Quaderni di Statistica
Vol. 7, 2005

Common long memory in the Italian stock market

Maria Simona Andreano
Dipartimento di Teoria Economica, Università degli Studi di Roma "La Sapienza"
E-mail: s.andreano@dte.uniroma1.it

Summary: The mixture of distribution hypothesis originally introduced by Clark (1973) and Epps and Epps (1976) suggests that returns and trading volume are driven by the same underlying latent information flow. According to this finding, Bollerslev and Jubinski (1999) have recently shown that the latent aggregate information-arrival process hitting the stock market should be fractionally integrated. This long-run dependency passes to both the trading volume and the volatility series, and their long-run decay rates are the same. This paper shows the existence of such common memory behaviour in the Italian stock market. The multivariate spectral method of Bollerslev and Jubinski is applied on the trading volume and absolute returns series of the individual firms composing the MIB30 index.

Keywords: Fractional integration, Return volatility, Common dependencies.

1. Introduction

Over the past decade, temporal dependencies in financial market volatility have increasingly attracted the attention of empirical literature. Numerous empirical findings have demonstrated a positive contemporaneous volume-absolute price change correlation (Karpoff, 1987, Stickel and Verrecchia, 1994). An explanation of this correlation comes from research on the distribution of speculative daily prices sampled from a set of distributions with different variances. According to this "mixture of distribution hypothesis" (MDH), Epps and Epps (1976) derive a MDH model in which the variance of the price change on a single transaction is conditioned by the volume of that transaction. Clark (1973) posits a joint dependence of returns and volume on an underlying latent event or information flow variable. The volume is driven by the identical factors that generate return volatility.

Early tests of the implications of the MDH have been supportive of the model (Harris, 1987) but recent studies have produced contradictory evidence (Lamoureux and Lastrapes, 1994, Heimstra and Jones, 1994). Modifications to the standard MDH model have been carried out by Tauchen and Pitts (1983), Andersen (1996) and Liesenfeld (1998). Tauchen and Pitts generalize the previous version of the model allowing for serial correlation in the information arrival process. Andersen assumed that the failure of the MDH could be due to the artificial distributional assumptions and let the information flow be a stochastic volatility process. Andersen found that the modified MDH model with a Poisson-distributed trading volume was not rejected when applied on series that firmly refused the more traditional version of the MDH.

The conflicting empirical results based on the standard MDH and the assumption of a single latent information arrival process may also reflect the existence of a more complex dynamic relationship between trading volume and volatility: the responses to the news are not necessarily the same in the two variables. On this idea Liesenfeld has recently proposed the bivariate MDH using a bivariate mixture model with two information-arrival variables.

Recent analyses have fully exploited the dynamic implications of the MDH by explicitly parameterizing the process for news arrivals. Bollerslev and Jubinski (1999), henceforth BoJu, test the implications of the MDH as a long-run hypothesis, according to which the aggregate latent information arrival process has long-memory characteristics. The authors show that volatility and volume exhibit the same degree of long-memory dependence. Using a semi nonparametric log-periodogram regression, they found that fractionally integrated processes succeeded in describing the long-run temporal dependencies in equity trading volume and volatility for the individual firms composing the Standard & Poor's 100 index.

These findings are consistent with a modified version of the MDH, in which the volume-volatility relationship is determined by a latent news-arrival structure with long-memory characteristics.

In this paper the hypothesis of a fractional common latent information arrival is verified for the Italian daily stock market series using the BoJu methodology. Namely stock-market volume exhibits long memory and return volatility and volume share the same long-memory parameter d.

The remainder of the paper is organized as follows. Section 2 discusses the method applied to investigate common long-memory characteristics in fractionally integrated processes. Section 3 reports on the results of the empirical analysis. Some brief conclusions follow.

2. Fractional integration and long-memory dependencies

We denote with $R_{j,t}$ the daily continuously compounded return on stock j, $j = 1, 2, ..., N$, and for day $t = 1, 2, ..., T$. The corresponding daily trading volume is denoted by $V_{j,t}$.

The intradaily price changes are likely to vary across time, depending on the number of "news" arrivals, $K_{j,t}$, that occur during the day. Conditional on this latent intensity process, the distribution for the daily returns can be expressed as

$$R_{j,t}|K_{j,t} \sim N(0, \sigma_j^2 \cdot K_{j,t}). \tag{1}$$

Andersen (1996) showed that the resulting daily trading volume, conditional on the information-arrival process, are approximately Poisson distributed

$$V_{j,t}|K_{j,t} \sim Po(\mu_{j,0} + \mu_{j,1} \cdot K_{j,t}) \tag{2}$$

where $\mu_{j,0}$ and $\mu_{j,1}$ are normalizing constants related to the importance of liquidity and information-based trading, respectively.

Parke (1996) argumented that, under reasonable assumptions about the survival probabilities of news, the latent aggregate information-arrival process is fractionally integrated. Therefore, as shown by BoJu, the MDH implies that

$$corr(\,|R_{j,t}|\,,|R_{j,t-\tau}|\,) \sim \tau^{2d_j-1} \tag{3}$$

and

$$corr(V_{j,t}, V_{j,t-\tau}) \sim \tau^{2d_j-1}. \tag{4}$$

Then, they proposed a semi nonparametric log-periodogram regression in order to test Equations (3) and (4).

This is a multivariate approach that specifies both estimators and test statistics for the hypothesis of a common long-run hyperbolic decay rate of absolute returns and trading volume.

The bivariate covariance stationary time series is denoted with $X_{j,t} \equiv (|R_{j,t}|\,, V_{j,t})'$. If each of these series is fractionally integrated, it follows that, at frequencies λ close to 0

$$\{f_j(\lambda)\}_{gg} \approx C_{j,g} \cdot \lambda^{-2d_{j,g}}, \tag{5}$$

where $f_j(\lambda)$ is the spectral density at frequency λ, $g = 1, 2; C_{j,1}$ and $C_{j,2}$ are two scaling constants, and $d_{j,1}$ and $d_{j,2}$ are the orders of integration of $|R_{j,t}|$ and $V_{j,t}$ respectively. The sample periodogram for $\{X_{j,t}\}_g$ is denoted by

$$I_{j,g}(\lambda) \equiv (2 \cdot \pi \cdot T)^{-1} \left| \sum_{t=1}^{T} \{X_{j,t}\}_g \exp(i \cdot t \cdot \lambda) \right|^2. \tag{6}$$

Following Robinson (1995), define the $(m - l) \times 2$ matrix, $\mathbf{Y}_j$, with the (k, j)th element equal to the log-periodogram at the kth Fourier frequency, that is

$$\{Y_j\}_{kg} \equiv \log\left[I_{j,g}(\lambda_k)\right] \tag{7}$$

for $k = l{+}1, l{+}2, ..., m$ with l and m the trimming and truncation parameters.

The authors suggest the following least squares estimator for $\mathbf{d}_j \equiv (d_{j,1}; d_{j,2})$

$$\widehat{\mathbf{d}}_j{=}\mathbf{Y}_j'\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{e}_2 \tag{8}$$

where $\mathbf{e}_2 \equiv (0,1)'$and the elements in the $(m-l) \times 2$ matrix of the explanatory variables are defined by $\{Z_j\}_{k,1} = 1$ and $\{Z_j\}_{k,2} = -2 \cdot \log(\lambda_k)$ for $k = l+1, ..., m.$

Although the residual vectors in the regression to derive $\widehat{\mathbf{d}}_j$ do not satisfy usual regularity conditions, asymptotic properties of this least squares estimate have been derived by Robinson (1995)

$$\left[\mathbf{e}_2'(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{e}_2\widehat{\mathbf{\Omega}}_j\right]^{-1/2}\left(\widehat{\mathbf{d}}_j{-}\mathbf{d}_j\right) \longrightarrow N(0, I) \tag{9}$$

where

$$\widehat{\mathbf{\Omega}}_j{=}\ (m-l)^{-1}{\cdot} \sum_{k=l+1}^{m} \widehat{\mathbf{u}}_{j,k}\ \widehat{\mathbf{u}}_{j,k}' \tag{10}$$

and $\mathbf{u}_{j,k}$ is the approximation of the errors of the estimated regression.

Similarly, a two-sided test for the hypothesis $d_{j,1} = d_{j,2}$, a common long-run hyperbolic rate of decay across volatility and trading volume series exists, can be based on the asymptotic chi-squared statistic

$$(\widehat{\mathbf{d}}_j'\,\mathbf{f})^2{\cdot}\mathbf{e}_2'(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{e}_2{\cdot}\mathbf{f}'\widehat{\mathbf{\Omega}}_j\mathbf{f}{\longrightarrow}\ \chi_1^2 \tag{11}$$

where $\mathbf{f}$ denotes the 2×1 vector $(1, -1)'$.

3. Data and empirical results

In this section we formally test for the existence of common long-memory dependency in volume and return series in the Italian stock market components of the MIB30 index. The series share the period from 1/1/1999 to 8/10/2004, for a maximum of 1453 observations for each series. The starting date has been chosen to avoid misspecifications problems due to the Euro introduction. Two series have been eliminated from the analysis because they were defined over a too short period.

In Figure 1 and 2 the average series of the absolute returns and trading volume are shown, with

$$\left|\overline{R}_t\right| = \frac{1}{N_t}\sum_{j=1}^{30}\left|R_{j,t}\right|\cdot I_{j,t} \qquad \overline{V}_t = \frac{1}{N_t}\sum_{j=1}^{30} V_{j,t}\cdot I_{j,t} \tag{12}$$

and the indicator function $I_{j,t}$ defined as

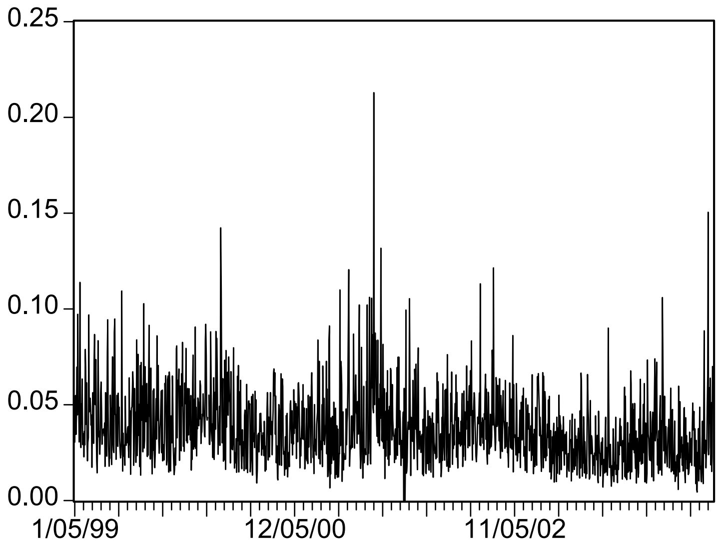

Figure 1. Average Absolute Returns

$$I_{j,t} = \begin{cases} 1 & \text{if } \mathrm{V}_{j,t} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

and $N_t = \sum_{j=1}^{30} I_{j,t}$.

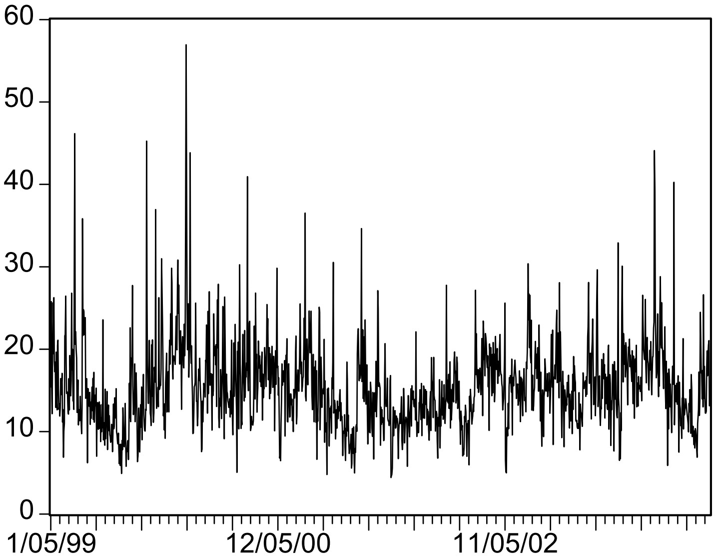

Figure 2. Average Trading Volume (in millions)

The estimates are based on the trimming parameter $l = 1$ and the truncation parameter $m_j = (T_j)^{1/2}$, where T_j is the number of observations for the j-th firm. In the table below we report some statistics for the averaged series of absolute returns and trading volume.

Table 1. Statistics of Absolute-Returns and Volume (in average)

Statistics	$\|R\|$	V
Skewness	1.583	1.601
Kurtosis	5.921	6.654
Mean	0.038	15435609
Q_5	123.73*	983.74*
Q_{20}	369.39*	1844.3*
d	0.316	0.558

Legenda: (*) significance at 5% level

Starting from similar results, Regùlez and Zarraga (2002) test whether stock return and trading volume share normality, skewness, kurtosis, serial correlation and seasonality as common features.They found a common factor driving returns and volume in terms of serial dynamics, seasonality and skewness, confirming common behaviour in the latent information arrivals.

Table 2. Estimates of the parameter d

	$d_{\|R\|}$	d_V	$d_{V^{DT}}$	d_1	d_2
Mean	0.377	0.403	0.346	0.376	0.351
Median	0.352	0.417	0.360	0.393	0.347
Min	0.035	0.103	0.073	0.085	0.059
Max	0.718	0.657	0.752	0.574	0.613

Legenda: $d_{V^{DT}}$ estimates for detrended volume series

d_1 restricted estimates when $d_{|R|}=d_V$

d_2 restricted estimates when $d_{|R|}=d_{V^{DT}}$

Table 1 shows the bivariate estimates of the long-range parameter d, equal to 0.316 and 0.558, both significantly different from 0 (the limiting standard errors are 0.110 and 0.111 respectively).

The statistics on the single series components of the MIB30 index are similar to those reported for the averaged series. The estimates of d for the absolute returns and the trading volume are summarized in Table 2.

BoJu's multivariate spectral methodology has been applied on both the raw trading volume and the detrended series. The results are quite similar in terms of d estimation and hypothesis testing. The linear detrending procedure does not seem to significantly affect our results. Graph and tests are reported for the linearly detrended series.

Table 2 shows the estimates of the fractional integration parameter for the absolute returns, the trading volume and the detrended trading volume. The value of d for the absolute returns varies between 0.035 and 0.718. However the proof of Robinson (1995) implied in (8) is valid only when $-0.5 < d_j < 0.5$. For the case where $d_{|R|} = 0.718$, we difference the series and re-estimate the d parameter. In this way, the value fall down at 0.607. For the trading volume series: the $\max d_{V^{DT}}$ is 0.591, approaching the stationarity region assumed by Robinson's findings.

The mean and median values of the estimated d are very similar, ranging between 0.346 and 0.417. Ray and Tsay (1998) and BoJu, using the firms in the S&P500 and S&P100 index respectively, reported estimates close to those obtained in this paper.

The estimates d_1 and d_2 are obtained by imposing the restrictions $d_{|R|} = d_V$ and $d_{|R|} = d_{V^{DT}}$. The standard errors of d_1 and d_2 are lower than those of the original estimates: the restrictions imply a gain in efficiency.

The presence of common long memory in stock returns and trading volume is represented by similar estimates of $d_{|R|}$ and $d_{V^{DT}}$. Therefore we expect a thickening of the points $(d_{|R|}; d_{V^{DT}})$ around the bisector line in the scatterplot of Figure 3. In fact, we note an underlying behaviour of the points along the 45° line, although some outliers are present.

Finally we report the results of the tests to statistically verify the presence of common long memory in the 28 firms examined in the present paper. Table 3 shows the test-statistics and the corresponding p-value for the test $H_0 : d_{|R|} = d_{V^{DT}}$. The statistics have a chi-squared distribution with one degree of freedom. We cannot reject the null at the 5% level in most cases, whilst for six series we reject H_0 at the 5% significant level. If we consider the non detrended volume series, only five firms do not show common long memory.

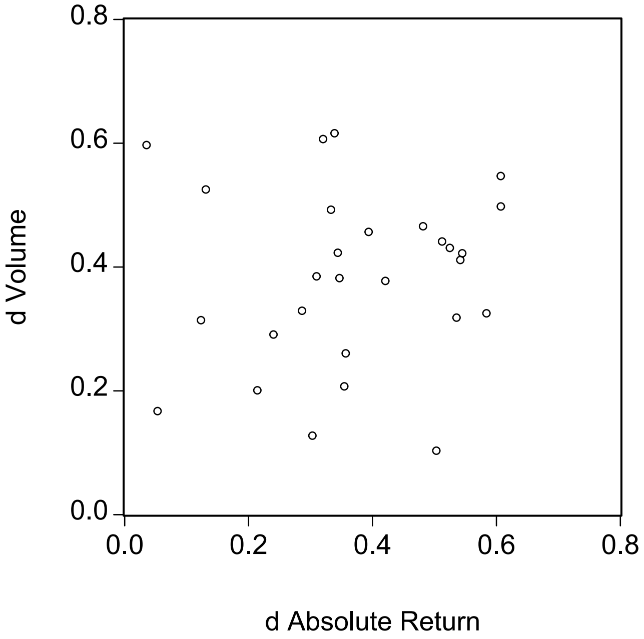

Figure 3. Estimated $d_{|R|}$ *and* $d_{V^{DT}}$

Table 3. Tests for common long-run dependence

Firm	Test-stat	p-value	Firm	Test-stat	p-value
AL	4.139	0.042	GA	0.512	0.474
AG	0.913	0.339	LU	0.180	0.671
AS	0.256	0.613	MD	5.476	0.019
BA	0.199	0.655	ME	0.003	0.952
BM	1.175	0.278	MS	3.371	0.066
BN	0.371	0.542	PG	7.264	0.007
BP	0.002	0.963	RA	0.918	0.338
BV	0.601	0.438	SI	0.050	0.824
CP	6.475	0.011	SP	0.788	0.375
CR	0.668	0.414	SM	0.231	0.631
ER	0.106	0.744	RR	0.469	0.493
EN	0.004	0.947	ST	6.602	0.010
EI	19.383	0.000	TI	0.812	0.367
FI	0.049	0.824	TT	0.091	0.763

It is interesting to note that in four out of the five cases in which H_0 is rejected $d_{|R|} > d_{V^{DT}}$, whilst in one $d_{|R|} = 0$.

The chi-squared statistic is less than the 5% critical value (3.841) in 78% of the firms included in the MIB30 index.

All the analyses have been carried out with a truncation parameter $m = (T_j)^{1/2}$ in the periodogram regression in (7). However, a sensitivity analysis has been made in order to verify the robustness of the results. The cases $m = 0.5\ (T_j)^{1/2}$ and $m = 1.5\ (T_j)^{1/2}$ have been examined. The bivariate estimated values of d, the restricted estimates and the test-statistics are similar and the conclusions are the same.

4. Conclusions

In this article the BoJu method for carrying out statistical inference on the long memory parameters of stock market trading volume and volatility have been examined for the Italian case. Applications indicate that for the stocks in the MIB30 index the trading volume exhibits long memory and that it shares the same long-memory parameter of the volatility process for most of the stocks considered. Contrarily to the critiques raised by Lobato and Velasco (2000) we have found that the results on the common long-run dependency is not affected by the application of the linearly detrending procedure on the trading volume series. However, a limited number of stock market series have been analysed. Further research should be carried out to confirm the presence of a latent information arrival that exhibits common long-run dependency in the Italian market.

References

Andersen T. G. (1996), Return volatility and trading volume: an information flow interpretation of stochastic volatility, *Journal of Finance*, 51, 169-204.

Bollerslev T. and Jubinski D. (1999), Equity trading volume and volatility:

latent information arrivals and common long-run dependencies, *Journal of Business and Economic Statistics*, 17, 9-21.

Clark P. K. (1973), A subordinated stochastic process model with finite variance for speculative prices, *Econometrica*, 41, 135-155.

Epps T. W. and Epps M. L. (1976), The stochastic dependence security price changes and transaction volumes: implications for the mixture-of-distributions hypothesis, *Econometrica*, 44, 305-321.

Harris L. (1987), Transaction data tests of the mixture of distributions hypothesis, *Journal of Financial and Quantitative Analysis*, 22, 127-141.

Heimstra C. and Jones J. D. (1994), Testing for linear and nonlinear Granger causality in the stock price-volume relation, *Journal of Finance*, 49, 1639-1664.

Karpoff J. M. (1987), The relation between price changes and trading volume: a survey, *Journal of Financial and Quantitative Analysis*, 22, 109-126.

Lamoureux C. G. and Lastrapes W. D. (1994), Endogenous trading volume and momentum in stock-return volatility, *Journal of Business and Economic Statistics*, 12, 253-260.

Liesenfeld R. (1998), Dynamic bivariate mixture models: modeling the behaviour of prices and trading volume, *Journal of Business and Economic Statistics*, 16, 101-109.

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

Parke W. R. (1996), What is a fractional unit root?, *unpublished manuscript*, Dept. of Economics, North Carolina University.

Ray B. K. and Tsay R. S. (1998), Common long-range dependence in daily stock volatility, *unpublished manuscript*, Graduate School of Business, University of Chicago.

Regùlez M. and Zarraga A. (2002), Common features between stock return and trading volume, *Applied Financial Economics*, 12, 885-893.

Robinson P. M. (1995), Semiparametric analysis of long-memory time series, *The Annals of Statistics*, 22, 515-539.

Stickel S. and Verrecchia R. (1994), Evidence that trading volume sustains stock price changes, *Financial Analysts*, 29, 101-116.

Tauchen G. E. and Pitts M. (1983), The price variability-volume relationship on speculative markets, *Econometrica*, 51, 485-505.

Forum