



SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY



CREATES
Center for Research in Econometric
Analysis of Time Series

CREATES Research Paper 2009-31

A No Arbitrage Fractional Cointegration Analysis Of The Range Based Volatility

Eduardo Rossi and Paolo Santucci de Magistris

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

A No Arbitrage Fractional Cointegration Analysis Of The Range Based Volatility *

Eduardo Rossi[†]

Paolo Santucci de Magistris[‡]

July 15, 2009

Abstract

The no arbitrage relation between futures and spot prices implies an analogous relation between futures and spot volatilities as measured by daily range. Long memory features of the range-based volatility estimators of the two series are analyzed, and their joint dynamics are modeled via a fractional vector error correction model (FVECM), in order to explicitly consider the no arbitrage constraints. We introduce a two-step estimation procedure for the FVECM parameters and we show the properties by a Monte Carlo simulation. The out-of-sample forecasting superiority of FVECM, with respect to competing models, is documented. The results highlight the importance of giving fully account of long-run equilibria in volatilities in order to obtain better forecasts.

Keywords. Range-based volatility estimator, Long memory, Fractional cointegration, Fractional VECM, Stock Index Futures.

J.E.L. classification. C32, C13, G13.

*We would like to thank Giuseppe Cavaliere, Katarzyna Łasak for very helpful and constructive comments on an earlier version of this paper. We also thank the participants to the CREATES seminar.

[†]Dipartimento di economia politica e metodi quantitativi, Via San Felice 5, University of Pavia, Italy. Corresponding author. Tel.: +39 0382/986207 Fax: +39 0382/304226. *E-mail address*: erossi@eco.unipv. Financial support from PRIN 2006 is gratefully acknowledged.

[‡]Dipartimento di economia politica e metodi quantitativi, University of Pavia, Italy. Paolo Santucci de Magistris thanks CREATES, funded by the Danish National Research Foundation, for hospitality.

Introduction

The relationship between spot and futures prices is a widely studied topic in the financial literature. Since the introduction of future contracts during the 1970s, a great effort has been made by academics and practitioners in understanding the pricing of future contracts.

Considering forward contracts, the no arbitrage assumption implies that, in a frictionless market, the spot and the forward prices, under risk neutral probability, are related by:

$$F_{t|t-k} = S_{t-k} \cdot e^{k \cdot r_{t|t-k}} \quad (1)$$

where $r_{t|t-k}$ is the return of a risk free asset that expires in period t and $e^{k \cdot r_{t|t-k}}$ is referred to as the cost of carry premium.

We investigate to what extent the no arbitrage condition in (1) induces a long run relationship between spot and futures volatilities, when, as in Parkinson (1980), Garman and Klass (1980), Wiggins (1992) and Alizadeh et al. (2002), we adopt as a measure of daily integrated volatility the daily range. We also show how the information implicit in the future contracts could be exploited in forecasting the volatility of spot prices.

Daily range is equal to the difference between the highest and lowest log price of a given day

$$R_t = \max_{\tau} \log P_{\tau} - \min_{\tau} \log P_{\tau} \quad (2)$$

where P_{τ} is the asset price at time $t - 1 < \tau \leq t$. As noted by Andersen and Bollerslev (1998), the accuracy of the high-low estimator is near that provided by the realized volatility estimator based on 2 or 3 hours returns. Daily range-based volatility estimator is still 5 times more efficient than the traditional daily squared returns, see Christensen and Podolskij (2007).

Combining equation (1) and (2) and assuming that the risk free rate is constant with respect to τ , we obtain a no arbitrage equilibrium relationship between the forward and spot ranges,

$$R_t^F = R_t^S. \quad (3)$$

Under the hypothesis that spot log price evolves as a random walk in continuous time, with the diffusion parameter ω_{τ} such that $\omega_{\tau} = \omega_t \forall \tau \in (t - 1, t)$, an unbiased estimator of daily

integrated volatility, σ_t^2 , is equal to

$$\sigma_t^2 = 0.361 \cdot R_t^2. \quad (4)$$

where, as demonstrated in Parkinson (1980), $E_{t-1}[\sigma_t^2] = \omega_t^2$ and $MSE_{t-1}[\sigma_t^2] = 0.4073\omega_t^4$, which is approximately one-fifth of the MSE of the daily squared return. We can then recast equation (3) as

$$\log \sigma_{t,F} = \log \sigma_{t,S}. \quad (5)$$

The equality in (5) is implied by the no-arbitrage pricing relationship in (1).

However, as pointed out in Andersen et al. (2001), Andersen et al. (2003), and Lieberman and Phillips (2008), volatilities can be assumed neither as an I(0) nor as an I(1) process, but we should consider that they can be best approximated by a fractionally integrated process, or I(d), where the parameter d can take any real value. In particular, when we model the relation between log-ranges we have to allow for the possibility that they can be fractionally cointegrated. In fact, the traditional cointegration concept only allows for an integer order of integration in the equilibrium error process. This assumption appears to be too restrictive here. Fractional cointegration refers to a generalization of the concept of cointegration, that is linear combinations of I(d) processes are I($d - b$), with $0 < b \leq d$. This corresponds to the idea that there exists a common stochastic trend, that is integrated of order $b > 0$, while the short period departures from the long run equilibrium are integrated of order $d - b \geq 0$. b stands for the fractional order of reduction obtained by the linear combination of I(d) variables. Christensen and Nielsen (2006) applied the concept of fractional cointegration in examining the relationship between implied and realized volatility, estimating the cointegration vector in a regression setup. Cheung and Lai (1993) applied the concept of fractional cointegration in studying the purchase power parity hypothesis with a parametric model, and Duecker and Startz (1998) investigated the relationship between bond rates estimating a bivariate long memory model, restricted in order to allow for cointegration. Robinson (1994) showed inconsistency of OLS estimator in the usual cointegration regression when the regressors are fractionally integrated, and proposed an estimation approach based on frequency domain least squares. Christensen and Nielsen (2006) derived the asymptotic distribution of the estimators in the stationary case, i.e. $d < \frac{1}{2}$, while Robinson and Marinucci (2003) provided the limiting distribution

in the case $d > \frac{1}{2}$. Christensen and Nielsen (2006) studied the fractional cointegration relationship between realized and implied volatilities, finding that their common fractional order of integration is reduced by a linear combination of them. More recently, many authors focused on the estimation of the cointegration rank. In particular, Robinson and Yajima (2002) derived a formal semiparametric test for the cointegration rank based on the spectral representation of the system. Nielsen and Shimotsu (2007), extended the analysis of Robinson and Yajima (2002), proposing a fractional cointegration testing procedure based on the exact Whittle estimator for both stationary and nonstationary processes. Breitung and Hassler (2002) provided a test based on a multivariate trace test, similar to that proposed by Johansen (1988), that is based on the solution of a generalized eigenvalue problem. However, they consider only the cointegration relation between non-stationary variables, such that $d > 0.5$.

In this paper, we jointly model the dynamics of future and spot volatility via the fractional cointegration system outlined in Granger (1986), that is a generalization of the vector ECM. The fractional VECM (FVECM hereafter), allows for a flexible characterization of the cointegration relation, in the sense that the integration orders of the endogenous variables, d , and the cointegration residuals, $d - b$, are not restricted to assume values 1 and 0, respectively. On the other hand, the estimation procedure presents an additional difficulty with respect to the standard VECM, since two additional parameters, d and b , need to be estimated. Lasak (2008) suggests a profile likelihood procedure that allows to firstly estimate b , while d is set to 1. In this paper, we provide an extension of the estimation procedure outlined in Lasak (2008), carrying out a new inference method that maximizes the profile likelihood in terms of both d and b , obtaining a joint estimate of them. A Monte Carlo simulation assesses the ability of the profile maximum likelihood procedure to correctly estimate the true parameters of the model, showing that the estimated parameters converge to the true values when the sample size increases. An out-of-sample forecasting exercise is then carried out in order to verify the capacity of FVECM, which includes an error correction term based on a no arbitrage restriction, to lead to significant improvements on models which do not account for the long run equilibrium. As expected, we find that FVECM consistently outperforms the competing models in terms of accuracy of forecasts. This evidence clearly confirms that imposing the no arbitrage condition produces superior long-horizon forecasts.

The paper is organized as follows. Section 1 presents a brief description of the data and the analysis of the long memory property of range-based volatility estimator, assessing the equality of the integration orders between spot and future volatility and showing that the two series have to be considered fractionally cointegrated. Given the evidence provided in section 1, section 2 introduces the FVECM, as in Granger (1986) and Davidson et al. (2006), in order to jointly model the dynamics of fractionally cointegrated series. Section 3 introduces the estimation method adopted, and section 4 reports the estimation results. Section 5 provides evidence in favor of the FVECM in terms of forecasting ability and section 6 concludes.

1 Data and Preliminary Analysis

The data used in this paper consist of the high and low daily spot and future prices on the S&P500 index for the period 27th November 1998 to 5th September 2008 for a total of 2450 trading days. The future prices are relative to the 3 months future contracts. Given the long period of time under analysis, it seems natural to analyze the possible presence of structural breaks in the series of log daily volatilities. In fact, as pointed out by Granger and Hyung (2004), the long memory property of volatility could be induced spuriously by the presence of structural breaks. For this reason, we search for structural changes in the series under analysis, following the procedure outlined by Bai and Perron (2003). In the following analysis we will refer to the residuals of the Bai-Perron procedure as the demeaned series while the original series will be called raw. The temporal dynamics of the series are very close each other, and the breakpoints are found to have the same intensity in correspondence of the same observations.¹. It is well known that volatility series are clustered and present a slow (hyperbolic) decay of the autocorrelation functions, see Andersen et al. (2001), that refer to this feature as induced by the presence of long memory. Long memory is defined in terms of decay rates of long-lag autocorrelations, or in the frequency domain in terms of rates of explosion of low frequency spectra. A long-lag autocorrelation definition of long memory is

$$\gamma(\tau) = c\tau^{2d-1} \quad \tau \rightarrow \infty \quad (6)$$

the correlations of long memory process decay with a hyperbolic rate. They are not summable. An alternative, although not equivalent, definition of long range dependence can be obtained in terms of the spectral density $f(\lambda)$ of the process:

$$\lim_{\lambda \rightarrow 0^+} \frac{f(\lambda)}{c_f |\lambda|^{-2d}} = 1 \quad 0 < c_f < \infty. \quad (7)$$

The spectral density $f(\lambda)$ has a pole and behaves like a constant c_f times λ^{-2d} at the origin. If $|d| \in (0, 1/2)$ the process is stationary. In particular, if $d \in (0, 1/2)$, it presents long memory; instead, if $d \in (-1/2, 0)$ the process is antipersistent with short memory. A popular approach to the modeling of long memory is represented by the ARFIMA class introduced by Granger and Joyeux (1980) and Hosking (1981).

Figure 1 displays the autocorrelation functions of the original volatility series with respect to the residual series from the breakpoint searching analysis. The evidence from the ACF is clearly supportive of the long memory property of the volatility series since the autocorrelation functions decay very slowly. However, the reduction in the persistence obtained by removing the breaks from the original series is evident. Even if the sensible reduction in the persistence, the residual series still display a certain level of long memory. Given the equilibrium relation stated in equation (5), we test for the possibility of fractional cointegration considering both the raw series and the demeaned series. Robinson and Yajima (2002) discuss a semi-parametric procedure for determining the cointegration rank, focusing on stationary series. Nielsen and Shimotsu (2007) extend the analysis of Robinson and Yajima (2002), in order to consider cointegration for both stationary and non-stationary variables. In particular, they apply the exact local Whittle analysis in a multivariate setup, see Shimotsu and Phillips (2005), and estimate the rank of spectral cointegration of the d th differenced process by examining the rank of the spectral matrix, G , around the origin. As pointed out by the authors, their approach does not require the estimation of any cointegration vector, but it relies on the choice of bandwidths and threshold parameters.

Since the presence or absence of cointegration is not known when the fractional integration order is estimated, they propose, as in Robinson and Yajima (2002), a test statistic for the equality of integration orders that is informative in both circumstances, in the bivariate case

$$\hat{T}_0 = m(S\hat{d})' \left(S \frac{1}{4} \hat{D}^{-1} (\hat{G} \odot \hat{G}) \hat{D}^{-1} S' + h(T)^2 \right)^{-1} (S\hat{d}) \quad (8)$$

where \odot denotes the Hadamard product, $S = [1, -1]'$, $h(T) = \log(T)^{-k}$ for $k > 0$, $D = \text{diag}(G_{11}, G_{22})$, while $\hat{G} = \frac{1}{m} \sum_{j=1}^m \text{Re}(I_j)$ and I_j is the coperiodogram at the frequency λ_j (see Nielsen and Shimotsu (2007) for more details). The parameter \hat{d} is the exact Local Whittle estimator of d , introduced by Shimotsu and Phillips (2005). If the variables are not cointegrated, that is the cointegration rank r is zero, $\hat{T}_0 \rightarrow \chi_1^2$, while if $r \geq 1$, $\hat{T}_0 \rightarrow 0$. A significantly large value of \hat{T}_0 , with respect to χ_1^2 , can be taken as an evidence against the equality of the integration orders.

Moreover, the estimation of the cointegration rank r is obtained by calculating the eigenvalues of the estimated matrix \hat{G} . The estimator \hat{G} uses a new bandwidth parameter n . Let $\hat{\delta}_i$ the i th eigenvalue of \hat{G} , it is possible to apply a model selection procedure to determine r . In particular,

$$\hat{r} = \arg \min_{u=0,1} L(u) \quad (9)$$

where

$$L(u) = v(T)(2-u) - \sum_{i=1}^{2-u} \hat{\delta}_i \quad (10)$$

for some $v(T) > 0$ such that

$$v(T) + \frac{1}{n^{1/2}v(T)} \rightarrow 0. \quad (11)$$

Tables 1 and 2 show the results of the Nielsen and Shimotsu (2007) fractional cointegration analysis, with two different choices for the bandwidths, m , used in the estimation of d 's in the exact local Whittle estimation, and n used in the estimation of G and $L(u)$. The estimates of the long memory parameter, d , are close to $1/2$. Otherwise, when we consider the residuals of the Bai and Perron (2003) procedure for the breaks (demeaned series), the estimated d falls into the stationarity region for all bandwidths. The \hat{T}_0 statistic takes values close to 0 for both the raw data and demeaned series. The analysis of the cointegration rank, in Panel B, confirms the presence of cointegration, in fact $\hat{r} = 1$ in all cases. Since the 95% critical value of a χ_1^2 is 3.841, we cannot reject the null of equality of the integration orders in all cases. Interestingly, the series are fractionally cointegrated even if the presence of structural breaks is removed.

The result of the test of Nielsen and Shimotsu (2007) confirms that volatility of spot and future prices have the same fractional integration order and are cointegrated.

Moreover, given the equality of the integration orders, we estimate the cointegrating vector

in a regression setup, as suggested by Engle and Granger (1987). Given the equilibrium relation stated in equation (5), it seems to be natural to test empirically the difference

$$\log \sigma_{t,F} - \log \sigma_{t,S} = z_t \sim I(d - b) \quad (12)$$

where $b > 0$ and z_t is a stationary fractionally integrated variable with fractional order less than d , i.e. $d_z = d - b < d$. In this context, we call *full* cointegration the case in which $b = d$, that is the case in which the order of integration of the residual is zero. The typical case considered in empirical works is $d = b = 1$, that is X_t are $I(1)$ and z_t is $I(0)$. Cointegration requires that z_t is mean reverting, that is a long run restriction on the dynamics of X_t .

A simple way to test this hypothesis is examining the degree of persistence of the residuals of the following regression

$$\log \sigma_{t,F} = \beta \log \sigma_{t,S} + z_t \quad (13)$$

under the assumption that spot and future volatility have the same d . The parameter β is estimated with the Frequency Domain Least Squares (FDLS), as suggested by Robinson and Marinucci (2003). Since the series of demeaned series are stationary and present long memory (the estimated d is between 0 and 0.5), we follow the approach suggested by Christensen and Nielsen (2006). The parameter β is estimated with the Narrow Band Frequency Domain Least Squares (NB-FDLS, hence after).² As noted by Robinson (1994) and Robinson and Marinucci (2003), when x_t is stationary, the $\hat{\beta}_m$ is consistent for β due to dominance of the spectrum of $\log \sigma_{t,S}$ over that of z_t near zero frequency. Christensen and Nielsen (2006) have derived the asymptotic distribution of $\hat{\beta}_m$ when $0 < d < \frac{1}{2}$ and $0 < b < d$. In particular, in the simple case of two variables, this is equal to

$$\sqrt{m} \lambda_m^{d_z - d} (\hat{\beta}_m - \beta) \rightarrow N \left(0, \frac{g_e(1 - 2d_x)^2}{2g_x(1 - 2d_x - 2d_z)} \right) \quad (14)$$

where g_e and g_x correspond to $\text{var}(\Delta^b z_t)$ and $\text{var}(\Delta^d x_t)$. The results of the procedure of Christensen and Nielsen (2006) for different bandwidth, m , are presented in table 3. When the demeaned series are taken into account, the estimates of β and d from table 3 strongly supports the idea of fractional cointegration among the two series. In fact, the fractional order of the residuals is close to zero, so that $b \approx d$, and we get full cointegration. Moreover, the estimated cointegration parameter, $\hat{\beta}_m$, is very close to 1, in particular when the NB-

FDLS are used instead of OLS.

On the other hand, when the cointegration regression is implemented on the raw data, we are not able to compute the standard error of β_m , since $\hat{d} + \hat{d}_z > 0.5$.³ Nevertheless, the point estimate of β_m is in all cases very close to the theoretical value 1, but the fractional reduction of the integration order in this case is not complete, since $\hat{d}_z \approx 0.2$. This result can be explained by the presence of structural breaks, that are a source of non stationarity. For this reason, the following analysis will be conducted on the demeaned series.

2 The Model

Given the analysis in the previous chapter, we propose to jointly model the dynamics of the spot and future volatilities through a fractionally integrated vector error correction model that accounts for their equilibrium relation. Differently from Davidson (2002) and Duecker and Startz (1998), our series are stationary. We consider the Granger (1986) representation of fractionally cointegrated systems, that is the Fractionally Integrated Vector ECM (FVECM)⁴, given by

$$\Delta^d X_t = (1 - \Delta^b)(\Delta^{d-b} \alpha \beta' X_t) + \sum_{j=1}^{k-1} \Gamma_j \Delta^d X_{t-j} + \epsilon_t \quad t = 1, \dots, T \quad (15)$$

where $X_t = (\log \sigma_{t,F}, \log \sigma_{t,S})$, and Γ_j are the short run matrices of parameters. ϵ_t is a independent identically distributed vector with mean 0 and covariance matrix Ω . α and β are the error correction and cointegration matrices. Our setup slightly differs from the bivariate model proposed by Duecker and Startz (1998). Their model is a bivariate ARFIMA process, with the additional restriction that the fractional difference of a linear combination of the two series is $d - b \geq 0$; the cointegrating parameter β is treated as an additional unknown parameter in constructing the Gaussian likelihood function (see Sowell (1989) and Sowell (1992)). An alternative representation of fractionally cointegrated systems is presented in Johansen (2008).

In the FVECM model, the element α_{ij} of the parameters matrix α measures the single period response of variable i to the shock on the equilibrium relation j . In our case, $j = 1$ with $\alpha = (\alpha_1, \alpha_2)'$, thus α_1 should be negative in order to move toward the unique long run relationship implied by the no arbitrage assumption. The vector coefficient α has a

clear interpretation as a short term adjustment coefficient and represents the proportion by which the long run disequilibrium in the spot (futures) volatility is being corrected in each period.

Omitting the vector autoregressive terms, our model is defined as

$$\begin{aligned}(1-L)^d \log \sigma_{t,F} &= \alpha_1((1-L)^{(d-b)} - (1-L)^d)(\log \sigma_{t,F} - \beta \log \sigma_{t,S}) + \epsilon_{1t} \\ (1-L)^d \log \sigma_{t,S} &= \alpha_2((1-L)^{(d-b)} - (1-L)^d)(\log \sigma_{t,F} - \beta \log \sigma_{t,S}) + \epsilon_{2t}\end{aligned}$$

where the $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})$ are assumed to be Gaussian with mean zero and variance Ω .

3 Estimation

As pointed out by Lasak (2008), the estimation procedure of the fractional vector error correction model (15) presents an additional source of uncertainty with respect to the standard VECM (where d and b are restricted to be equal to 1), since the additional unknown parameters d and b need to be estimated. A solution to this problem is provided by Lasak (2008), that suggests to impose $d = 1$ and concentrate the likelihood function with respect to b . In particular, the model (15) is estimated via a maximum likelihood technique analogous to that developed by Johansen (1991) for the standard VECM, where the initial step consists of maximizing the profile likelihood with respect to b . In our case, we do not restrict the parameter d to be equal to 1, but we propose to jointly estimate the parameters d and b , by maximizing the profile likelihood with respect to both parameters at the initial stage.

The estimation procedure begins defining $Z_{0,t} = \Delta^d X_t$ and $Z_{1,t} = (\Delta^{d-b} - \Delta^d)X_t$ while $Z_{k,t} = (\Delta^d X_{t-1}, \dots, \Delta^d X_{t-k+1})$, so that the system (15) can be written as

$$Z_{0,t} = \alpha\beta'Z_{1,t} + \Gamma Z_{k,t} + \epsilon_t, \quad (16)$$

where $\Gamma = (\Gamma_1, \dots, \Gamma_{k-1})$. We define the matrix of cross product matrix as

$$M_{ij} = T^{-1} \sum_{t=1}^T Z_{i,t} Z'_{j,t} \quad i, j = 0, 1, k, \quad (17)$$

so that the residual of the regression of $Z_{0,t}$ and $Z_{1,t}$ on $Z_{k,t}$ can be defined as

$$R_{i,t} = Z_{i,t} - M_{ik}M_{kk}^{-1}Z_{k,t} \quad i = 0, 1. \quad (18)$$

The residual sum of squares results to be:

$$S_{ij} = M_{ij} - M_{ik}M_{kk}^{-1}M_{kj}, \quad (19)$$

where S_{ij} depends on d and b , except when $i = j = 0$. For fixed d , b , α and β , Γ is estimated as

$$\Gamma(d, b, \alpha, \beta) = (M_{0k} - \alpha\beta' M_{1k})M_{kk}^{-1}, \quad (20)$$

For fixed d , b and β , α and Ω are estimated as

$$\hat{\alpha}(d, b, \beta) = S_{01}\beta(\beta' S_{11}\beta)^{-1}, \quad (21)$$

and

$$\hat{\Omega}(d, b, \beta) = S_{00} - \hat{\alpha}(\beta' S_{11}\beta)\hat{\alpha}'. \quad (22)$$

Plugging this estimates into the likelihood we get

$$L(d, b, \hat{\alpha}, \beta, \hat{\Omega}) = |S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10}|, \quad (23)$$

that is maximized by the eigenvector corresponding to the r maximum eigenvalues that solve the problem

$$|\lambda S_{11} - S_{01}S_{00}^{-1}S_{01}| = 0 \quad (24)$$

The vector β is estimated as the r -dimensional space spanned by the eigenvectors, v_j for $j = 1, \dots, r$, corresponding to the r largest eigenvalues of $S_{11}^{-1}S_{01}S_{00}^{-1}S_{01}$.

Given this choice of β , the likelihood function is maximized only with respect to d and b , that is

$$(\hat{d}, \hat{b}) = \arg_{d,b} \max L(d, b), \quad (25)$$

where

$$L_T(d, b) = \left[|S_{00}| \prod_{i=1}^r (1 - \lambda_i) \right]^{-\frac{T}{2}} \quad (26)$$

Once d and b are estimated, as the values that maximize the function $L(b, d)$, all the other parameter of model (15) are obtained as functions of \hat{d} and \hat{b} .

Note that $\hat{\beta}$, $\hat{\alpha}$, $\hat{\Gamma}$ and $\hat{\Omega}$ are all consistent estimator given that \hat{d} and \hat{b} are consistent estimators of d and b , since when d and b are known to be estimated consistently, $\hat{\beta}$, $\hat{\alpha}$, $\hat{\Gamma}$ and $\hat{\Omega}$ are also consistent estimates, given the Continuous Mapping Theorem and the results in Johansen (1995). On the other hand, we conjecture that \hat{d} and \hat{b} exist and are consistent estimates of d and b if the objective function in equation (26) is uniquely maximized at the true point (d, b) , that is $L_T(d, b)$ being concave. In appendix, we show, in Monte Carlo setup, the consistency of all the estimators.

4 Results

System (15) for spot and future range-based volatility has been estimated using the procedure described in section 3. Since the number of lags included plays an important role in this context, we first implement two information-based criteria for the model selection. The Schwarz and Hannan-Quinn information criterion functions are minimized for $p = 1$ and $p = 2$ for respectively. The following analysis is based on $p = 1$.

The estimation results are reported in table 4.⁵ Table 4 also reports the bootstrapped confidence interval of the parameters. The bootstrapped sample has been generated with 1000 replications of *wild* bootstrap, that is robust to heteroskedastic effects. Most of the confidence intervals of the parameters does not contain the value 0, meaning that they are statistically different from zero. $\hat{\beta}$ is very close to the NB-FDLS estimates obtained in the previous section, and the theoretical value $\beta = 1$ is contained in the 90% confidence interval. Moreover, the estimates of d and b are close to the values obtained in the semi-parametric analysis in section 1. From a visual inspection of figure 2, it clearly appears that the residual component of the cointegration relation does not have any long memory feature, as confirmed by the semiparametric estimate of the parameter d , that is close to 0. The vector α highlights the speed of adjustment to the long run equilibrium; the volatility of the spot price returns faster to the equilibrium. Moreover, the bootstrapped 95th quantile of α_1 is positive, so that α_1 cannot be considered statistically different from zero, and there are no changes in the future volatility due to shocks in the cointegration relation. All the corrections to the equilibrium are made by changes in the spot volatility. This means

that future volatility has to be considered *weakly exogenous* with respect to spot volatility in this context, that is future volatility leads spot volatility to the equilibrium, implicitly confirming the Cox (1976)'s hypothesis on the efficiency of the future market in processing the new information.

5 Forecast

If the forecasting performances, at large horizons, depend on the model specification, then allowing for a mechanism restoring the long-run equilibrium between future and spot volatility, as in the fractional VECM could improve the forecasts. We evaluate the accuracy of our model in forecasting volatility and we compare the out-of-sample forecasts of FVECM with those provided by alternative model specifications. In particular, we consider

- Vector Autoregression model with 4 lags, $VAR(4)$;
- Univariate HAR model, $UHAR$, proposed by Corsi (2009), where the observed log volatility is regressed on its own daily, weekly and monthly past values

$$\log \sigma_t = \omega + \beta_1 \log \sigma_{t-1} + \beta_2 W_{t-1} + \beta_3 M_{t-1} + u_t$$

where $W_{t-1} = \frac{1}{5} \sum_{i=1}^5 \log \sigma_{t-i}$ and $M_{t-1} = \frac{1}{22} \sum_{i=1}^{22} \log \sigma_{t-i}$;

- Bivariate HAR, $BHAR$, where we include past values of the future (spot) log volatility in the equation of the spot (future) volatility:

$$\begin{aligned} \log \sigma_t^F &= \omega_1 + \beta_{11} \log \sigma_{t-1}^F + \beta_{12} W_{t-1}^F + \beta_{13} M_{t-1}^F + \beta_{14} \log \sigma_{t-1}^S + \beta_{15} W_{t-1}^S + \\ &\quad + \beta_{16} M_{t-1}^S + u_{t,1} \\ \log \sigma_t^S &= \omega_2 + \beta_{21} \log \sigma_{t-1}^F + \beta_{22} W_{t-1}^F + \beta_{23} M_{t-1}^F + \beta_{24} \log \sigma_{t-1}^S + \beta_{25} W_{t-1}^S + \\ &\quad + \beta_{26} M_{t-1}^S + u_{t,2} \end{aligned}$$

The system's parameters are estimated by least squares.

- ARFIMA(0,d,0) model, $ARFIMA$, that accounts for the long memory feature of the data;

$$(1 - L)^d \log \sigma_t = u_t \tag{27}$$

where $u_t \sim iiN(0, v^2)$;

- FIVAR(p,d) model, $FIVAR$,:

$$\Phi(L)(1-L)^d \log \sigma_t = u_t \quad (28)$$

where $\Phi(L)$ is an autoregressive matrix polynomial so that the FIVAR could be considered a VAR calculated on the fractionally differenced series.⁶

The forecasts are based on parameter estimates from rolling samples with fixed sample size of 1350 days. For every date $t \geq 1350$ in the sample, we estimate the parameters of each specification over the 1350 observations up to and including date t . Then, we consider the forecasts of the log volatilities of both assets over the period $t+1, \dots, t+s$, where s is equal to 1, 5, and 22, respectively for daily, weekly and monthly horizon forecasts, so that we have both short-term and long-term forecasts. We avoid the presence of overlapping observations, meaning that the forecast sample is constituted by 50 monthly, 200 weekly and 1100 daily forecasts (this approach has been followed for instance by Brandt and Jones (2006)). The forecasts refer to the integrated log volatilities over the period $(t+1, t+s)$, these are approximated by the daily log volatility, and by the averages over a week and a month of the log volatilities as measured by the the log-adjusted range, that is

$$\bar{V}_{t+1, t+s} = \frac{1}{s} \sum_{\tau=1}^s \log \sigma_{t+\tau}. \quad (29)$$

The MSE, the RMSE, and the MAE statistics in Table 5 clearly depict a situation where the FVECM outperforms the alternative models. We also evaluate the unbiasedness of the estimates regressing the actual integrated log volatilities on a constant and the corresponding out-of-sample forecast, the so-called Mincer and Zarnowitz (1969) regression:

$$\bar{V}_{t+1, t+s} = \alpha + \beta \widehat{V}_{t+1, t+s}^j + u_{t+1, t+s}, \quad (30)$$

where $\widehat{V}_{t+1, t+s}^j$ is the average of model j forecasts over the period $(t+1, t+s)$. In Table 6 the coefficient estimates (Panel (a) and (b)) and the regression adjusted R^2 (Panel (c)) are reported. We also compute test statistic for the restriction that $\alpha = 0$ and $\beta = 1$. Table 6 reports the results of the tests based on the regression in (30). It clearly emerges that the FVECM, differently from alternative specifications, provides unbiased forecasts of the out-

of-sample log-volatility for all the different choices of s . This result turns out to be more evident when considering longer forecast horizon, confirming that allowing for fractional cointegration improves the forecasts. Table 6 reports also the adjusted R^2 that is, in all cases, higher than that obtained with other specifications, demonstrating superiority of the FVECM in forecasting volatility.

We also test for the forecasting superiority of FVECM in the Diebold and Mariano (1995) framework, focusing here on the mean squared error (MSE) of the forecasts, where the error of model i at date t is defined as the difference, $\epsilon_{i,t}$, between the sample average of the log volatility in the period $(t+1, t+s)$ and the corresponding forecast provided by model i . Specifically, we are interested in measuring the relative forecasting performance of the different model specifications, testing the superiority of model i over model j with a t -test of the μ_{ij} coefficient in

$$\epsilon_{i,t}^2 - \epsilon_{j,t}^2 = \mu_{ij} + \eta_t \quad (31)$$

where a positive estimate of μ_{ij} indicates support for model i . In our case, we evaluate all the pairwise tests with respect to the FVECM model for all the choices of s . Table 6 reports the t -statistics for the estimates of μ_{ij} . It is clear that the forecasts based on the FVECM specification provides in many cases an improvement with respect to the alternatives, since the value of the t -test is always positive and, in many cases, significant. Moreover, it is interesting to note the superiority of forecasts of the spot volatility at weekly and monthly frequency, that is probably due to the convergence of spot volatility to the long run equilibrium that is implicit in the FVECM. These results suggest that, properly accounting for the long-run relation between volatilities, implicit in the no-arbitrage pricing, provides a significant forecast improvement, since future volatility, given the speculative nature of future contracts, leads spot volatility.

6 Conclusions

This paper focused on a no arbitrage cointegration relationship between two range based volatility measures. Given the long memory property of the volatility series, the analysis is carried out in terms of fractional cointegration so that the dynamic behavior of the two series has been modeled by a fractional VECM model, as defined by Granger (1986). The cointegrated system is estimated, implementing a new procedure, based on the profile

likelihood, that allows to jointly estimate d and b . This technique extends the estimation method proposed in Lasak (2008). We confirm the presence of a common stochastic trend with long memory that captures the total persistence of the system, so that the error correction term is integrated of order 0. Moreover the parameter β is close to the theoretical value 1, while spot volatility converges faster toward the long run equilibrium than the volatility of the future price. This evidence suggests that future volatility is the driving factor in the volatility process, since futures contracts are more efficient in processing the new information. Allowing the long range dependence, between spot and future volatility, improves significantly the out-of-sample forecasts, given the equilibrium mechanism that is incorporated in the model for fractional cointegration.

A Monte Carlo Simulation

The following Monte Carlo simulation is intended to assess the ability of the ML procedure, outlined in section 3, to correctly estimate the true parameters of the model. In other words, we wish to evaluate the consistency of our estimation procedure assuming that the cointegration rank is known. We generate two fractionally cointegrated processes, from the representation in Granger (1986), as

$$\begin{aligned} Y_t &= \alpha_1(\Delta^{-b} - 1)(Y_{t-1} - \beta X_{t-1}) + (1 - L)^{-d}\epsilon_{1t} \\ X_t &= \alpha_2(\Delta^{-b} - 1)(Y_{t-1} - \beta X_{t-1}) + (1 - L)^{-d}\epsilon_{2t} \end{aligned} \quad (32)$$

We chose $d = 0.4$, $\beta = 1$ and $\alpha = (-0.5, 0.5)$. The parameter b assumes values 0.4 and 0.3. The infinite moving-average representation of the long memory process, u_t , is given by

$$u_{i,t} = (1 - L)^{-d}\epsilon_{i,t} = \sum_{i=0}^{\infty} \psi_i \eta_{i,t-i} \quad (33)$$

where $\psi_i = i^{d-1}/(d-1)$ as $i \rightarrow \infty$, see Hosking (1981).

From a practical point of view, we consider a truncated version of ((33)), that is

$$u_{i,t}^+ = (1 - L)^{-d}\epsilon_{i,t} = \sum_{i=0}^{t-1} \psi_i \epsilon_{i,t-i} \quad (34)$$

where the presample values are assumed to be equal to zero.

We generate 1000 time series with $T = 500, 1000, 2000$ observations each. ϵ_{1t} and ϵ_{2t} are randomly generated from bivariate normal distribution with mean 0 and variance 1, with correlation equal to 0.9.⁷ All the parameters are then estimated following the method presented in section 3.⁸ The precision in the estimation of the parameters increases dramatically with the sample size. This is due to the long memory feature of the series under exam. Moreover, the *RMSE* highlights the idea, already noted by Lasak (2008), that the parameter dispersion increases with the difference between d and b . We obtain particularly good estimates of β , also for moderate sample size. On the other hand, the estimates of the vector α appear more volatile; this is due to the fact that α is function of d , b and β , and it is sensible to the estimation uncertainty present in the previous steps. As shown in graph 4, the estimates of α are positively skewed. This is due to the presence of outliers in correspondence of values of b that are negative. Note that the skewness tends to zero when $b = d$ and $T \rightarrow \infty$; that is when b is more precisely estimated and does not take negative values. Since volatility of financial series often present GARCH effects, see for example Corsi et al. (2005), we check for the robustness of our estimation procedure running again the previous analysis, adding a multivariate GARCH component to the errors of equation ((32)). In particular, we define the conditional volatility of $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})'$ as the Constant Conditional Correlation model, presented by Bollerslev (1990). The constant correlation parameter, ρ , is set equal to 0.95, while the conditional GARCH(1,1) variances are

$$\sigma_{i,t}^2 = \omega_i + \delta_i \sigma_{i,t-1|t-2}^2 + \gamma_i \epsilon_{i,t-1}^2 \quad i = 1, 2 \quad (35)$$

so that ϵ_t is distributed as a $N(0, \Sigma_{ij,t})$, where $\Sigma_{ij} = \rho \sigma_{i,t} \sigma_{j,t}$ for $i \neq j$ and $\Sigma_{ij,t} = \sigma_{i,t}^2$ for $i = j$. We select $\gamma_i = 0.1$, $\omega_i = 0.05$ and $\delta_i = 0.8$ for $i = 1, 2$. The results from the Monte Carlo exercise in presence of GARCH effects are also presented in table 9. The results seems to confirm the robustness of our estimation procedure with respect to GARCH effects.

References

- Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, Vol. 57, No. 3:pp. 1047–1091.
- Andersen, T. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of stock return volatility. *Journal of Financial Economics*, 61:43–76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.
- Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.
- Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch approach. *Review of Economics and Statistics*, 72:498–505.
- Brandt, M. W. and Jones, C. S. (2006). Volatility forecasting with range-based egarch models. *Journal of Business & Economic Statistics*, Vol. 24, No. 4:470–486.
- Breitung, J. and Hassler, U. (2002). Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics*, 110(2):167–185.
- Cheung, Y. and Lai, K. (1993). A fractional cointegration analysis of purchasing power parity. *Journal of Business and Economic Statistics*, 11:103–122.
- Christensen, B. J. and Nielsen, M. . (2006). Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation. *Journal of Econometrics*.
- Christensen, K. and Podolskij, M. (2007). Realized range based estimation of integrated variance. *Journal of Econometrics*, 141:323–349.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196.

- Corsi, F., Kretschmer, U., Mittnik, S., and Pigorsch, C. (2005). The volatility of realized volatility. CFS Working Paper Series 2005/33, Center for Financial Studies. available at <http://ideas.repec.org/p/cfs/cfswop/wp200533.html>.
- Cox, C. C. (1976). Futures trading and market information. *The Journal of Political Economy*, 84:1215–1237.
- Davidson, J. (2002). A model of fractional cointegration, and tests for cointegration using the bootstrap. *Journal of Econometrics*, 110(2):187–212.
- Davidson, J., Peel, D., and Byers, J. (2006). Support for governments and leaders: Fractional cointegration analysis of poll evidence from the uk, 1960-2004. *Studies in Nonlinear Dynamics & Econometrics*, 10(1):1345–1345.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.
- Duecker, M. and Startz, R. (1998). Maximum likelihood estimation of fractional cointegration with an application to us and canadian bond rates. *Review of Economics and Statistics*, 80:420426.
- Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55, issue 2:251–76.
- Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Source: The Journal of Business*, Vol. 53, No. 1:67–78.
- Granger, C. W. and Hyung, N. (2004). Occasional structural breaks and long memory with application to the s&p 500 absolute stock returns. *Journal of Empirical Finance*, 11:399–421.
- Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48(3):213–28.
- Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 4:221–238.
- Hosking, J. (1981). Fractional differencing. *Biometrika*, 68:165–76.

- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.
- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.
- Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory*, Vol 24 ,3.
- Lasak, K. (2008). Maximum likelihood estimation of fractionally cointegrated systems. Technical report, CREATES Research Paper 2008-53.
- Lieberman, O. and Phillips, P. C. B. (2008). Refined inference on long-memory in realized volatility. *Econometric Reviews*, 27(1):254–267.
- Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. Nber working papers, National Bureau of Economic Research, Inc.
- Nielsen, M. . and Shimotsu, K. (2007). Determining the cointegration rank in nonstationary fractional system by the exact local whittle approach. *Journal of Econometrics*, 141:574–596.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, 53:61–65.
- Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. *The Annals of Statistics*, 22(1):515–539.
- Robinson, P. M. and Marinucci, D. (2003). Semiparametric frequency domain analysis of fractional cointegration. In Robinson, P. M., editor, *Time Series with Long Memory*, pages 334–373. Oxford University Press.
- Robinson, P. M. and Yajima, Y. (2002). Determination of cointegrating rank in fractional systems. *Journal of Econometrics*, 106:217–241.
- Santucci de Magistris, P. and Christensen, B. (2009). Fractional cointegration and level shifts in the realized and implied volatility relation. Technical report, Aarhus University.

- Shimotsu, K. and Phillips, P. (2005). Exact local whittle estimation of fractional integration. *Annals of Statistics*, 33:1890-1933.
- Sowell, F. (1989). Maximum likelihood estimation of fractionally integrated time series models. Technical report, unpublished manuscript Carnegie-Mellon University.
- Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53(1-3):165–188.
- Wiggins, J. B. (1992). Estimating the volatility of s&p 500 futures prices using the extreme-value method. *Journal of Futures Markets*, v12(3):265–274.

Notes

¹In a recent paper, Santucci de Magistris and Christensen (2009) note that the presence of a common level shifts process among two or more $I(0)$ series induces spurious fractional cointegration. Our purpose, in this section, is to show that fractional cointegration between spot and future volatility is not induced spuriously by the presence of common shifts but it is due to their common stochastic trend. However, we are aware of the fact that a more efficient inference technique could be implemented, providing a synthesis between the concept of fractional cointegration and the idea of structural breaks, in order to carry out a new testing procedure to distinguish between the two sources of common persistence. This is left for future research.

²Christensen and Nielsen (2006) used $m = 3, 6, 9, 15$.

³This correspond to the case I in Robinson and Marinucci (2003), for which a asymptotic distribution for β is not explicitly defined.

⁴Johansen (2008) proposes an alternative parametrization that allows for a Granger representation of co-fractional systems, where the short run terms are written in terms of the new lag operator $L_b = 1 - (1 - L)^b$. On the other hand, the ECM term is the same in both models and the Granger (1986) representation has been already considered by Lasak (2008) with inference purposes.

⁵The model also includes Γ_{10} . In this way, the model accounts for a significant spike in the autocorrelations at the 10-th lag.

⁶Even if a correct, but very slow, procedure has been implemented to estimate the FIVAR, see Sowell (1992), we deal with a faster two step estimation method that consists of estimating first the fractional parameter d and then calculating the VAR on the fractionally filtered series.

⁷This choice is motivated by the wish to resemble the statistical properties of the observed data.

⁸Note that all the parameters are left free to vary over all \mathbb{R} .

Raw Data			
	$m = T^{0.7} = 235$	$m = T^{0.6} = 108$	$m = T^{0.5} = 49$
$\log \sigma_{t,F}$	0.4797 (0.0639)	0.5715 (0.0943)	0.5188 (0.1400)
$\log \sigma_{t,S}$	0.4909 (0.0639)	0.5636 (0.0943)	0.5375 (0.1400)
\hat{T}_0	0.1653	0.0452	0.1247
Demeaned series			
	$m = T^{0.7} = 235$	$m = T^{0.6} = 108$	$m = T^{0.5} = 49$
$\log \sigma_{t,F}$	0.4147 (0.0639)	0.4851 (0.0943)	0.2715 (0.1400)
$\log \sigma_{t,S}$	0.4313 (0.0639)	0.4753 (0.0943)	0.3101 (0.1400)
\hat{T}_0	0.3600	0.0690	0.5300

Table 1: Fractional integration estimation with exact local Whittle estimator (standard error in parenthesis). The \hat{T}_0 test statistic is calculated with $h(T) = \log(T)$.

Raw Data				Demeaned series			
$L(u)$	$v(T) = n^{-0.45}$	$v(T) = n^{-0.35}$	$v(T) = n^{-0.25}$	$L(u)$	$v(T) = n^{-0.45}$	$v(T) = n^{-0.35}$	$v(T) = n^{-0.25}$
$m = 235, n = 109$							
$L(0)$	-1.7956	-1.6607	-1.4368	$L(0)$	-1.7956	-1.6607	-1.4368
$L(1)$	-1.8457	-1.7762	-1.6682	$L(1)$	-1.8452	-1.7778	-1.6688
\hat{r}	1	1	1	\hat{r}	1	1	1
$m = 108, n = 50$							
$L(0)$	-1.7099	-1.5545	-1.3158	$L(0)$	-1.7099	-1.45545	-1.3158
$L(1)$	-1.8326	-1.7549	-1.6355	$L(1)$	-1.8321	-1.7523	-1.6350
\hat{r}	1	1	1	\hat{r}	1	1	1
$m = 49, n = 22$							
$L(0)$	-1.5883	-1.4118	-1.1655	$L(0)$	-1.5853	-1.4118	-1.1655
$L(1)$	-1.7836	-1.6968	-1.5737	$L(1)$	-1.7826	-1.6959	-1.5727
\hat{r}	1	1	1	\hat{r}	1	1	1

Table 2: Fractional cointegration estimation. The table reports the value of the function $L(u)$ for different choices of m and n .

Demeaned series				
Bandwidth	$\hat{\beta}_m$	s.e.($\hat{\beta}_m$)	\hat{d}_z	s.e.(\hat{d}_z)
$m = T - 1$	0.8938	--	0.1147	(0.0483)
$m = 20$	0.9845	0.0310	0.0118	(0.0483)
$m = 15$	0.9885	0.0332	0.0103	(0.0483)
$m = 9$	0.9754	0.0420	0.0165	(0.0483)
$m = 6$	1.0147	0.0517	0.0088	(0.0483)
Raw Series				
Bandwidth	$\hat{\beta}_m$	s.e.($\hat{\beta}_m$)	\hat{d}_z	s.e.(\hat{d}_z)
$m = T - 1$	0.9475	--	0.2895	(0.0483)
$m = 20$	1.0373	--	0.2034	(0.0483)
$m = 15$	1.0415	--	0.2024	(0.0483)
$m = 9$	1.0435	--	0.2021	(0.0483)
$m = 6$	1.0494	--	0.2018	(0.0483)

Table 3: Fractional Cointegration Analysis: the estimation of β is performed with $m = T - 1, 20, 15, 9, 6$, while d_z is obtained with the Whittle estimator with the bandwidth equal to $T^{0.6}$

\hat{d}	0.3717 (0.3357, 0.4080)	$\hat{\alpha}_1$	-0.1979 (-0.4757, 0.07039)
\hat{b}	0.3717 (0.2326, 0.5253)	$\hat{\alpha}_2$	0.7436 (0.4753, 1.0229)
$\hat{\beta}$	-1.0081 (-1.0387, -0.9795)		
$\hat{\gamma}_{11}^1$	-0.1631 (-0.2602, -0.0690)	$\hat{\gamma}_{11}^{10}$	0.0753 (0.0026, 0.1426)
$\hat{\gamma}_{12}^1$	-0.0599 (-0.1557, 0.0346)	$\hat{\gamma}_{12}^{10}$	-0.0026 (-0.0703, 0.0656)
$\hat{\gamma}_{21}^1$	-0.0820 (-0.1815, 0.0153)	$\hat{\gamma}_{21}^{10}$	0.0932 (0.0193, 0.1681)
$\hat{\gamma}_{22}^1$	-0.1689 (-0.2672, -0.0677)	$\hat{\gamma}_{22}^{10}$	-0.0238 (-0.0967, 0.0478)
JB_F	0.2343	JB_S	0.7993
LM_F	0.1690	LM_S	0.5771

Table 4: Estimation Results. Table reports the estimated parameter values. In parenthesis the 5th and 95th bootstrapped quantiles. JB_F and JB_S are the p -values of Jarque-Bera test of normality for spot and future volatility, while LM_F and LM_S are p -values of Godfrey test of heteroschedasticity for spot and future volatility.

(a) <i>Futures</i>									
	<i>MSE</i>			<i>RMSE</i>			<i>MAE</i>		
	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22
<i>VAR</i> (4)	0.178	0.067	0.078	0.423	0.260	0.279	0.341	0.206	0.213
<i>UHAR</i>	0.174	0.065	0.059	0.418	0.254	0.244	0.339	0.201	0.183
<i>BHAR</i>	0.176	0.065	0.060	0.419	0.255	0.245	0.340	0.202	0.183
<i>ARFIMA</i>	0.187	0.068	0.069	0.432	0.261	0.262	0.350	0.211	0.199
<i>FIVAR</i>	0.187	0.068	0.069	0.433	0.260	0.263	0.350	0.210	0.200
<i>FVECM</i>	0.168	0.063	0.055	0.410	0.251	0.235	0.333	0.201	0.182

(b) <i>Spot</i>									
	<i>MSE</i>			<i>RMSE</i>			<i>MAE</i>		
	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22	<i>s</i> = 1	<i>s</i> = 5	<i>s</i> = 22
<i>VAR</i> (4)	0.184	0.085	0.057	0.429	0.291	0.238	0.345	0.226	0.188
<i>UHAR</i>	0.182	0.061	0.056	0.426	0.247	0.237	0.344	0.198	0.175
<i>BHAR</i>	0.181	0.062	0.054	0.425	0.248	0.233	0.344	0.199	0.175
<i>ARFIMA</i>	0.209	0.068	0.069	0.457	0.261	0.262	0.369	0.208	0.204
<i>FIVAR</i>	0.204	0.068	0.068	0.452	0.261	0.261	0.365	0.207	0.204
<i>FVECM</i>	0.180	0.058	0.051	0.424	0.242	0.227	0.343	0.194	0.171

Table 5: Table reports the MSE, the RMSE and the MAE of the alternative forecasts of the futures (Panel (a)) and spot (Panel (b)) integrated log-volatilities $\frac{1}{s} \sum_{\tau=1}^s \log \sigma_{t+\tau}$

(a) $H_0: \alpha = 0$ vs $H_1: \alpha \neq 0$							(b) $H_0: \beta = 1$ vs $H_1: \beta \neq 1$						
	$s = 1$	$\log \sigma_t^F$ $s = 5$	$s = 22$	$s = 1$	$\log \sigma_t^S$ $s = 5$	$s = 22$		$s = 1$	$\log \sigma_t^F$ $s = 5$	$s = 22$	$s = 1$	$\log \sigma_t^S$ $s = 5$	$s = 22$
<i>VAR</i> (4)	-0.3017 (-1.315)	-0.2520 (0.3324)	-1.8054 (-2.960)	-0.2885 (1.301)	-0.2370 (-0.776)	-1.6327 (-2.898)	<i>VAR</i> (4)	0.9421 (1.3072)	0.9511 (0.767)	0.7341 (2.342)	0.9445 (1.291)	0.9538 (0.874)	0.6819 (2.981)
<i>UHAR</i>	-0.1517 (-0.672)	-0.2008 (-0.604)	-1.2324 (-2.049)	-0.1325 (-0.528)	-0.1649 (-0.776)	-1.0958 (-1.981)	<i>UHAR</i>	0.9705 (0.678)	0.9612 (0.610)	0.7616 (2.095)	0.9741 (0.594)	0.9680 (0.534)	0.7873 (2.034)
<i>BVAR</i>	-0.2111 (-0.935)	-0.2234 (-0.674)	-1.3733 (-2.291)	-0.2483 (1.144)	-0.2639 (-0.871)	0.5226 (-2.338)	<i>BVAR</i>	0.9593 (0.934)	0.9571 (0.675)	0.7341 (2.342)	0.9519 (1.146)	0.9492 (0.786)	0.7627 (2.397)
<i>ARFIMA</i>	-1.300 (-7.453)	-0.6913 (-2.287)	-1.5134 (-2.514)	-1.535 (-8.816)	-0.8024 (-2.782)	-1.4613 (-2.554)	<i>ARFIMA</i>	0.7482 (7.508)	0.8655 (2.290)	0.7055 (2.575)	0.7024 (8.888)	0.8439 (2.831)	0.7156 (2.614)
<i>FIVAR</i>	-1.297 (-8.053)	-0.6978 (-2.281)	-1.5221 (-2.545)	-1.415 (-8.053)	-0.8009 (-2.788)	-1.4656 (-2.583)	<i>FIVAR</i>	0.7487 (7.524)	0.8644 (2.316)	0.7038 (2.606)	0.7256 (8.109)	0.8444 (2.831)	0.7147 (2.646)
<i>FVECM</i>	0.0194 (0.088)	0.0989 (0.289)	-0.5693 (-0.860)	-0.243 (-1.139)	0.0768 (0.245)	-0.4640 (-0.761)	<i>FVECM</i>	1.003 (0.0812)	1.0185 (0.284)	0.8888 (0.887)	0.9524 (1.152)	1.0144 (0.240)	0.9089 (0.789)

(c) \bar{R}^2 of the Mincer-Zarnowitz regression.							(d) $H_0: \alpha = 0 \cap \beta = 1$ vs $H_1: \alpha \neq 0 \cup \beta \neq 1$						
	$s = 1$	$\log \sigma_t^F$ $s = 5$	$s = 22$	$s = 1$	$\log \sigma_t^S$ $s = 5$	$s = 22$		$s = 1$	$\log \sigma_t^F$ $s = 5$	$s = 22$	$s = 1$	$\log \sigma_t^S$ $s = 5$	$s = 22$
<i>VAR</i> (4)	0.3013	0.5233	0.3312	0.3133	0.5596	0.3850	<i>VAR</i> (4)	0.8713 (0.4186)	0.3056 (0.73694)	5.3746 (0.0078)	0.8553 (0.4254)	0.3204 (0.7261)	5.1663 (0.0093)
<i>UHAR</i>	0.3155	0.5391	0.4575	0.3186	0.5691	0.5047	<i>UHAR</i>	0.2332 (0.7920)	0.1909 (0.8263)	2.5324 (0.0900)	0.18114 (0.8343)	0.1489 (0.8617)	2.4580 (0.0963)
<i>BVAR</i>	0.3120	0.5367	0.4353	0.3229	0.5750	0.5035	<i>BVAR</i>	0.4368 (0.6461)	0.2283 (0.7960)	2.50294 (0.0924)	0.6570 (0.5185)	0.3823 (0.6827)	2.6480 (0.0811)
<i>ARFIMA</i>	0.2974	0.5302	0.3903	0.2667	0.5404	0.4280	<i>ARFIMA</i>	28.298 (0.0000)	2.8167 (0.0621)	3.7847 (0.0298)	0.85531 (0.4254)	4.2786 (0.0150)	3.842 (0.0283)
<i>FIVAR</i>	0.2984	0.5316	0.3895	0.2761	0.5424	0.4247	<i>FIVAR</i>	28.400 (0.0000)	2.8412 (0.0605)	3.86473 (0.0278)	32.976 (0.000)	4.8149 (0.0089)	3.9590 (0.0256)
<i>FVECM</i>	0.3407	0.5509	0.4602	0.3263	0.5923	0.5157	<i>FVECM</i>	0.0111 (0.9889)	0.0575 (0.9440)	0.4470 (0.6421)	0.6792 (0.5072)	0.0349 (0.9656)	0.5280 (0.5931)

Table 6: Panel (a) and (b) report estimates of the intercept and slope coefficients, α and β , in the regression (30). The t -statistics, in parenthesis, are computed using NeweyWest standard errors. Bold character means rejection of the null hypothesis ($\alpha = 0$ or $\beta = 1$) at 5% of significance. Panel (c) reports the regression adjusted R^2 , while Panel (d) reports F test for the joint hypothesis $\alpha = 0 \cap \beta = 1$, the p -value is in parenthesis. Bold character means rejection of the null at 10% of significance.

	$s = 1$	$\log \sigma_t^F$ $s = 5$	$s = 22$	$s = 1$	$\log \sigma_t^S$ $s = 5$	$s = 22$
<i>VAR</i> (4)	3.980^a	0.5233	1.649^c	1.155	3.021^a	2.024^a
<i>UHAR</i>	2.984^a	0.752	1.024	0.650	2.076^b	0.893
<i>BVAR</i>	3.454^a	0.859	1.277	0.361	1.759^c	0.562
<i>ARFIMA</i>	4.355^a	2.121^b	1.434	5.447^a	2.705^a	2.863^a
<i>FIVAR</i>	4.305^a	1.355	2.143^b	4.741^a	2.702^a	2.790^a

Table 7: Table reports the t -statistic of the estimate of $\mu_{i,j}$ in the regression $\epsilon_{i,t}^2 - \epsilon_{FVECM,t}^2 = \mu_{i,j} + \eta_t$, where $\epsilon_{i,t}$ is the forecast error of model i in period t . a, b and c stands for 1%, 5% and 10% significance level of the corresponding t -ratio test.

T	2000		1000		500	
b	0.3	0.4	0.3	0.4	0.3	0.4
$Q_{50,\hat{d}}$	0.4011	0.4015	0.4015	0.4029	0.4056	0.4014
$Q_{5,\hat{d}}$	0.3702	0.3697	0.3572	0.3570	0.3399	0.3435
$Q_{95,\hat{d}}$	0.4299	0.4306	0.4420	0.4404	0.4577	0.4583
$RMSE$	0.0178	0.0184	0.0261	0.0252	0.0361	0.0357
$Q_{50,\hat{b}}$	0.3072	0.4043	0.3076	0.4086	0.3244	0.4233
$Q_{5,\hat{b}}$	0.1860	0.3092	0.1297	0.2761	0.0507	0.2145
$Q_{95,\hat{b}}$	0.4268	0.5076	0.4771	0.5426	0.5920	0.6329
$RMSE$	0.0733	0.0594	0.1087	0.0822	0.1695	0.1293
$Q_{50,\hat{\beta}}$	-1.0000	-1.0002	-1.0005	-0.9996	-0.9978	-0.9996
$Q_{5,\hat{\beta}}$	-1.0215	-1.0152	-1.0388	-1.0203	-1.0530	-1.0301
$Q_{95,\hat{\beta}}$	-0.9772	-0.9867	-0.9681	-0.9794	-0.9490	-0.9691
$RMSE$	0.0134	0.0085	0.0206	0.0124	0.0311	0.0183
$Q_{50,\hat{\alpha}_1}$	-0.4837	-0.4908	-0.4927	-0.4894	-0.4329	-0.4600
$Q_{5,\hat{\alpha}_1}$	-1.0273	-0.8458	-1.6025	-1.1012	-1.9003	-1.4107
$Q_{95,\hat{\alpha}_1}$	-0.1509	-0.2085	-0.0396	-0.1284	0.3624	0.0388
$RMSE$	0.2567	0.1835	0.4801	0.2500	0.7043	0.4184
$Q_{50,\hat{\alpha}_2}$	0.5033	0.4868	0.4663	0.4894	0.4490	0.4698
$Q_{5,\hat{\alpha}_2}$	0.1408	0.2099	-0.0710	0.1303	-0.2989	-0.0591
$Q_{95,\hat{\alpha}_2}$	1.0428	0.8339	1.4173	1.0015	1.9363	1.3114
$RMSE$	0.2718	0.1814	0.4402	0.2485	0.7103	0.3953

Table 8: Table reports the median (Q_{50}), the 5th (Q_5) and 95th (Q_{95}) percentile of estimators \hat{d} , \hat{b} , $\hat{\beta}$ and $\hat{\alpha}$ for $T = 2000, 1000, 500$ observations. In the simulation, $d = 0.4$, $\alpha = (-0.5, 0.5)$ and $\beta = -1$. The values of b used in the Monte Carlo are reported.

T	2000		1000		500	
b	0.3	0.4	0.3	0.4	0.3	0.4
$Q_{50,\hat{d}}$	0.4017	0.4011	0.4016	0.4004	0.4044	0.4010
$Q_{5,\hat{d}}$	0.3750	0.3659	0.3533	0.3505	0.3288	0.3327
$Q_{95,\hat{d}}$	0.4286	0.4351	0.4470	0.4498	0.4700	0.4677
$RMSE$	0.0162	0.0215	0.0292	0.0301	0.0426	0.0414
$Q_{50,\hat{b}}$	0.3007	0.4022	0.3071	0.4089	0.3224	0.4137
$Q_{5,\hat{b}}$	0.1980	0.3005	0.1255	0.2577	0.0344	0.1647
$Q_{95,\hat{b}}$	0.3874	0.5102	0.4966	0.5598	0.5839	0.6329
$RMSE$	0.0570	0.0617	0.1111	0.0912	0.1747	0.1418
$Q_{50,\hat{\beta}}$	-0.9992	-1.0005	-1.0012	-1.0003	-0.9984	-1.0001
$Q_{5,\hat{\beta}}$	-1.0184	-1.0165	-1.0401	-1.0243	-1.0611	-1.0362
$Q_{95,\hat{\beta}}$	-0.9790	-0.9843	-0.9644	-0.9777	-0.9416	-0.9650
$RMSE$	0.0118	0.0096	0.0233	0.01409	0.0374	0.0223
$Q_{50,\hat{\alpha}_1}$	-0.5077	-0.5037	-0.4983	-0.4995	-0.4575	-0.4693
$Q_{5,\hat{\alpha}_1}$	-0.9109	-0.8899	-1.4939	-1.0717	-2.5199	-1.5050
$Q_{95,\hat{\alpha}_1}$	-0.2135	-0.1942	0.0223	-0.0982	0.4741	0.0560
$RMSE$	0.2081	0.1964	0.4619	0.3123	0.9924	0.4671
$Q_{50,\hat{\alpha}_2}$	0.5098	0.4992	0.4914	0.4845	0.4467	0.4999
$Q_{5,\hat{\alpha}_2}$	0.2182	0.1942	-0.0282	0.0897	-0.3841	-0.0522
$Q_{95,\hat{\alpha}_2}$	0.9275	0.8407	1.5519	1.0320	2.5715	1.4114
$RMSE$	0.2114	0.1946	0.4839	0.3171	0.9318	0.4359

Table 9: Table reports the median (Q_{50}), the 5th (Q_5) and 95th (Q_{95}) percentile of estimators \hat{d} , \hat{b} , $\hat{\beta}$ and $\hat{\alpha}$ for $T = 2000, 1000, 500$ observations with constant conditional correlation errors. In the simulation, $d = 0.4$, $\alpha = (-0.5, 0.5)$ and $\beta = -1$. The values of b used in the Monte Carlo are reported.

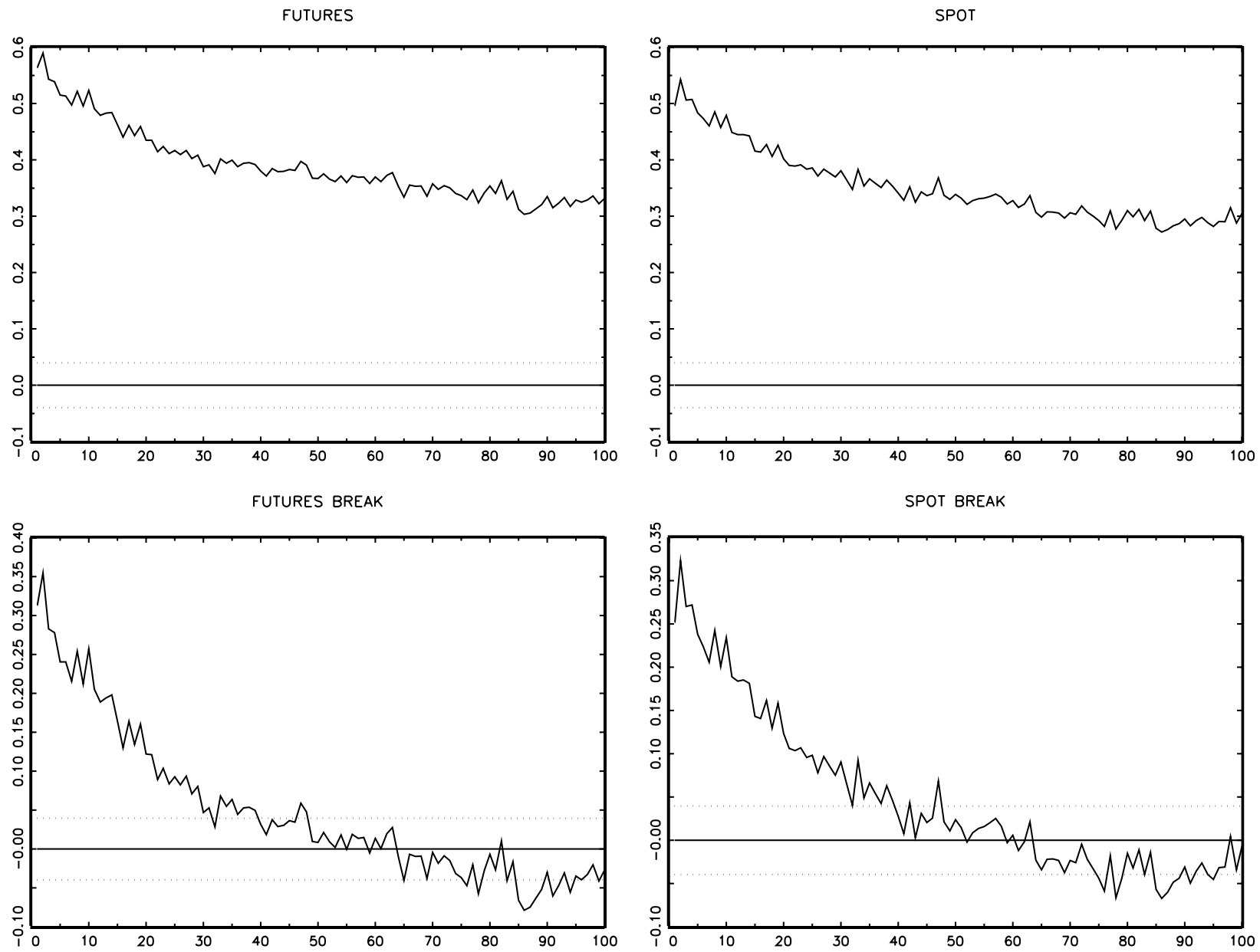


Figure 1: Autocorrelogram of $\log \sigma_{t,F}$ and $\log \sigma_{t,S}$.

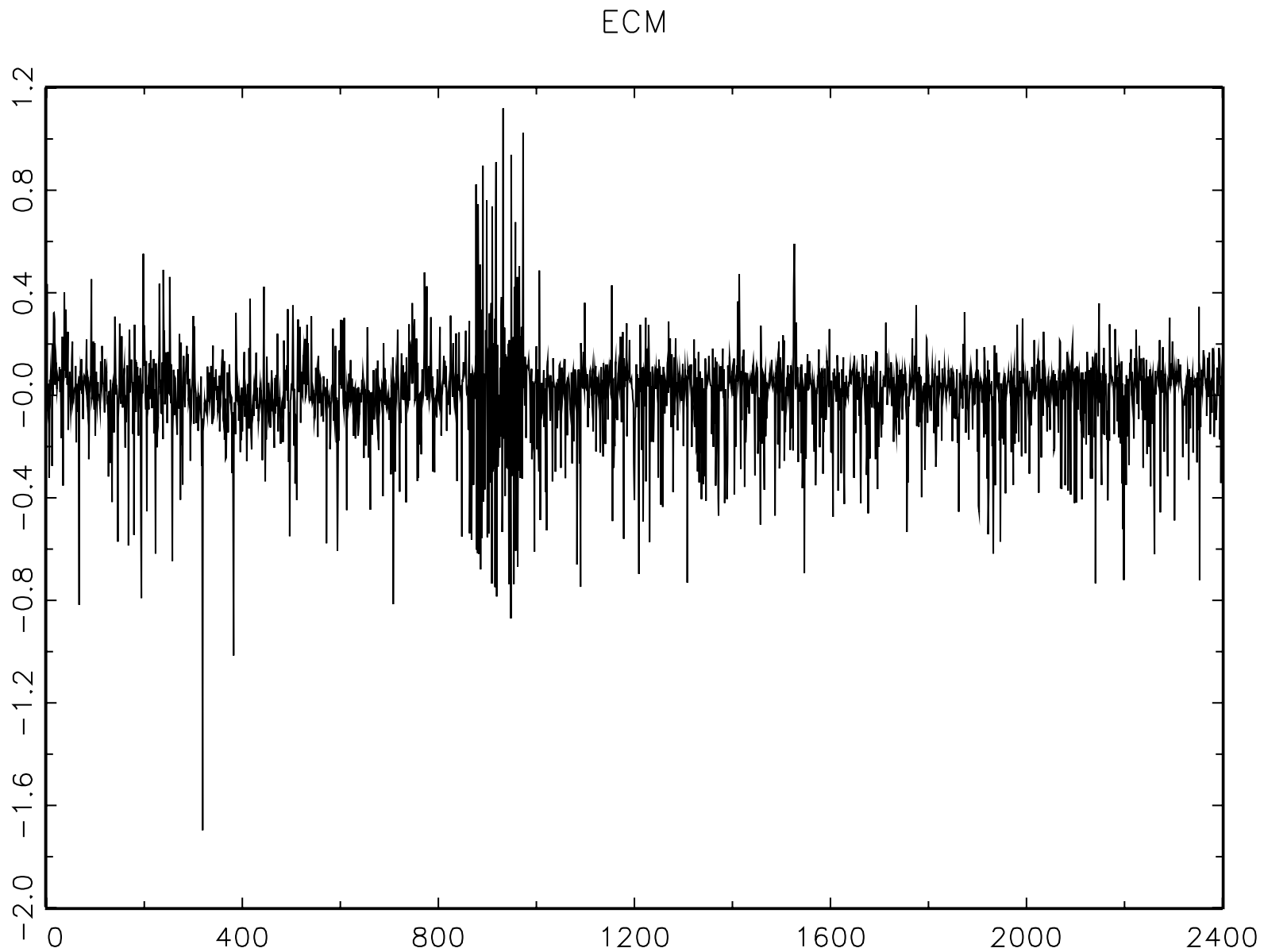


Figure 2: Error Correction Term: Figure plots the error term given by $\log(\sigma_t^F) + \hat{\beta} \log(\sigma_t^F)$. The fractional integration order of the error term, estimated with the exact local Whittle estimator, is equal to 0.0007.

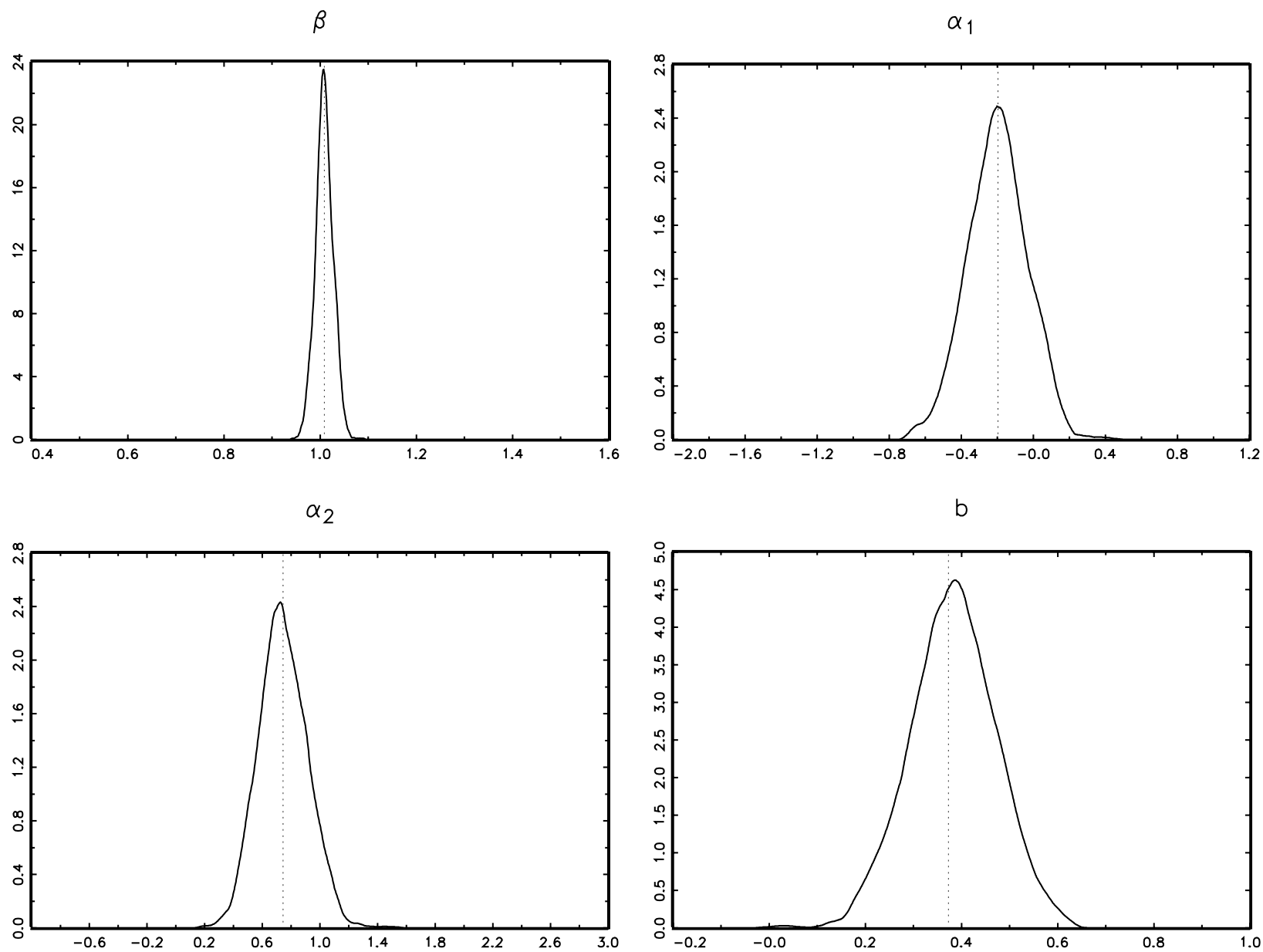


Figure 3: Kernel Densities of $\hat{\beta}$, $\hat{\alpha}$ and \hat{b} .

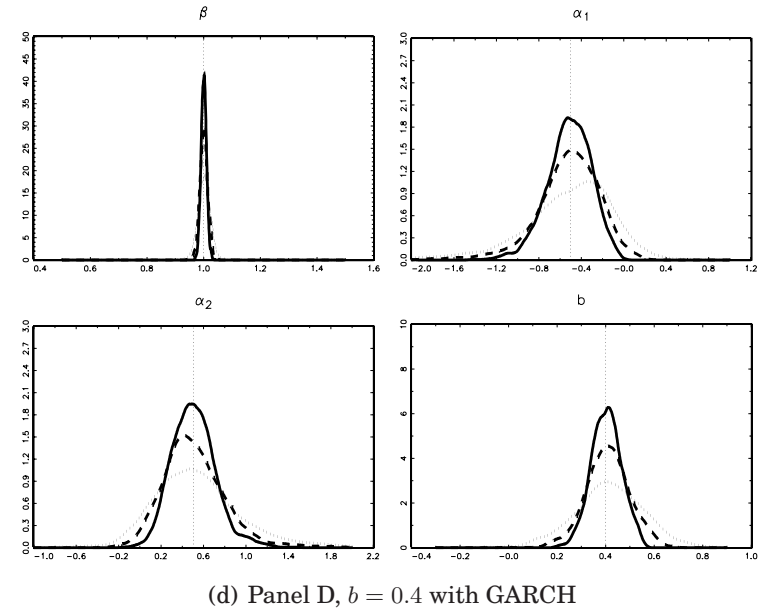
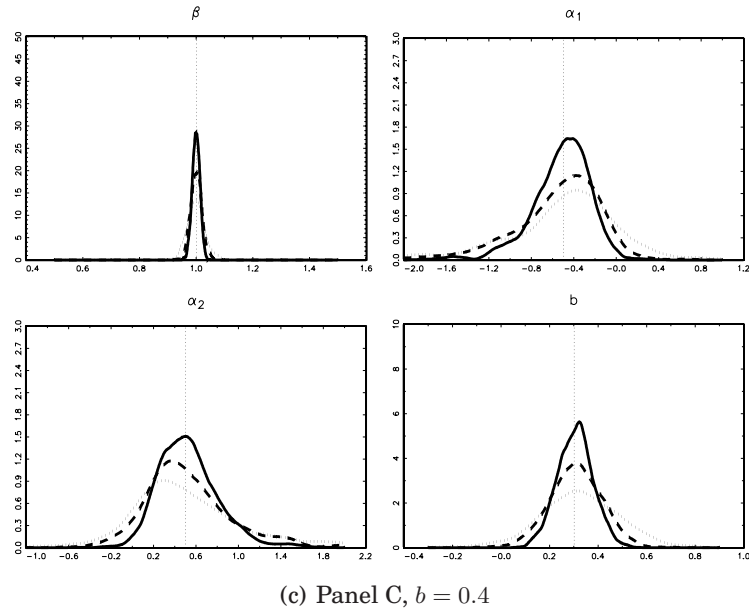
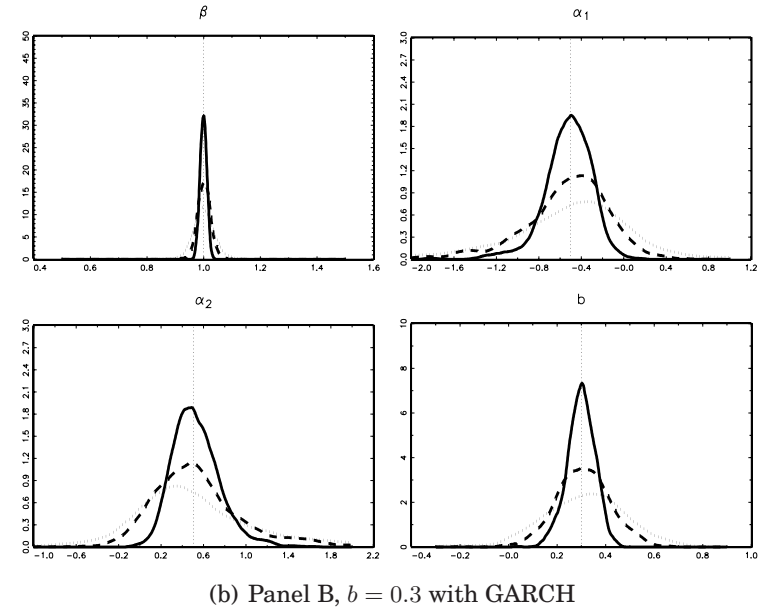
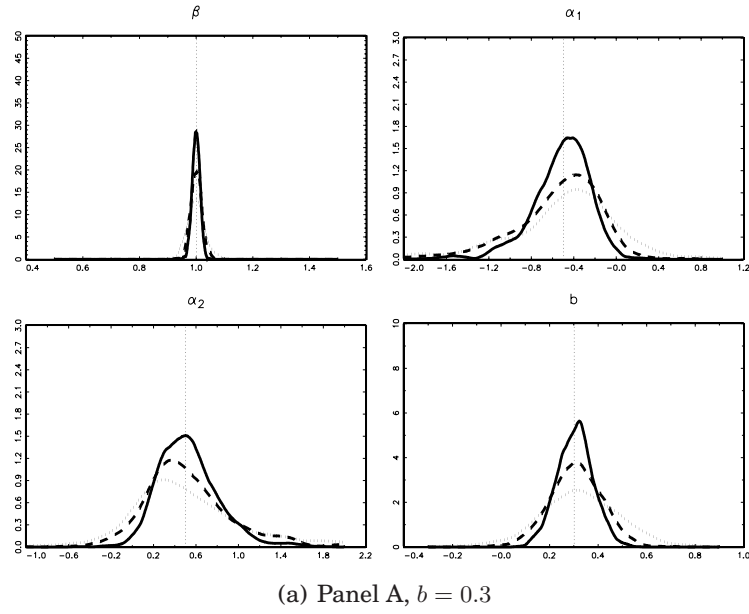


Figure 4: Kernel Densities of the estimated parameters for $b = 0.3$ and $b = 0.4$. Right panel presents the kernel densities for the estimated parameter of the model with GARCH effects.

- 2009-17: Tom Engsted: Statistical vs. Economic Significance in Economics and Econometrics: Further comments on McCloskey & Ziliak
- 2009-18: Anders Bredahl Kock: Forecasting with Universal Approximators and a Learning Algorithm
- 2009-19: Søren Johansen and Anders Rygh Swensen: On a numerical and graphical technique for evaluating some models involving rational expectations
- 2009-20: Almut E. D. Veraart and Luitgard A. M. Veraart: Stochastic volatility and stochastic leverage
- 2009-21: Ole E. Barndorff-Nielsen, José Manuel Corcuera and Mark Podolskij: Multipower Variation for Brownian Semistationary Processes
- 2009-22: Giuseppe Cavaliere, Anders Rahbek and A.M. Robert Taylor: Co-integration Rank Testing under Conditional Heteroskedasticity
- 2009-23: Michael Frömmel and Robinson Kruse: Interest rate convergence in the EMS prior to European Monetary Union
- 2009-24: Dominique Guégan: A Meta-Distribution for Non-Stationary Samples
- 2009-25: Ole E. Barndorff-Nielsen and Almut E. D. Veraart: Stochastic volatility of volatility in continuous time
- 2009-26: Tim Bollerslev and Viktor Todorov: Tails, Fears and Risk Premia
- 2009-27: Kim Christensen, Roel Oomen and Mark Podolskij: Realised Quantile-Based Estimation of the Integrated Variance
- 2009-28: Takamitsu Kurita, Heino Bohn Nielsen and Anders Rahbek: An I(2) Cointegration Model with Piecewise Linear Trends: Likelihood Analysis and Application
- 2009-29: Martin M. Andreasen: Stochastic Volatility and DSGE Models
- 2009-30: Eduardo Rossi and Paolo Santucci de Magistris: Long Memory and Tail dependence in Trading Volume and Volatility
- 2009-31: Eduardo Rossi and Paolo Santucci de Magistris: A No Arbitrage Fractional Cointegration Analysis Of The Range Based Volatility