + K_2 \int_1^{n^{3/4}} \left(\frac{1}{n} + \frac{1}{x^{4/3}}\right) dx \le 1 + 3K_2. \blacksquare
\end{aligned}$$

Proof of Lemma 2. Set $g_1(x) = m^2(x),\ g_2(x) = xm(x),\ g_3(x) = x^2$, and

$$I_j = \frac{1}{N} \sum_{t=N+1}^{n} g_j(Y_{t-1}), \qquad j = 1, 2, 3.$$

Clearly, for $m \in \mathcal{M}$ the functions $g = g_j$ satisfy (30). Also, we have

$$\sup_{0<\vartheta<1} |\zeta_n(m,\vartheta) - d_n(m,\vartheta)| \le |I_1 - \mathbf{E}_m I_1| + 2|I_2 - \mathbf{E}_m I_2| + 2|I_3 - \mathbf{E}_m I_3|. \tag{42}$$

This and (32) with $g = g_j$, $j = 1, 2, 3$, $t_0 = 0$ and $\tau_n = \{N, \ldots, n\}$ yield Lemma 2. ∎

Proof of Lemma 3. Acting as in (42) and applying (32) to the Markov chain (4) starting at $t_0 = N - 1$ with $y_0 = Y_{N-1}$ and $\tau_n = \{N, \ldots, n\}$ we find

$$\begin{aligned}
&\mathbf{P}_m\{\sup_{\vartheta\in[a,b]} \sqrt{n}\,|\zeta_n(m,\vartheta) - \widetilde{d}_n(m,\vartheta)| > C \,|\, \mathcal{F}_1^{N-1}\} \\
&\qquad \le \mathbf{P}_m\Big\{2\sum_{j=1}^{3} |I_j - \mathbf{E}_m(I_j \,|\, \mathcal{F}_1^{N-1})| > C \,|\, \mathcal{F}_1^{N-1}\Big\} \\
&\qquad = \mathbf{P}_m\Big\{2\sum_{j=1}^{3} |I_j - \mathbf{E}_m(I_j \,|\, Y_{N-1})| > C \,|\, Y_{N-1}\Big\} \\
&\qquad \le \frac{C_3}{C^2}\,(|Y_{N-1}|^{4+\delta} + 1),
\end{aligned}$$

where $C_3 = C_3(c, d, p(\cdot))$ is a constant. Taking expectations and using (31) we obtain

$$\mathbf{P}_m\{\sup_{\vartheta\in[a,b]} \sqrt{n}\,|\zeta_n(m,\vartheta) - \widetilde{d}_n(m,\vartheta)| > C\} \le \frac{C_3}{C^2}\,(C_1 + 1),$$

which together with Lemma 2 gives Lemma 3. ∎

Proof of Lemma 4. Since $\widehat{\vartheta}_n$ is $\mathcal{F}_1^{N-1}$-measurable we have

$$(43)\quad \mathbf{P}_m\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}};\, |R_n| \geq \frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n)\Big\}$$
$$= \mathbf{E}_m\bigg(\mathbf{P}_m\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}};\, |R_n| \geq \frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n)\,\Big|\,\mathcal{F}_1^{N-1}\Big\}\bigg)$$
$$= \mathbf{E}_m\bigg(I\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}}\Big\}\mathbf{P}_m\Big\{|R_n| \geq \frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n)\,\Big|\,\mathcal{F}_1^{N-1}\Big\}\bigg).$$

Note that $\widehat{\nu}_t$ is conditionally independent of $(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})$, $t = N+1,\ldots,n$, for fixed $\mathcal{F}_1^{N-1}$, and $\mathbf{E}_m((m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})\widehat{\nu}_t \,|\, \mathcal{F}_1^j) = 0$ for $N-1 \leq j \leq t-2$. This entails

$$(44)\quad \mathbf{E}_m(R_n^2 \,|\, \mathcal{F}_1^{N-1}) \leq \mathbf{E}_m\bigg(\frac{1}{N}\sum_{t=N+1}^{n}(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})^2\widehat{\nu}_t^2\,\bigg|\,\mathcal{F}_1^{N-1}\bigg)$$
$$+ \frac{2}{N}\sum_{t=N+2}^{n}|\mathbf{E}_m\{(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})\widehat{\nu}_t(m(Y_{t-2}) - \widehat{\vartheta}_n Y_{t-2})\widehat{\nu}_{t-1} \,|\, \mathcal{F}_1^{N-1}\}|$$
$$\leq \frac{9}{N}\,\mathbf{E}_m(\widehat{\nu}_t^2 \,|\, \mathcal{F}_1^{N-1})\mathbf{E}_m\Big(\sum_{t=N+1}^{n}(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})^2\,\Big|\,\mathcal{F}_1^{N-1}\Big)$$
$$\leq C_4\mathbf{E}_m\bigg(\frac{1}{N}\sum_{t=N+1}^{n}(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})^2\,\bigg|\,\mathcal{F}_1^{N-1}\bigg)$$

where C_4 is a finite constant depending on σ^2, η^2 only.

Applying the Chebyshev inequality to the last conditional probability in (43) and using (44) we get

$$\mathbf{P}_m\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}};\, |R_n| \geq \frac{\sqrt{N}}{2}\,\widetilde{d}_n(m,\widehat{\vartheta}_n)\Big\}$$
$$\leq \frac{4C_4}{N}\,\mathbf{E}_m\bigg(I\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}}\Big\}$$
$$\times\,\frac{\mathbf{E}_m\{N^{-1}\sum_{t=N+1}^{n}(m(Y_{t-1}) - \widehat{\vartheta}_n Y_{t-1})^2 \,|\, \mathcal{F}_1^{N-1}\}}{\widetilde{d}_n^2(m,\widehat{\vartheta}_n)}\bigg)$$
$$= \frac{4C_4}{N}\,\mathbf{E}_m\bigg([\widetilde{d}_n(m,\widehat{\vartheta}_n)]^{-1} I\Big\{\widetilde{d}_n(m,\widehat{\vartheta}_n) > \frac{Q}{\sqrt{n}}\Big\}\bigg) \leq \frac{4C_4}{Q}\,\frac{\sqrt{n}}{N}.$$

This completes the proof of the lemma.

6. A simulation study. In this section, we provide simulation evidence of the finite sample behavior of the test statistic derived in Section 3.

We consider the following function m:

$$m(x) = \vartheta x + \frac{\lambda}{n^{1/4}} \sin(2\pi x),$$

where $\lambda > 0$ is a parameter which determines the deviation from linearity. Clearly, $|m(x)| \leq |x|\,|\vartheta| + \lambda/n^{1/4}$, and therefore $m \in \mathcal{M}(c,d)$ for some $c \in (0,1), d > 0$, if $|\vartheta| < 1$.

We generated 1000 replications of the series

$$\begin{aligned} &Z_t = Y_t + \xi_t, \\ (45) \qquad &Y_t = \vartheta Y_{t-1} + \frac{\lambda}{n^{1/4}} \sin(2\pi Y_{t-1}) + \varepsilon_t, \qquad t = 1, \ldots, n, \end{aligned}$$

where ξ_t and ε_t are independent i.i.d. $\mathcal{N}(0,1)$ random variables, $n = 10000$, $6561, 4096$ and $\vartheta = 0.95$. The numbers n were chosen to obtain simple values for the sensitivity coefficient $\lambda/n^{1/4}$. We have not included a constant shift ω into (45), so that we directly calculate parameter estimates and test statistics without the iterative procedure of Section 4. The constants $\sigma^2, \eta^2, \mu, \nu$ needed for computation of M_n and B_n are explicitly known in view of the normality of the errors. We do not split the sample, i.e. apply the summation until $t = n$ in (6) and use

$$T_n = \frac{1}{\sqrt{nB_n}} \sum_{t=2}^{n} \{(Z_t - \widehat{\vartheta}_n Z_{t-1})^2 - M_n\}$$

instead of (7).

Summary statistics of the T_n test statistic are given in Table 1. The estimates $\widehat{\vartheta}_n$ were always very close to the true value of 0.95, so they are not reported.

Table 1. Summary statistics of simulated test statistics T_n. The first rows of each row-triple gives the value of λ, the second the mean of T_n for 1000 replications, the third the standard deviation.

n	10000	6561	4096
λ	0.000	0.000	0.000
mean	−0.048	0.019	0.014
std.dev.	1.502	1.551	1.490
λ	0.500	0.450	0.400
mean	0.383	0.365	0.357
std.dev.	1.553	1.569	1.418
λ	1.000	0.900	0.800
mean	1.801	1.418	1.052
std.dev.	1.521	1.520	1.492

n	10000	6561	4096
λ	1.500	1.350	1.200
mean	4.251	3.445	2.747
std.dev.	1.545	1.664	1.643
λ	2.000	1.800	1.600
mean	8.486	6.868	5.377
std.dev.	1.903	1.864	1.823
λ	2.500	2.250	2.000
mean	15.406	12.445	9.881
std.dev.	3.306	3.974	3.332

The distributions of T_n for $n = 4096$ and $n = 10000$ are depicted in Figures 1 and 2, respectively, for $\lambda = 0$ to $\lambda = 1.5$. The distributions move to the right when λ increases, which shows the consistency of the test. We also present the power functions for the levels $\alpha = 0.05$ and $\alpha = 0.1$ in Figure 3 (for $n = 4096$) and Figure 4 (for $n = 10000$). We see that the power converges fast to 1 as λ grows.

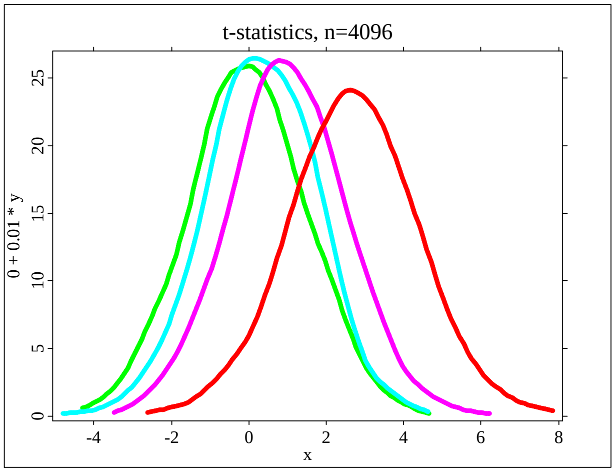

Fig. 1. The distribution of T_n for $n = 4096$. From left to right: $\lambda = 0, 0.5, 1, 1.5$.

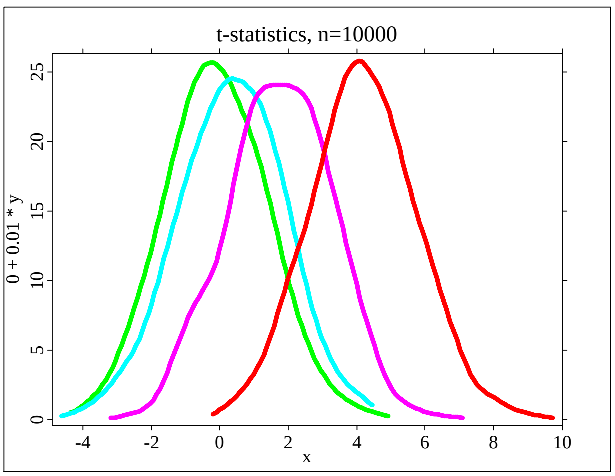

Fig. 2. The distribution of T_n for $n = 10000$. From left to right: $\lambda = 0, 0.5, 1, 1.5$.

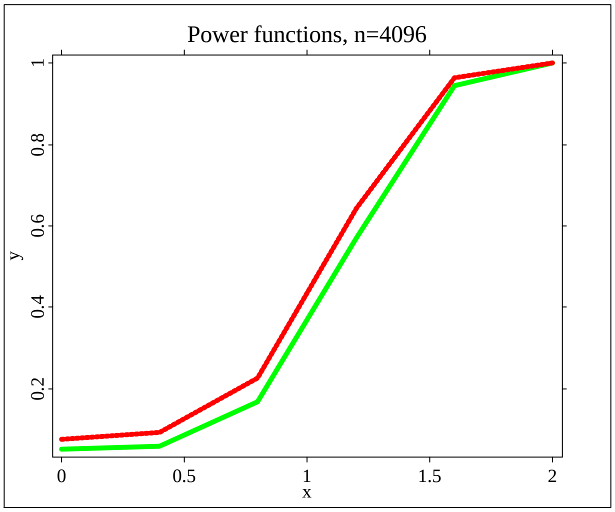

Fig. 3. Power functions of T_n for $n = 4096$. The abscissa represents the parameter λ. Under the null hypothesis, $\lambda = 0$. The black curve is the power function for $\alpha = 0.05$, the grey curve is the power function for $\alpha = 0.1$.

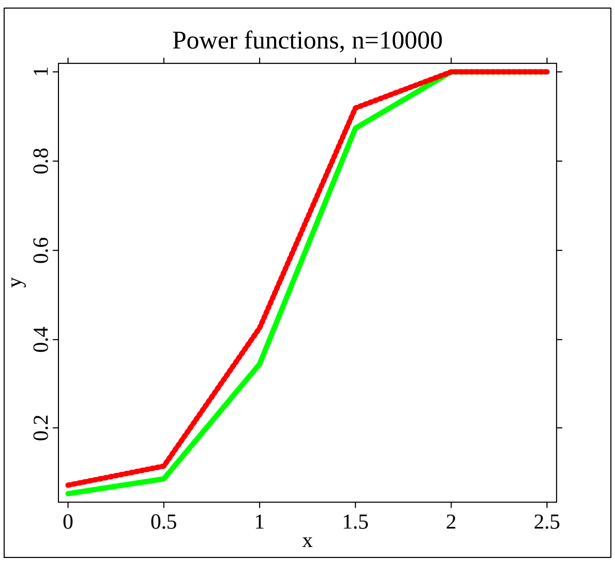

Fig. 4. Power functions of T_n for $n = 10000$. The abscissa represents the parameter λ. Under the null hypothesis, $\lambda = 0$. The black curve is the power function for $\alpha = 0.05$, the grey curve is the power function for $\alpha = 0.1$.

7. Application to the HFDF96 data set. We extracted two stock price indices, the Dow Jones Industrial Average and the Standard & Poors 500, and the DEM/USD exchange rate from the HFDF96 data set, provided by Olsen & Associates. The data are half-hourly sampled index values. For the stock indices, we skipped the intervals corresponding to non-trading hours at the New York Stock Exchange. For DEM/USD, we skipped those intervals for which the time of the last quote was more than half an hour behind. This left us with 3680 observations for the stock indices and 14234 observations for DEM/USD. The time series are depicted in Figures 5, 6 and 7.

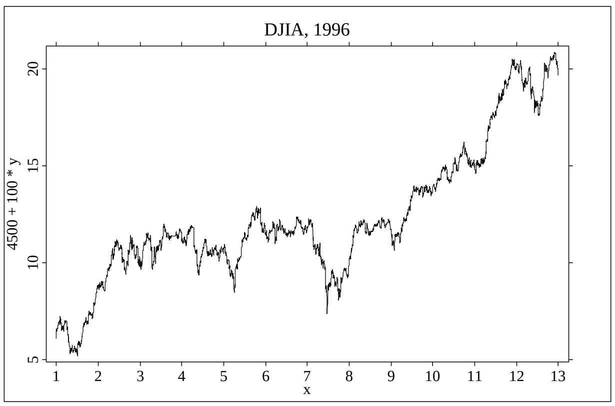

Fig. 5. The Dow Jones Index

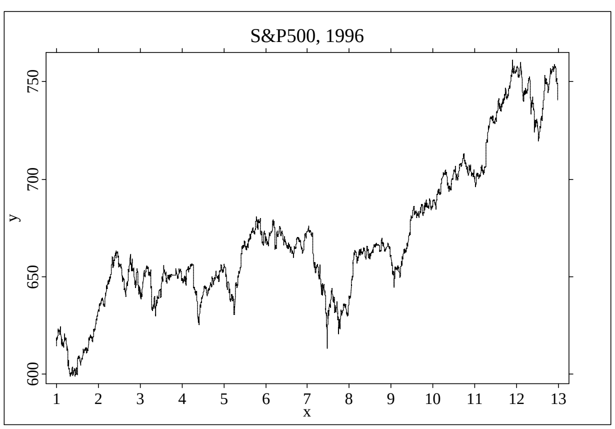

Fig. 6. The S&P500 Index

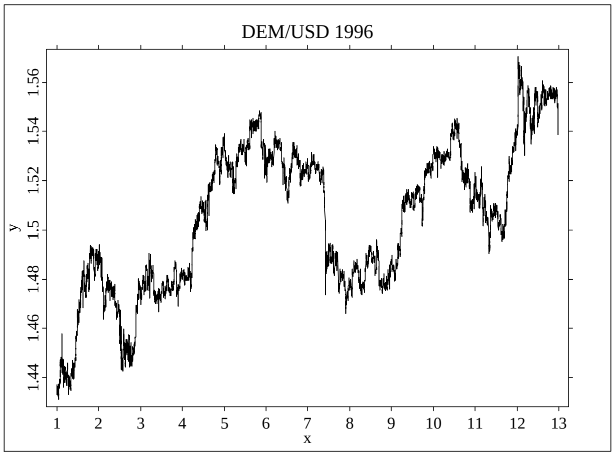

Fig. 7. The DEM/USD exchange rate

First, we estimated ϑ under the null hypothesis as described above, and obtained $\overline{\vartheta}_{0n} = 0.9004$ for DEM/USD, $\overline{\vartheta}_{0n} = 1.0153$ for DJIA, and $\overline{\vartheta}_{0n} = 0.9241$ for S&P500. These results confirm previous results of SV models for high frequency financial data (see e.g. Mahieu and Schotman (1997)). The AR parameter is usually found to be close to one, implying a high persistence of shocks in volatility. For DJIA we have even $\overline{\vartheta}_{0n} > 1$, which could mean a nonstationary volatility, and therefore a nonstationary return process. To rigorously apply our theoretical results we should project this $\overline{\vartheta}_{0n}$ on an interval $[a, b]$ with $0 < a < b < 1$. However, the choice of a and b is not clear in practice. Also we believe that, for such a large sample size, the value $\overline{\vartheta}_{0n} > 1$ is not due to a random error and reflects the fact that the system is indeed nonstationary. Therefore, we keep the value $\overline{\vartheta}_{0n} = 1.0153$ in computations.

Let us turn to the test statistic $\widetilde{T}_n$ in (13). In our real data situation, the moments of the random errors are unknown. Therefore, the values B_n and M_n cannot be computed, but one can try to estimate them. We use the following estimates $\widehat{M}_n$ and $\widehat{B}_n$ for M_n and B_n, respectively:

$$\widehat{M}_n = \frac{1}{n-1}\sum_{t=2}^{n}(\widetilde{z}_t - \overline{\vartheta}_{0n}Z_{t-1})^2, \qquad \widehat{B}_n = \frac{1}{n-1}\sum_{t=2}^{n}(\widetilde{z}_t - \overline{\vartheta}_{0n}Z_{t-1})^4.$$

It is clear that now there is no sense to use the test statistic $\widehat{T}_n$ which is obtained by replacing M_n and B_n in (13) by $\widehat{M}_n$ and $\widehat{B}_n$ because, obviously, $\widehat{T}_n = 0$. To avoid this problem, we divide the sample $(Z_1, \ldots, Z_n)$ into k groups $(Z_1, \ldots, Z_{n_1})$, $(Z_{n_1+1}, \ldots, Z_{n_2})$, $\ldots$, $(Z_{n_{k-1}+1}, \ldots, Z_n)$ of equal size and study the behavior of k test statistics $T_n^{(j)}, j = 1, \ldots, k$, defined as follows:

$$T_n^{(j)} = \sqrt{\frac{k}{n\widehat{B}_n}} \sum_{t=n_{j-1}+1}^{n_j} \{(\widetilde{z}_t - \overline{\vartheta}_{0n} Z_{t-1})^2 - \widehat{M}_n\}.$$

In particular, one can take $k = 2$. However, the use of a larger number of subsamples appears to be reasonable because in this case we have an additional information on how many times the hypothesis is accepted or rejected. On the other hand, k should not be too large, since then the number of observations in the subsamples may become too small. Thus, we obtain k test statistics, and k decisions to accept or reject the null hypothesis at level α. Also, we can estimate ϑ for each subsample. It should be noted that most of these estimates were very close to the estimates reported above for the entire sample.

Table 2 gives the number of rejections for selected k. Ideally, under the null hypothesis we would expect to reject αk times. Especially for the stock indices this holds closely for $k < 100$. Note that for $k = 100$ there are only 36 observations in each subsample for the stock indices. For DEM/USD, we reject slightly more often than one would expect under linearity. However, recall the still moderate sizes of the subsamples and the slow rate of the test.

Table 2. Number of rejections for k subsamples, each of size n/k, for two levels, $\alpha = 0.05$ and $\alpha = 0.1$

	DEM/USD		DJIA		S&P500	
k	α=0.05	α=0.1	α=0.05	α=0.1	α=0.05	α=0.1
2	0	0	0	0	0	0
3	0	0	0	0	1	1
4	0	0	1	1	0	0
5	1	1	1	1	0	0
6	0	0	1	1	2	2
10	1	1	1	1	0	2
12	1	1	1	1	1	2
20	2	3	1	1	1	3
50	5	7	2	3	4	6
100	12	15	10	13	9	9

To summarize, the hypothesis of a linear AR(1) structure in log volatility is confirmed by our results. This is surprising, at least for the stock indices, since in the ARCH literature very often nonlinearities were found for stock volatility. But recall that our sample period 1996 does not cover any major crashes of the markets, so volatility exhibits a rather smooth behavior. It would be interesting to apply the test to other time periods.

Appendix

Proof of Proposition 2. Let us prove (31) first. For $m \in \mathcal{M}(c,d)$, the process Y_t satisfies the recursive inequalities

$$|Y_t| \le c|Y_{t-1}| + d + |\varepsilon_t|.$$

Since $c \in (0,1)$, this easily entails

$$|Y_t| \le |y_0|c^t + \frac{d}{1-c} + \sum_{k=0}^{t-1} c^k |\varepsilon_{t-k}|,$$

and therefore

$$\mathbf{E}_m\{|Y_t|^{4+\delta}\} \le 2^{3+\delta}\Big\{\Big(|y_0| + \frac{d}{1-c}\Big)^{4+\delta} + \mathbf{E}\Big(\Big(\sum_{k=0}^{t-1} c^k|\varepsilon_{t-k}|\Big)^{4+\delta}\Big)\Big\}. \tag{46}$$

Set $X_k = c^k|\varepsilon_{t-k}|$. Since $\mathbf{E}|\varepsilon_{t-k}|^{4+\delta} < \infty$ and $0 < c < 1$ there exists a finite constant $C_* > 0$ such that

$$\max\Big(\sum_{k=0}^{\infty}\mathbf{E}|X_k|,\ \sum_{k=0}^{\infty}\mathbf{E}|X_k|^2,\ \sum_{k=0}^{\infty}\mathbf{E}|X_k|^{4+\delta}\Big) \le C_*.$$

Using this, the convexity inequality and Rosenthal's inequality (Petrov (1995), p. 59), we obtain

$$\begin{aligned}
\mathbf{E}\Big(\Big|\sum_{k=0}^{t-1} X_k\Big|^{4+\delta}\Big) &\le 2^{3+\delta}\Big[\mathbf{E}\Big(\Big|\sum_{k=0}^{t-1}(X_k - \mathbf{E}X_k)\Big|^{4+\delta}\Big) + \Big(\sum_{k=0}^{t-1}\mathbf{E}|X_k|\Big)^{4+\delta}\Big] \\
&\le 2^{3+\delta}\Big[c(\delta)\Big(\sum_{k=0}^{t-1}\mathbf{E}|X_k - \mathbf{E}X_k|^{4+\delta} + \Big(\sum_{k=0}^{t-1}\mathbf{E}|X_k - \mathbf{E}X_k|^2\Big)^{2+\delta/2}\Big) + C_*^{4+\delta}\Big] \\
&< \infty,
\end{aligned}$$

where $c(\delta) > 0$ is a constant depending only on δ. This, together with (46), yields

$$\mathbf{E}_m\{|Y_t|^{4+\delta}\} \le C_5(1 + |y_0|^{4+\delta}), \tag{47}$$

where $C_5 > 0$ depends only on $c, d, p(\cdot)$. This proves (31).

Let us prove (32). Set $V_t = g(Y_t) - \mathbf{E}_m\{g(Y_t)\}$. From Davydov's (1968) inequality,

$$\text{(48)}\quad \mathbf{E}_m\Big\{\Big(\sum_{t\in\tau_n}(g(Y_t)-\mathbf{E}_m\{g(Y_t)\})\Big)^2\Big\} = \sum_{t,l\in\tau_n}\mathrm{Cov}(V_t,V_l)$$
$$\le 2\sum_{t,l\in\tau_n}(\alpha^{t,l})^{\delta/(4+\delta)}[\mathbf{E}_m(|V_t|^{2+\delta/2})\mathbf{E}_m(|V_l|^{2+\delta/2})]^{1/(2+\delta/2)},$$

where

$$\alpha^{t,l} = 2\sup_{A\times B\in\sigma(Y_t)\times\sigma(Y_l)}|\mathrm{Cov}(I\{A\},I\{B\})|$$

is the α-mixing coefficient between the σ-fields generated by the random variables Y_t and Y_l under the measure $\mathbf{P}_m$, and $I\{A\}$ is the indicator function of the set A.

Now we evaluate the α-mixing coefficient. It is easy to show that there exist finite positive constants c', d', a' such that the function

$$f(x) = \max(c'|x| - d', a')$$

is a Lyapunov function for the Markov chain Y_t, i.e.

$$\mathbf{E}(f(Y_t)\,|\,Y_{t-1}=x) \le c'' f(x) - \gamma_0, \qquad \forall |x| \ge x_0,$$

for some $0 < c'' < 1$, $\gamma_0 > 0$, $x_0 > 0$. As follows from Meyn and Tweedie (1992) (see also Doukhan (1995), p. 92, Remark 2), under the assumptions **A1**–**A3** we have

$$\text{(49)}\qquad \sup_{m\in\mathcal{M}}\alpha^{t,l} \le C_6(1+f(y_0))\varrho^{|t-l|} \le C_7(1+|y_0|)\varrho^{|t-l|}$$

where $C_6 = C_6(c,d,p(\cdot))$, $C_7 = C_7(c,d,p(\cdot))$ and $\varrho = \varrho(c,d,p(\cdot)) < 1$ are finite constants.

Using (30) and (47), we get

$$\mathbf{E}_m(|V_t|^{2+\delta/2}) \le 2^{2+\delta/2}\mathbf{E}_m(|g(Y_t)|^{2+\delta/2}) \le |g_0|^{2+\delta/2}2^{3+\delta}(1+\mathbf{E}_m|Y_t|^{4+\delta})$$
$$\le C_8(1+|y_0|^{4+\delta}),$$

where $C_8 = C_8(c,d,p(\cdot),g_0)$ is a finite constant. This together with (48) and (49) entails

$$\mathbf{E}_m\Big\{\Big(\sum_{t\in\tau_n}(g(Y_t)-\mathbf{E}_m\{g(Y_t)\})\Big)^2\Big\}$$
$$\le C_9(1+|y_0|)^{\delta/(4+\delta)}(1+|y_0|^{4+\delta})^{4/(4+\delta)}\sum_{t,l\in\tau_n}\varrho^{|t-l|\delta/(4+\delta)}$$
$$\le C_{10}(1+|y_0|^{4+\delta})\sum_{t,l\in\tau_n}\varrho^{|t-l|\delta/(4+\delta)},$$

where the constants C_9, C_{10} do not depend on y_0. Since $\varrho < 1$, the off-diagonal terms in the last sum are exponentially decreasing, and this sum is of order card τ_n. Finally, we obtain (32) by applying Chebyshev's inequality.

References

Y. A. Davydov (1968), *Convergence of distributions generated by stationary stochastic processes*, Theory Probab. Appl. 13, 691–696.

P. Doukhan (1995), *Mixing*, Lecture Notes in Statist. 85, Springer, New York.

J.-C. Duan (1995), *The GARCH option pricing model*, Math. Finance 5, 13–32.

R. Engle (1982), *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation*, Econometrica 50, 987–1008.

C. Gouriéroux (1997), *ARCH Models and Financial Applications*, Springer, New York.

C. Gouriéroux and A. Monfort (1992), *Qualitative threshold ARCH models*, J. Econometrics 52, 159–199.

C. Gouriéroux, A. Monfort and C. Tenreiro (1995), *Kernel M-estimators and functional residuals plots*, CREST DP 9546.

L. de Haan (1990), *Fighting the ARCH-enemy with mathematics*, Statist. Neerlandica 44, 45–68.

C. Hafner (1998), *Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility*, Physica-Verlag, Heidelberg.

W. Härdle and A. Kneip (1999), *Testing a regression model when we have smooth alternatives in mind*, Scand. J. Statist. 26, 221–238.

W. Härdle and E. Mammen (1993), *Comparing nonparametric versus parametric regression fits*, Ann. Statist. 21, 1926–1947.

W. Härdle, V. Spokoiny and S. Sperlich (1997), *Semiparametric single-index versus fixed link function modelling*, ibid. 25, 212–243.

W. Härdle and A. Tsybakov (1997), *Local polynomial estimators of the volatility function in nonparametric autoregression*, J. Econometrics 81, 223–242.

W. Härdle, A. Tsybakov and L. Yang (1998), *Nonparametric vector autoregression*, J. Statist. Plann. Inference 68, 221–245.

A. Harvey, E. Ruiz and N. Shephard (1994), *Multivariate stochastic volatility models*, Rev. Econom. Stud. 61, 247–267.

F. Leblanc and O. Lepski (1996), *Test for symmetry of regression curves*, SFB373 DP 51, Humboldt-Univ. Berlin.

R. Mahieu and P. Schotman (1997), *An empirical application of stochastic volatility models*, DP Maastricht Univ.

S. P. Meyn and R. L. Tweedie (1992), *Stability of Markovian processes I: Criteria for discrete-time chains*, Adv. Appl. Probab. 24, 542–574.

V. V. Petrov (1995), *Limit Theorems of Probability Theory*, Clarendon Press, Oxford.

N. Shephard (1996), *Statistical aspects of ARCH and stochastic volatility*, in: *Time Series Models in Econometrics, Finance and Other Fields*, D. R. Cox *et al.* (eds.), Monographs Statist. Appl. Probab., 65, Chapman and Hall, 1–67.

V. Spokoiny (1997), *Testing a linear hypothesis using Haar transform*, Preprint WIAS, No. 314, Berlin.

S. J. Taylor (1986), *Modelling Financial Time Series*, Wiley, Chichester.

A. N. Tikhomirov (1980), *On the convergence rate in the central limit theorem for weakly dependent random variables*, Theor. Probab. Appl. 25, 790–809.

C. G. de Vries (1994), *Stylized facts of nominal exchange rate returns*, in: F. van der Ploeg (ed.), *Handbook of International Macroeconomics*, Blackwell, Oxford, 348–389.

D. Feldmann, W. Härdle, C. Hafner
SFB 373 and
Institut für Statistik und Ökonometrie
Wirtschaftswissenschaftliche Fakultät
Humboldt-Universität zu Berlin
Spandauer Str. 1
D-10178 Berlin, Germany
E-mail:

O. Lepski
Centre des Mathématiques
et d'Informatique
Université, Aix-Marseille 1
39, rue F. Joliot-Curie
F-13453 Marseille, France
E-mail:

M. Hoffmann
Laboratoire de probabilités
et modèles aléatoires
Université Paris 7
pl. Jussieu 2
F-75252 Paris, France
E-mail:

A. Tsybakov
Laboratoire de probabilités
et modèles aléatoires
Université Paris 6
pl. Jussieu 4, BP 188
F-75252 Paris, France
E-mail:

Received on 27.6.2002 (1645)