Modelling financial time series with SEMIFAR-GARCH model

Feng, Yuanhua and Beran, Jan and Yu, Keming

Heriot-Watt University, University of Konstanz and Brunel University

2006

Modelling financial time series with SEMIFAR-GARCH model

Yuanhua Feng
Department of Actuarial Mathematics and Statistics and
Maxwell Institute for Mathematical Sciences, Heriot-Watt University, UK

Jan Beran
Department of Mathematics and Statistics, University of Konstanz, Germany

and Keming Yu
Department of Mathematical Sciences, Brunel University, UK

Abstract

A class of semiparametric fractional autoregressive GARCH models (SEMIFAR-GARCH), which includes deterministic trends, difference stationarity and stationarity with short- and long-range dependence, and heteroskedastic model errors, is very powerful for modelling financial time series. This paper discusses the model fitting, including an efficient algorithm and parameter estimation of GARCH error term. So that the model can be applied in practice. We then illustrate the model and estimation methods with a few of different finance data sets.

Keywords: Financial time series, GARCH model, SEMIFAR model, parameter estimation, kernel estimation, asymptotic property.

1 Introduction

For some financial time series, several "trend generating" mechanisms may occur simultaneously. Semiparametric fractional autoregressive models (SEMIFAR) (Beran and Feng, 2002a, 2002b) has been introduced for modelling different components in the mean function of a financial time series simultaneously, such as nonparametric trends, stochastic nonstationarity, short- and long-range dependence as well as antipersistence. Let $d \in (-0.5, 0.5)$ be the fractional differencing parameter, $m \in \{0, 1\}$ be the integer differencing parameter, B be the backshift operator, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\psi(B) = 1 + \psi_1 B + \cdots + \psi_q B^q$ be polynomials in B with no common factors and all roots outside the unit circle, and ϵ_t be while noise, then SEMIFAR can be defined as

$$\phi(B)(1-B)^d\{(1-B)^m Y_t - g(x_t)\} = \psi(B)\epsilon_t.$$

SEMIFAR includes ARIMA $(p, m, 0)$ model and the fractional autogressive process (Hosking, 1981, Granger and Joyeux, 1980). However, the assumption of while noise on ϵ of SEMIFAR ignores possible heteroskedasticity of financial time series. Often financial time series exhibit conditional heteroskedasticity, i.e. the volatility (or conditional variance) of a financial process often depends on the past information but the mean may not. Well known models for modelling conditional heteroskedasticity are the autoregressive conditional heteroskedastic (ARCH, Engle, 1982) and generalized ARCH (GARCH, Bollerslev, 1986) models. Since then many extensions of the ARCH and GARCH models are introduced into the literature. Engle, Lilien and Robins (1987) extended the ARCH model to the ARCH in mean (or ARCH-M) model, where the conditional standard deviation also effects the mean of the observations. The ARCH-M model can be analogously generalized to a GARCH-M model. Another well known extension of the GARCH model is the exponential GARCH (EGARCH) introduced Nelson (1991), where the GARCH property is defined for the log-transformation of the volatility. A FARIMA-GARCH model to model long memory in the mean and conditional heteroskedasticity in the volatility is introduced by Ling and Li (1997).

However, there is little research on SEMIFAR-GARCH model except Beran and Feng (2001) which describes the model and derives the asymptotic normality of trend term estimation only. Some important problems for the practical implementation of this model, e.g., estimation of the unknown parameters and the development of a data-driven algorithm, were not discussed. In this paper we provide a full implementation of a SEMIFAR-GARCH model for financial time series. We will extend the SEMIFAR model to a SEMIFAR-GARCH model, which is the same as SEMIFAR model but under the additional assumption that the innovation process $\{\epsilon_i\}$ follows a GARCH model. Section 2 describes the SEMIFAR-GARCH model, which extends the FARIMA-GARCH model. Stochastic nonstationarity is also considered in the model. Section 3 designs a three-stage fitting algorithm for the SEMIFAR-GARCH mode. Section 4 discusses the asymptotic normality and consistency of the model functionals and parameters estimation. Section 5 provides an comprehensive algorithm for SEMIFAR-GARCH model. Section 6 illustrate the model and estimation methods by the analysis of three different finance data sets and examples. Section 7 concludes the paper.

2 The Model

In the following, the notation SEMIFAR also stands for a slight generalization of the SEMIFAR model with an additional MA (moving average) component in the short-range dependence part. Similarly to the SEMIFAR model, the SEMIFAR-GARCH model is defined by

$$\phi(B)(1-B)^d\{(1-B)^m Y_t - g(x_t)\} = \psi(B)\epsilon_t \tag{2.1}$$

with

$$\epsilon_t = z_t h_t^{\frac{1}{2}}, \quad h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j h_{t-j}, \tag{2.2}$$

where $d \in (-0.5, 0.5)$ is the fractional differencing parameter, $m \in \{0, 1\}$ is the integer differencing parameter, $x_t = t/n$ is the re-scaled time, $g : [0, 1] \to \Re$ is a smooth function, z_t are i.i.d. standard normal random variables, $\alpha_0 > 0$, $\alpha_1, ..., \alpha_r, \beta_1, ..., \beta_s \geq 0$, $d \in (-0.5, 0.5)$, B is the backshift operator, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\psi(B) = 1 + \psi_1 B + \cdots + \psi_q B^q$ are polynomials in B with no common factors and all roots outside the unit circle. The fractional differencing operator $(1-B)^d$ is the same as defined before. For $m = 0$, model (2.1) and (2.2) may be thought of as an extension of model (7) and (8) in Ling and Li (1997) by replacing the constant mean with a nonparametric trend function.

As in the SEMIFAR model, the two differencing parameters m and d may be summarized in one parameter $\delta = m + d$. The innovation process defined in (2.2) follows a GARCH model (Bollerslev, 1986). It is assumed that $\sum_{j=1}^{r} \alpha_j + \sum_{j=1}^{s} \beta_j < 1$, which ensures that there exists a strictly and second order stationary solution ϵ_t of (2.2) with variance

$$\sigma_\epsilon^2 = \mathrm{var}\,(\epsilon_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{r} \alpha_i - \sum_{j=1}^{s} \beta_j} \tag{2.3}$$

(see Theorem 2 in Bollerslev, 1986).

For the derivation of the asymptotic properties it is further assumed that $E(\epsilon_i^4) < \infty$, which implies the above condition $\sum_{j=1}^{r} \alpha_j + \sum_{j=1}^{s} \beta_j < 1$. A necessary and sufficient condition which guarantee the existence of the $2m$-th moments for the special case of a GARCH(1, 1) model with normal innovations z_t was also found in Bollerslev (1986). Necessary and sufficient conditions which guarantee the existence of higher order moments of a GARCH model in more general cases may be found in Ling and McAleer (2002) (see also Ling and Li 1997, Chen and An 1998 and He and Teräsvirta 1999a). Model (2.1) and (2.2) are a variety of the SEMIFAR model by replacing the i.i.d. innovations there with the GARCH innovations defined in (2.2).

Denote by $\theta = (\sigma_\epsilon^2, \delta, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q, \alpha_0, \alpha_1, ..., \alpha_r, \beta_1, ..., \beta_s)^{\mathrm{T}} = (\eta^{\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}}$ the parameter vector, where $\eta = (\sigma_\epsilon^2, \delta, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q)^{\mathrm{T}}$ is the parameter vector for the FARIMA part of the process and $\lambda = (\alpha_0, \alpha_1, ..., \alpha_r, \beta_1, ..., \beta_s)^{\mathrm{T}}$ for the GARCH part. Also denote by $\theta^* = (\sigma_\epsilon^2, d, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q, \alpha_0, \alpha_1, ..., \alpha_r, \beta_1, ..., \beta_s)^{\mathrm{T}}$ and $\eta^* = (\sigma_\epsilon^2, d, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q)^{\mathrm{T}}$, which are the same as θ and η respectively but with δ being replaced by d. Denote the unknown value of θ^* by $\theta_0^* = (\phi_1^0, ..., \phi_p^0, \psi_1^0, ..., \psi_q^0, d^0, \alpha_0^0, \alpha_1^0, ..., \alpha_r^0, \beta_1^0, ..., \beta_s^0)^{\mathrm{T}}$, which is assumed to be in the interior of a compact set Θ^*. Note however, under model (2.1) and (2.2), σ_ϵ^2 is determined by (2.3) and is hence not an independent unknown parameter. We aim to develop efficient estimation for g (or its derivatives) and θ from the observations $y_1, ..., y_n$. This will be discussed in details next section.

3 The Semiparametric Estimation Procedure

In this section we propose to estimate the SEMIFAR-GARCH model in three stages: Firstly, estimate the trend function g nonparametrically; secondly, estimate the FARIMA parameter vector η from the residuals; and thirdly, to estimate the GARCH parameter vector λ from the inverted innovations obtained from the residuals by means of a FARIMA model with $\hat{\eta}$. Under the three steps, nonparametric estimators of $g^{(\nu)}$, the ν-th derivatives of g, can also be carried out after Step 2 by replacing the unknown parameter vector η with $\hat{\eta}$.

This semiparametric estimation procedure is proposed based on the following lemmas.

Lemma 1. *Assume that Y_t is a stationary FARIMA-GARCH process as defined in Ling and Li (1997), i.e. model (2.1) and (2.2) holds with $m = 0$ and $g(x) \equiv \mu$, where μ is an unknown constant. Assume further that var$(\epsilon_t) = E(\epsilon_t^2) = \sigma_\epsilon^2$ is an unknown constant and that $E(\epsilon_t^4) < \infty$. Then the asymptotic properties of the MLE of η are independent of the unknown GARCH parameter vector λ.*

Lemma 1 is a straightforward consequence of Theorem 3.2 of Ling and Li (1997). This lemma shows that η and λ can be estimated separately from the original data. Hence the FARIMA parameters in the FARIMA-GARCH model can be estimated at first using a proper package such as S+GARCH (e.g. the S-PLUS function *arima.fracdiff*). Let $\hat{g}$ be the estimator of g, the following lemma shows that the asymptotic properties of $\hat{g}(x)$ are independent of λ.

Lemma 2. *Assume that $g(x)$ in (2.1) is at least $(p+1)$-times differentiable and that the other conditions of Lemma 1 hold. Then the asymptotic properties of a local polynomial estimator of $g^{(\nu)}$ $(\nu \leq p)$ are independent of the unknown GARCH parameter vector λ.*

Lemma 2 is a consequence of the results in Theorems 5 and 6 of Beran and Feng (2001). See also Theorem 2 below. Lemmas 1 and 2 together show that

Lemma 3. *Let Y_t be defined by (2.1) and (2.2) with the GARCH innovations such that var$(\epsilon_t) = E(\epsilon_t^2) = \sigma_\epsilon^2$ and $E(\epsilon_t^4) < \infty$. Assume that other regularity conditions on the FARIMA model, the bandwidth and the smoothness of g hold. Then a SEMIFAR algorithm can be directly used for estimating $g^{(\nu)}$ and η in the SEMIFAR-GARCH model without changing the asymptotic properties of these estimators.*

Assume that g is at least $(p+1)$-times differentiable. Following Beran and Feng (2001), we propose to estimate g with a p-th order local polynomial or a k-th order kernel method with $k = p+1$. Detailed description on this approach may be found in that work and will be omitted here to save space. The trend function can also be estimated following other nonparametric approaches, e.g., smoothing splines. Note that the SEMIFAR-GARCH model may be rewritten as a semiparametric regression model with the FARIMA-GARCH error process. Following (2.1) and (2.2) we have, for $m = 0$

$$Y_t = g(x_t) + \xi_t, \quad t = 1, ..., n, \tag{3.1}$$

and for $m = 1$

$$U_t = g(x_t) + \xi_t, \quad t = 2, ..., n, \tag{3.2}$$

where $U_t = Y_t - Y_{t-1}, \quad t = 2, ..., n$, and

$$\xi_t = (1 - B)^{-d} \phi^{-1}(B) \psi(B) \epsilon_t \tag{3.3}$$

is a FARIMA-GARCH process, where ϵ_t are the GARCH innovations as defined in (2.2).

Let $\hat{g}(x_t) = \hat{g}(x_t; m)$ denote the kernel estimator of g obtained from (3.1) or (3.2) for $m = 0$ and $m = 1$ respectively with the bandwidth h, where it is assumed that $h \to 0$, $nh \to \infty$ as $n \to \infty$. Consider now ϵ_t as a function of η. For given p, q and a trial value of $\eta = (d, \phi_1, ..., \phi_p, \psi_1, ..., \psi_q)^{\mathrm{T}}$, where $\delta = m + d$. Denote by

$$e_t(\eta) = \sum_{i=0}^{t-m-2} a_i(\eta)[c_i(\eta)Y_{t-i} - \hat{g}(x_{t-i}; m)] \tag{3.4}$$

the (approximate) residuals. Although ϵ_t in the SEMIFAR-GARCH model is non-Gaussian, the approximate maximum likelihood estimator proposed by Beran (1995) also applies under the assumption $E(\epsilon_t^4) < \infty$, because now ϵ_t and $\epsilon_t^2 - h_t$ are both martingale-differences (see also the results on the parameter estimation in the FARIMA-GARCH model given by Ling and Li, 1997). In this case the estimator is indeed a quasi maximum likelihood estimate. For given p and q, $\hat{\eta}$ is estimated from e_t by minimizing

$$S_n(\eta) = \frac{1}{n} \sum_{i=m+2}^{n} e_i^2(\eta) \tag{3.5}$$

w.r.t. $\hat{\eta}$. Let $\hat{m} = [\hat{\delta} + 0.5]$ and $\hat{d} = \hat{\delta} - \hat{m}$. This procedure can be carried out for $p = 0, 1, ..., P$ and $q = 0, 1..., Q$, where P and Q are the maximal orders of the AR and MA parts, which will be considered here. Then p and q can be selected following the BIC rule

$$\hat{p} = \arg\min\{\mathrm{BIC}(p, q); p = 0, 1, ..., P\}$$

and

$$\hat{q} = \arg\min\{\mathrm{BIC}(p, q); q = 0, 1, ..., Q\},$$

where

$$\mathrm{BIC}(p, q) = n \log(\hat{\sigma}^2(p, q) + (\log n)(p + q), \tag{3.6}$$

and $\hat{\sigma}_\epsilon^2$ is the estimate of var (ϵ_t) given by

$$\hat{\sigma}_\epsilon^2 = \frac{1}{n} \sum_{t=m+2}^{n} e_t^2(\hat{\eta}). \tag{3.7}$$

It is well known that $\hat{p}$ and $\hat{q}$ obtained in this way are consistent; see the details from Beran, Bhansali and Ocker (1998).

Now, assume that $\hat{\eta}$ is a consistent estimator of η, then the $e_t(\hat{\eta})$ are approximations of the unobservable GARCH innovations ϵ_t. The parameter vector λ can be estimated from $e_t(\hat{\eta})$

following the standard maximum likelihood method for a GARCH model. Following Bollerslev (1986), the (unobservable) conditional Gaussian log-likelihood function based on ϵ_t is given by (ignoring constants)

$$L^*(\lambda) = \frac{1}{n}\sum_{t=1}^{n} l_t, \qquad \text{where} \qquad l_t = -\frac{1}{2}\ln(h_t(\epsilon;\lambda)) - \frac{\epsilon_t^2}{2h_t(\epsilon;\lambda)}. \tag{3.8}$$

Denote by λ^* the maximizer of $L^*(\lambda)$. Note however that λ^* is not available. Hence we define the approximate log-likelihood function in the current context by

$$\hat{L}(\lambda) = \frac{1}{n}\sum_{t=1}^{n} l_t, \qquad \text{where} \qquad l_t = -\frac{1}{2}\ln(h_t(e(\hat{\eta});\lambda)) - \frac{e_t^2}{2h_t(e(\hat{\eta});\lambda)}, \tag{3.9}$$

where $\hat{\eta}$ is as defined above. Similar to the estimation of $\hat{\eta}$, the proposed approximate MLE of λ is $\hat{\lambda}$, the maximizer of $\hat{L}(\lambda)$. The symbols $h_t(\epsilon;\lambda)$ and $h_t(e(\hat{\eta});\lambda)$ are used to indicate that, for given value of λ, these functions also depend on the innovations or their approximations. Given e_t, $\hat{\lambda}$ can be calculated using a standard package for estimating the GARCH model simply by replacing ϵ_t with e_t. In this work the S+GARCH package will be used.

4 Asymptotic Results

The asymptotic behavior of $\hat{g}^{(\nu)}$ under model (3.1) were studied by Beran and Feng (2001). Part of the asymptotic results on $\hat{g}^{(\nu)}$ are represented in the following. The analysis given in the following involves infinite past history of Y_t and ϵ_t. For simplicity, we assume that the presample values of Y_t and ϵ_t are zero, and choose the presample values of h_t and ϵ_t^2 to be $\sum_{m+2}^{n}\hat{e}_t^2/n$. This simplification will not affect the asymptotic properties of the proposed estimators.

Theorem 1. *Let Y_i follow the semiparametric regression model (3.1), where the errors ξ_t are generated by (3.3) with innovation process $\{\epsilon_t\}$ following the GARCH model (2.2), which is assumed to be strictly stationary such that $E(\epsilon_t^4) < \infty$. Let kernel K be a symmetric probability density having compact support [-1 1]. $\hat{g}^{(\nu)}(t)$ $(\nu \le p)$ is obtained by solving the locally weighted least squares problem*

$$Q = \sum_{i=1}^{n}\left\{Y_i - \sum_{j=1}^{p} b_j(t_i - t)^j\right\}^2 K\left(\frac{t_i - t}{b}\right).$$

Under the regularity conditions on continuity of $g^{(\nu)}$, the following results hold.

i) Let $t = ch$ with $0 \le c \le 1$. For all $d \in (-0.5, 0.5)$, assume that $nb^{(2k+1-2d)/(1-2d)} \to \Delta^2$ as $n \to \infty$, for some $\Delta > 0$, then

$$(nb)^{1/2-d} b^{\nu} (\hat{g}^{(\nu)}(t) - g^{(\nu)}(t)) \xrightarrow{\mathcal{D}} N(\Delta D, V(c,d)), \tag{4.1}$$

where $D = \frac{g^{(k)}(t)\beta_c}{k!}$, and b is the bandwidth. $V(c,d) = lim_{n\to\infty} V_n(c, \delta, b)$ exists with

$$V_n(c,\delta,b) = (nb)^{-1-2\delta} \sum_{n_0-n_c}^{n_0+n_1} K_{(\nu,k,c)}(\frac{t_i - t}{b}) K_{(\nu,k,c)}(\frac{t_i - t}{b}) \gamma(i-j),$$

and $K_{(\nu,k,c)}$ is the asymptotically equivalent boundary kernel for estimating $g^{(\nu)}$ (Ruppert and Wand, 1994). $\beta_c = \int h^j K_{(\nu,j,c)}(u)du$.

ii) The asymptotically optimal bandwidth that minimizes the asymptotic MISE is given by

$$h_{\mathrm{A}} = C_{\mathrm{A}} n^{(2d-1)/(2k+1-2d)} \tag{4.2}$$

with

$$C_{\mathrm{A}} = \left[\frac{2\nu + 1 - 2d}{2(k-\nu)} \frac{[k!]^2 V}{I(g^{(k)})\beta^2_{(\nu,k)}} \right]^{1/(2k+1-2d)}, \tag{4.3}$$

where it is assumed that $I(g^{(k)}) > 0$ and $V = V(1,d) = V(d)$ and $\beta = \beta_c$ with $c = 1$.

We now check that some sufficient conditions on the asymptotic normality are fulfilled. This is ensured by the following lemma.

Lemma 4. *Let ξ_t be generated by (3.3) with $d \in (-0.5, 0.5)$. Assume that the innovation process $\{\epsilon_t\}$ is generated by the GARCH model (2.2), which is strictly stationary such that $E(\epsilon_t^4) < \infty$. And assume further that $\phi(B)$ and $\psi(B)$ have no common factors and all roots of them lie outside of the unit circle. Then, for the sample mean $\bar{\xi}$ of ξ_t, we have*

$$n^{1/2-d}\bar{\xi} \xrightarrow{\mathcal{D}} N(0, V(d)),$$

where

$$V(d) = \sigma_\epsilon^2 \frac{|\psi(1)|^2}{|\phi(1)|^2} \frac{\Gamma(1-2d)}{(2d+1)} \frac{\sin(\pi d)}{\pi d}. \tag{4.4}$$

Lemma 4 shows that the sample mean $\bar{\xi} = \frac{1}{n}\sum_{t=1}^{n} \xi_t$ of a FARIMA-GARCH process defined in (3.3) is asymptotically normal, if $E(\epsilon_t^4) < \infty$, which extends the results of Theorem 8 *ii*) in Hosking (1996) to nonstationary processes. Under the condition $E(\epsilon_t^4) < \infty$, we have $\sum_{j=1}^r \alpha_j + \sum_{k=1}^s \beta_k < 1$ (see Lemma 2.2 in Chen and An 1998) and that ϵ_t is a square integrable martingale-difference w.r.t $(\mathcal{F}_t, t \in \{-\infty, ..., -1, 0, 1, ..., \infty\})$, where $\mathcal{F}_t$ is the σ-field generated by the information in the past. And hence ϵ_t is an uncorrelated white noise. The autocovariance function $\gamma_\xi(k)$ of the FARIMA-GARCH process ξ_t is given in Beran (1994). Furthermore, He and Teräsvirta (1999a) showed that, under the condition $E(\epsilon_t^4) < \infty$, the autocorrelation function of the squared process ϵ_t^2 decays exponentially. This is easy to understand, because now the squared process ϵ_i^2 is itself a second order stationary process having an ARMA representation with all roots of its characteristic polynomials lying outside the unit circle. See equations (6) and (7) in Bollerslev (1986). More detailed results on this topic may be found in He and Teräsvirta (1999b) for second order GARCH models.

The asymptotic properties of the estimation of the FARIMA parameter vector η in the SEMIFAR-GARCH model are the same as those of the corresponding parameter estimates in an extended SEMIFAR model with a MA component in the short-range dependent part. Further we could prove the consistency of $\hat{m}$ as well as $\hat{\eta}$,

Theorem 2. *Assume that $\{\epsilon_t\}$ is a GARCH process defined by (2.2) with $E(\epsilon_t^4) < \infty$ and that the conditions of Theorem 1 hold. Then we have*

i) $\hat{m} \xrightarrow{p} m^0$, *provided* $b = O(n^\alpha)$ *with* $0 < \alpha < 1$ *such that* $(p+1)\alpha + d > 0$ *and*

ii) $\sqrt{n}(\hat{\eta}^* - \eta_*^0) \xrightarrow{\mathcal{D}} N(0, \Sigma)$, *if* $0 < \alpha < 1/2$ *such that* $(p+1)\alpha + d > 1/4$, *where*

$$\Sigma = 2D^{-1} \tag{4.5}$$

is as defined in Theorem 1 in Beran (1995).

Theorem 2 shows that, under suitable conditions on the bandwidth and other regularity conditions, $\hat{\eta}$ is always $\sqrt{n}$-consistent. In this case the effect of $\hat{\eta}$ on $\hat{\lambda}$ is negligible. In the following, we will assume that the stronger conditions on the bandwidth as stated in Theorem 2 *ii*) hold, so that the error in $\hat{\eta}$ does not have any effect on $\hat{\lambda}$. Under this condition, $\hat{\eta}$ can be simply replaced by the true unknown vector η^0 to simplify the representation given below. We define

$$\tilde{e}_t(\eta^0) = \sum_{i=0}^{t-1} a_i(\eta^0)[(1-B)^{m^0} Y_{t-i} - g(x_{t-i})], \tag{4.6}$$

which are also not observable, because g is unobservable. Let

$$\Omega_\lambda = E\left[\frac{1}{2h_t^2(\tilde{e}(\eta^0);\lambda)} \frac{\partial h_t(\tilde{e}(\eta^0);\lambda)}{\partial \lambda} \left(\frac{\partial h_t(\tilde{e}(\eta^0);\lambda)}{\partial \lambda}\right)^{\mathrm{T}}\right] \tag{4.7}$$

and Ω_0, the value of Ω_λ at $\lambda = \lambda^0$, denote the information matrix. For the proposed approximate MLE of the GARCH parameter vector, $\hat{\lambda}$, we have

Theorem 3. *Assume that the conditions of Theorem 2 ii) hold. Then we have*

i) There exists a MLE $\hat{\lambda}$ *satisfying* $\partial \hat{L}(\lambda)/\partial \lambda = 0$ *and* $\hat{\lambda} \xrightarrow{p} \lambda^0$ *as* $n \to \infty$.

ii) $\sqrt{n}(\hat{\lambda} - \lambda^0) \xrightarrow{\mathcal{D}} N(0, \Omega_0^{-1})$, *where* Ω_0 *is as defined above.*

Now, define

$$\tilde{L}(\lambda) = \frac{1}{n}\sum_{t=1}^{n} l_t, \qquad \text{where} \qquad l_t = -\frac{1}{2}\ln(h_t(\tilde{e};\lambda)) - \frac{\tilde{e}_t^2}{2h_t(\tilde{e};\lambda)}. \tag{4.8}$$

Denote by $\tilde{\lambda}$ the maximizer of $\tilde{L}(\lambda)$, which is again not available, since $\tilde{e}_t(\eta^0)$ are unknown. Following the results in Ling and Li (1997), $\tilde{\lambda}$ is $\sqrt{n}$-consistent. Hence results given in Theorem 3 will hold, if we can show that $\hat{\lambda} - \tilde{\lambda} = o_p(n^{-1/2})$. Note that the conditions on the GARCH model ensure that λ^0 is in the interior of a compact set Λ. To prove Theorem 3, we will introduce the following lemmas, which are required to calculate the difference between $\hat{\lambda}$ and $\tilde{\lambda}$.

Lemma 5. *Under the assumptions of Theorem 3 we have*

$$h_t(e(\eta^0);\lambda) - h_t(\tilde{e}(\eta^0);\lambda) \doteq O_p(e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0)) \quad \forall\, \lambda \in \Lambda. \tag{4.9}$$

Lemma 5 gives an interesting results for quantifying the order of magnitude of the difference between the estimates of the conditional variance with the two approximations of the innovations, $h_t(e(\eta^0);\lambda) - h_t(\tilde{e}(\eta^0);\lambda)$, which shows that this order is the same as that of $(e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0))$.

The following lemma extends the results of Lemma 5 to quantify the order of magnitude of the difference between the first derivatives of h_t obtained using the two different approximations of ϵ_t, i.e. e_t and $\tilde{e}_t$, respectively.

Lemma 6. *Under the assumptions of Theorem 3 we have,* $\forall\, \lambda \in \Lambda$*, the first element of*

$$\frac{\partial h_t(e(\eta^0); \lambda)}{\partial \lambda} - \frac{\partial h_t(\tilde{e}(\eta^0); \lambda)}{\partial \lambda}$$

is zero and the other elements of it are all of the order $O_p(e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0))$*.*

5 Data-driven Algorithms

Based on the asymptotic results obtained in the last section the following algorithm in S-PLUS is proposed for the practical implementation of the SEMIFAR-GARCH model.

1. Carry out one of the data-driven SEMIFAR algorithms, e.g., AlgB in Beran and Feng (2002b), to the observations to obtain $\hat{g}(x_t)$ and $\hat{\eta}$;
2. Calculate the residuals $r_t = y_t - \hat{g}(x_t)$ and invert r_t using $\hat{\eta}_t$ into $\hat{\epsilon}_t$, the approximations of ϵ_t;
3. For $r = 0, 1, ..., r_{\max}$ and $s = 0, 1, ..., s_{\max}$, estimate $\hat{\lambda}(r, s)$ using S+GARCH and calculate $\text{BIC}(r, s)$.
4. Choose the couple $\{\hat{r}, \hat{s}\}$ that minimizes the BIC. We obtain the fitted GARCH model.

Where the BIC will be used to select the orders of the GARCH model, while the definition of the BIC in S+GARCH will be used, which is given by

$$\text{BIC}(r, s) = -2 \log(\text{maximized likelihood}) + (\log n)(r + s + 2). \tag{5.1}$$

Remark 1. *The estimated parameter vectors for the FARIMA and the GARCH models are asymptotically independent. In the case without a trend function, these two models can hence be selected either separately or jointly. In the SEMIFAR-GARCH model it is however inconvenient, if we want to select the two models at the same time. Hence they are selected separately.*

It is easy to show that the results of Theorems 1 through 3 hold for $\hat{g}(x_t)$, $\hat{\eta}$ and $\hat{\lambda}$ obtained following the above algorithm. Furthermore, all results on the selected bandwidth as given in theorems in Beran and Feng (2001) hold for the bandwidth selected following this algorithm, since these results are independent of the GARCH parameter vector λ. Details on these results will be omitted to save space. Simulation studies on this algorithm were also not carried out, because the first step of this algorithm is exactly a SEMIFAR algorithm and the other steps are simply a procedure for fitting a parametric GARCH model from the approximated innovations. If for a data set it happens to be $\hat{r} = 0$ and $\hat{s} = 0$, then the fitted model reduces to a SEMIFAR model.

6 Applications

In this section the proposal will be applied to modelling some well known financial time series. For all examples log-transform of the original observations will be used. By doing this, the residuals of the SEMIFAR model stand automatically for the (trend adjusted) log-returns. The proposed algorithm is applied in this section to some data examples. For estimating the SEMIFAR model, the AlgB in Beran and Feng (2002b) is used. The trend is estimated by local linear regression using the Epanechnikov kernel as weight function. For the short-memory part, only an AR component is considered as in the original SEMIFAR model. The AR model is chosen from $p = 0, 1, ..., 5$, and the GARCH model from $r = 0, 1, 2$ and $s = 0, 1, 2$, by means of the BIC.

Figure 1(a) shows that the log-transformation of the time series of the daily world copper price from January 03, 1995 to September 30, 2003, downloads from the web site of the London Metal Exchange. It is expected that the errors of such a price time series are antipersistent. The selected order of the autoregressive part is $\hat{p} = 0$, i.e. there is no significant short-range dependence in this time series. The fitted SEMIFAR results show that this time series is integrated with a significant nonparametric drift (Figure 1(a)). And the residuals are significantly antipersistent. Figure 1(b) shows the estimated innovations ($\hat{\epsilon}_t$) obtained by inverting the residuals. We can see that there is clear conditional heteroskedasticity in this series. Further calculations show that all fitted GARCH models are strongly significant. A GARCH(1, 2) model with

$$\hat{h}_t = 1.082 \cdot 10^{-5} + 0.1238\hat{\epsilon}_{t-1}^2 + 0.2507\hat{h}_{t-1} + 0.5695\hat{h}_{t-2} \quad (6.1)$$

was selected following the BIC. The estimated conditional standard deviations ($\hat{h}_t^{1/2}$) and the standardized innovations ($\hat{z}_t := \hat{\epsilon}_t/\hat{h}_t^{1/2}$) are shown in Figures 1(c) and 1(d). The series in Figure

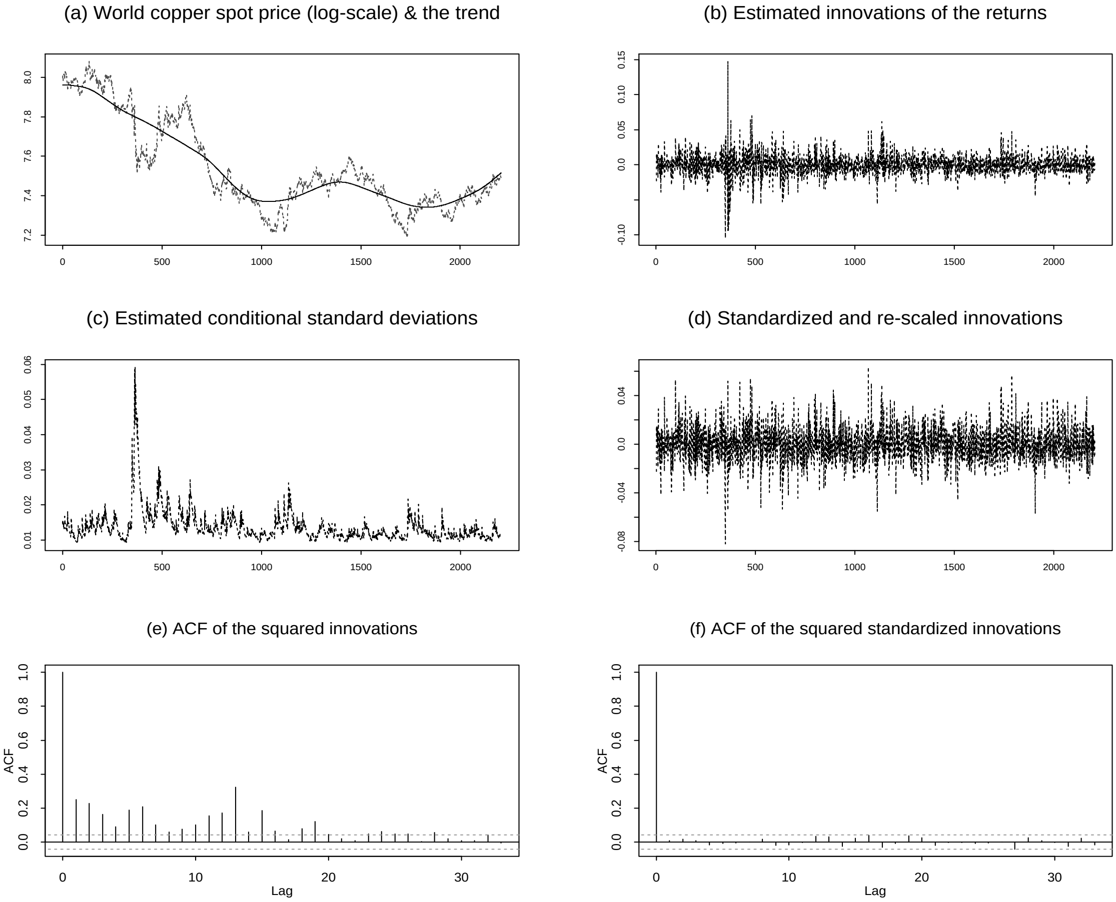

Figure 1: The daily world copper price (log-scaled) and the trend (a), the inverted innovations (b), the GARCH conditional SD (c), the standardized innovations (d) and acf's of the squared, nonstandard and standardized innovations (e and f).

1(d) is re-scaled with corresponding sample standard deviation so that it is comparable with that shown in Figure 1(b). Figures 1(e) and 1(f) show the autocorrelations of the squared series $\hat{\epsilon}_t^2$ and $\hat{z}_t^2$. We see that $\hat{\epsilon}_t^2$ are clearly correlated but $\hat{z}_t^2$ are almost uncorrelated, which shows the goodness of the fitted model.

The other two examples are the log-transformed series of the daily Standard and Poor 500 (S&P 500) Index from January 01, 1997 to August 23, 2000 and the series of the daily exchange rates between Euro and US Dollar (Euro/USD) from January 04, 1999 to October 31, 2003. For the S&P 500 series only observations in a relatively short time period are used to avoid possible nonstationarity in the variance/covariance in this series. The fitted results show that both of these two series, like for the first example, are integrated with a significant nonparametric trend. The selected order of the autoregressive part is again $\hat{p} = 0$. The long-range dependence in the third example is not significant and just slightly significant in the second example. For both series, a GARCH(1, 1) model was selected from the estimated innovations. The fitted GARCH conditional variance is

$$\hat{h}_t = 1.132 \cdot 10^{-5} + 0.0948\hat{\epsilon}_{t-1}^2 + 0.8308\hat{h}_{t-1} \tag{6.2}$$

for the S&P 500 series, and

$$\hat{h}_t = 6.387 \cdot 10^{-7} + 0.0196\hat{\epsilon}_{t-1}^2 + 0.9649\hat{h}_{t-1} \tag{6.3}$$

for the Euro series. Figures 2 and 3 show the same results as those given in Figure 1 for these two examples respectively. From Figure 3 we see that the GARCH effect in the Euro series is not clear. This means that the Euro/USD exchange rates can well be modelled by a SEMIFAR model with no short- or long-range dependence but with a clearly significant, nonparametric trend. Furthermore, it can be shown that the marginal distribution of the Euro/USD exchange rates series is not far from a normal distribution.

The selected bandwidth $\hat{h}$, the estimates $\hat{m}$, $\hat{d}$ together with the 95%-confidence intervals of them and other statistics are summarized in Table 1.

Table 1: Estimation results for all examples

Series	$\hat{h}$	$\hat{m}$	$\hat{d}$ & 95%-CI	$\hat{p}$	$\hat{r}$	$\hat{s}$	trend
Copper	0.1405	1	-0.0819 [-0.1146, -0.0492]	0	1	2	sign.
S&P 500	0.2592	1	-0.0590 [-0.1088, -0.0091]	0	1	1	sign.
Euro/USD	0.1279	1	-0.0007 [-0.0448, 0.0433]	0	1	1	sign.

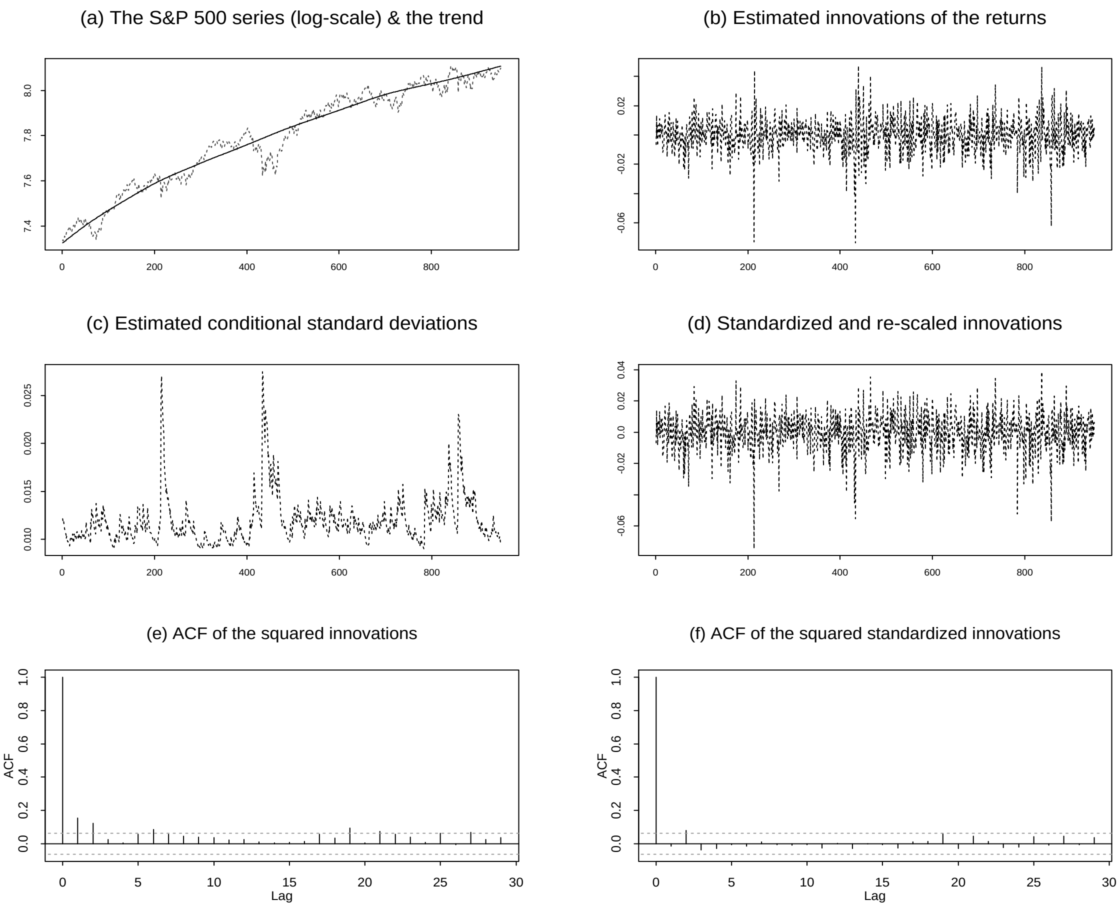

Figure 2: The same results as shown in Figure 1 but for the S&P 500 series.

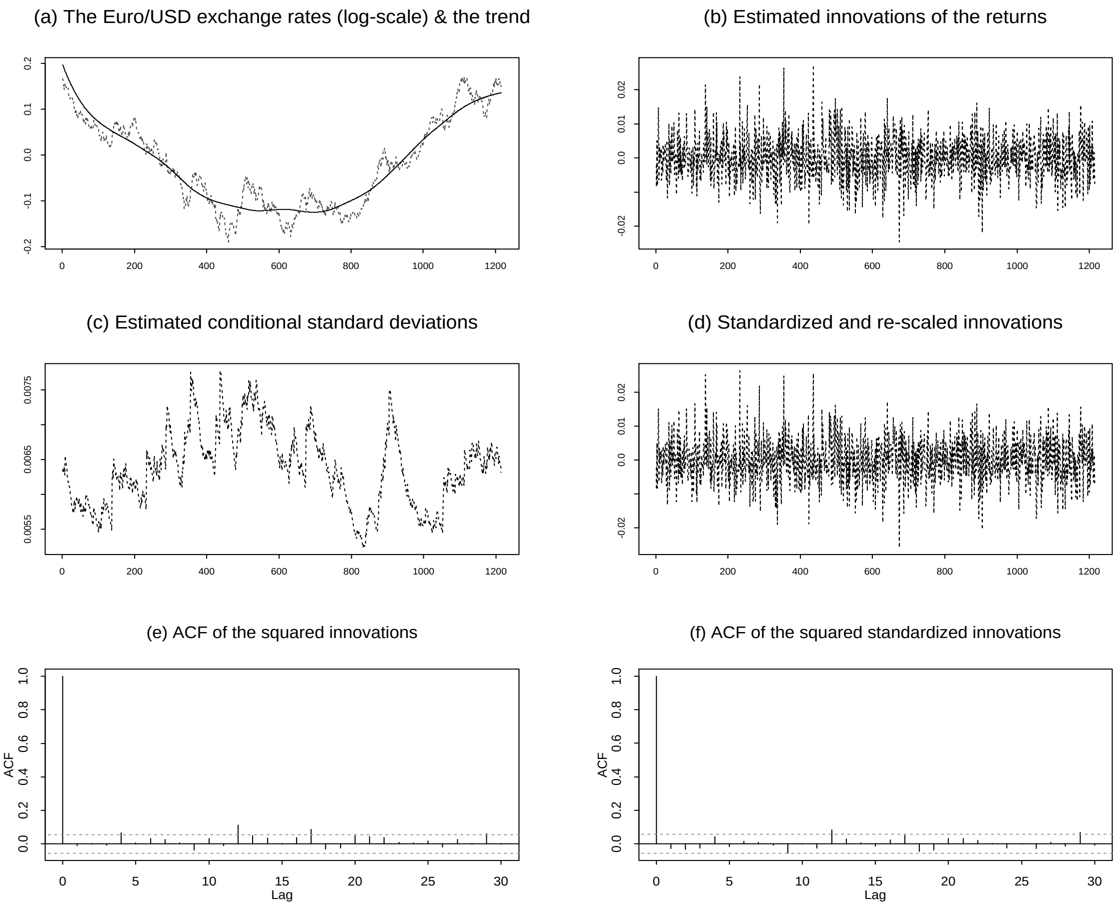

Figure 3: The same results as shown in Figure 1 but for the Euro/USD series.

7 Conclusions

This paper extends the SEMIFAR model to a SEMIFAR-GARCH model, so that conditional heteroskedasticity in financial time series can also be modelled by the SEMIFAR model. A semiparametric estimation procedure is proposed. Asymptotic results on the SEMIFAR model are extended to the current proposal. It is shown in particular that the same asymptotic results obtained in Beran and Feng (2001) for the SEMIFAR model with i.i.d. normal innovations hold for the SEMIFAR-GARCH model under the much weaker condition that the GARCH innovation process has finite fourth moments. These theoretical results and the important property that the estimates of the FARIMA and GARCH parameter vectors are independent of each other, allow us to apply the data-driven SEMIFAR algorithms to estimate the trend and the FARIMA parameters in the SEMIFAR-GARCH model. It is proposed to estimate the GARCH parameter from the approximated GARCH innovations calculated by inverting the final residuals. Data examples show that the proposed algorithm works well. Further extensions of the SEMIFAR model are also possible. For instance, a seasonal component can also be introduced into the mean function to model daily periodicity in high-frequency financial data.

Note that the SEMIFAR-GARCH model only has long memory in the mean but does not have long memory in the volatility. Bailie et al. (1995, 1996) introduced the FIGARCH (fractionally integrated GARCH) process for modelling long memory in the volatility. However, the FIGARCH is not second order stationary and is not considered as error process in this work. A stationary process with long memory in the volatility is the fractional LARCH (linear ARCH, Robinson, 1991 and Giraitis, et al., 2004) model. Hence nonparametric regression with fractional LARCH errors should be studied so that long memory in the volatility of a financial time series can be modelled.

Appendix: Proofs of results

Proof of Lemma 4. The formula of the asymptotic variance of $\bar{\xi}$ remains unchanged from case to case, if only the ϵ_t are uncorrelated $(0, \sigma^2)$ random variables. Hence it is the same as that for i.i.d. innovations given by Theorems 1 and 8 of Hosking (1996), i.e. var $(\bar{\xi}) = n^{2d-1}V(d)$ for $-\frac{1}{2} < d < \frac{1}{2}$, where

$$V(d) = \sigma_\epsilon^2 \frac{|\psi(1)|^2}{|\phi(1)|^2} \frac{\Gamma(1-2d)}{(2d+1)} \frac{1}{\Gamma(1+d)\Gamma(1-d)}.$$

Using the relationships $\Gamma(1+d) = d\Gamma(d)$ and $\Gamma(d)\Gamma(1-d) = \frac{\pi}{\sin(\pi d)}$ (for $d \in (-0.5, 0.5) \setminus \{0\}$), we obtain the alternative representation of $V(d)$

$$V(d) = \sigma_\epsilon^2 \frac{|\psi(1)|^2}{|\phi(1)|^2} \frac{\Gamma(1-2d)}{(2d+1)} \frac{\sin(\pi d)}{\pi d},$$

which is used in this work.

Since ξ_t defined in (3.3) is a zero mean FARIMA process with innovations ϵ_i following a GARCH model, we have

$$X_i = \sum_{k=0}^{\infty} c_k \epsilon_{i-k} \tag{A.1}$$

with $c_k \sim \frac{|\psi(1)|}{|\phi(1)|} k^{d-1}$ as $n \to \infty$ (see Beran, 1994). Hence, for $-0.5 < d < 0.5$, $\sum_{k=0}^{\infty} c_k^2 < \infty$. This shows that X_i fulfills the conditions of Theorem 4 of Beran and Feng (2001), and so

$$(\xi_1 + \cdots + \xi_n)/\sigma_n \xrightarrow{\mathcal{D}} N(0, 1).$$

Observe that

$$[n^{1/2-d}\bar{\xi} - (\xi_1 + \cdots + \xi_n)/\sigma_n] \xrightarrow{p} 0,$$

following Theorem 4 of Beran and Feng (2001) we have

$$n^{1/2-d}\bar{\xi} \xrightarrow{D} N(0, 1).$$

◇

Proof of Theorem 1 *i***).** Following Lemma 4 and noting that the weights w_i of a local polynomial estimator satisfy the conditions of Theorem 5 of Beran and Feng (2001), the asymptotic normality of $\hat{g}^{(\nu)}(x)$ follows from there.

A sketched proof of Theorem 2. *i*). Note in particular that the necessary condition so that the consistency of $\hat{m}$ shown in the proof of Theorem 7.2 in Feng (2004) holds is that $\hat{\eta}$

is consistent in the case with $m = m^0$. This is ensured by the condition on the bandwidth in Theorem 2 *i*) and the further assumption $E(\epsilon^4) < \infty$. This shows that $\hat{m}$ is consistent under the assumptions of Theorem 2.

ii). To show the results given in *ii*) of Theorem 2 one has to show that the error in $\hat{\eta}$ caused by $\hat{e}_t - \epsilon_t$ is of the order $o_p(n^{-1/2})$. This holds following the same arguments used in the proof of Theorem 5.2 *ii*) in Feng (2004), because the orders of magnitude of $\hat{e}_t - \epsilon_t$ are the same for i.i.d. and GARCH innovations. ◇

Proof of Lemma 5.

For any trial value $\lambda = (\alpha_0, \alpha_1, ..., \alpha_r, \beta_1, ..., \beta_s)' \in \Lambda$, one can rewrite $h_t(e(\eta^0); \lambda)$ as

$$h_t(e(\eta^0); \lambda) = \alpha_0 \left(1 - \sum_{j=1}^{s} \beta_j\right)^{-1} + \left(\sum_{j=1}^{r} \alpha_j B^j\right) \left(1 - \sum_{k=1}^{s} \beta_k B^k\right)^{-1} e_t^2(\eta^0)$$

and $h_t(\tilde{e}(\eta^0); \lambda)$ as

$$h_t(\tilde{e}(\eta^0); \lambda) = \alpha_0 \left(1 - \sum_{j=1}^{s} \beta_j\right)^{-1} + \left(\sum_{j=1}^{r} \alpha_j B^j\right) \left(1 - \sum_{k=1}^{s} \beta_k B^k\right)^{-1} \tilde{e}_t^2(\eta^0).$$

This leads to

$$\begin{aligned} h_t(e(\eta^0); \lambda) - h_t(\tilde{e}(\eta^0); \lambda) &= \left(\sum_{j=1}^{r} \alpha_j B^j\right) \left(1 - \sum_{k=1}^{s} \beta_k B^k\right)^{-1} (e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0)) \\ &= \left(\sum_{j=1}^{\infty} a_j B^j\right) (e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0)) \\ &\doteq O_p(e_t^2(\eta^0) - \tilde{e}_t^2(\eta^0)), \end{aligned} \tag{A.2}$$

where a_j are obtained by matching the powers in B, which decay exponentially. Lemma 5 is proved. ◇

Proof of Lemma 6.

Following (21) in Bollerslev (1986) we have

$$\frac{\partial h_t(\tilde{e}(\eta^0); \lambda)}{\partial \lambda} = \tilde{z}_t + \sum_{j=1}^{s} \beta_j \frac{\partial h_{t-j}(\tilde{e}(\eta^0); \lambda)}{\partial \lambda}, \tag{A.3}$$

where $\tilde{z}_t = (1, \tilde{e}_t^2(\eta^0), ..., \tilde{e}_t^2(\eta^0), h_{t-1}(\tilde{e}(\eta^0); \lambda), ..., h_{t-s}(\tilde{e}(\eta^0); \lambda))^{\mathrm{T}}$. Analogously, we have

$$\frac{\partial h_t(e(\eta^0); \lambda)}{\partial \lambda} = z_t + \sum_{j=1}^{s} \beta_j \frac{\partial h_{t-j}(e(\eta^0); \lambda)}{\partial \lambda}, \tag{A.4}$$

where $z_t = (1, e_t^2(\eta^0), ..., e_t^2(\eta^0), h_{t-1}(e(\eta^0); \lambda), ..., h_{t-s}(e(\eta^0); \lambda))^{\mathrm{T}}$. Denoting by $Bz_t = z_{t-1}$, $B\tilde{z}_t = \tilde{z}_{t-1}$,

$$B\frac{\partial h_t(e(\eta^0); \lambda)}{\partial \lambda} = \frac{\partial h_{t-1}(e(\eta^0); \lambda)}{\partial \lambda}$$

and

$$B\frac{\partial h_t(\tilde{e}(\eta^0); \lambda)}{\partial \lambda} = \frac{\partial h_{t-1}(\tilde{e}(\eta^0); \lambda)}{\partial \lambda},$$

we have

$$\left(1 - \sum_{j=1}^{s} \beta_j B^j\right) \frac{\partial h_t(\tilde{e}(\eta^0); \lambda)}{\partial \lambda} = \tilde{z}_t$$

and

$$\left(1 - \sum_{j=1}^{s} \beta_j B^j\right) \frac{\partial h_t(e(\eta^0); \lambda)}{\partial \lambda} = z_t.$$

This leads to

$$\begin{aligned} \frac{\partial h_t(e(\eta^0); \lambda)}{\partial \lambda} - \frac{\partial h_t(\tilde{e}(\eta^0); \lambda)}{\partial \lambda} &= \left(\sum_{j=0}^{\infty} c_j B^j\right)(z_t - \tilde{z}_t) \\ &\doteq O_p(z_t - \tilde{z}_t). \end{aligned} \tag{A.5}$$

Again, the c_j decay exponentially. The first element of $z_t - \tilde{z}_t$ is obviously zero. Results of Lemma 6 follow from (A.5) and Lemma 5. ◇

Proof of Theorem 3.

i) Following the proofs of Theorem 3.1 and 3.2 in Ling and Li (1997), the conditions of Lemma 5.1 in Feng (2004) hold for $\tilde{L}(\lambda)$ under the conditions of Theorem 3. Under these conditions we also have $e_t(\eta^0) \xrightarrow{p} \tilde{e}_t(\eta^0)\ \forall\ \lambda \in \Lambda$. Following Lemmas 5 and 6 we have $\hat{L}(\lambda) \xrightarrow{p} \tilde{L}(\lambda)\ \forall\ \lambda \in \Lambda$. Following Lemma 5 in Feng (2004) there exists a consistent approximate MLE $\hat{\lambda}$ satisfying the equation $\partial\hat{L}(\lambda)/\partial\lambda = 0$ such that

$$(\hat{\lambda} - \tilde{\lambda}) = O_p(\hat{L}'(\tilde{\lambda})). \tag{A.6}$$

ii) To show the results in this part we have to show $\hat{L}'(\tilde{\lambda}) = o_p(n^{-1/2})$.

Note that

$$\hat{L}'(\tilde{\lambda}) = \frac{1}{n}\sum_{t=1}^{n} \frac{1}{2h_t(e(\eta^0);\tilde{\lambda})} \left.\frac{\partial h_t(e(\eta^0);\lambda)}{\partial \lambda}\right|_{\lambda=\tilde{\lambda}} \left(\frac{e_t^2(\eta^0)}{h_t(e_t^2(\eta^0);\tilde{\lambda})} - 1\right). \tag{A.7}$$

By means of Taylor series expansion and using the results of Lemmas 5 and 6 we have

$$\begin{aligned}\frac{1}{2h_t(e(\eta_0);\tilde{\lambda})} &\doteq \frac{1}{2h_t(\tilde{e}(\eta_0);\tilde{\lambda})} + O_p(h_t(e(\eta_0);\tilde{\lambda}) - h_t(\tilde{e}(\eta_0);\tilde{\lambda}))\\ &\doteq \frac{1}{2h_t(\tilde{e}(\eta_0);\tilde{\lambda})} + O_p(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0)),\end{aligned}$$

$$\left.\frac{\partial h_t(e(\eta_0);\lambda)}{\partial\lambda}\right|_{\lambda=\tilde{\lambda}} \doteq \left.\frac{\partial h_t(\tilde{e}(\eta_0);\lambda)}{\partial\lambda}\right|_{\lambda=\tilde{\lambda}} + O_p(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0)),$$

where O_p denote the order of magnitude of a random vector, and

$$\frac{e_t^2(\eta_0)}{h_t(e(\eta_0);\tilde{\lambda})} \doteq \frac{e_t^2(\eta_0)}{h_t(e(\eta_0);\tilde{\lambda})} + O_p(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0)).$$

Furthermore, note that

$$L'(\tilde{\lambda}) = \frac{1}{n}\sum_{t=1}^{n} \frac{1}{2h_t(\tilde{e}(\eta_0);\tilde{\lambda})} \left.\frac{\partial h_t(\tilde{e}(\eta_0);\lambda)}{\partial \lambda}\right|_{\lambda=\tilde{\lambda}} \left(\frac{\tilde{e}_t^2(\eta_0)}{h_t(\tilde{e}(\eta_0);\tilde{\lambda})} - 1\right) = 0.$$

Inserting these results into (A.7), we obtain

$$\begin{aligned}\hat{L}'(\tilde{\lambda}) &\doteq \frac{1}{n}\left[\sum_{i=1}^{n} \frac{1}{2h_t(\tilde{e}(\eta_0);\tilde{\lambda})} \left.\frac{\partial h_t(\tilde{e}(\eta_0);\lambda)}{\partial \lambda}\right|_{\lambda=\tilde{\lambda}} \left(\frac{\tilde{e}_t^2(\eta_0)}{h_t(\tilde{e}(\eta_0);\tilde{\lambda})} - 1\right) + O_p(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0))\right]\\ &=: L'(\tilde{\lambda}) + T \qquad (A.8)\\ &= T,\end{aligned}$$

where the random vector

$$T = O_p\left(\frac{1}{n}\sum_{i=1}^{n}(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0))\right). \tag{A.9}$$

Using calculations similar to those given in the proof of Theorem 5.2 in Feng (2004) we have

$$\begin{aligned} T &= O_p\left(\frac{1}{n}\sum_{i=1}^{n}(e_t^2(\eta_0) - \tilde{e}_t^2(\eta_0))\right)\\ &= o_p(n^{-1/2}). \qquad (A.10)\end{aligned}$$

Theorem 3 is proved. ◇

References:

Bailie, R.T., Bollerslev, T, and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 74, 3-30.

Bailie, R.T., Chung, C.F. and Tiles, M.A. (1995). Analyzing inflation by the fractionally integrated ARFIMAR-GARCH model. *J. Appl. Econometrics*, 11, 23–40.

Beran, J. (1994). Statistics for long-memory processes. Chapman & Hall, New York.

Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short- and long-memory ARIMA models. J. Roy. Statist. Soc. B, 57, 672-695.

Beran, J., Bhansali, R.J., Ocker, D. (1998). On unified model selection for stationary and nonstationary short-and long-memory autoregressive processes. *Biometrika,* 85, 921–934.

Beran, J. and Feng, Y. (2001). Local polynomial estimation with a FARIMA-GARCH error process. *Bernoulli*, 7, 733–750.

Beran, J. and Feng, Y. (2002a). SEMIFAR models - A semiparametric framework for modelling trends, long-range dependence and nonstationarity. *Computat. Statist. Data Anal.*, 40, 393–419.

Beran, J. and Feng, Y. (2002b). Iterative plug-in algorithms for SEMIFAR models - definition, convergence and asymptotic properties. *J. Computat. Graph. Statist.*, 11, 690–713.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Hetreoskedasticity. *J. Econometrics*, 31, 307-327.

Chen, M. and An, H.Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica*, 8, 505–510.

Engel, R.F. (1982). Autoregressive conditional heteroskedasticity with estimation of U.K. inflation. *Econometrica*, 50 987–1008.

Engle, R.F., Lilien, D.M. and Robins, R.P. (1987). Estimating time-varying risk premia in the term structure: the ARCH-M model. *Econometrica*, 55, 391–407.

Feng, Y. (2004). *Non- and Semiparametric Regression with Fractional Time Series Errors -*

Theory and Applications to Financial Data. Habilitation Monograph, University of Konstanz.

Giraitis, L., Leipus, R., Robinson, P.M. and Surgailis, D. (2004). LARCH, leverage, and long memory. *J. Financ. Econometr.*, 2, 177–210.

Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal*, 1, 15-30.

He, C. and Teräsvirta, T. (1999a). Forth moment structure of the GARCH(p, q) process. *Econometric Theory,* 15, 824–846.

He, C. and Teräsvirta, T. (1999b). properties of autocorrelation function of squared observations for second-order GARCH processes under two sets of parameter constrains. *J. Time Ser. Anal* 20, 23–30.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.

Hosking, J. R. M. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econometrics*, 73, 261–284.

Ling, S. and Li, W.K. (1997). On fractional integrated autoregressive moving-average time series models with conditional heteroskedasticity. *J. Amer. Statist. Assoc.,* 92, 1184–1194.

Ling, S. and McAleer, M. (2002). Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r,s) models. *Econometric Theory*, 18, 722–729.

Ling, S.-Q., Li, W.K. (1997). Fractional ARIMA-GARCH time series models. *J. Amer. Statist. Assoc.,* 92, 1184–1194.

Nelson, D.B. (1991). Conditional herteroskedasticity in Asset Returns: A new Approach. *Econometrica*, 59, 347–370.

Robinson, P.M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. Econometr.*, 47, 6784.

Ruppert, D. and Wand, M.P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, 22, 1346–1370.