Department of
Economics and Finance

Working Paper No. 13-16

Economics and Finance Working Paper Series

Aris Kartsaklas and Menelaos Karanasos

Long Run Dependencies in Stock Volatility and Trading Volume

May 2013

http://www.brunel.ac.uk/economics

Long Run Dependencies in Stock Volatility and Trading Volume.

Aris Kartsaklas , Menelaos Karanasos
Brunel University, London, UK

Draft: April 2013

Abstract

This paper provides empirical evidence on the degree of long run dependence of volatility and trading volume in the Korean Stock Exchange using the semiparametric estimators of Robinson (1994, 1995a). The results of testing for long memory support the argument for long run dependence in both Garman-Klass volatility and trading volume (turnover). Total and domestic trading volume exhibit very similar long memory characteristics for all sample periods. The degree of long memory in foreign volume is significantly lower than that experienced in domestic volume. Interestingly, the results for trading volume are not influenced by structural breaks in the mean of the series. On the other hand, the long range dependence in volatility is quite sensitive to the different sample periods considered and comparable to foreign volume. Furthermore, the null hypothesis that volatility and volume share a common long memory parameter is only accepted for foreign volume and Garman-Klass volatility in all three subperiods. This result is consistent with a modified version of the mixture of distributions hypothesis in which volatility and volume have similar long memory characteristics as they are both influenced by an aggregate information arrival process displaying long range dependence. Finally, we find no evidence that foreign volume and volatility share a common long memory component.

Keywords: futures markets; range-based volatility; financial crisis; foreign investors; trading volume

JEL classification: C32, C52, G12, G15.

I thank Andrea Carriero, George Kapetanios and Chrisostomos Florackis, for their valuable suggestions. I have also benefited from the comments given by seminar participants at Queen Mary University of London, University of York and Brunel University..

*

1 Introduction

The analysis of long-run dependence in time series has provided a wealth of statistical tools, parametric and non parametric, in order to test and measure the persistence of macroeconomic and …nancial processes. One of the most popular statistics used to describe long run dependence is the long memory parameter d

volume in the Korean Stock Exchange. We employ semiparametric analysis in the frequency domain and estimates of the long memory parameter are reported for the whole sample as well as for subsamples subject to prior investigation for structural break in the mean of the two series. The same analysis is performed for domestic and foreign investors' trading volume. We also test whether volatility and trading volume have the same degree of long memory as a modi…ed version of the mixture of distributions hypothesis suggests. Finally, we examine if both processes are driven by the same long-memory component in case both volume and volatility possess the same long-memory parameter.

Our results support the argument that long run dependence is evident in both Garman-Klass volatility and trading volume. The degree of long memory in total and domestic trading volume ranges from 0.55 to 0.65 while across di¤erent sample periods similar long memory characteristics are experienced. The degree of long range dependence in foreign volume is signi…cantly lower (almost half) than that experienced in domestic volume and no signi…cant change is evident for the di¤erent periods considered. The long range dependence in Garman-Klass volatility for the whole sample is 0.50 and diminishes to 0.25 for the pre-crisis period and 0.38 for the post crisis one. As we can see, neglecting the structural break in the mean of Garman-Klass volatility may overestimate the degree of long memory. This result is consistent with Granger and Hyung (2004) who …nd that the volatility series may show the long memory property because of the presence of neglected breaks. Moreover, when we test for a common long memory parameter the null hypothesis is only accepted for foreign volume and Garman-Klass volatility in all three subperiods. Therefore, it appears that there is a close correspondence between the estimated degrees of fractional integration as predicted by the modi…ed MDH (see Andersen and Bollerslev, 1997, Bollerslev and Jubinksi, 1999). Finally, we …nd no evidence that foreign volume and volatility share a common long memory component.

Section 2 reviews the several versions of the mixture of distributions hypothesis that give rise to common long run dependencies in volatility and volume. Moreover, some empirical evidence is provided.

ditionally, he finds that trading volume, used as a proxy for the latent information variable, contains significant explanatory power for return volatility while his inference is mainly univariate and based on the assumption that trading volume is exogenous. Tauchen and Pitts (1983) suggest that price changes and trading volume are jointly determined by an information arrival process functioning as a common mixing variable. Both studies assume that the information process is serially independent, and as a result this argument cannot explain the well known empirical fact that return volatility exhibits highly persistent autoregressive behavior.

Andersen (1996) suggest a mixture of distributions model that explains the joint distribution of return volatility and trading volume under the market microstructure setting of Glosten and Milgrom (1985).

volatility process which is characterised by a highly persistent, though stationary, autocorrelation function

$v_t; j \quad j^{2d-1}$

$j \quad Cov \ X_t; X_{t+j} \quad c_x \ j^{2d \square 1}$

where c_x is a slowly varying function at infinity and positive and ' ' indicates that the ratio of left and right hand sides tends to . The parameter d governs the intensity at which the autocorrelation function decays and summarizes the degree of long range dependence of the series X_t.

In the frequency domain long range dependence is replicated in the spectral density f_x of X_t, defined by

$j \quad \int_{\square} f_x \quad e^{ij} \ d \quad , \quad j \quad ; \quad ; ::: ,$

where f_x asymptotically converges to $G_x j \ j^{\square 2d}$ as ! for some finite constant $G_x >$. The spectral density has a pole at zero frequency when d > ,

$f_X \quad \frac{1}{2} \sum_{j=\square} \ _x j \quad 1$

and this indicates the increasing contribution of low frequency components to the variance decomposition of X_t. When d , the series is weakly dependent and f_x is bounded and positive. In addition, the above asymptotic relations do not provide any information about the short run, seasonal or cyclical behavior of X_t. Robinson (2003) argues that semiparametric definitions of the long range dependence indicate that short-run modeling is almost irrelevant at very low frequencies and very long lags, where d dominates.

3.2 Estimation of long memory parameter d

To test the hypothesis of long-memory we follow Robinson's (1995) semiparametric bivariate approach. To this end, let the sample periodogram for y_{it}, i v; g, at the r-th Fourier frequency, vector of the residuals u_r be given by

$$u_{ir} \quad I_i \quad _r \quad \log c_i \quad ' \quad \hat{d}_{mi} \quad \log \quad _r \; ;$$

with estimated variance-covariance matrix

$$' \; n \quad ^{-1} \sum_{r= \; + \; '} u_r u_r :$$

A test of whether the two variables, $d_{mv;}$ d_{mg}, have the same degree of fractional integration, d_m, is given by

$$W \quad \hat{d} \; f \; ^2 e_2 \; Z \; Z \; ^{-1} e_2 f \quad f \; ^a \; ^2_1; \tag{2}$$

where f denotes the vector ; .

Finally, the corresponding restricted least squares estimator that imposes this commonality on the fractional orders of integration is expressed as

$$d_m \quad \frac{\sum_{r= \; + \; '}^{n} \; _r \quad ^{-1} \quad _r}{^{-1} \quad \sum_{r= \; + \; '}^{n} \; _r^2}; \tag{3}$$

where is a vector of ones, $_r$ is the r-th row of and $_r$, $_r$ $' = n$ $\sum_{r= \; + \; '}^{n}$ $_r$.

where u and c are the differences in the natural logarithms of the high and low, and of the closing and opening prices respectively. Garman-Klass (1980) show that their volatility estimator is about eight times more efficient than using the close to close prices to measure volatility. Moreover, Alizadeh et al. (2002) and Chen and Daigler (2006) argue in favor of using range based volatility measures due the bias introduced by microstructure effects. Shu and Zhang (2006) find that the range estimators quite close to the daily integrated variance.[1]

As regards trading volume disaggregated data concerning domestic and foreign investors' trading activity is also available. We use turnover as a measure of volume. This is the ratio of the value of shares traded to the value of shares outstanding (see, Campbell et al., 1993; Bollerslev and Jubinski, 1999; Lo and Wang, 2000). Because trading volume is nonstationary several detrending procedures for the volume data have been considered in the empirical finance literature (see, for details, Lobato and Velasco, 2000).[2] We form a trend-stationary time series of turnover (y_{vt}) by incorporating the procedure used by Campbell et al. (1993) that uses a 100-day backward moving average

$$y_{vt} \quad \frac{VLM_t}{\frac{1}{100}\sum_{i=1}^{100} VLM_{t-i}};$$

where VLM denotes volume. This metric produces a time series that captures the change in the long run movement in trading volume (see, Brooks, 1998; Fung and Patterson, 1999). The moving average procedure is deemed to provide a reasonable compromise between computational ease and effectiveness. We also extract a linear trend from the volume series. As detailed below, the results for the linearly detrended volume series are very similar to those reported for the moving average detrending procedure. In what follows, we will denote volume by $y_{vt}^{(s)}$ (s totaltota260Td[n2(follo)mhshaTdes2Td[1tota260Td[n2(follo)rted(h)-318

The structural break tests for volatility reveal two change points. The first break is detected on the 15th of October 1997 and, thus, we break the entire sample into two sub-periods. The first subperiod runs from 3rd January 1995 to 15th October 1997, the pre-Asian crisis sample (or Sample A). The second subperiod runs from 16th October 1997 to 26th October 2005, the post-Asian crisis sample (or Sample B). The second change-point for volatility is detected on the 6th of October 2000. For total/domestic volume the testing procedure reveals the existence of a single change-point that is detected on the 20th of January 1999. A single structural break is also detected for foreign volume and it coincides with the first break in volatility. That is, the results of the structural break tests do not support the null hypothesis of homogeneity in the two variables. In order to ensure that the results of this study are not influenced by the break in volume, we also examine the post-crisis period excluding data from 16th of October 1997 to 20th of January 1999 (Sample C). [3]

Table 1. Testing for long-memory

Long memory tests	Volume			Volatility
	Total	Domestic	Foreign	Garman-Klass
KPSS	: □□□	: □□□	: □□□	: :

4.2 Long run dependence in volatility and volume

In this section we are interested in exploring the long run dependence of Garman-Klass volatility as well as that of domestic and foreign investors' trading volume. We employ semiparametric analysis in the frequency domain and estimates of the long memory parameter are reported for the series under study.

Table 2. Semiparametric estimates of d_{mi}

Sample	Volume			Volatility
	Total	Domestic	Foreign	Garman-Klass
Total Sample	: (0:04)	: (0:04)	: (0:04)	: (0:04)
Sample A	: (0:04)	: (0:04)	: (0:03)	: (0:04)
Sample B	: (0:03)	: (0:03)	: (0:03)	: (0:03)
Sample C	: (0:03)	: (0:03)	: (0:04)	: (0:03)

Notes: In all cases and ' . For sample Total (A) we use

$n \quad T^{:7} \; T^{:85}$. For the two post-crisis periods we use $n \quad T^{:8}$.

The numbers in parentheses are standard errors.

It is worth mentioning the empirical results in Granger and Hyung (2004). They suggest that there is a possibility that, at least, part of the long.-memory may be caused by the presence of neglected breaks in the series (see also Diebold and Inoue, 2001). However, the fractional integration parameters are estimated for the various sub-periods, after taking into account the presence of breaks. The long-memory character of the di¤erent volume series remains strongly evident while in the case of volatility there is a signi…cant reduction in the long memory parameter d for the pre and post crisis periods.

4.3 Common long run dependence in volatility and volume

In this section we test whether the Garman-Klass volatility and trading volume have the same degree of long memory as a modi…ed mixture of distributions model may suggest. Empirical results in favor of this common long memory property are reported in Bollerslev and Jubinski (1999) and Lobato and Velasco (2000) for individual stocks. It is very appealing to see whether this property holds for an emerging market's stock index volatility and its trading volume. Additionally, we are interested in investigating whether di¤erent types of investors' trading volume show quite similar long memory characteristics.

Because it has been repeatedly shown that a main feature of return volatility and volume is the presence of long-memory, it is of interest to test if the two variables share the same stochastic properties. The results of Robinson's ^{2} test for a common long-memory parameter in volatility and any of the three volumes are reported in Table 3. A formal test of this hypothesis is available in equation 2. In all four samples the total and domestic volumes produce chi-squared statistics that are higher than the 5% chi-squared critical value of 3.841. In sharp contrast, in all three sub-periods, the null hypothesis that the volatility and foreign volume share a common long-memory parameter cannot be rejected at any conventional signi…cance level. Therefore, it appears that there is a close correspondence between the estimated degrees of fractional integration for the two series as predicted by the MDH (see Bollerslev and Jubinski, 1999). Restricting the value of the d_m to be the same across the volatility and the foreign volume, as in equation 3, results in estimates of d_m: 0.42, 0.27, 0.37 and 0.35 (see the last column of Table 3).

Table 3. Test for equality of d_{mi} estimates

Sample	$d_m^{(T)}$? d_{mg}	$d_m^{(D)}$? d_{mg}	$d_m^{(F)}$? d_{mg}	$d_{mv}^{(F)}$ d_{mg}
Total Sample	: [0:02]	: [0:01]	: [0:01]	: (0:03)
Sample A	: [0:00]	: [0:00]	: [0:72]	: (0:02)
Sample B	: [0:00]	: [0:00]	: [0:62]	: (0:02)
Sample C	: [0:00]	: [0:00]	: [0:08]	: (0:02)

Notes: The table reports Robinson's (1995) ^{2} test statistic for the null hypothesis that volume and volatility have the same long-memory meanparameter d_m. The last column reports the restricted long-memory parameter d_m for foreign volume and volatility. The numbers in [ ] are p-values. The numbers in parentheses are standard errors.

The semiparametric estimates and test statistics also point toward a remarkable commonality in the degree of fractional integration for foreign volume and volatility.[6]

4.4 Fractional cointegration and a common long-memory component

Because it appears that both foreign volume and volatility possess the same long-memory parameter, it is of interest to examine if both processes are driven by the same long-memory component. One way of doing that is to examine whether the two variables are fractionally cointegrated. Fractional cointegration has received much attention lately. Following Davidson (2002) we attempt a fractional bivariate analysis. We employ two versions (the generalised and the regular one) of the fractionally cointegrating vector error correction model (FVECM). General cointegration as de…ned in Davidson et al. (2006) is the case where the cointegrating variables may be fractional di¤erences of the observed series. The generalised FVECM is given by

$$L \qquad L \quad L \; y_t \qquad {}_t;$$

[6] These results are in line with those obtained from the fully parametric bivariate ccc AR-FI-GARCH model.

where is a vector given by , 1 2 , is a vector given by , , and L is a diagonal matrix polynomial with diagonal elements $L^{d^*_{mi}}$, i v; g, with d^*_{mi} d_{mi}. $"_t$ i:i:d: ; with , 1 2 . In the case of regular cointegration linear combinations of fractionally integrated variables are integrated to lower order. Since there are just two variables in the system their order of integration must be equal: d_{m1} d_{m2} d_m. This implies that the orders of integration of the error correction terms must also match to ensure cointegration: d^*_{m1} d^*_{m2} d^*_m. In the generalised

those reported for the 100-day moving average detrending. Further work using Gaussian semiparametric estimators and fractional cointegration analysis as suggested by Robinson (1995b) and Robinson and Marrinucci (2001, 2003) for an Index as well as its constituent individual securities is a subject of future research.

References

[1] Alizadeh, S., Brandt, M., and Diebold, F., (2002). Range-based estimation of stochastic volatility models. Journal of Finance, 1047-1091.

[2] Andersen, T., (1994). Stochastic autoregressive volatility: A framework for volatility modeling. Mathematical Finance, 4, 75-102.

[3] Andersen, T., (1996). Return volatility and trading volume: an information flow interpretation of stochastic volatility. Journal of Finance, 1, 169-204.

[4] Andersen, T., and Bollerslev, T., (1996). Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns. Journal of Finance, 52, 975-1005.

[5] Andersen, T., and Bollerslev, T., (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. International Economic Review, 39, 885-905.

[6] Andersen, T., and Bollerslev, T., Diebold, F., and Ebens, H., 2001. The distribution of realized stock return volatility. Journal of Financial Economics, 61, 43-76.

[7] Baillie, R., Bollerslev, T., and Mikkelsen, H., (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 74, 3–30.

[8] Bai, J., and Perron, P., (1998). Estimating and testing linear models with multiple structural changes. Econometrica, 66, 47-78.

[9] Bai, J., and Perron, P., (2003a). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18, 1-22.

[10] Bai, J., and Perron, P., (2003b). Critical values for multiple structural change tests. Econometrics Journal, 6, 72-78.

[11] Bollerslev, T., and Jubinski, D., (1999). Equity trading volume and volatility: latent information arrivals and common long-run dependencies. Journal of Business & Economic Statistics, 1, 9-21.

[12] Breidt, F. J., Crato, N., and de Lima, P., (1997). On the detection and estimation of long memory in stochastic volatility. Journal of Econometrics, 83, 325-348.

[13] Brooks, C., (1998). Predicting stock index volatility: can market volume help? Journal of Forecasting, 17, 59-80.

[14] Campbell, J. Y., Grossman, S. J., and Wang, J., (1993). Trading volume and serial correlation in stock returns. Quarterly Journal of Economics, 108, 905-939.

[15] Chen, Z., Daigler, R., and Parhizgari, A., (2006). Persistence of volatility in future markets. Journal of Futures Markets, 26, 571-594.

[16] Clark, P., (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. Econometrica, 41, 135-155.

[17] Davidson, J., (2002). A model of fractional cointegration, and tests for cointegration using the bootstrap. Journal of Econometrics, 110, 187-212.

[18] Davidson, J., Peel, D., and David Byers, J. (2006). Support for Government and Leaders: Fractional cointegration analysis of poll evidence from the UK, 1960-2004. Studies in Nonlinear Dynamics and Econometrics, 10, 1-21

[19] Davidson, J. (2006) Time Series Modelling 4.19, Main Document, http://www.timeseriesmodelling.com.

[20] Diebold, F. and Inoue, A. (2001). Long memory and regime switching. Journal of Econometrics, 105, 131— 159.

[21] Ding, Z., Granger, C. W., and Engle, R. F., (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1, 83-106.

[22] Fung, H., and Patterson, G., (1999). The dynamic relationship of volatility, volume, and market depth in currency futures markets. Journal of International Financial Markets, Institutions and Money, 9, 33-59.

[23] Garman, M., and Klass, M., (1980). On the estimation of security price volatilities from historical data. Journal of Business, 1, 67-78.

[24] Glosten, L., and Milgrom, P., (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed tradrers. Journal of Financial Economics, 13, 71-100.

[25] Granger, C., and Hyung, N., (2004). Occasional structural breaks and long memory with an application to the S&P 500 absnl(op)te stock returns. Journal of Empirical Finance, 11, 399-421.

[26] Geweke, J. and Porter-Hudak, S., (1983). The estimation and application of long memory time series models. Journal of Time Series Analysis, 4, 221-238.

[27] Harris, D., McCabe, B., and Leybourne, S., (2008). Testing for long memory. Econometric Theory, 24, 143-175.

[28] Hurst, H., 1951, Long Term Storage Capacity of Reservoirs. Transactions of the American Society of Civil Engineers, 116, 770-799.

[29] Hurvich, C.M., R.S. Deo and J. Brodsky (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. Journal of Time SeriesAnalysis 19, 19-46.

[30] Hurvich, C., and Soulier, P., (2002). Testing for long memory in volatility. Econometric Theory, 18, 1291-1308.

[31] Kirman, A., and Teyssiere, G., (2002). Microeconomic models for long memory in the volatility of …nancial time series. Studies in Nonlinear Dynamics and Econometrics, 5, 281-302.

[32] Kunch, H. R. (1986) Discrimination between monotonic trends and long range dependence. Journal of Applied Probability, 23, 1025-1030.

[33] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y., (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Journal of Econometrics, 54, 159-178.

[34] Lavielle, M. and Moulines, E., (2000). Least-squares estimation of an unknown number of shifts in a time series, Journal of Time Series Analysis, 21, 33-59.

[35] Lee, D., and Schmidt, P., (1996). On the power of the KPSS test of stationarity against fractionally-integrated alternatives. Journal of Econometrics, 73, 285-302.

[36] Lo, A. W., (1991). Long term memory in stock market prices. Econometrica, 59, 1279-1313.

[37] Lo, A. W., and Wang, J., (2000). Trading volume: de…nitions, data analysis, and implications of portfolio theory. Review of Financial Studies, 13, 257-300.

[38] Lobato, I., and Velasco, C., (2000). Long memory in stock-market trading volume. Journal of Business & Economic Statistics, 18, 410-427.

[39] Luu, J., and Martens, M., (2003). Testing the mixture-of-distributions hypothesis using "realized" volatility. Journal of Future Markets, 7, 661-679.

[40] Newey, W. K., and West, K. D., (1987). A simple positive de…nite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703-708.

[41] Perron, P., and Qu, Z., (2004). An analytical evaluation of the log-periodogram estimate in the presence of level shifts and its implications for stock returns volatility. Working Paper Series, Department of Economics, Boston University, USA.

[42] Robinson, P. M., (1994). Semiparametric analysis of long memory time series. Annals of Statistics, 22, 515-539.

[43] Robinson, P. M., (1995a). Log-periodogram regression of time series with long-range dependence. Annals of Statistics, 23, 1048-1072.

6 Appendix

6.1 Testing for Long Memory

In order to test for long memory we use the Lo's modified R=S test (Lo, 1991), the KPSS test (Kwiatkowski et al., 1992), and the 'HML' test (Harris et al. 2008). Lo (1991) proposed a modified version Hurst's (1951) 'rescaled range' or 'R=S' statistic. The 'R=S' statistic is the range of partial sums, S_k, of deviations of a time series from its mean, S_k $\sum_{j=1}^{k} Y_j$ Y_n, rescaled by its standard deviation, $_n$, and is defined as

$$R=S \quad \frac{1}{n} \; {}_{1 \; k \; n} S_k \quad {}_{1 \; k \; n} S_k .$$

Lo's modified version of the 'rescaled range' statistic differs from the 'R=S' defined above only in

itF1084F52(g)S5ned above only in1(a)1(IeN2d[(’)-129(d)-1(eTd.7295)124(ned)-407(00(’)-129(d)-1(eTd.72911.494ns.7295)12