Forecasting financial market activity using a semiparametric fractionally integrated Log-ACD

Yuanhua Feng[1] and Chen Zhou

University of Paderborn

Abstract

This paper discusses forecasting of long memory and a nonparametric scale function in nonnegative financial processes based on a fractionally integrated Log-ACD (FI-Log-ACD) and its semiparametric extension (Semi-FI-Log-ACD). Necessary and sufficient conditions for the existence of a stationary solution of the FI-Log-ACD are obtained. Properties of this model under log-normal assumption are summarized. A linear predictor based on the truncated AR(∞) form of the logarithmic process is proposed. It is shown that this proposal is an approximately best linear predictor. Approximate variances of the prediction errors for an individual observation and for the conditional mean are obtained. Forecasting intervals for these quantities in the log- and the original processes are calculated under log-normal assumption. The proposals are applied to forecasting daily trading volumes and daily trading numbers in financial market.

Keywords: Approximately best linear predictor, FI-Log-ACD, financial forecasting, long memory time series, nonparametric methods, Semi-FI-Log-ACD

JEL Codes: C14, C51

[1]Corresponding author with E-mail: yuanhua.feng@wiwi.upb.de; Tel: +49 5251 60 3379

1 Introduction

Modeling and forecasting of short- and long memory, and a possible nonparametric scale function in financial time series is of great interest. Well known models in this context with short memory are e.g. the ARCH (autoregressive conditional heteroskedasticity, Engle, 1982) and GARCH (generalized ARCH, Bollerslev, 1986) for returns and the ACD (autoregressive conditional duration, Engle & Russell, 1998) for transaction durations. The ACD can also be used for modeling trading volume (Manganelli, 2005). Models based on logarithmic transformation are also proposed, including the Log-GARCH (Geweke, 1986, and Pantula, 1986) and the (first type) Log-ACD (Bauwens & Giot, 2000, Bauwens et al., 2008, Karanasos, 2008). Now, the log-data can be modeled by well known linear time series approaches. For instance, the Log-ACD is indeed equivalent to an ARMA for the log-data (Allen et al., 2006). And the resulting estimates of the original data are always nonnegative. Modeling of a smooth scale function in volatility caused by changing macroeconomic environment was investigated by Feng (2004), and Engle & Rangel (2008).

Well known long memory volatility models are the FIGARCH (fractionally integrated GARCH, Baillie et al., 1996), the LM-GARCH (long memory GARCH, Karanasos et al., 2004), the FIACD (Jasiak, 1998) and the LM-ACD (Karanasos, 2004). So far as we know, estimation of a nonparametric scale function in volatility models with long memory is not yet well studied. Most recently, Beran et al. (2012) proposed to model short memory, long memory and a nonparametric scale function in financial time series based on the log-transformation. They found in particular that the log-normal distribution is a suitable marginal distribution for daily average transaction durations and proposed to model the stochastic component of the log-data by a Gaussian FARIMA (fractional autoregressive integrated moving average, Hosking, 1981 and Beran, 1994). The log-data themselves are analyzed by a SEMIFAR (semiparametric fractional autoregressive, Beran & Feng, 2002a). Their proposals are hence called an EFARIMA (exponential FARIMA) and an ESEMIFAR, respectively, which can be easily estimated using existing software packages.

In this paper the EFARIMA and ESEMIFAR are first redefined as a FI-Log-ACD and a Semi-FI-Log-ACD, respectively. Necessary and sufficient conditions for the existence of a stationary solution of the FI-Log-ACD are obtained. Detailed properties of this model under log-normal assumption are summarized. In particular, now the long memory parameter is not affected by the log-transformation and the processes cannot exhibit antipersisitence (see also Dittmann & Granger, 2002). Our main focus is on the forecasting using the Semi-FI-Log-ACD. Now, the log-process follows a semiparametric regression with Gaussian FARIMA errors. Best linear predictor for the SEMIFAR model was proposed by Beran and Ocker (1999). In this paper we propose to use a simple, truncated linear predictor based on the AR(∞) form, because the sample size is very large. This idea is often applied to ARMA models (e.g. Brockwell & Davis, 2006, p. 184). Properties of the proposal are investigated in detail. It is shown that, in the presence of long memory the proposed predictor is still an approximately best linear predictor. Asymptotic variances of the prediction errors for an individual observation and for the conditional mean are obtained. Calculation of approximate forecasting intervals under log-normal assumption is discussed. Effect of the errors in the estimated trend on the asymptotic properties of the proposed predictor is also investigated. The Semi-FI-Log-ACD is then applied for modeling and forecasting daily trading volumes and daily trading numbers. The results indicate that this model is widely applicable and the proposed linear predictor works very well in practice. It is also shown that the log-normal distribution is a suitable choice for different kinds of aggregated financial data.

The paper is organized as follows. Definitions of the models are given in Section 2. Section 3 describes the properties and estimation of these models. The linear predictor is proposed and studied in Section 4. Section 5 reports the application results. Final remarks in Section 6 conclude the paper. Proofs of results are put in the appendix.

2 The FI-Log-ACD and Semi-FI-Log-ACD

A well known model for a stationary nonnegative financial time series, X_t^*, $t = 1, ..., n$, is the MEM (multiplicative error model, Engle, 2002) defined by

$$X_t = \nu \lambda_t \eta_t, \tag{1}$$

where $\nu > 0$ is a scale parameter, $\lambda_t > 0$ is the conditional mean of $X_t^* = X_t/\nu$ and $\eta_t \geq 0$ are i.i.d. random variables. In this paper we propose the use of a semiparametric MEM model by replacing ν in (1) with a nonparametric smooth scale function $\nu(\tau) > 0$:

$$X_t = \nu(\tau_t) X_t^* = \nu(\tau_t) \lambda_t \eta_t, \tag{2}$$

where $\tau_t = t/n$ denotes the rescaled time. Let $\varepsilon_t = \ln(\eta_t)$, $Y_t = \ln(X_t)$, $Z_t = \ln(X_t^*)$, $\mu(\tau_t) = \ln[\nu(\tau_t)]$ and $\zeta_t = \ln(\lambda_t)$, where $\mu(\tau_t)$ and ζ_t are the local and conditional means of Y_t, respectively. Following Beran et al. (2012), we assume that $E(\varepsilon_t) = 0$, var $(\varepsilon_t) = \sigma_\varepsilon^2$ and the stochastic component Z_t follows a FARIMA

$$(1-B)^d \phi(B) Z_t = \psi(B)\varepsilon_t, \tag{3}$$

where $d \in (-0.5, 0.5)$ is the fractional differencing parameter, $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ and $\psi(B) = 1 + \psi_1 B + \ldots + \psi_q B^q$ are the MA- and AR-characteristic polynomials with no common factor and all roots outside the unit circle. The model defined by (1) and (3) is called an exponential FARIMA (EFARIMA), which is a nonnegative process whose log-transformation follows a FARIMA. The model defined by (2) and (3) will be called an ESEMIFAR, because $Y_t = \ln(X_t) = Z_t + \mu(\tau_t)$ follows a SEMIFAR (Beran & Feng, 2002a) with the integer integration parameter $m = 0$ and an additional MA part.

Beran et al. (2012) indicated that the EFARIMA model can be written as a fractionally integrated generalization of the (first type) Log-ACD model. The reason is that, similar to Eq. (7) in Bauwens et al. (2008), the conditional mean of Z_t can be represented as

$$\zeta_t = \ln \lambda_t = \sum_{i=1}^{\infty} \pi_i \ln \lambda_{t-i} + \sum_{j=1}^{\infty} \omega_j \ln(\eta_{t-j}), \tag{4}$$

where π_i are coefficients of $\pi(B) = (1-B)^d \phi(B) = 1 - \sum_{i=1}^{\infty} \pi_i B^i$ with $\pi_i \approx c_\pi i^{d-1}$ for large i, $\omega_j = \pi_j + \psi_j$ for $1 \leq j \leq q$, and $\omega_j = \pi_j$ for $j > q$. The model defined by (1) and (4) will be called a FI-Log-ACD. And Eq. (2) and (4) define a semiparametric generalization of the FI-Log-ACD, called a Semi-FI-Log-ACD. The Log-ACD (p, q) model is the special case with $d = 0$. Moreover, note that ζ_t can also be rewritten as

$$\zeta_t = \sum_{i=1}^{\infty} \pi_i Z_{t-i} + \sum_{j=1}^{q} \psi_j \varepsilon_t = [\pi(B) - 1] Z_t + [\psi(B) - 1]\varepsilon_t. \tag{5}$$

The relationship between the FI-Log-ACD and the EFARIMA is given below.

Proposition 1. *The EFARIMA model defined by* (1) *and* (3)*, and the FI-Log-ACD model defined by* (1) *and* (4) *are equivalent to each other.*

Proof of Proposition 1 is omitted. This result means that the proposed models are the application of the well known FARIMA and SEMIFAR models to the log-process. Hence, the log-transformation of a nonlinear (nonnegative) process following the FI-Log-ACD is assumed to be a linear process. The original process X_t^* is hence a log-linear process.

3 Properties and estimation of the models

3.1 The stationary solutions

Firstly, some results in Beran et al. (2012) under log-normal assumption will be extended to more general distributions. Let $\alpha(B) = (1-B)^{-d}\phi^{-1}(B)\psi(B) = 1 + \sum_{i=1}^{\infty} \alpha_i B^i$. It is well known that the stationary solution of the FARIMA process Z_t is given by

$$Z_t = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i} \tag{6}$$

with $\alpha_i \approx c_\alpha i^{d-1}$ for large i, and, for large k, the autocorrelation (ACF) of Z_t is given by

$$\rho_Z(k) \approx c_\rho^Z |k|^{2d-1}, \tag{7}$$

where c_ρ^Z is a constant. Note that $c_\rho^Z > 0$ for $d > 0$ and now Z_t has long memory.

Now, let $\alpha_{\max} = \max(\alpha_i)$ and $\alpha_{\min} = \inf(\alpha_i)$, where $\alpha_{\max} \geq 1$ and $\alpha_{\min}$ may be negative. Conditions for the existence of a stationary solution of X_t^* in the current case with $2u$-th finite moment are similar as those given by Karanasos (2008).

A1. Z_t is a stationary and invertible FARIMA process as defined in (3).

A2. Both $E(\eta_t^{2u\alpha_{\max}})$ and $E(\eta_t^{2u\alpha_{\min}})$ are finite for some $u > 0$.

Now, the stationarity solution of X_t^* is given by

$$X_t^* = \prod_{i=0}^{\infty} \eta_{t-i}^{\alpha_i}. \tag{8}$$

Lemma 1. *The solution of X_t^* given in* (8) *is strictly stationary with finite $2u$-th moment, if and only if A1 and A2 hold. If A2 holds for $u \geq 1$, X_t^* is also weakly stationary.*

The proof is similar to that of Lemmas 1 and 2 in Karanasos (2008) and is omitted.

A2 ensures that all of the terms in the product in (8) exist. A1 implies that $\sum_{i=0}^{\infty} \alpha_i^2 < \infty$ and $E(\varepsilon_t) = 0$. This together with A2 ensures the convergence of X_t^* defined in (8). Note that the condition $E(\varepsilon_t) = 0$ is indeed a restriction on the distribution of η_t. For instance, if η_t is exponentially distributed with the density $f(u) = \mu_\eta^{-1}\exp(-u/\mu_\eta)$, this leads to $\mu_\eta = \exp(\gamma) \approx 1.781$, where γ is the Euler constant. Thus, μ_η is now fixed. The possible change in the scale is reflexed by ν or $\nu(\tau)$. If a two-parameter family of nonnegative distributions is considered, the restriction $E(\varepsilon_t) = 0$ means that only one of the two parameters is free. The restriction $E(\varepsilon_t) = 0$ is obviously fulfilled e.g. by:

Example 1. The log-normal innovations η_t with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $\sigma_\varepsilon^2 > 0$,

Example 2. The log-logistic innovations η_t with $\varepsilon_t \sim Lo(0, b)$ and $b > 0$ or

Example 3. The log-Laplace innovations η_t with $\varepsilon_t \sim La(0, b)$ and $b > 0$.

Note that A2 may or may not be affected by d. Whether A2 is fulfilled or not, is jointly determined by the distribution of η_t, the value of u and the FARIMA coefficients. In Ex. 1 above, A2 is always fulfilled and X_t^* is strictly and weakly stationary. In Ex. 2 and 3, X_t^* is only weakly stationary, if b is small enough.

Furthermore, the stationary solution of the conditional mean of Z_t is given by

$$\zeta_t = \sum_{i=1}^{\infty} \alpha_i \varepsilon_{t-i}. \tag{9}$$

Under the same assumptions we obtain the stationary solution of λ_t:

$$\lambda_t = \prod_{i=1}^{\infty} \eta_{t-i}^{\alpha_i}.$$

The forecasts of the FARIMA process Z_t and its conditional mean ζ_t to propose later are based on their AR(∞) representations, respectively. For Z_t we have

$$Z_t = \sum_{j=1}^{\infty} \beta_j Z_{t-j} + \varepsilon_t, \tag{10}$$

where β_j are the coefficients of $\beta(B) = (1-B)^d \phi(B)\psi^{-1}(B) = 1 - \sum_{j=1}^{\infty} \beta_j B^j$. For large j, we have $\beta_j \approx c_\beta j^{-d-1}$ with $c_\beta > 0$. This yields the representation of ζ_t based on Z_t

$$\zeta_t = \sum_{j=1}^{\infty} \beta_j Z_{t-j}. \tag{11}$$

The stationary solutions of X_t^* and λ_t, respectively, can be rewritten as

$$X_t^* = \eta_t \prod_{j=1}^{\infty} (X_{t-j}^*)^{\beta_j} \text{ and } \lambda_t = \prod_{j=1}^{\infty} (X_{t-j}^*)^{\beta_j}. \tag{12}$$

3.2 Properties under log-normal assumption

Beran et al. (2012) found that when aggregated financial data are considered, the log-normal assumption is usually a suitable choice. They hence studied the properties of the proposed models with log-normally distributed innovations in detail. In the following, their results will be first summarized briefly. Then we will focus on discussing the application of the Semi-FI-Log-ACD model. We will see that now all of Z_t, X_t^*, ζ_t and λ_t exhibit long memory. In particular, the authors showed that the process X_t^* is log-normally distributed, $X_t^* \sim LN(0, \sigma^2)$ with $\sigma^2 = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \alpha_i^2$, if ε_t are i.i.d. $N(0, \sigma_\varepsilon^2)$ random variables. Closed form formula of the ACF of X_t^* can be obtained. Furthermore, in the presence of long memory it holds

$$\rho_{X^*}(k) \approx c_\rho^{X^*} |k|^{2d-1}$$

for large k, where $0 < c_\rho^{X^*} < c_\rho^Z$. We see that X_t^* is a long memory process with the same memory parameter d, if Z_t has long memory. This confirms the well known fact that the long memory parameter in Z_t and that in X_t^* under log-normal assumption is the same (see e.g Dittmann & Granger, 2002). The reason is that the Hermite rank of the exponential function is one. However, the constant in the asymptotic formula of $\rho_{X^*}(k)$ is smaller than that in $\rho_Z(k)$. If Z_t is a FARIMA with $-0.5 < d \leq 0$, it can be shown that $\sum \rho_{X^*}(k) > 0$. We see that X_t^* does not have antipersistence, even if Z_t is antipersistent (see also Dittmann & Granger, 2002). This leads to the very interesting fact:

Proposition 2. *A FI-Log-ACD process X_t^* with log-normal marginal distribution cannot exhibit antipersistence.*

In financial econometrics, long memory property of the conditional means ζ_t in Z_t and λ_t in X_t^* is also of great interest. The ACF of ζ_t with $d > 0$ is given by

$$\rho_\zeta(k) \approx c_\rho^\zeta |k|^{2d-1} \tag{13}$$

for large k, where $c_\rho^\zeta > c_\rho^Z$. From (13) we see that ζ_t also has long memory with the same memory parameter d. However, the constant in the asymptotic formula of $\rho_\zeta(k)$ is larger than that in $\rho_Z(k)$. And the ACF of λ_t for large k is given by

$$\rho_\lambda(k) \approx c_\rho^\lambda |k|^{2d-1},$$

where $0 < c_\rho^\lambda < c_\rho^\zeta$. Again, the long memory parameter in λ_t is d. But the constant in the asymptotic formula of the ACF after the exponential transformation is reduced.

3.3 Estimation of the models

Let $K(u)$ be a symmetric density with compact support $[-1, 1]$ and h be the bandwidth. The trend can estimated by local polynomial regression with FARIMA errors (Beran & Feng, 2002b). Now, $\hat{\mu}(\tau)$ is obtained by solving the weighted least squares problem

$$Q = \sum_{t=1}^{n} \left\{ Y_t - \sum_{j=0}^{p} b_j(\tau_t - \tau)^j \right\}^2 K\left(\frac{\tau_t - \tau}{h}\right) \Rightarrow \min .$$

From now on we will mainly consider the model of Z_t without the MA part, so that the Semi-FI-Log-ACD can be estimated using existing software packages. But the theoretical discussion holds in the case, when Z_t follows a general FARIMA model. Now, let $\theta = (\sigma_\varepsilon^2, d, \phi_1, \ldots, \phi_p)^{\mathrm{T}}$ denote the unknown parameter vector of the SEMIFAR model. Under the normal assumption of ε_t, θ can be estimated by approximate Gaussian MLE from $\hat{Z}_t = Y_t - \hat{\mu}(\tau_t)$. The AR order p can be selected consistently by the BIC. A crucial point by the practical implementation of the Semi-FI-Log-ACD model is the selection of the bandwidth h. This can e.g. be done by means of the package FinMetrics in S+, where however only a kernel estimator of $\mu(\tau)$ is built-in. In this paper that package is improved slightly to include local polynomial estimator of $\mu(\tau)$.

4 Forecasting based on the Semi-FI-Log-ACD

Now, we will discuss forecasting based on the Semi-FI-Log-ACD model, which is equivalent to the ESEMIFAR. Note that the ESEMIFAR is a SEMIFAR applied to the log-transformed data, the ESEMIFAR forecasting hence consists of two stages: 1) The forecasting based on the SEMIFAR model applied to the log-data, and 2) The calculation of the forecasts for the original data through exponential transformation. The former consists again of two parts: the extrapolation of the estimated trend function $\hat{\mu}(\tau_n)$ and the prediction of the stochastic part Z_{n+k}. This will be discussed in the following separately.

4.1 Extrapolation of the trend function

In this paper the trend estimated by local linear regression will be used, because a higher order local polynomial estimator may be instable at the endpoint and is hence not suitable for forecasting. A great advantage of the local linear estimator compared to a kernel estimator is that $\hat{\mu}$ has automatic boundary correction, i.e. the bias of $\hat{\mu}(\tau_n)$ at the endpoint of the time series is of the order $O(h^2)$, while the bias of a kernel estimator at τ_n is of the order $O(h)$. We propose to forecast the trend $\mu(\tau_{n+k})$ in the future by linear extrapolation of $\hat{\mu}(\tau_n)$. Let $\Delta\mu = \hat{\mu}(\tau_n) - \hat{\mu}(\tau_{n-1})$. By means of the linear extrapolation, the forecasted trend $\hat{\mu}(\tau_{n+k})$ is given by

$$\hat{\mu}(\tau_{n+k}) = \hat{\mu}(1) + k\Delta\mu. \tag{14}$$

The following assumptions are required for further discussion.

A3. In A1 assume further that $d \in (0, 0.5)$, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, and $q = 0$ for simplicity.

A4. The weighting function $K(u)$ is a symmetric density on the compact support $[-1, 1]$.

A5. The trend function μ is at least secondly continuously differentiable.

A6. The bandwidth h is selected by a consistent data-driven algorithm.

Assumptions A4 and A5 are standard assumptions in nonparametric regression. A3 through A6 ensure that the model can be estimated using some existing SEMIFAR algorithm. A6 ensures that the used bandwidth $\hat{h} \approx h_A = O(n^{(2d-1)/(5-2d)})$, where h_A is the asymptotically optimal bandwidth, and that $\hat{\mu}$ achieves the optimal rate of convergence.

4.2 The best linear and approximately best linear predictors

Let $Z_1, \ldots, Z_n$ denote the past observations. The best linear predictor of Z_{n+k} under the SEMIFAR model was proposed by Beran and Ocker (1999):

$$\breve{Z}_{n+k} = \sum_{j=1}^{n} \beta_{k,j}^o Z_j, \tag{15}$$

where $\beta_{k,o} = (\beta_{k,1}^o, \ldots, \beta_{k,n}^o)^{\mathrm{T}}$ is as given in Theorem 1 of Beran and Ocker (1999), which minimizes the mean squared prediction error (MSE). Furthermore, $\breve{Z}_{n+k}$ satisfies

$$E[(Z_{n+k} - \breve{Z}_{n+k})Z_t] = 0, \quad t = 1, \ldots, n. \tag{16}$$

Eq. (16) implies that the prediction error of $\breve{Z}_{n+k}$ is orthogonal to any of the observations.

It is however not easy to use $\breve{Z}_{n+k}$ defined in (15), because $\beta_{k,o}$ has to be solved repeatedly at each forecasting step. Note that in the current context n is very large. For simplicity, we hence propose to use an approximately best linear predictor based on the truncated part of the AR(∞) representation of Z_t. This idea is often employed to carry out forecasting based on an ARMA model, when the sample size is large (Brockwell & Davis, 2006, p. 184). Hence, an approximately best linear predictor based on $Z_1, \ldots, Z_n$, by means of the AR(∞) representation (10) of the FARIMA process is defined by

$$\hat{Z}^*_{n+k} = \sum_{j=1}^{k-1} \beta_j \hat{Z}^*_{n+k-j} + \sum_{j=k}^{n+k-1} \beta_j Z_{n+k-j}, \tag{17}$$

where $\hat{Z}^*_{n+k-j}$ are the previously predicted values. For the practical implementation, we propose to use the following linear predictor

$$\hat{Z}_{n+k} = \sum_{j=1}^{k-1} \hat{\beta}_j \hat{Z}_{n+k-j} + \sum_{j=k}^{n+k-1} \hat{\beta}_j \hat{Z}_{n+k-j}, \tag{18}$$

where $\hat{\beta}_j$ are the estimated coefficients in the AR(∞) form of Z_t, $\hat{Z}_{n+k-j}$ for $j = k, ..., n + k - 1$, are the obtained residuals and $\hat{Z}_{n+k-j}$, $j = 1, ..., k-1$, are the previously predicted values. The linear predictor in (18) is our proposal to use in practice. To our knowledge, in the literature the above defined approximately best linear predictor is not yet proposed in the presence of long memory. The relationship between $\hat{Z}_{n+k}$ and $\hat{Z}^*_{n+k}$ is given by

Lemma 2. *Under Assumptions A3 through A6, the two linear predictors $\hat{Z}_{n+k}$ and $\hat{Z}^*_{n+k}$ are asymptotically equivalent to each other.*

Proof of Lemma 2 is given in the appendix. Lemma 2 indicates that the asymptotic properties of $\hat{Z}^*_{n+k}$ defined based on the unobservable quantities β_i and Z_t, and those of $\hat{Z}_{n+k}$ defined using $\hat{\beta}_i$ and $\hat{Z}_t$ are the same.

Now, the best linear predictor given infinite past observations $Z_n, \ldots, Z_1, Z_0, Z_{-1}, \ldots$, is introduced. Similar to Eq. (5.5.3) in Brockwell & Davis (2006), this linear predictor based on the AR(∞) form of the FARIMA model (10) is defined by

$$\tilde{Z}_{n+k} = \sum_{j=1}^{k-1} \beta_j \tilde{Z}_{n+k-j} + \sum_{j=k}^{\infty} \beta_j Z_{n+k-j}. \tag{19}$$

Properties of $\hat{Z}_{n+k}$ are stated in the following theorem.

Theorem 1. *Under the same assumptions of Lemma 2, the proposed linear predictor $\hat{Z}_{n+k}$ is an approximately best linear predictor in the following sense:*

i) $E[(\tilde{Z}_{n+k} - \hat{Z}_{n+k})^2] = o(1)$ *and*

ii) $E[(Z_{n+k} - \hat{Z}_{n+k})Z_t] = o(1)$, $t = 1, \dots n$.

Proof of Theorem 1 is given in the appendix. Theorem 1 i) shows that $\hat{Z}_{n+k}$ converges to $\tilde{Z}_{n+k}$ in mean squared error. Theorem 1 ii) shows that the prediction error of $\hat{Z}_{n+k}$ is approximately orthogonal to all of the observations. Note that $\tilde{Z}_{n+k}$ is the best linear predictor based on infinite past observations. Hence, its MSE is no larger than that of $\breve{Z}_{n+k}$, because the σ-algebra generated by $Z_n, \dots, Z_1, Z_0, Z_{-1}, \dots$ includes that generated by $Z_n, \dots, Z_1$ as a subset. Moreover, the MSE of $\hat{Z}_{n+k}$ is no smaller than that of $\breve{Z}_{n+k}$. Thus, Theorem 1 i) ensures that the MSE's of all of the above mentioned linear predictors are asymptotically the same. This leads to the following corollary.

Corollary 1. *Under Assumptions A3 to A6, the linear predictor $\hat{Z}_t$ is asymptotically equivalent to the (exactly) best linear predictor $\breve{Z}_t$ proposed by Beran and Ocker (1999).*

4.3 Approximate forecasting intervals

Now, we will discuss the interval forecasting of an individual observation, the conditional mean and the total mean. Note that the point forecasting for ζ_{n+k} is the same as that for Z_{n+k}, i.e. $\hat{\zeta}_{n+k} = \hat{Z}_{n+k}$. The variance of $Z_{n+k} - \hat{Z}_{n+k}$ and that of $\zeta_{n+k} - \hat{Z}_{n+k}$ can be easily obtained by adapting known results in the literature.

Theorem 2. *Under the same conditions of Theorem 1 we have*

i) $var(Z_{n+k} - \hat{Z}_{n+k}|Z_n, \dots, Z_1) \approx V_{Z_{n+k}}$, *where* $V_{Z_{n+k}} = \sigma_\varepsilon^2 \sum_{i=0}^{k-1} \alpha_i^2$,

ii) $var(\zeta_{n+k} - \hat{Z}_{n+k}|Z_n, \dots, Z_1) \approx V_{\zeta_{n+k}}$, *where* $V_{\zeta_{n+k}} = \sigma_\varepsilon^2 \sum_{i=1}^{k-1} \alpha_i^2$.

Proof of Theorem 2 is given in the appendix. The result in Theorem 2 i) is well known. Note however that, in the current case $V_{Z_{n+k}}$ tends to $\operatorname{var}(Z_t)$ very slowly. Moreover, so far as we know, the result in Theorem 2 ii) on the variance of the prediction error for the conditional mean is usually not discussed in the literature. This is however an interesting topic in financial econometrics. For example, it helps us to understand the accuracy of the forecasted volatility or the forecasted conditional mean duration.

The point forecasting for an individual future observation is $\hat{Y}_{n+k} = \hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k}$. The length of the forecasting interval for Y_{n+k} is the same as that for Z_{n+k}, because the error in $\hat{\mu}(\tau_{n+k})$ is asymptotically negligible compared to that in $\hat{Z}_{n+k}$. Assume that ε_t are i.i.d. $N(0, \sigma_\varepsilon^2)$. The approximate $100(1-\alpha)\%$-forecasting interval for Y_{n+k} is given by

$$Y_{n+k} \in \left(\hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k} - q_{\alpha/2}\sqrt{V_{Z_{n+k}}}, \hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k} + q_{\alpha/2}\sqrt{V_{Z_{n+k}}}\right) \tag{20}$$

and, for $k > 1$, the approximate $100(1-\alpha)\%$-forecasting interval of ζ_{n+k} is given by

$$\zeta_{n+k} \in \left(\hat{Z}_{n+k} - q_{\alpha/2}\sqrt{V_{\zeta_{n+k}}}, \hat{Z}_{n+k} + q_{\alpha/2}\sqrt{V_{\zeta_{n+k}}}\right), \tag{21}$$

where $q_{\alpha/2}$ is the upper $\alpha/2$-quantile of $N(0,1)$. Furthermore, let $m(\tau_t) = \mu(\tau_t) + \zeta_t$ and $g(\tau_t) = \exp[m(\tau_t)]$ be the total means in Y_t and X_t, respectively. We have $\hat{m}(\tau_{n+k}) = \hat{Y}_{n+k}$. But the prediction error for $m(\tau_{n+k})$ is approximately equal to that for ζ_{n+k}. Thus, the approximate $100(1-\alpha)\%$-forecasting interval for $m(\tau_{n+k})$, $k > 1$, is given by

$$m(\tau_{n+k}) \in \left(\hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k} - q_{\alpha/2}\sqrt{V_{\zeta_{n+k}}}, \hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k} + q_{\alpha/2}\sqrt{V_{\zeta_{n+k}}}\right). \tag{22}$$

Note that the prediction errors in $\hat{\zeta}_{n+1}$ and $\hat{m}(\tau_{n+1})$ are both asymptotically negligible.

Our main purpose is to achieve suitable forecasting for X_{n+k}, λ_{n+k} and $g(\tau_{n+k})$. Taking the exponential transformation of $\hat{Z}_{n+k}$ and $\hat{m}(\tau_{n+k}) = \hat{Y}_{n+k}$, respectively, we have

$$\hat{\lambda}_{n+k} = \exp(\hat{Z}_{n+k}) = \prod_{j=1}^{n+k-1} \hat{Z}_{n+k-j}^{\hat{\beta}_j}, \tag{23}$$

$$\hat{X}_{n+k} = \hat{g}(\tau_{n+k}) = \exp[\hat{\mu}(\tau_{n+k}) + \hat{Z}_{n+k}] = \hat{\nu}(\tau_{n+k})\hat{\lambda}_{n+k}. \tag{24}$$

The approximate $100(1-\alpha)\%$-forecasting intervals for X_{n+k}, λ_{n+k} and $g(\tau_{n+k})$ can be obtained based on (20) to (22), respectively, through exponential transformation.

5 Application

Modeling and forecasting trading volumes and trading numbers in the financial market is of great interest, because these quantities play a critical role and can be used as indicators for market activity. Manganelli (2005) employed the ACD to model the conditional mean volume and called it an ACV. We found that the ACD is also a useful framework for analyzing trading numbers. In the following, we will therefore apply the Semi-FI-Log-ACD to the daily trading volumes and trading numbers of BMW and Air France (AF) from Jan 02, 2006 to Jun 30, 2012. It is found that daily trading volumes and trading numbers may exhibit significant short memory, long memory as well as a significant slowly changing trend at the same time (see Table 1 given later). Let $\hat{Z}_t = Y_t - \hat{\mu}(\tau_t)$ denote the residuals of the log-data. Histograms of the standardized values of $\hat{Z}_t$ and those of their exponential values are shown in Fig. 1 for all examples. We see that the distribution of $\hat{Z}_t$ in all cases is nearly normal. This indicates that the normal assumption on ε_t is a suitable choice for analyzing and forecasting these quantities.

Fig. 2(a) shows the (aggregated) daily trading volumes of BMW together with the point and interval forecasts ($\alpha = 5\%$) for an individual observation for the next 50 days, obtained through exponential transformation of (20). From Fig. 2(a) we can see that the higher the scale function, the larger the variation in the observations, which reflexes the fact that X_t has time varying unconditional variance $\text{var}\,(X_t) = \nu^2(\tau_t)\text{var}\,(X_t^*)$. To this end see also the other examples. This provides an evidence for the use of the log-transformation, which transfers the multiplicative nonparametric regression to an additive nonparametric regression. Fig. 2(b) displays the log-transformed data together with the estimated trend $\hat{\mu}(\tau_t)$ (solid line) and the corresponding forecasts for Y_{n+k} obtained from (20). In this paper the Epanechnikov kernel is used as the weighting function. To ensure the stability of $\hat{h}$, the bandwidth is selected based on 95% of the observations in the middle part. From Fig. 2(b) we can see that after the log-transformation the data becomes more concentrated and distributed symmetrically around the trend. And the level of variation keeps now nearly unchanged over the whole observation period.

The estimated conditional means of the log-data together with the corresponding point and interval forecasts for ζ_{n+k} calculated from (21) are given in Fig. 2(c). The estimated conditional means look quite stationary. The difference between the forecasting interval of Y_{n+k} in Fig. 2(b) and that of ζ_{n+k} in Fig. 2(c) is that the former is affected by ε_{n+k}, but the latter not. The estimated total means in the original data $\hat{g}(\tau_t)$ together with the point forecasts $\hat{g}(\tau_{n+k})$ and their forecasting intervals calculated through exponential transformation of (22) are displayed in Fig. 2(d), which reflex the total dynamics of the daily trading volumes of BMW caused by past information and slowly changing macroeconomic environment. Results on $\hat{\lambda}_t$ and $\hat{m}(\tau_t)$ are omitted. The same results for daily trading volumes of Air France are displayed in Fig. 3.

Similar results for daily trading numbers of BMW and those of Air France are displayed in Fig. 4 and 5. Fig. 2 to 5 indicate that both of the Semi-FI-Log-ACD model and the proposed approximately best linear predictor work in practice very well. From Fig. 5(b) we can see that at the end of the time series the point forecasts are much lower than the estimated trend, but will tend to the average level in the near future. This reflexes a well known feature of a long memory process clearly, i.e. long memory may cause spurious local trends. From Fig. 4(d) and 5(d) we can see that trading numbers and the volatility of trading numbers of both companies increased strongly in the last years and will possibly increase in the future. This fact seems also to be true for trading volumes of Air France, but not for trading volumes of BMW. Finally, comparing Fig. 1(c) with Fig. 3(c), or Fig. 2(c) with Fig. 4(c), it seems that the conditional means of the daily trading volumes and trading numbers of the same company are strongly correlated.

The selected bandwidth, the estimated long memory parameter, the selected AR order by the BIC together with the estimated short memory parameter, if applicable, in all cases are listed in Table 1, where the 95%-confidence intervals for the corresponding parameters and the significance test of the trend are also listed. For instance, for daily trading volumes of BMW we have $\hat{h} = 0.146$ and $\hat{p} = 1$, where $\hat{\mu}(\tau)$, $\hat{d}$ and $\hat{\phi}_1$ are all significant at $\alpha = 5\%$. This means that daily trading volumes of BMW can be well modeled by a nonparametric regression model with long memory errors. After removing the fitted trend, the residuals follow a FARIMAR(1, 0.299, 0) model. For daily trading volumes of Air France an EFARIMA(1, 0.393, 0) model is fitted, but now the trend is insignificant and the selected bandwidth $\hat{h} = 0.243$ is very large. That is the log-data of daily trading volumes of Air France can be simply modeled by a FARIMA model.

6 Final remarks

In this paper necessary and sufficient conditions for the existence of a stationary solution of the FI-Log-ACD model are obtained. Short and middle-term forecasting of a nonnegative process with long memory and a nonparametric scale function based on the Semi-FI-Log-ACD model is studied. A simple linear predictor is proposed. It is shown that the proposal is an approximately best linear predictor. Application to real data sets shows that the proposed linear predictor works very well in practice. It is also confirmed that the Semi-FI-Log-ACD model is very useful for modeling different kinds of financial data, in particular aggregated financial data. Furthermore, simultaneous estimation of the nonparametric trend and the long memory error process will improve the quality of the forecasting. The more important reasons for this are as follows. On the one hand, if possible long memory in the conditional mean of a process is not considered, the selected bandwidth will be much smaller than it should be and the formula for calculating the asymptotic variance is also wrong. This will lead to a significant trend, even if the underlying process is indeed stationary. On the other hand, if an existing nonparametric scale function is not considered, it will be misinterpreted as very strong long memory.

Finally, the proposed models can be extended in different ways. Firstly, it is of interest to extend the data-driven algorithms used in this paper to include the MA-part in Z_t. Furthermore, properties of the FI-Log-ACD model under the log-logistic, log-Laplace and other suitable distributions of η_t should also be studied in detail. Under these distributions, the nonparametric trend can be estimated similarly. The unknown FARIMA parameters can be estimated from the residuals by QMLE. Now, the selection of the most suitable distribution is also an important topic. These problems will be studied elsewhere.

Appendix. Proofs of results

Proof of Lemma 2. It is well known that, under assumptions A3 through A6, the local linear estimator $\hat{\mu}(\tau)$ with fractional times series error achieves the optimal convergence rate of the order $O(n^{-2(1-2d)/(5-2d)})$ (Feng and Beran, 2013). This results in the fact that the difference between $\hat{Z}_t$ and Z_t is also of the order $O(n^{-2(1-2d)/(5-2d)})$. Moreover, when $n \to \infty$ and $d > 0$, the effect of the estimated trend function on the estimation of the unknown parameter vector θ is negligible (Beran and Feng, 2002a), and $\hat{\theta}$ is now still $\sqrt{n}$-consistent. Using Taylor expansion it can be shown that $\hat{\beta}_j - \beta_j = \beta_j O_p(n^{-1/2})$. Detailed discussion on this point is omitted to save space.

In the following, we will only show the result of Lemma 2 for $k = 1$ in detail. Note that $\hat{Z}^*_{n+1} = \sum_{j=1}^{n} \beta_j Z_{n+1-j} = O_p(1)$ and $\sum_{j=1}^{n} |\beta_j| < \sum_{j=1}^{\infty} |\beta_j| < \infty$ for $d > 0$. We have

$$
\begin{aligned}
\hat{Z}_{n+1} - \hat{Z}^*_{n+1} &= \sum_{j=1}^{n} \hat{\beta}_j \hat{Z}_{n+1-j} - \sum_{j=1}^{n} \beta_j Z_{n+1-j} \\
&= \sum_{j=1}^{n} \hat{\beta}_j \hat{Z}_{n+1-j} - \sum_{j=1}^{n} \hat{\beta}_j Z_{n+1-j} - \sum_{j=1}^{n} \beta_j Z_{n+1-j} + \sum_{j=1}^{n} \hat{\beta}_j Z_{n+1-j} \\
&= \sum_{j=1}^{n} \hat{\beta}_j (\hat{Z}_{n+1-j} - Z_{n+1-j}) - \sum_{j=1}^{n} (\beta_j - \hat{\beta}_j) Z_{n+1-j} \\
&= \sum_{j=1}^{n} \hat{\beta}_j O_p(n^{-2(2d-1)/(5-2d)}) - \sum_{j=1}^{n} \beta_j Z_{n+1-j} O_p(n^{-1/2}) \\
&\approx \sum_{j=1}^{n} \hat{\beta}_j O_p(n^{-2(2d-1)/(5-2d)}) \qquad \text{(A.1)} \\
&\leq O_p(n^{2(2d-1)/(5-2d)}) \sum_{j=1}^{n} |\hat{\beta}_j| = o_p(1).
\end{aligned}
$$

Similarly, this result can be proved for $k > 1$. Lemma 2 is proved. $\diamond$

Proof of Theorem 1. Following Lemma 2, the results of Theorem 1 will be proved by replacing $\hat{Z}_{n+k}$ with $\hat{Z}^*_{n+k}$. Under the same conditions of Lemma 2 we have

i) For $k = 1$:

$$E[(\tilde{Z}_{n+1} - \hat{Z}^*_{n+1})^2] = E\left[\left(\sum_{j=1}^{\infty} \beta_j Z_{n+1-j} - \sum_{j=1}^{n} \beta_j Z_{n+1-j}\right)^2\right]$$

$$
\begin{aligned}
&= E\left[\left(\sum_{j=n+1}^{\infty} \beta_j Z_{n+1-j}\right)^2\right] \\
&= \sum_{i=n+1}^{\infty} \beta_i \sum_{j=n+1}^{\infty} \beta_j E[Z_{n+1-i} Z_{n+1-j}] \\
&= \sum_{i=n+1}^{\infty} \beta_i \sum_{j=n+1}^{\infty} \beta_j \gamma(i-j) \qquad \text{(A.2)} \\
&\approx \sum_{i=n+1}^{\infty} C_\beta i^{-d-1} \sum_{j=n+1}^{\infty} C_\beta j^{-d-1} \gamma(i-j) \\
&\leq \gamma(0) \sum_{i=n+1}^{\infty} |C_\beta| i^{-d-1} \sum_{j=n+1}^{\infty} |C_\beta| j^{-d-1} \\
&= \gamma(0) O[(n+1)^{-2d}] = o(1).
\end{aligned}
$$

More details of the proof above will be clarified by the remark given later.

Now, let $k > 1$ and assume that the results are proved for $i = 1, \ldots, k-1$. We have

$$
\begin{aligned}
E[(\tilde{Z}_{n+k} - \hat{Z}^*_{n+k})^2] &= E\left[\left(\sum_{j=1}^{k-1} \beta_j \tilde{Z}_{n+k-j} + \sum_{j=k}^{\infty} \beta_j Z_{n+k-j} - \sum_{j=1}^{k-1} \beta_j \hat{Z}^*_{n+k-j} - \sum_{j=k}^{n+k-1} \beta_j Z_{n+k-j}\right)^2\right] \\
&= E\left[\left(\left(\sum_{j=1}^{k-1} \beta_j \tilde{Z}_{n+k-j} - \sum_{j=1}^{k-1} \beta_j \hat{Z}^*_{n+k-j}\right) + \sum_{j=n+k}^{\infty} \beta_j Z_{n+k-j}\right)^2\right] \\
&= E\left[T_1^2 + 2T_1T_2 + T_2^2\right] \\
&= E[T_1^2] + 2E[T_1T_2] + 2E[T_2^2],
\end{aligned}
$$

where $T_1 = \sum_{j=1}^{k-1} \beta_j (\tilde{Z}_{n+k-j} - \hat{Z}^*_{n+k-j})$ and $T_2 = \sum_{j=n+k}^{\infty} \beta_j Z_{n+k-j}$.

Since all of the terms in T_1 are of the order $o_p(1)$, T_1 is hence an $o_p(1)$ term. Similarly as for $k = 1$, it can be shown that T_2 is also of the order $o_p(1)$. This leads to the conclusion that $E[(\tilde{Z}_{n+k} - \hat{Z}_{n+k})^2] = o(1)$ holds for $k > 1$.

ii) For $k = 1$ and any $t = 1, ..., n$:

$$
\begin{aligned}
E[(Z_{n+1} - \hat{Z}^*_{n+1})Z_t] &= E\left[\left(\sum_{i=1}^{\infty} \beta_i Z_{n+1-i} + \varepsilon_{n+1} - \sum_{i=1}^{n} \beta_i Z_{n+1-i}\right)\left(\sum_{j=1}^{\infty} \beta_j Z_{t-j} + \varepsilon_t\right)\right] \\
&= E\left[\left(\sum_{i=n+1}^{\infty} \beta_i Z_{n+1-i} + \varepsilon_{n+1}\right)\left(\sum_{j=1}^{\infty} \beta_j Z_{t-j} + \varepsilon_t\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= E\left[\sum_{j=1}^{\infty}\beta_j Z_{t-j}\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i} + \varepsilon_{n+1}\sum_{j=1}^{\infty}\beta_j Z_{t-j}\right.\\
&\quad\left. +\varepsilon_t\left(\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i} + \varepsilon_{n+1}\right)\right]\\
&= E\left[\sum_{j=1}^{\infty}\beta_j Z_{t-j}\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i}\right] + E\left[\varepsilon_{n+1}\sum_{j=1}^{\infty}\beta_j Z_{t-j}\right]\\
&\quad + E\left[\varepsilon_t\left(\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i} + \varepsilon_{n+1}\right)\right].
\end{aligned}
$$

Since $E\left[\varepsilon_{n+1}\sum_{j=1}^{\infty}\beta_j Z_{t-j}\right] = 0$ and $E\left[\varepsilon_t\left(\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i} + \varepsilon_{n+1}\right)\right] = 0$,

$$
\begin{aligned}
E[(Z_{n+1} - \hat{Z}^*_{n+1})Z_t] &= E\left[\sum_{j=1}^{\infty}\beta_j Z_{t-j}\sum_{i=0}^{\infty}\beta_{n+1+i}Z_{-i}\right] \qquad (A.3)\\
&= \sum_{i=0}^{\infty}\beta_{n+1+i}\sum_{j=1}^{\infty}\beta_j E[Z_{-i}Z_{t-j}]\\
&\le \sum_{i=0}^{\infty}|\beta_{n+1+i}|\sum_{j=1}^{\infty}\mid\beta_j\mid|\gamma(t+i-j)|\\
&\le \gamma(0)\sum_{i=0}^{\infty}|C_\beta|(n+1+i)^{-d-1} = O[(n+1)^{-d}] = o(1).
\end{aligned}
$$

Now, let $k > 1$ and assume that the results are proved for $i = 1, \ldots, k-1$, we have

$$
\begin{aligned}
E[(Z_{n+k} - \hat{Z}^*_{n+k})Z_t] &= E\left\{\left[\sum_{i=1}^{k-1}\beta_i(Z_{n+k-i} - \hat{Z}^*_{n+k-i}) + \sum_{i=n+k}^{\infty}\beta_i Z_{n+k-i} + \varepsilon_{n+k}\right]Z_t\right\}\\
&= E[T_3 + T_4 + T_5], \qquad (A.4)
\end{aligned}
$$

where $T_3 = Z_t\sum_{i=1}^{k-1}\beta_i(Z_{n+k-i} - \hat{Z}^*_{n+k-i})$, $T_4 = Z_t\sum_{i=n+k}^{\infty}\beta_i Z_{n+k-i}$ and $T_5 = \varepsilon_{n+k}Z_t$ with $E(T_5) = 0$. It is clear that $E(T_3) = o(1)$, because the results hold for $i = 1, \ldots, k-1$. The fact that $E(T_4) = o(1)$ can be proved similarly as for $k = 1$. Insert these results into (A.4) we obtain

$$
E[(Z_{n+k} - \hat{Z}^*_{n+k})Z_t] = o(1), t = 1, ..., n, \qquad (A.5)
$$

for any $k > 1$. Theorem 1 is proved. $\diamond$

Remark 1. *Some techniques used in the proof only apply to $d > 0$, while for $d < 0$ other approaches should be used. It is very common that some conclusions hold only for* *long memory errors but not for antipersistent errors. For instance, for $d > 0$ we have $\sum_{i=1}^{\infty} \beta_i = 1$. For $d < 0$, β_i are however not summable. Furthermore, the approximate formula of $\gamma(k)$ does not apply to $\gamma(i-j)$ in the fourth line of Eq.* (A.2)*. The reason is that although both i and j tend to infinity, their difference may be very small. Hence, here the fact that $|\gamma(k)| \leq \gamma(0)$ is simply employed. Detailed analysis of the second sum there may lead to more accurate result. This is however omitted to simplify the proof.*

Proof of Theorem 2. For a causal stationary and invertible ARMA model, the predictor $\tilde{Z}_{n+k}$ defined in Eq. (19) can be represented as a MA(∞) form (see e.g. Theorem 5.5.1 of Brockwell & Davis, 2006)

$$\tilde{Z}_{n+k} = \sum_{i=k}^{\infty} \alpha_i \varepsilon_{n+k-i}. \tag{A.6}$$

It can be shown that this fact also holds, if Z_t is a causal stationary and invertible FARIMA model considered in this paper. The difference between Z_{n+k} and $\tilde{Z}_{n+k}$ is:

$$Z_{n+k} - \tilde{Z}_{n+k} = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{n+k-i} - \sum_{i=k}^{\infty} \alpha_i \varepsilon_{n+k-i} = \sum_{i=0}^{k-1} \alpha_i \varepsilon_{n+k-1}. \tag{A.7}$$

The variance of $Z_{n+k} - \tilde{Z}_{n+k}$ is therefore

$$\mathrm{var}\left(Z_{n+k} - \tilde{Z}_{n+k}\right) = \sigma_\varepsilon^2 \sum_{i=0}^{k-1} \alpha_i. \tag{A.8}$$

Note that the point forecasting for the conditional mean, $\tilde{\zeta}_{n+k}$, is the same as $\tilde{Z}_{n+k}$. The difference between ζ_{n+k} and $\tilde{Z}_{n+k}$ is given by

$$\zeta_{n+k} - \tilde{Z}_{n+k} = \sum_{i=1}^{\infty} \alpha_i \varepsilon_{n+k-i} - \sum_{i=k}^{\infty} \alpha_i \varepsilon_{n+k-i} = \sum_{i=1}^{k-1} \alpha_i \varepsilon_{n+k-i} \tag{A.9}$$

with the variance

$$\mathrm{var}\left(\zeta_{n+k} - \tilde{Z}_{n+k}\right) = \sigma_\varepsilon^2 \sum_{i=1}^{k-1} \alpha_i^2. \tag{A.10}$$

In Theorem 1 it is shown that $\hat{Z}_{n+k} \approx \tilde{Z}_{n+k}$. Thus, $Z_{n+k} - \hat{Z}_{n+k} \approx Z_{n+k} - \tilde{Z}_{n+k}$ and $\zeta_{n+k} - \hat{Z}_{n+k} \approx \zeta_{n+k} - \tilde{Z}_{n+k}$. Consequently, $\mathrm{var}\left(Z_{n+k} - \hat{Z}_{n+k}\right) \approx \mathrm{var}\left(Z_{n+k} - \tilde{Z}_{n+k}\right)$ and $\mathrm{var}\left(\zeta_{n+k} - \hat{Z}_{n+k}\right) \approx \mathrm{var}\left(\zeta_{n+k} - \tilde{Z}_{n+k}\right)$. Theorem 2 is proved. $\diamond$

Reference

Allen, D.E., Chan, F., McAleer, M., & Peiris, S. (2006). Finite sample properties of the QMLE for the Log-ACD model: Application to Australian Stocks. Preprint, *Edith Cowan University*.

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

Bauwens, L., Galli, F., & Giot, P. (2008). The moments of Log-ACD models. *Quantitative and Qualitative Analysis in Social Sciences*, 2, 1-28.

Bauwens, L., & Giot, P. (2000). The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annales d'Économie et de Statistique*, 60, 117-149.

Beran, J. (1994). Statistics for Long-Memory Processes. New York: Chapman & Hall.

Beran J., & Feng, Y. (2002a). SMIFAR models- a semiparametric approach to modelling trends, long-range dependence and nonstationarity. *Computational Statistics & Data Analysis*, 40, 393-419.

Beran J., & Feng, Y. (2002b). Local polynomial fitting with long-memory short-memory and antipersistent errors. *Ann. Institute of Statistical Mathematics*, 54, 291-311.

Beran, J., Feng, Y., & Ghosh, S. (2012). Exponential FARIMA and SEMIFAR models and their application. Preprint, University of Konstanz.

Beran, J., & Ocker, D. (1999). SEMIFAR forecasts, with applications to foreign exchange rates. *Journal of Statistical Planning and Inference* 80, 137-153.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

Brockwell, P.J., & Davis, R.A. (2006). Time Series: Theory and Methods (2nd edition). New York : Spring

Dittman, I. and Granger, C. (2002). Properties of nonlinear transformations of fractionally integrated processes. *Journal of Econometrics*, 110, 113-133.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007

Engle, R.F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17, 425-446.

Engle, R.F., & Rangel, J.G. (2008). The Spline-GARCH model for low frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, 21, 1187-1222.

Engle, R.F., & Russell, J.R. (1998). Autoregressive conditional duration: A new approach for irregularly spaced transaction data. *Econometrics*, 66, 1127-1162.

Feng, Y.(2004). Simultaneously modelling coditional heteroskedasticity and scale change. *Econometric Theory*, 20, 563-596.

Feng, Y., & Beran J. (2013). Optimal convergence rates in non-parametric regression with fractional time series errors. *Journal of Time Series Analysis*, 34, 30-39.

Geweke, J. (1986). Modeling the persistence of conditional variances: A comment. *Econometric Review*, 5, 57-61.

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

Jasiak, J. (1998). Persistence in intertrade durations. *Finance*, 19, 166-195.

Karanasos, M. (2004). The statistical properties of long-memory ACD models. *WSEAS Transactions on Business and Economics*, 2, 169-175.

Karanasos, M. (2008). The statistical properties of exponential ACD models. *Quantitative and Qualitative Analysis in Social Sciences*, 2, 29-49.

Karanasos, M., Psaradakis, Z., & Sola, M. (2004). On the autocorrelation properties of long-memory GARCH processes. *Journal of Time Series Analysis*, 25, 265-281.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets*, 8, 377 - 399.

Pantula, S. G. (1986). Modelling the persistence of conditional variances: A comment. *Econometric Reviews*, 5, 71-73.

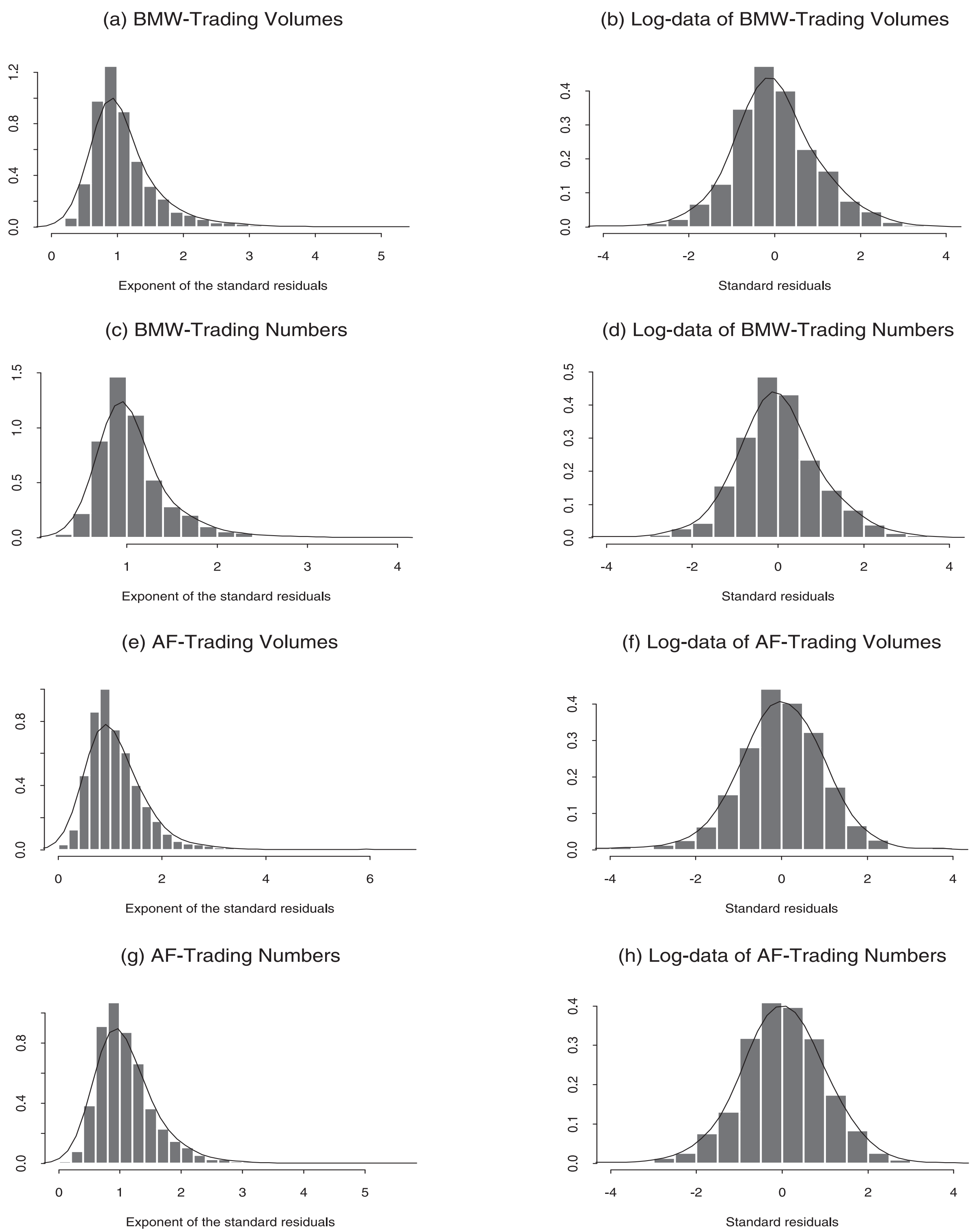

Figure 1: Histograms of the standardized residuals of the SEMIFAR model and their exponential transformation for all examples.

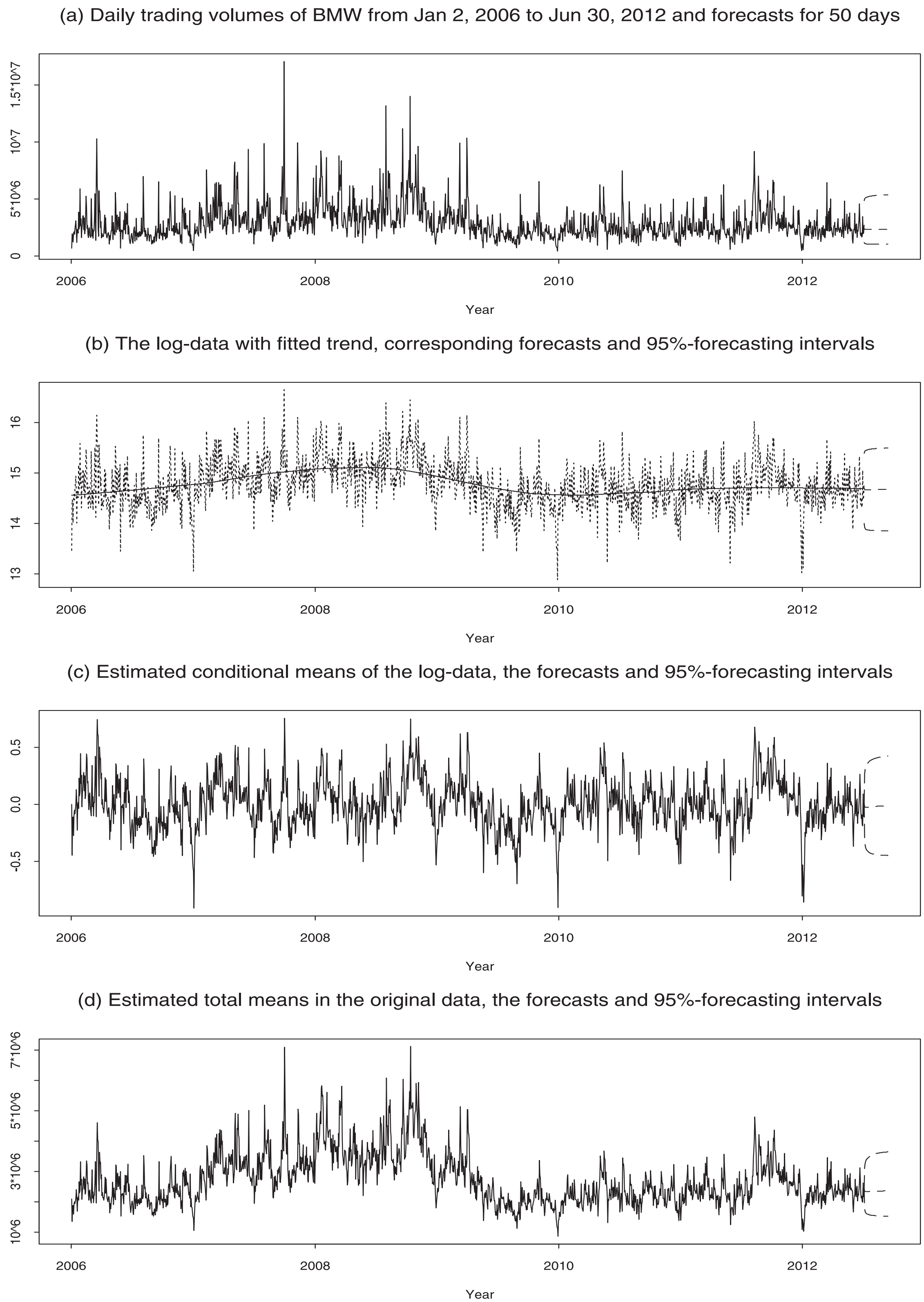

Figure 2: Estimation and forecasting results of the daily trading volumes of BMW from Jan 02, 2006 to Jun 30, 2012, obtained by the Semi-FI-Log-ACD model.

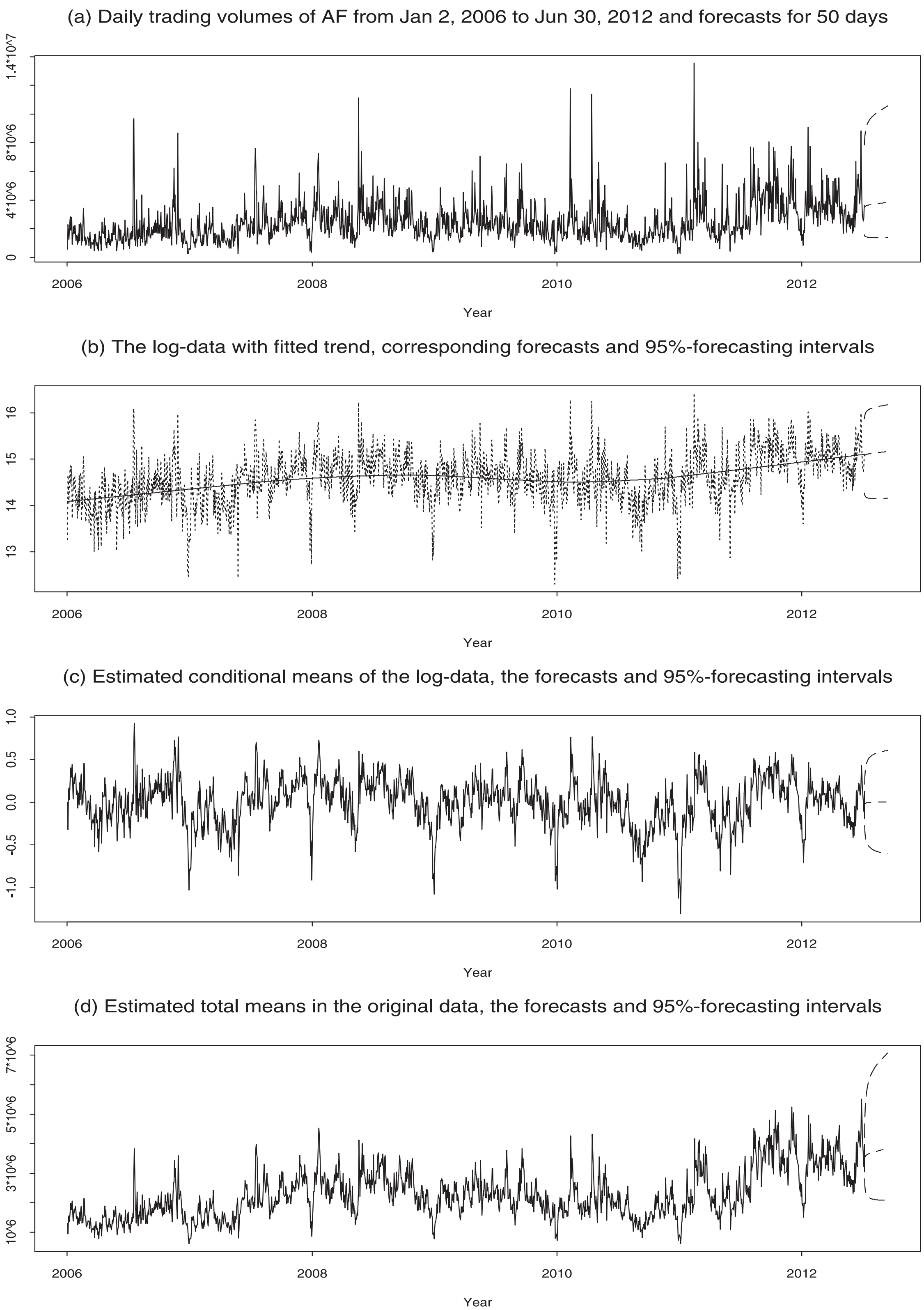

Figure 3: The same results as given in Fig. 2 for daily trading volumes of Air France.

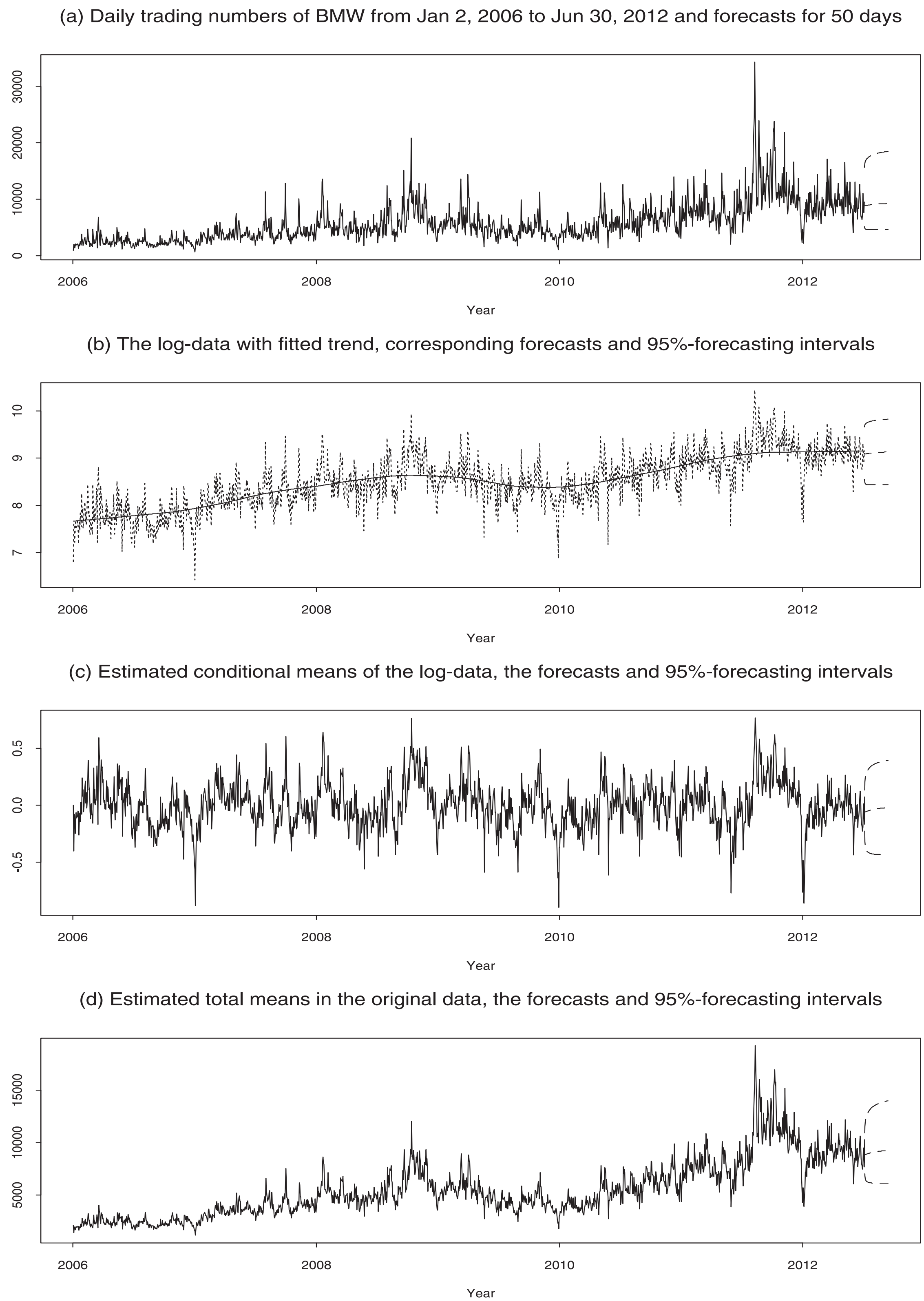

Figure 4: The same results as given in Fig. 2 for daily trading numbers of BMW.

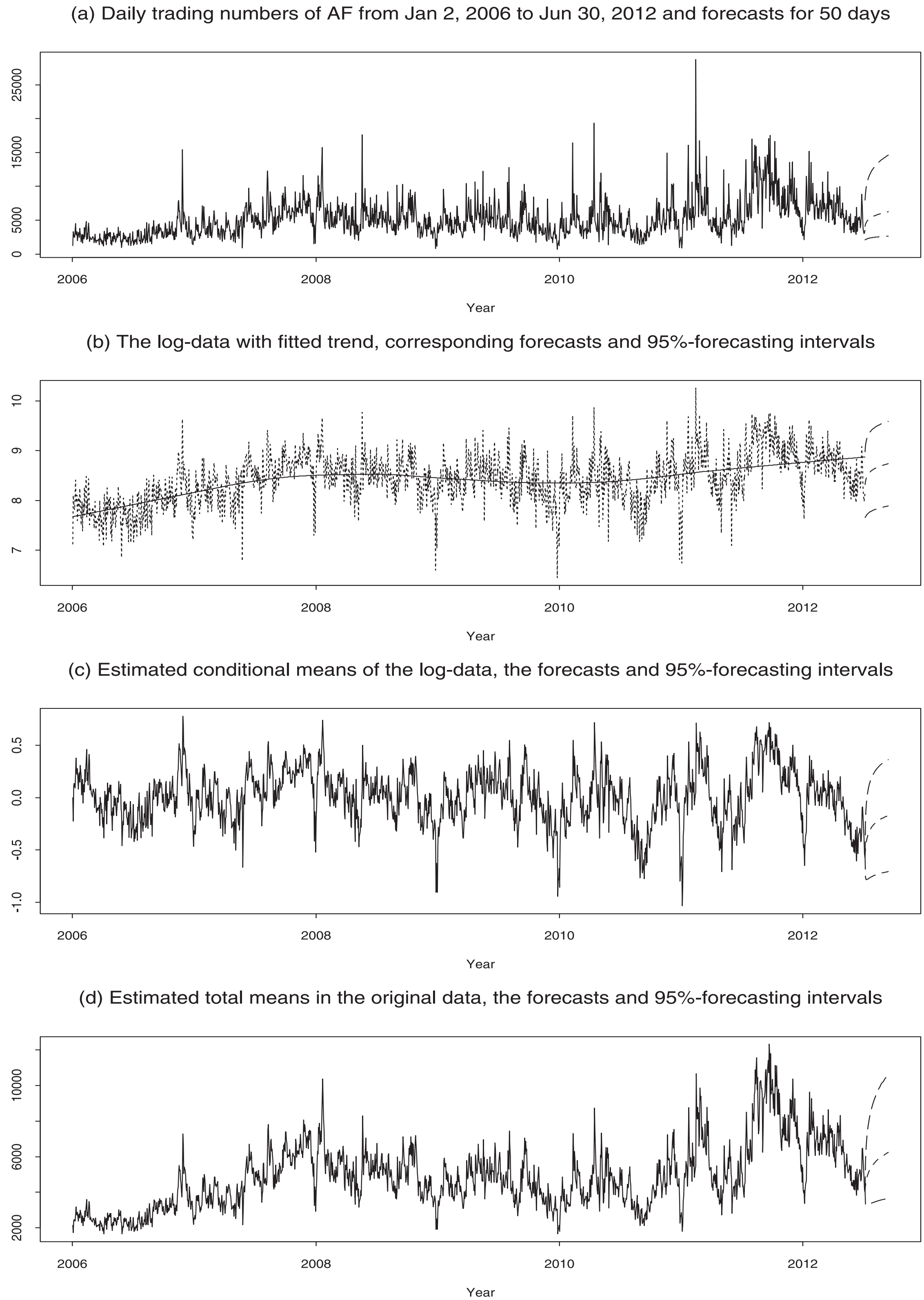

Figure 5: The same results as given in Fig. 2 for daily trading numbers of Air France.

Table 1: Results of ESEMIFAR models for the four datasets

Series		$\hat{h}$	$\hat{d}$ & 95%-CI	$\hat{p}$	$\hat{\phi}_1$ & 95%-CI	trend
VOL	BMW	0.146	0.299 [0.233, 0.366]	1	0.109 [0.024, 0.193]	sign.
	AF	0.243	0.393 [0.355, 0.431]	0	—	insign.
TrN	BMW	0.124	0.329 [0.260, 0.398]	1	0.142 [0.055 0.228]	sign.
	AF	0.207	0.409 [0.372, 0.447]	0	—	sign.