Statistical properties of trading volume depending on size

Maria Pasquale [a]

[a] *Dipartimento di Fisica, Università di Pisa*

Roberto Renò [b,1]

[b] *Dipartimento di Economia Politica, Università di Siena*
Pia zza S.Francesco 7, 53100, Siena, Italy

Abstract

By analyzing high frequency data of transactions on the Italian stock index futures, we show that the statistical properties of average volume depend on the size of transactions, defined as the number of contracts transacted in a single operation. In particular, we find that large transactions are less correlated with market volatility and display a longer memory.

1 Introduction

In financial markets, it is important to empirically determine the impact of information on price movements. This task has typically been accomplished by studying the volume-volatility relation, see e.g. [1].

Typically, share volume statistical properties have been studied irrespective of the size of single trades, see e.g. [2,3]. Neverthless, it emerges in literature [4–9] that the volume-volatility relation may depend on the type of trader. Moreover, in competitive models with asymmetric information [10–12] the *size* of trades, that is the number of shares traded in a single transaction, is positively correlated to the quality or precision of information possessed by informed traders.

In this note, we study the properties of trading volume when conditioned on size on the Italian futures market. In particular, we show that the relation

[1] Corresponding author. E-mail: reno@unisi.it.

Preprint submitted to Elsevier Preprint *13 May 2004*

volume-volatility and the memory of volume are different when we consider volumes of different size. We find that the average volume of larger trasactions is more anticorrelated with volatility, and that it has a longer memory than expected. Our results can be compared with those of [4], who find that better informed traders induce a negative volume-volatility relation. From this perspective, the size could be interpreted as a proxy for information of a trader. Moreover, a recent theoretical model [13] predicts that large volumes are due to large traders, and that they could be less correlated with the rest of the market, which is in line with our empirical findings. To our knowledge, this is the first empirical study on the statistical properties of trading volume when we include the size as a distinctive property.

2 Data

Our data set consists of all tick-by-tick transactions of the futures on the Italian stock index MIB30, named FIB30, in 2000-2002, with quarterly expirations. We focus on next-to-expiration contracts, that provides us more than $8,500,000$ transactions. Every transaction is recorded with its size. In Figure 1 we report the intraday time series of the size of transactions in a typical trading day, September 2^{nd}, 2000. It emerges that the size is usually smaller than few contracts (in the full sample, 80.06% of transactions has a size of 1 contract and the average size is 1.5; the nominal value of one FIB30 contract is 5 euros times the futures price), but there are transactions with more than 180 contracts (in the full sample, 0.67% of transactions have size larger or equal than 10).

Our volatility estimation technique makes use of an algorithm proposed in [14] and developed in [15,16]. This method is based on the Fourier analysis of the asset price time series and employs all the unevenly spaced observations. Estimating precisely daily volatility with intraday data is crucial for the subsequent analysis. This algorithm was first applied to the transactions of the FIB30 in 2000-2001 [17]. We extended the analysis to 2002. The whole time series of Fourier daily volatility σ_t is displayed in Figure 2. The maximum corresponds to September 11^{th}, 2001.

We defined three proxy for trading volume: number of transactions N_t, total share volume V_t and average volume traded per transaction, $A_t = V_t/N_t$. The subscript t indexes the trading day. The time series of average volume A_t, shown in Figure 3, exhibits periodical peaks, corresponding to expiration days, due to *rollovers* [18]. In order to avoid spurious correlations, we discard the last four days before expiration from the sample. In total, we are left with 703 days.

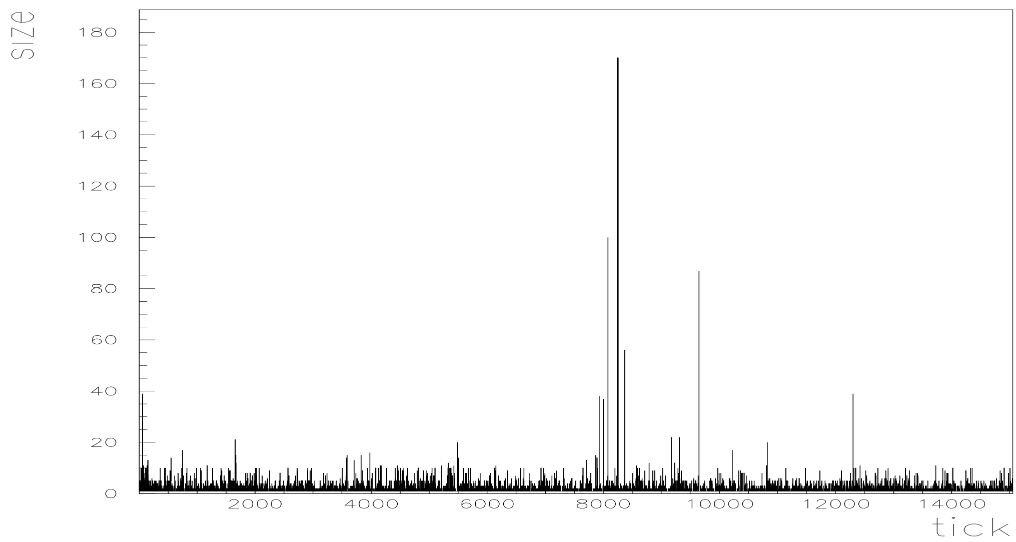

Fig. 1. Intraday time series of the size of transactions in one trading day (September 2^{nd}, 2000).

As expected, Figure 4 shows that there is a strong correlation between daily logarithmic volatility and daily number of transactions N_t, as suggested by Clark [1]. The same relation holds between daily logarithmic volatility and daily total share volume V_t. It appears that the number of contracts and daily volume are strongly correlated (the linear correlation coefficient between the two in the sample is 0.90 ± 0.07). In what follows, we analyze the volume only, since we find identical results when using the number of contracts instead. On the other side, the relation between volatility and average volume A_t is not straightforward to analyze.

Our purpose is to analyze whether and how the properties of the different volume measures change when we calculate the volume using only the transactions characterized by a size over a fixed threshold. As the threshold increases, the number of transactions considered decreases and the statistical analysis on time series may present a non negligible white noise contribution. So we will always compare the results achieved for those subsamples with those achieved for subsamples of equal dimension (same number of transactions) but randomly sampled (bootstrapping). By iterating the random sampling of transactions, we are able to estimate confidence intervals.

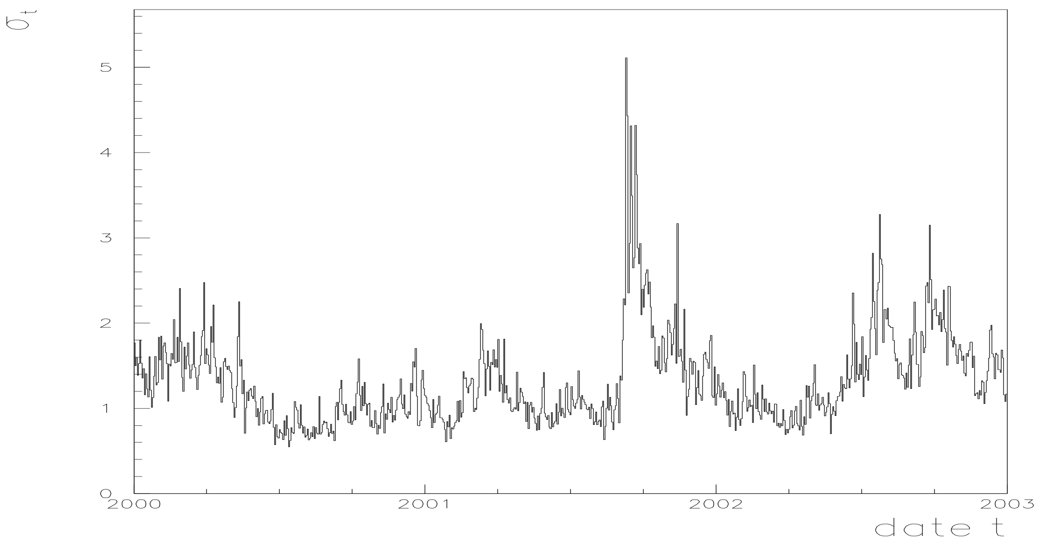

Fig. 2. Time series of daily volatility computed with the Fourier method.

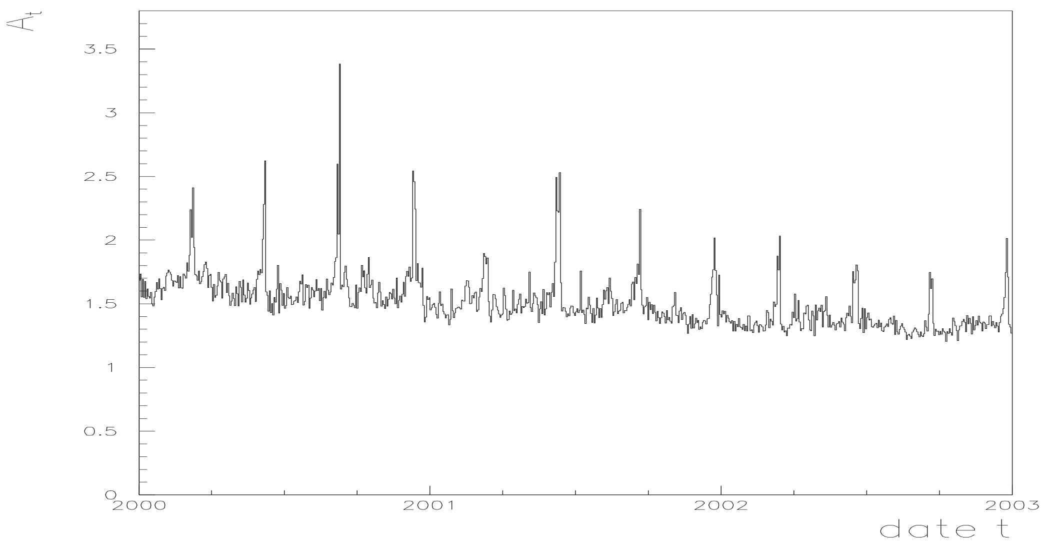

Fig. 3. Time series of average volume A_t. Periodical peaks are observed, due to *rollovers* as maturity approches.

3 Volume-volatility correlation

Daigler and Wiley [4] suggest that the volume-volatility relation depends on the degree of information of the trader. In particular, they find a negative relation for more informed traders, as opposed to the usual positive relation. If the size is a proxy of the degree of information of a trade, volumes of different sizes may have a different relation with volatility.

We then study the correlation between daily logarithmic volatility $\log \sigma_t$ and the volume V_t^s for $t = 1, \ldots, 703$, where V_t^s is defined as the volume of transactions with size $\geq s$. Then V_t^1 is the total daily volume. We define in the same way the number of contracts N_t^s and the average volume $A_t^s = V_t^s / N_t^s$.

Figure 5 shows that the correlation of volumes with volatility decreases with increasing size, but this effect can be ascribed to the reduction of the sample.

On the other hand, we observe a decrease of the correlation of average volume with volatility, which becomes significantly more negative when we exclude transactions with size smaller than 3. For the larger transactions, we cannot disentangle this effect from the reduction effect, maybe due to the dramatic drop in statistics (only 5.07% of the trades have size larger or equal than 4).

4 Serial dependence

If volumes of different sizes have different dynamic properties, they may display different serial dependence. Typically, share volumes display long memory [19,20].

We study the memory of V_t^s and A_t^s for different values of s via spectral analysis and Detrended Fluctuation Analysis.

Spectral analysis is linked to the correlation function through the *Wiener-Khintchine* theorem [21],

$$S(f) = \mathcal{F}[Corr(g(t))] \tag{1}$$

where $\mathcal{F}$ denotes Fourier transform and $g(t)$ is the time series under study. We investigate the relation $S(f) \sim f^{-\eta}$, and estimate η by regressing $\log S(f)$ on $\log(f)$.

Detrended fluctuation analysis (*DFA*) is, instead, based on the idea that a correlated time series can be mapped to a self-similar process by integration [22–24]. Therefore, the measurement of self-similar features reveals indirect informations about the correlation properties. The advantages of *DFA* over the other methods are that it allows the detection of long-range correlations embedded in a non-stationary time series and also avoids the spurious detection of apparent long-range correlations that are an artifact of non-stationarities. Denoting by l the time window length and by $\sigma(l)$ the detrended variance of the time series in the window, we look for a relation $\sigma(l) \sim l^{\alpha}$ (see [23,24] for details).

For a long-range process, characterized by an autocorrelation function $Corr(t) \sim t^{\xi-1}$ with $0 < \xi < 1$, we have a power spectrum with $\eta = \xi$ and $\alpha = \frac{\xi+1}{2}$. For a random walk, we have $\eta = 0$ and $\alpha = 0.5$. Examples of the estimates of η and α for V_t^1 are reported in Figures 6 and 7 respectively.

We first investigate the memory properties of V_t^s and A_t^s via spectral analysis. Figure 8 shows how η changes depending on the cut in size. Volumes and average volumes of larger size transactions show lower α and η. However, it is not possible to disentangle any effect, since the same behavior is displayed by random subsamples of same dimension. For the total volume ($s = 1$), we find the well known result $\eta = 1.0 \pm 0.2$, in line with the result of [19] on US stocks, while for the average volume of all transactions A_s^1 we find $\eta = 1.5 \pm 0.1$.

Figure 9 reports the results of DFA analysis. The α of V_t^s decreases with s in the same fashion of a random subsample, thus no peculiar behavior is observed. For the total volume, we have $\alpha = 0.98 \pm 0.03$, slightly larger than [20] on US stocks. Things change with the average volume A_t^s. The estimate for A_t^1 is $\alpha = 0.91 \pm 0.02$. We observe that, while the memory coefficient α decreases with increasing s, it is well above its estimate on a random subsample. Then, we conclude that the memory of the average volume of large size transactions is higher then expected.

5 Conclusions

In this note, we studied the properties of volume when conditioned on the size of the trades in the Italian stock index futures market.

Keeping in mind what emerged from the study of volume-volatility correlation, we conclude that the process of average volume of large transactions changes significantly with size: it is more anticorrelated with volatility and exhibits a longer memory than the average volume of random transactions.

Our results indicate that the size of transactions is not included in the information set which includes prices and volume. Then, it is important to assess the precise impact of size on price movements as well as its role as a proxy for information. We think that this could be a relevant topic for future research.

6 Acknowledgements

We wish to acknowledge Borsa Italiana, and especially Concetta Ricciardi and Ada De Roma, for their assistance. We wish to acknowledge Rosario Rizza, Carmelo Genovese and Simone Bianco for research assistance, as well as Carlo Lucheroni and Michele Trenti for their remarks.

References

[1] P. K. Clark. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41(1):135–156, 1973.

[2] V. Plerou, P. Gopikrishnan, X. Gabaix, L. A. N. Amaral, and H. E. Stanley. Price fluctuations, market activity and trading volume. *Quantitative Finance*, 1:262–269, 2001.

[3] V. Plerou, P. Gopikrishnan, L. A. N. Amaral, X. Gabaix, and H. E. Stanley. Economic fluctuations and anomalous diffusion. *Physical Review E*, 62(3):3023–3026, 2000.

[4] R. T. Daigler and M. K. Wiley. The Impact of Trader Type on the Futures Volatility-Volume Relation. *The Journal of Finance*, LIV(6):2297–2316, Dicembre 1999.

[5] M. D. D. Evans and R. K. Lyons. Order Flow and Exchange Rate Dynamics. *Journal of Political Economy*, 110:170–180, 2002.

[6] M. Grimblatt and M. Keloharju. How Distance, Language, and Culture Influence Stockholdings and Trades. *Journal of Finance*, 56:1053–1073, 2001.

[7] H. Bessembinder and P. J. Seguin. Futures-Trading Activity and Stock Price Volatility. *Journal of Finance*, 28:2015–2034, 1992.

[8] E. C. Chang, J. M. Pinegar, and B. Schachter. Intraday variations in volume, variance and participation of large speculatiors. *Journal of Banking and Finance*, 21:797–810, 1997.

[9] G. H. Bjonnes, D. Rima, and H. O. A. Solheim. Volume and Volatility in the FX-Market: Does it matter who you are? CESIFO working paper n.786.

[10] B. D. Grundy and M. McNichols. Trade and Revelation of Information through Prices and Direct Disclosure. *Review of Financial Studies*, 2:495–526, 1989.

[11] O. Kim and R. E. Verrecchia. Market Reaction to Anticipated Annuouncements. *Journal of Financial Economics*, 30:272–309, 1991.

[12] J. Wang. A Model of Competitive Stock Trading Volume. *Journal of Political Economy*, 102:127–168, 1994.

[13] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley. A theory of power-law distributions in financial market fluctuations. *Nature*, 423:267–270, 2003.

[14] P. Mallavin and M. Mancino. Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6:49–61, 2001.

[15] E. Barucci and R. Renò. On Measuring Volatility of Diffusion Processes with High Frequency Data. *Economics Letters*, 74:371–378, 2002.

[16] E. Barucci and R. Renò. On measuring volatility and the GARCH forecasting performance. *Journal of International Financial Markets, Institutions and Money*, 12:183–200, 2001.

[17] R. Renò and R. Rizza. Is volatility lognormal? Evidence from Italian futures. *Physica A*, 322:620–628, 2003.

[18] M. Pasquale. Analisi di volumi e volatilità nel mercato italiano dei futures. Master's thesis, Università degli studi di Pisa, 2003-2004.

[19] G. Bonanno, F. Lillo, and R. N. Mantegna. Dynamics of the number of trades of financial securities. *Physica A*, 280:136–141, 2000.

[20] P. Gopikrishnan, V. Plerou, X. Gabaix, and E. Stanley. Statistical properties of share volume traded in financial markets. *Physical Review E*, 62(4):4493, Ottobre 2000.

[21] M. B. Priestley. *Spectral Analysis and Time Series.* Academic Press, 1981.

[22] J. Beran. *Statistics for Long-Memory Processes.* Chapman & Hall, New York, 1994.

[23] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Physical Review E*, 49:1685–1689, 1994.

[24] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5:82–87, 1995.

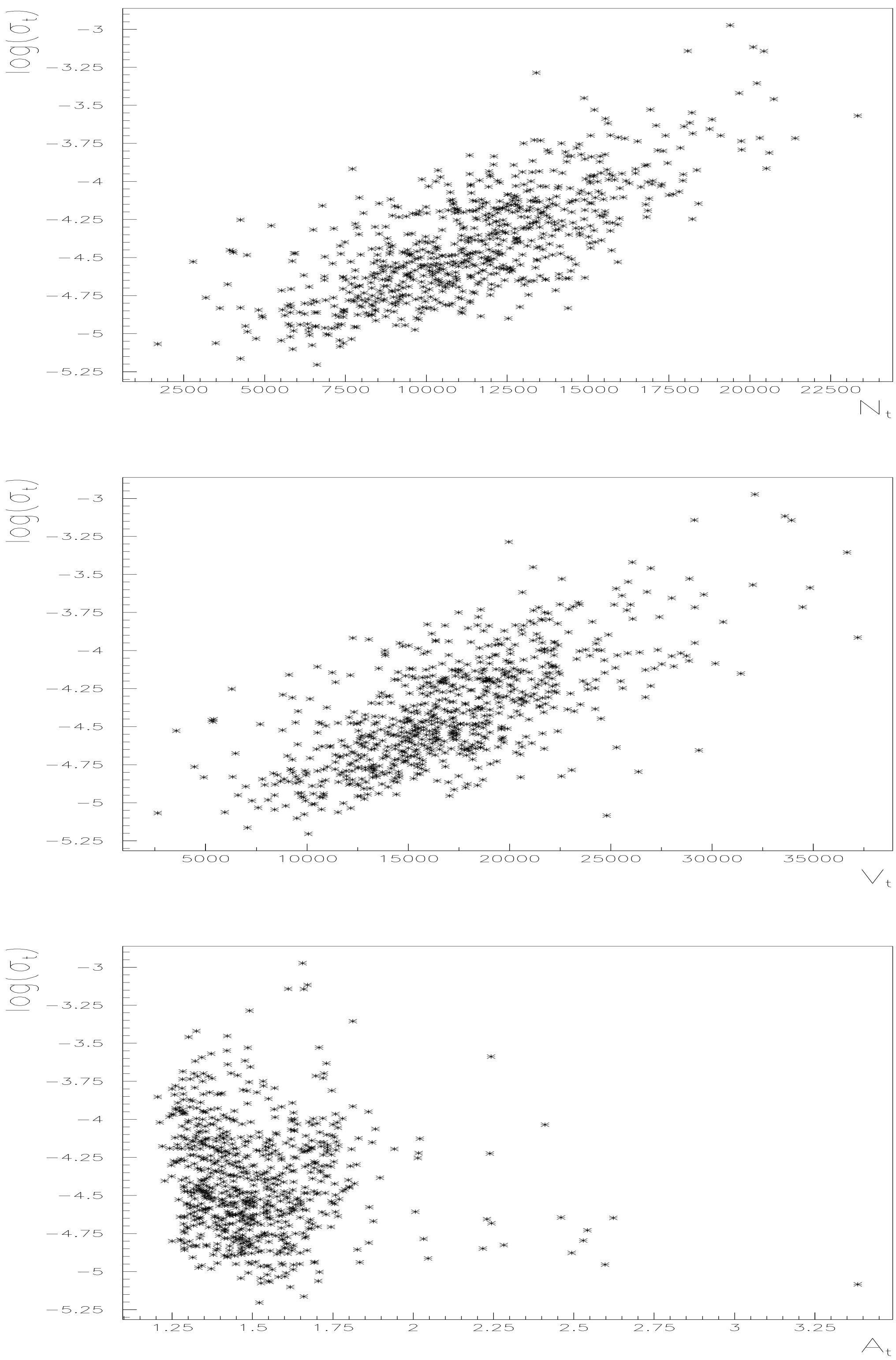

Fig. 4. Scatter-plot between daily logarithmic volatility and different measures of daily trading volume: number of transaction N_t (top), volume V_t (center) and average volume A_t (bottom).

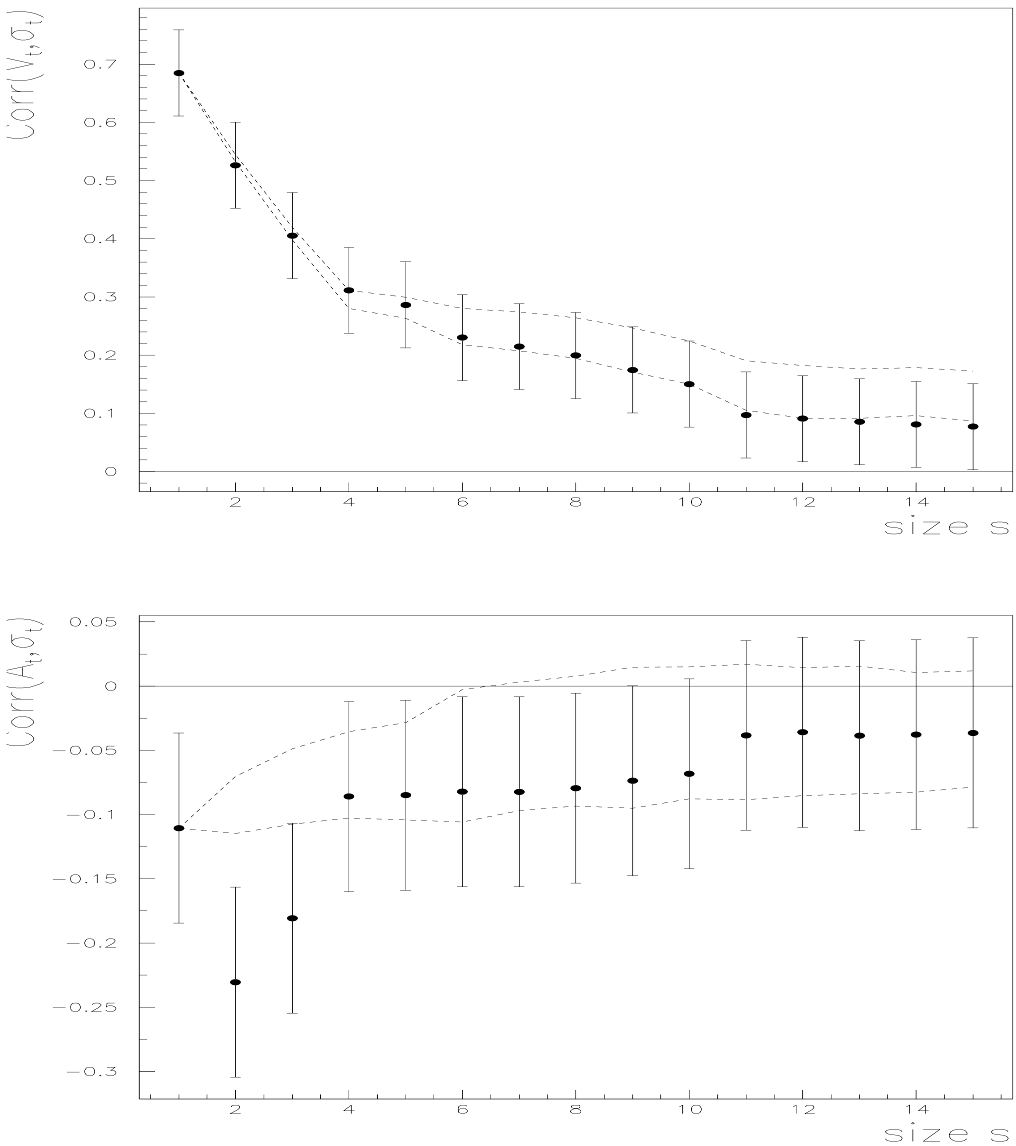

Fig. 5. Correlation of volatility $\log \sigma_t$ with V_t^s (top) and A_t^s (bottom) as a function of s. The dotted lines define the confidence interval at 95%, as computed with random subsamples with a number of transactions equal to actual transactions $\geq s$.

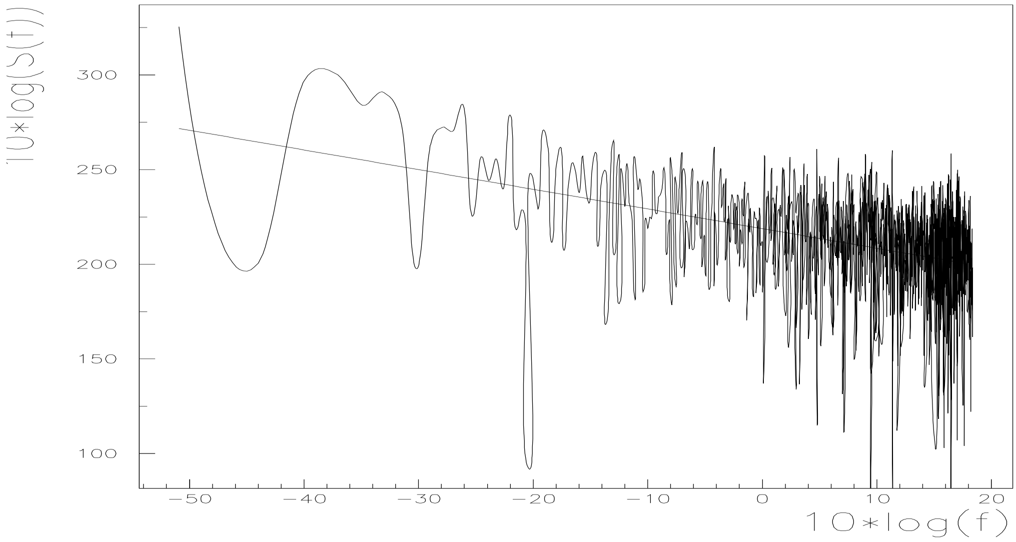

Fig. 6. Logarithm of power spectrum of daily total volume, as a function of the logarithm of the frequency. The straight line is the best linear fit by the least square method, drawn in the hypothesis of $S(f) \sim \frac{1}{f^\eta}$.

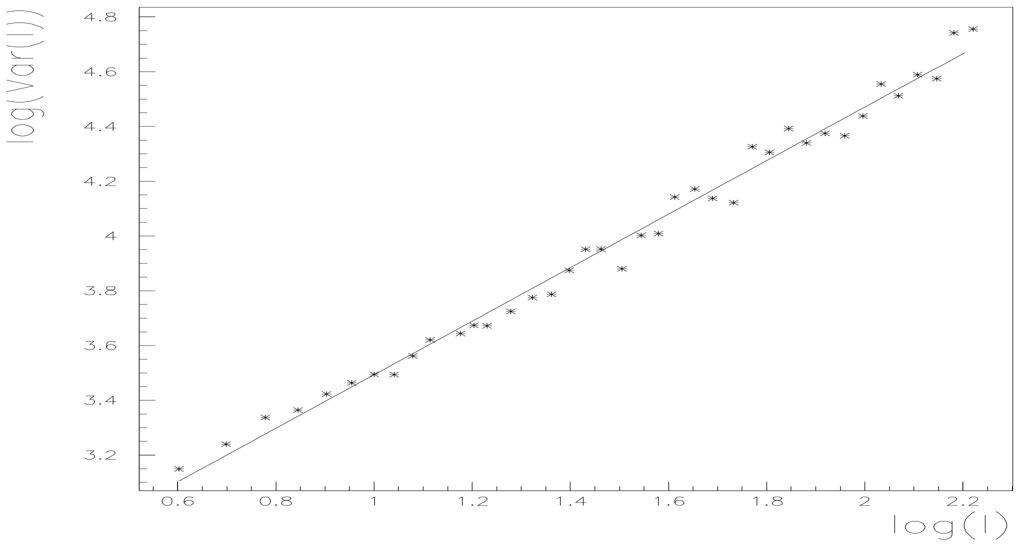

Fig. 7. In bilogarithmic scale, the scaling function $\sigma(l) \sim l^\alpha$ of the detrended variance for the daily total volume. The straight line is the best linear fit by the least square method; the slope coefficient gives an estimate of the scaling parameter.

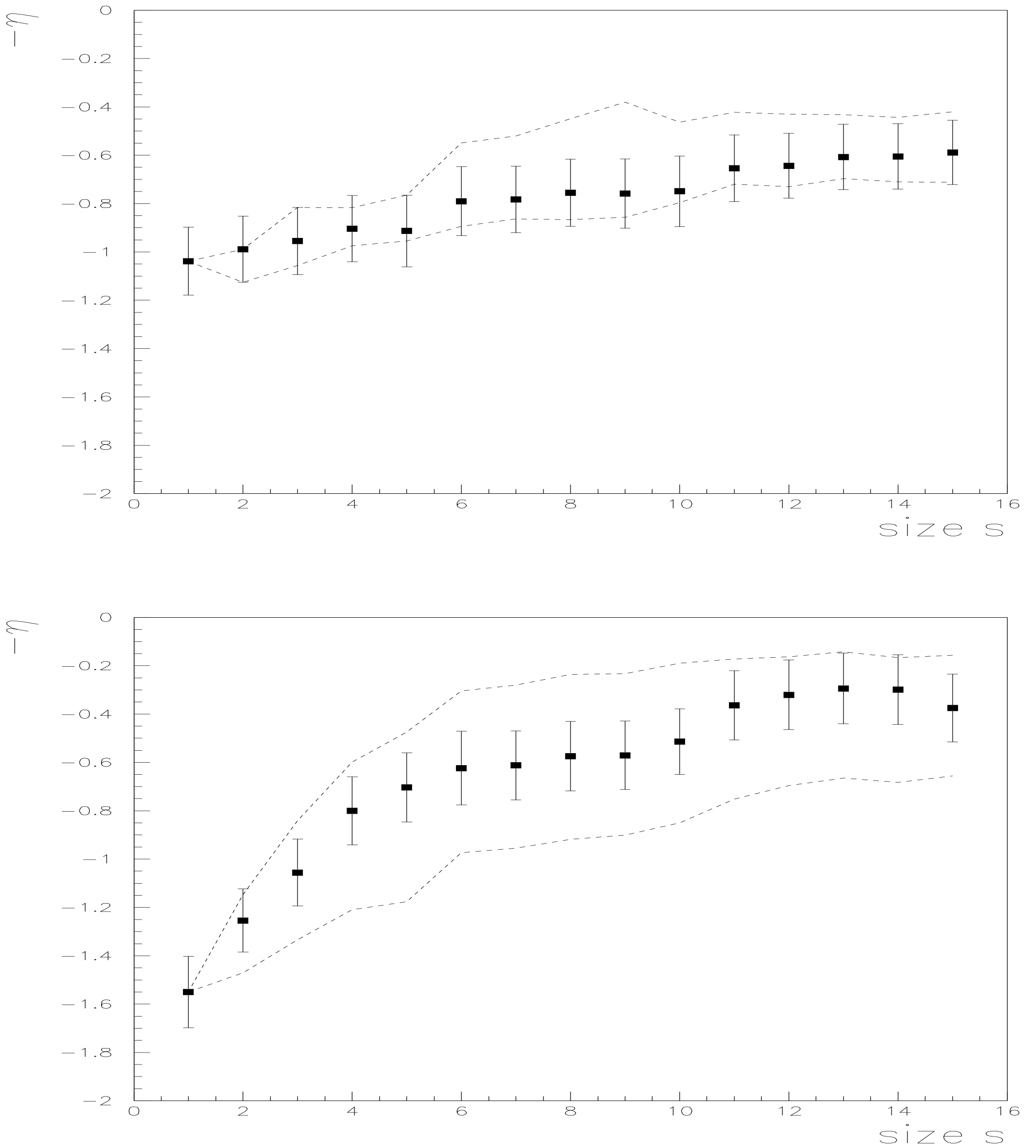

Fig. 8. Value of $-\eta$ from spectral analysis for V_t^s (top) and A_t^s (bottom) as a function of s. The dotted lines define the confidence interval at 95%, as computed with random subsamples with a number of transactions equal to actual transactions $\geq s$.

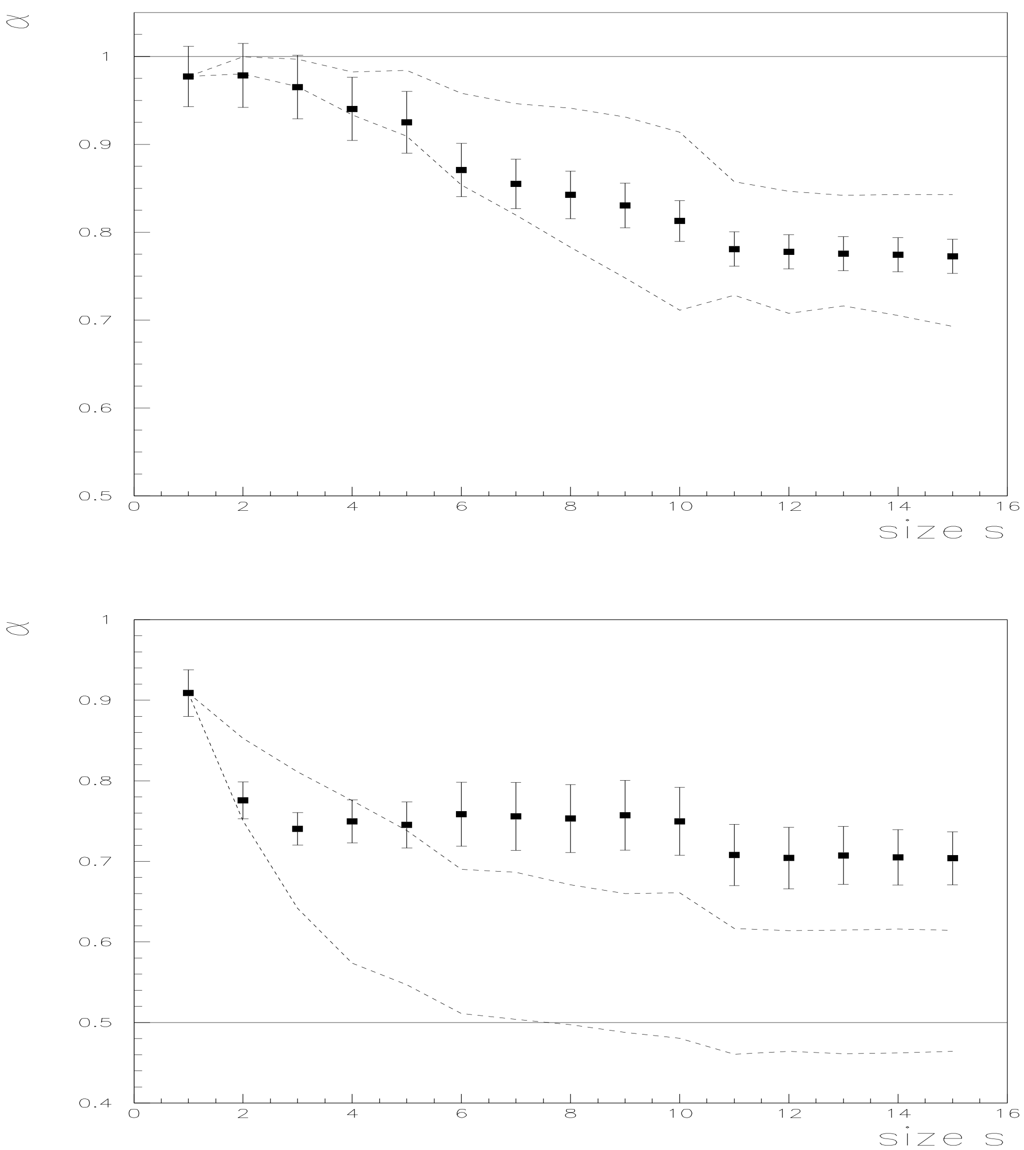

Fig. 9. Value of α from DFA for V_t^s (top) and A_t^s (bottom) as a function of s. The dotted lines define the confidence interval at 95%, as computed with random subsamples with a number of transactions equal to actual transactions $\geq s$.