UDT: 336.764/.768(430)
JEL Classification: G14, G35
Keywords: DAX 30; trading volume; univariate and bivariate long memory

Long Memory on the German Stock Exchange

Henryk GURGUL* – Tomasz WÓJTOWICZ**

1. Introduction

Long memory, also known as the long-term dependence property, describes the high-order correlation structure of a series. If a time series possesses long memory, there is a persistent temporal dependence between observations even considerably separated in time. The autocorrelation function (ACF) of series with long memory tails off hyperbolically. These series exhibit low-frequency spectral distributions. In contrast to long memory, the short-memory property is characterized by the low order correlation structure of a series. It is no difficult task to recognize these types of time series because they exhibit quickly declining autocorrelations and, in the spectral domain, demonstrate high-frequency distributions. It is clear that standard ARMA processes do not exhibit long memory. They can only exhibit short-run (high-frequency) properties.

The presence of long memory in financial data causes a number of both theoretical and empirical problems. The long-memory property is connected with nonlinearities in economic data. Martingale models of stock prices cannot follow from arbitrage, because new information cannot be entirely arbitraged away. The second problem in the case of long memory is pricing derivative securities with the martingale method. This method is usually not correct if the accompanying stochastic (continuous) processes exhibit long memory. Another problem concerns the standard testing procedures applied to asset pricing models. In the case of long memory in a series these procedures may not be relevant.

Some researchers raise doubts about the semi-strong market efficiency of stock markets where financial data exhibit long memory. This is because long memory is responsible for a nonlinear dependency in the first moment of the series distribution and therefore can be a reason for a time-series component which can be forecasted. In the case of the long-memory property even observations which are far apart can be significantly correlated. Therefore past returns can help to forecast future returns and speculative profits can be reached. This clearly violates the market efficiency assumption.

* Faculty of Management, University of Science and Technology, Krakow – corresponding author (h_gurgul@o2.pl)

** Faculty of Management, University of Science and Technology, Krakow

The present study concentrates on the role of trading volume in the process that generates stock returns and return volatilities on the German stock market, namely the stocks of companies listed in the DAX30 of *Deutsche Börse*. In contrary to most papers on this subject, we use individual stock data instead of index data.

The main goal of this paper is to show the existence of long memory in log-volume data and in the return volatility series of companies listed in the DAX30. In order to check the robustness of our results against the sample size we examine the log-volume and volatility properties of the largest German companies not only in the entire considered period from January 1994 to November 2005 but also in three subperiods: January 1994 to December 1997, January 1998 to December 2001 and January 2002 to November 2005.

The outline of the article is as follows: the most important contributions in the context of returns, return volatility, trading volume and finally long memory in finance are reviewed in the next section 2. Section 3 presents in detail the concept of long memory and estimation methods of the long-memory parameter. The data basis is characterized in Section 4, while empirical results are presented in Section 5. The final section 6 concludes the paper.

2. Trading Volume, Prices and Long Memory

In the past, researchers and investors concentrated their attention primarily on stock prices and their behavior over time. Taking into account a given set of information, stock prices reflect investors' expectations about the future development of a firm. Upcoming information is the main reason to expect changes in investors beliefs and therefore the main reason for price movements. There are situations when prices remain unaltered in spite of new, important upcoming information. This can occur when different investors interpret new information differently, or when they interpret new information identically but start from different initial expectations. As we see, changes in stock prices reflect the average of investors' beliefs caused by upcoming information. It is clear that stock price changes can be noticed if there is positive trading volume.

In applications there are several measures of trading activities for individual stocks:

a) number of trades per period,
b) share volume $X_{j,t}$,
c) value of shares traded (dollar volume) $P_{j,t}\,X_{j,t}$, where $P_{j,t}$ denotes price of j-th equity,
d) relative dollar volume $P_{j,t}X_{j,t} / \sum_j P_{j,t}X_{j,t}$,
e) share turnover (turnover ratio), i.e. the ratio of the number of shares traded and number of shares outstanding $\tau_{i,j} = X_{j,t} / N_{j,t}$,
f) dollar turnover $\nu_{j,t} = P_{j,t}X_{j,t} / P_{j,t}N_{j,t} = \tau_{j,t}$.

One can see that the last two measures are equal. The most common measures used in empirical investigations are given by b), c) and e). In order to measure aggregate trading activity similar measures can be defined.

In the literature an important question arises as to whether volume data are just a descriptive parameter of the trading process, or whether trading volume contains specific information that can be applied in modeling stock returns or return volatilities. As in the case of prices, volume and volume volatility depend on changes of the available set of relevant information on the market. In contrast to stock prices, a change of intraday expectations always leads to an enlargement of trading volume. Thus, trading volume incorporates the sum of investors decisions in reaction to news. This process increases trading volume. The differences between investors' reactions to the arrival of new information do not get lost as in the case of the averaging process that establishes prices. The joint observation of stock price behavior and trading volume enables us to determine the dynamic properties of stock markets and allows a better understanding of the impact of upcoming news on the market. The speculative motive, however, leads investors to trade even in the absence of new information. Volume data are regularly reported in the financial media together with price data.

Up to now, a considerable number of papers which examine in a theoretical framework the role of trading volume in return formation have been published. One of the first contributors to the subject was Clark (1973), who formulated the *Mixture of Distribution Hypothesis* (MDH). He claims that stock returns and trading volume are related because they are jointly dependent on an underlying latent information-flow variable. According to MDH, upcoming information is a source of price volatility. Clark suggested applying volume data as a proxy for the upcoming information stochastic process. The assumption of MDH implies strong positive contemporaneous (but not causal) linkages between volume and return volatility data, whereas return levels and volume data feature no interactions. The MDH hypothesis was extended, among others, by Andersen (1996), who argued that asymmetries and liquidity needs cause trading activities in response to the arrival of new information.

The second alternative hypothesis, known as the *sequential information flow model,* was formulated by Copeland (1976). He suggested that new information is disseminated sequentially rather than simultaneously to market participants. This results in a sequence of transitional price equilibriums that are accompanied by a persisting high trading volume. One important implication of Copeland's assumptions is the existence of positive contemporaneous as well as causal relations between price volatilities and trading activities.

In a framework where stock prices are assumed to follow *random walk*, some studies, e.g. (Blume et al., 1994) and (Suominen, 2001), try to prove the assumption that trading-volume data reveal unique information to the market, and that this information is not contained in prices. The Blume model assumes that informed traders transfer their private information to the market through trades, and uninformed traders learn from volume data about the precision and dispersion of an informational signal. Therefore return volatility and trading volume exhibit time persistence also in those cases when information arrivals do not. In a model by Suominen (2001), trading volume is used by uninformed traders as a signal of private information in the market and therefore it can help to overcome information asymmetries. It follows from these models that trading volume not only describes market behavior but also affects market development. The level of trading volume directly enters into the decision-making process of market participants. In this sense a strong relationship (contemporaneous as well as causal) between volume and return volatility can be expected.

In past decades, the hypotheses outlined above were confirmed by empirical studies concerning volume-price relations on capital markets. The relationship between trading volume and price changes, mainly using index data, was considered in contributions by Karpoff (1987), Hiemstra and Jones (1994), Brailsford (1996) and Lee and Rui (2002). Although these studies differ in detail, the contributors draw a common conclusion about a positive volume-price relationship. On the other hand, the relation between stock-return volatility and trading volume was the subject of contributions by Karpoff (1987), Bessembinder and Seguin (1993), Brock and LeBaron (1996), Avouyi-Dovi and Jondeau (2000) and Lee and Rui (2002). All these studies uniformly confirmed a strong relationship, contemporaneous and dynamic between return volatility and trading volume. The only exception is the study by Darrat et al. (2003) based on intraday data from DJIA stocks. The above-mentioned contributors find evidence that dynamic (causal) relations are significant. They neglected the contemporaneous correlation between return volatility and trading volume.

Lamoureux and Lastrapes (1990) were the first to apply stochastic time-series models of conditional heteroskedasticity (GARCH-type) to explore the contemporaneous relationship between volatility and volume data. The authors find that persistence in stock-return variance vanishes for the most part when trading volume is included in the conditional variance equation. If trading volume is considered to be an appropriate measure for the flow of information into the market, this finding is consistent with the MDH. However, one has to realize that the observation by Lamoureux and Lastrapes (1990) is mainly proof of the fact that trading volume and return volatility are driven by identical factors, leaving the question of the source of the joint process largely unresolved. This GARCH cum volume approach has been applied and extended in several studies, such as (Lamoureux – Lastrapes, 1994), (Andersen, 1996), (Brailsford, 1996), (Gallo – Pacini, 2000) and (Omran – McKenzie, 2000).

Contributions by Campbell et al. (1993) and McKenzie and Faff (2003) deal with linkages between trading volume and the autocorrelation pro-perties of daily stock returns. The authors established that trading volume is responsible for time-varying autocorrelations in stock returns. In the case of higher trading volume the contributors noticed a drop in return autocorrelation. Recently, Connolly and Stivers (2003) investigated the autocorrelation properties of stock returns in conjunction with abnormal turnover on a weekly basis. They found a contemporaneous dependence between stock returns and trading volume. Chordia and Swaminathan (2000) were concerned with the role of trading volume in the cross-autocorrelation patterns which stock returns exhibit. According to the contributors, returns of stocks with high trading volume precede returns of stocks with lower trading volume. This finding confirms the speed-of-adjustment hypothesis. According to this hypothesis high volume stocks adjust more quickly to new information than low volume stocks do. Chordia et al. (2001) found a negative cross-sectional relation between expected stock returns and both the level of and the changes in trading volume.

A persistence in autocorrelation can be observed in many financial time series. Loosely speaking this property is called long memory. The concept of long memory was introduced by the British hydrologist Hurst (1951). Early contributions to the subject of long memory in time series are those by Mandelbrot (1971) (which formalized Hurst's empirical findings using cumulative river-flow data), Geweke and Porter-Hudak (1983), and Hosking (1981). Granger and Joyeux (1980) introduced fractionally integrated ARMA models, which were more recently discussed by Sowell (1992), Beran (1992) and Baillie (1996).

The potential presence of long memory in financial data has been an important subject of both theoretical and empirical investigation by econometricians and finance researchers. A number of studies have focused on long memory (persistence) in financial asset returns. As we mentioned in the first section of this paper, the finding of long-term dependence in financial data might be in contradiction to the Efficient Markets Hypothesis of Fama (1970), which is based on the assumption of martingale behavior of financial market prices. The martingale theory requires an invariant stationarity and an independence from any innovations of historical price information sets, but it is difficult to show that this requirement is fulfilled either in a weak form or, even less so, in a semi-strong or strong form. The theory of the Fractional Market formulated by Peters (1994) is an application of the long-term dependence concept. This concept is more general than Fama's understanding of efficiency. The first contribution to this subject in finance is that by Greene and Fielitz (1977) who, by means of the rescaled-range (R/S) method of Hurst, found long memory in daily equity returns. This result was rejected by Lo (1991), who applied a more adequate form of the R/S method. Also, in the subsequent contributions by Crato (1994), Cheung et al. (1993), Cheung and Lai (1995), Barkoulas and Baum (1996) the presence of long memory in finance data could not be significantly confirmed. Beveridge and Oickle (1997) investigate long-memory dependence in Canadian daily stock returns using ARIMA models and find long-memory mean reversion.

In parallel, spot and futures foreign-exchange rates and commodity prices were investigated with respect to long memory. Contributions by Helms et al. (1984), Cheung and Lai (1993), Fang et al. (1994), and Barkoulas et al. (1997) confirm long memory in the above-mentioned kinds of foreign-currency rates.

In recent years researchers have come back to stock markets and started to investigate volatility (absolute values of returns or squared returns) and more recently trading volume, also with respect to long memory and bivariate long memory.

Estimation results by Bollerslev and Mikkelsen (1996) provide new evidence that the apparent long-run dependence in US stock market volatility is best described by a mean-reverting fractionally integrated process, so that a shock to the optimal forecast of the future conditional variance dissipates at a low hyperbolic rate.

Granger and Zhuanxin (1996) illustrate the relevance of long memory using returns from a daily stock-market index. The authors also point out that a number of other processes like generalized fractionally integrated models resulting from aggregation, time-changing coefficient models, and possibly nonlinear models can be long memory.

Koop et al. (1997) provide a Bayesian analysis of ARFIMA models and describe a test of ARFIMA against ARIMA alternatives.

Bollerslev and Jubinski (1999) examined the behavior of stock-trading volume and volatility for the individual firms composing the Standard & Poor 100 composite index. In line with the MDH hypothesis, they found that long-run hyperbolic decay rates appear to be common across each volume-volatility pair. In addition, they also established that fractionally integrated processes best describe the long-run temporal dependencies in volume and volatility series.

Lobato and Velasco (2000) investigated the properties of 30 equities in the DJIA with respect to long memory. They found that trading volume exhibits long memory, and that volatility and volume exhibit the same degree of long memory for most of the stocks. However, the contributors did not find a common long memory component for both processes.

Analogously to both the above studies, we use in our contribution individual stock data instead of index data.

In the next section we explain in detail the notion of long-memory parameter d.

3. Long Memory Estimators

A covariance stationary stochastic process exhibits long memory with memory parameter d when its spectral density function $f(\lambda)$ satisfies:

$$f(\lambda) \sim c\lambda^{-2d} \text{ as } \lambda \to 0^{+} \quad (1)$$

where c is a finite positive constant and the symbol "$\sim$" means that the ratio of the left- and right-hand sides tends to one at the limit. When the process satisfies condition (1) and $d > 0$ its autocorrelation function dies out at a hyperbolic rate (Granger – Joyeux, 1980), (Hosking, 1981), (Beran, 1994), i.e.

$$\rho_k \sim c_\rho k^{2d-1} \text{ as } k \to \infty$$

The parameter d determines the memory of the process. If $d > 0$ the spectral density is unbounded near the origin, and the process exhibits long memory. If $d = 0$ the spectral density is bounded at 0 and the process is called short memory. When $d < 0$ the spectral density is zero at the origin and the process is called anti-persistent and displays negative memory.

The most well-known class of long-memory processes satisfying (1) is the class of autoregressive fractionally integrated moving average (ARFIMA) processes introduced into econometrics by Granger and Joyeux (1980).

We say that x_t is an ARFIMA(p, d, q) process if:

$$\Phi(B)(1 - B)^d (x_t - \mu) = \Theta(B)\varepsilon_t$$

where $\Phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$ and $\Theta(z) = 1 - \theta_1 z - ... - \theta_q z^q$ are lag polynomials of order p and q respectively, in the backshift operator B with roots outside the unit circle, ε_t is iid $(0,\sigma^2)$, and $(1-B)^d$ is defined by binomial expansion:

$$(1-B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\,\Gamma(j+1)} B^j$$

In addition to the previously mentioned properties of memory, if $d > -0.5$ the ARFIMA process is invertible and possesses linear Wold representation and if $d < 0.5$ it is covariance stationary. Thus, if $0 < d < 0.5$ the process is stationary and exhibits long memory. Many non-stationary series can be transformed by integer integrating into stationary ones with spectral density satisfying (1).

There are several methods for the estimation of long-memory parameter d. We will review them briefly in the following subsections.

3.1 Maximum Likelihood Estimator

Maximum likelihood estimation (MLE) in the time domain needs an assumption about the exact form of the estimated ARFIMA model. Then the exact Gaussian likelihood function for the given sample $\{x_t\}_{t=1...T}$ is:

$$L(d, \phi, \theta, \sigma^2, \mu) = -\frac{T}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{x} - \mu\boldsymbol{1})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \mu\boldsymbol{1}) \qquad (2)$$

where $\boldsymbol{1} = (1,...,1)^T$, $\boldsymbol{x} = (x_1,...,x_T)^T$, φ and θ are the parameters of autoregression and moving average polynomials respectively, μ is the mean of the process, and $\boldsymbol{\Sigma}$ is its covariance matrix. Sowell (1992) proved that the exact maximum likelihood estimator (EML) obtained by maximising the likelihood function (2) is consistent and asymptotically normal, i.e.

$$\hat{d}_{EML} \sim N\left[d, (\pi^2 T / 6 - \mathrm{c})^{-1}\right]$$

where $c = 0$ when $p = q = 0$ and $c > 0$ otherwise.

Other properties of MLE and methods of solving some computational problems are discussed in (Sowell, 1992) and (Doornik – Ooms, 2003).

There are several modifications of exact maximum likelihood estimation, e.g. modified profile likelihood (see (Cox – Reid, 1987)) or conditional maximum likelihood (see (Tanaka, 1999)). The main drawback of such maximum likelihood estimators is their sensitivity to any model misspecification, and thus they can by easily influenced by any short-run dynamics.

3.2 GPH Estimator

Another class of estimators of the long-memory parameter d are semiparametric estimators based on the approximation (1) of the spectral density function near the origin. Among them the most popular is the log-periodogram regression method originally developed by Geweke and

Porter-Hudak (1983) and analyzed in detail by Robinson (1995a). Semiparametric estimators use only information from the periodogram for very low frequencies. Thus they are robust to short-run dynamics. Based on condition (1), after taking the logarithms and inserting sample quantities, the long-memory estimator is computed from the approximate regression relationship:

$$\ln\left[I(\lambda_j)\right] \approx const - 2d \ln(\lambda_j)$$

where $\lambda_j = \frac{2\pi j}{T}$ are the Fourier frequencies and $I(\lambda) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} x_t e^{it\lambda}\right|^2$ is the periodogram of the given sample $x_1,...,x_T$. The GPH estimator is then defined as the OLS estimator in the above regression using only $j = 1,..., m$ its first values, where $m = m(T)$ is a bandwidth parameter satisfying condition:

$$\frac{1}{m} + \frac{m}{T} \to 0 \text{ as } T \to \infty$$

Geweke and Porter-Hudak originally suggested choosing m equal to $\sqrt{T}$. For further considerations about the optimal bandwidth see (Hurvich et al., 1998) and (Henry – Robinson, 1996). The asymptotical normality of the GPH estimator was initially proved by Robinson (1995a) for $d \in (-1/2, 1/2)$, but recently Velasco (1999a) showed that it is consistent for $d \in (-1/2, 1)$ and has an asymptotically normal limit distribution for $d \in (-1/2, 3/4)$:

$$\hat{d}_{GPH} \sim N\left(d, \frac{\pi^2}{24m}\right)$$

There are several modifications of the GPH estimator. For example, Agiakloglou et al. (1993) suggested replacing the constant in the regression by the polynomial in order to reduce bias (see also (Andrews – Guggenberger, 2003)). Similarly, an estimator that allows a short-run component was proposed by Shimotsu and Phillips (2002a).

The univariate GPH estimator described above can be generalized for a multivariate case. Consider $\boldsymbol{x_t} = (x_{1,t}, ..., x_{N,t})^T$ a covariance stationary N-dimensional vector process with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Gamma}_j$ at lag j and a fractional integration vector $(d_1,...,d_N)^T$, i.e. each $x_{i,t}$ is integrated of order d_i. For any a, $b = 1,..., N$ and $\lambda_j = \frac{2\pi j}{T}$ define the crossperiodogram of the process $\boldsymbol{x_t}$:

$$\boldsymbol{I}_{ab}(\lambda) = \left(\frac{1}{(2\pi T)^{1/2}} \sum_{t=1}^{T} x_{a,t} e^{it\lambda}\right)\left(\frac{1}{(2\pi T)^{1/2}} \sum_{t=1}^{T} x_{b,t} e^{it\lambda}\right)^*$$

where the asterisk means complex conjugation. For a bandwidth parameter m we define $\boldsymbol{Y}_{kj} = \ln[\boldsymbol{I}_{kk}(\lambda_j)]$, $k = 1,..., N$, $j = 1,..., m$. Then the multivariate GPH estimator of fractional integration d_k is given by:

$$\hat{d}_k = -\frac{\sum_{j=1}^{m} \nu_j Y_{kj}}{2\sum_{j=1}^{m} \nu_j^2} \quad \text{where} \quad \nu_j = \ln\lambda_j - \frac{1}{m}\sum_{j=1}^{m} \ln\lambda_j \qquad (3)$$

For individual series $\{x_{j,t}\}_{t=1,..,T}$ this estimator is equivalent to the univariate GPH estimator previously described. Based on its asymptotic normality (Robinson (1995a)) a Wald-type test for the null hypothesis:

$$H_0 : \boldsymbol{Pd} = \boldsymbol{\rho}$$

for a $u \times N$ matrix $\boldsymbol{P}$ and $u \times 1$ vector $\boldsymbol{\rho}$ can be constructed. The test statistic:

$$4m(\boldsymbol{P}\hat{\boldsymbol{d}} - \boldsymbol{\rho})^T (\boldsymbol{P}\hat{\boldsymbol{\Omega}}\boldsymbol{P}^T)^{-1} (\boldsymbol{P}\hat{\boldsymbol{d}} - \boldsymbol{\rho})$$

has the limiting χ^2_u distribution, where $\hat{\boldsymbol{d}} = (\hat{d}_1, ..., \hat{d}_N)^T$ and $\hat{\boldsymbol{\Omega}}$ is a consistent estimate of the limiting variance of $2\sqrt{m}\,(\hat{\boldsymbol{d}} - \boldsymbol{d})$ [see (Robinson, 1995a)]. In the case of testing for a common long-memory parameter of the process, $\boldsymbol{\rho}$ is a vector of zeroes and $\boldsymbol{P} = (\boldsymbol{I}_{N-1} : \boldsymbol{0}) - (\boldsymbol{0} : \boldsymbol{I}_{N-1})$ is a $(N-1) \times N$ matrix, where $\boldsymbol{I}_{N-1}$ is the identity matrix of dimension $N-1$.

When the existence of a common order of integration d is assumed, the restricted least square estimator is given by:

$$\hat{d} = -\frac{1}{2} \frac{\sum_{j=1}^{m} \boldsymbol{1}_N^T \hat{\boldsymbol{\Omega}}^{-1} \boldsymbol{Y}_j \nu_j}{\boldsymbol{1}_N^T \hat{\boldsymbol{\Omega}}^{-1} \boldsymbol{1}_N \sum_{j=1}^{m} \nu_j^2} \tag{4}$$

where $\boldsymbol{Y}_j = (Y_{1j},..., Y_{Nj})^T$ and $\boldsymbol{1}_N$ is a $N \times 1$ vector of ones. Like the unrestricted estimates, the $\hat{d}$ is asymptotically normally distributed.

3.3 Whittle Estimator

Another class of semiparametric estimators are the narrow-band Gaussian or local Whittle estimators introduced by Künsch (1987), and developed by Robinson (1995b) and Lobato (1999). In the univariate case it is defined as a maximiser of the likelihood function:

$$Q(g,d) = -\frac{1}{m} \sum_{j=1}^{m} \left[\ln(g\lambda_j^{-2d}) + \frac{I(\lambda_j)}{g\lambda_j^{-2d}} \right] \tag{5}$$

The ranges of consistency and asymptotic normality of the local Whittle estimator are the same as those for the GPH estimator (see (Velasco, 1999b) and (Phillips – Shimotsu, 2004)) but the Whittle estimator is more efficient because asymptotically:

$$\hat{d}_{LW} \sim N\left(d, \frac{1}{4m}\right)$$

For further modifications of the local Whittle estimator, see for example (Shimotsu – Phillips, 2002b) or (Andrews – Sun, 2004).

As with the GPH estimator, the local Whittle estimator can be defined in the multivariate case. The corresponding (concentrating) likelihood function is:

$$Q(\boldsymbol{d}) = -\frac{2}{m}\sum_{i=1}^{N} d_i \sum_{j=1}^{m} \ln\lambda_j + \ln\left|\hat{\boldsymbol{R}}(\boldsymbol{d})\right| \quad (6)$$

where

$$\hat{\boldsymbol{R}}(\boldsymbol{d}) = \frac{1}{m}\sum_{j=1}^{m} \boldsymbol{\Lambda}_j \,\mathrm{Re}\left\{\boldsymbol{I}(\lambda_j)\right\}\boldsymbol{\Lambda}_j \quad (7)$$

with $\boldsymbol{\Lambda}_j = diag(\lambda_j^{d_1}, ..., \lambda_j^{d_N})$ and a crossperiodogram matrix $\boldsymbol{I}(\lambda)$. The estimator $\hat{\boldsymbol{d}} = (\hat{d}_1, ..., \hat{d}_N)^T$ is defined as a maximizer of the concentrating likelihood function (6). It can be computed in two ways: by numerical maximizing of (6) or using the two-step procedure proposed by Lobato (1999). The first step is to compute the univariate QMLE for every series (denoted that vector by $\hat{\boldsymbol{d}}^{(1)}$) and the second step is to compute the following expression:

$$\hat{\boldsymbol{d}}^{(2)} = \hat{\boldsymbol{d}}^{(1)} - \left(\frac{\partial^2 Q(\boldsymbol{d})}{\partial \boldsymbol{d} \partial \boldsymbol{d}^T}\Big|_{\hat{\boldsymbol{d}}^{(1)}}\right)^{-1} \left(\frac{\partial Q(\boldsymbol{d})}{\partial \boldsymbol{d}}\Big|_{\hat{\boldsymbol{d}}^{(1)}}\right)$$

As showed by Lobato (1999) the above two-step estimator has the same asymptotic distribution as the QMLE based on the equation (6), but it is straightforward to calculate. Under the reasonable assumption this estimator is normally distributed with parameters:

$$\hat{\boldsymbol{d}}^{(2)} \sim N\left(\boldsymbol{d}, \frac{1}{\sqrt{m}} \boldsymbol{E}^{-1}\right)$$

where $\boldsymbol{E} = 2(\boldsymbol{I}_N + \boldsymbol{R} \circ \boldsymbol{R}^{-1})$ and $\circ$ denotes the Hadamard product of two matrices.

Based on these asymptotic properties a test for the null hypothesis of a linear set of restrictions on d is available. Consider $\boldsymbol{P}$ which is $q \times N$ matrix, $q \times 1$ vector $\boldsymbol{\rho}$ and the null hypothesis

$$H_0 : \boldsymbol{Pd} = \boldsymbol{\rho}$$

Then the test statistic:

$$m(\boldsymbol{P}\hat{\boldsymbol{d}}^{(2)} - \boldsymbol{\rho})^T (\boldsymbol{P}\hat{\boldsymbol{E}}^{-1} - \boldsymbol{P}^T)^{-1} (\boldsymbol{P}\hat{\boldsymbol{d}}^{(2)} - \boldsymbol{\rho})$$

is asymptotically χ_q^2 distributed under the null hypothesis. It allows testing for a common long-memory parameter. In this case $\boldsymbol{\rho}$ is a vector of zeroes and $\boldsymbol{P} = (\boldsymbol{I}_{N-1} \vdots \boldsymbol{0}) - (\boldsymbol{0} \vdots \boldsymbol{I}_{N-1})$ is a $(N-1) \times N$ matrix. On the other hand, it allows testing if the vector process is $I(0)$ or $I(1)$. In this case $\boldsymbol{P} = \boldsymbol{I}_N$ and $\boldsymbol{\rho}$ is $q \times 1$ vector of zeroes or ones, respectively.

If the existence of a common order of integration is assumed, the estimator of d_* can be computed by maximizing the likelihood function

$$Q_*(d) = -\frac{2Nd}{m}\sum_{j=1}^{m} \ln\lambda_j + \ln\left|\hat{\boldsymbol{R}}(\boldsymbol{d}\boldsymbol{1}_N)\right| \quad (8)$$

The resulting QMLE $\hat{d}_*$ is asymptotically normally distributed:

$$\hat{d}_* \sim N(d_*, \frac{1}{4Nm})$$

3.4 Fractional Cointegration

We will consider the special but simplest case of the cointegration of two processes. Several definitions of fractional cointegration can be found in the literature (see (Robinson – Yajima, 2002)). The most common definition is as follows. We say that two fractionally integrated series x_t and y_t are cointegrated of order d if:

– x_t and y_t share the same long memory, i.e. $d_x = d_y$;
– there exists a constant β such that process $\varepsilon_t = y_t - \beta x_t$ has long-memory parameter $d < d_y$.

Estimation of parameter β can be done by means of the frequency domain least squares (FDLS) method. Based on the definition of the crossperiodogram from Subsection 3.2., for any Fourier frequencies λ_j define the averaged crossperiodogram:

$$\hat{\boldsymbol{F}}_{ab}(m) = 2\text{Re}\left\{\frac{2\pi}{T}\sum_{j=1}^{m}\boldsymbol{I}_{ab}(\lambda_j)\right\}$$

where $m < \left\lfloor \frac{T}{2} \right\rfloor$ is the bandwidth parameter. Then the FDLS estimate of β is given by:

$$\hat{\beta} = \hat{\boldsymbol{F}}_{xy}(m)\,\hat{\boldsymbol{F}}_{xx}^{-1}(m)$$

assuming that the inverse exists. The regularity conditions and asymptotic properties of estimates of β were considered in (Robinson – Marinucci, 2001).

Alternatively, based on local Whittle estimation methods, another way to check the existence of fractional cointegration between x_t and y_t is to test the necessary condition that the coherency between both series is 1 at zero frequency. Given the estimates matrix $\hat{\boldsymbol{R}}$, the squared coherency estimate is expressed by:

$$\hat{H}_{xy}^2(0) = \frac{\hat{\boldsymbol{R}}_{xy}^2}{\hat{\boldsymbol{R}}_{xx}\hat{\boldsymbol{R}}_{yy}}$$

4. Data Description

The data consist of the absolute daily rates of return, squared return and the natural logarithms of the trading-volume series for 22 companies listed in the DAX in the whole period from January 1994 to November 2005. For each firm calculations exclusively concentrate on the period of its DAX membership. Therefore, it was possible to extract 22 companies over the whole above-mentioned period. All time series were derived from Reuters. Con-

TABLE 1 Aggregated Summary Statistics for Stock Market Data of DAX Companies

Panel A: Daily Absolute Stock Returns

	01. 1994–11. 2005				01. 1994–12. 1997				01. 1998–12. 2001				01. 2002–11. 2005			
	Mean	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**
Min	0.013	0.012	1.85	7.97	0.009	0.009	1.53	5.46	0.014	0.013	1.34	5.08	0.011	0.010	1.50	6.2
1st Quartile	0.014	0.014	2.15	10.5	0.010	0.010	1.67	7.08	0.016	0.015	1.72	7.02	0.013	0.013	1.93	7.6
Median	0.015	0.015	2.31	11.5	0.011	0.010	2.07	9.29	0.018	0.017	1.95	8.94	0.016	0.015	2.06	8.6
3rd Quartile	0.016	0.016	2.44	13.7	0.012	0.011	2.44	14.6	0.020	0.018	2.32	12.6	0.017	0.019	2.41	12.6
Max	0.022	0.023	5.17	75.7	0.016	0.018	5.23	60.5	0.034	0.027	3.12	28.5	0.024	0.023	5.65	73.3

Panel B: Daily Squared Stock Returns

	01. 1994–11. 2005				01. 1994–12. 1997				01. 1998–12. 2001				01. 2002–11. 2005			
	Mean	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**
Min	0.0003	0.0007	5.49	46.78	0.0002	0.0003	3.41	17.42	0.0004	0.001	3.53	20.19	0.0002	0.0004	3.651	19.1
1st Quartile	0.0004	0.0010	6.63	76.42	0.0002	0.0004	4.78	35.45	0.0005	0.001	4.68	34.90	0.0003	0.0008	4.336	27.8
Median	0.0005	0.0012	8.00	102.6	0.0002	0.0005	5.81	51.87	0.0006	0.001	6.12	60.51	0.0005	0.001	5.208	40.8
3rd Quartile	0.0005	0.0014	9.98	179.2	0.0003	0.0006	8.71	124.9	0.0007	0.002	8.32	98.0	0.0006	0.0017	8.372	116
Max	0.0010	0.0029	37.2	1695	0.0006	0.0029	19.9	438.3	0.0018	0.003	19.6	507	0.0011	0.0037	25.57	740

Panel C: Daily Log-volume

	01. 1994–11. 2005				01. 1994–12. 1997				01. 1998–12. 2001				01. 2002–11. 2005			
	Mean	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**	**Mean**	**SD**	**Skew-ness**	**Kurtosis**
Min	11.9	0.56	–0.91	2.0	10.9	0.48	–0.87	2.8	11.9	0.45	–0.22	2.4	12.57	0.38	–0.21	3.44
1st Quartile	13.2	0.65	–0.35	2.5	12.1	0.50	–0.47	3.4	13.2	0.51	0.43	4 .1	13.56	0.41	0.05	3.82
Median	13.9	0.79	–0.13	3.1	12.9	0.60	–0.22	3.8	14.0	0.55	0.58	5.0	14.68	0.47	0.22	4.22
3rd Quartile	14.4	1.06	–0.04	3.6	14.1	0.70	–0.15	4.0	14.5	0.58	0.75	5 .6	15.16	0.53	0.42	4.96
Max	15.9	1.57	0.24	5.9	14.7	0.95	0.44	12.7	15.4	0.89	0.97	9.5	16.83	0.57	0.76	6.63

TABLE 2 Univariate Local Whittle Estimates of the Long Memory Parameter of Absolute Returns, Squared Returns and Log-volume in the Whole Period from January 1994 to November 2005 (in bold) and in Each of the Three Subperiods

	Absolute Returns				Squared Returns				Log-volume			
	01.94 –11.05	**01.94 –12.97**	**01.98 –12.01**	**01.02 –11.05**	**01.94 –11.05**	**01.94 –12.97**	**01.98 –12.01**	**01.02 –11.05**	**01.94 –11.05**	**01.94 –12.97**	**01.98 –12.01**	**01.02 –11.05**
Allianz	**0.44**	0.35	0.29	0.48	**0.38**	0.30	0.23	0.53	**0.66**	0.29	0.49	0. 55
BASF	**0.34**	0.25	0.25	0.38	**0.25**	0.17	0.22	0.38	**0.46**	0.35	0.39	0.32
Bayer	**0.34**	0.27	0.25	0.32	**0.20**	0.25	0.15	0.14	**0.47**	0.35	0.43	0.29
BMW	**0.42**	0.38	0.30	0.42	**0.40**	0.33	0.32	0.43	**0.45**	0.38	0.39	0.30
CommB	**0.42**	0.38	0.37	0.40	**0.38**	0.29	0.41	0.36	**0.49**	0.38	0.28	0.40
Daimler	**0.43**	0.29	0.33	0.43	**0.35**	0.21	0.34	0.46	**0.53**	0.45	0.44	0. 36
DBank	**0.42**	0.34	0.39	0.47	**0.38**	0.25	0.36	0.45	**0.42**	0.40	0.39	0.34
DTelekom	**0.45**	0.35	0.35	0.45	**0.41**	0.29	0.27	0.45	**0.65**	0.34	0.62	0.31
Henkel	**0.38**	0.31	0.30	0.32	**0.32**	0.30	0.24	0.26	**0.53**	0.48	0.35	0.22
HVB	**0.45**	0.29	0.45	0.45	**0.44**	0.19	0.46	0.48	**0.65**	0.38	0.54	0.43
Linde	**0.39**	0.37	0.31	0.32	**0.33**	0.34	0.24	0.27	**0.57**	0.40	0.36	0.18
Lufthansa	**0.35**	0.26	0.41	0.36	**0.32**	0.22	0.51	0.27	**0.56**	0.42	0.36	0.31
MAN	**0.36**	0.15	0.23	0.38	**0.37**	0.17	0.21	0.40	**0.54**	0.34	0.26	0.44
Metro	**0.35**	0.26	0.20	0.38	**0.30**	0.16	0.20	0.30	**0.54**	0.35	0.38	0.26
MuRe	**0.45**	0.32	0.36	0.46	**0.42**	0.24	0.28	0.41	**0.66**	0.28	0.53	0.50
RWE	**0.40**	0.31	0.29	0.38	**0.35**	0.26	0.23	0.38	**0.57**	0.22	0.47	0.36
SAP	**0.36**	0.14	0.30	0.41	**0.25**	0.15	0.22	0.31	**0.65**	0.41	0.64	0.30
Schering	**0.30**	0.29	0.20	0.34	**0.23**	0.21	0.19	0.21	**0.40**	0.35	0.23	0.25
Siemens	**0.46**	0.40	0.32	0.45	**0.40**	0.32	0.25	0.40	**0.43**	0.32	0.44	0.34
Thyssen	**0.35**	0.21	0.23	0.34	**0.24**	0.16	0.11	0.29	**0.40**	0.24	0.23	0.34
TUI	**0.39**	0.25	0.28	0.38	**0.36**	0.22	0.35	0.31	**0.60**	0.41	0.28	0.57
VW	**0.44**	0.27	0.37	0.43	**0.42**	0.21	0.41	0.40	**0.36**	0.26	0.33	0.39

tinuously compounded stock returns were calculated from daily closing prices, adjusted for the effects of dividend payouts and stock splits. In order to examine the robustness of estimated parameters across time and with respect to sample size, the whole period has been divided into three subperiods: from January 1994 to December 1997, from January 1998 to December 2001 and from January 2002 to November 2005.

As can be seen from panels A–C of *Table,* the stylized fact of a 'fat-tailed and high-peaked' distribution, widely reported for return-volatility series, is mostly present in our data, especially in case of squared daily returns. The median of stock squared return kurtosis in the whole period is 102.6 and ranges from 46.78 (MAN) to 1694 (Bayer). Positive skewness additionally confirms non-normality of returns volatility. Unlike returns volatility, trading volume displays rather negative skewness in the whole period. However, in second and third subperiods a change in sign of skewness can be observed.

5. Empirical Results

Long-memory properties of volatility (measured either as absolute or squared returns) and log-volume series were examined by means of the methods described in Section 3.

As a first step, the individual long-memory parameter of each series was estimated. In order to examine the robustness of the estimates, three different values of bandwidth m were considered ($T^{0.5}$, $1.5*T^{0.5}$ and $T^{0.65}$) and several long-memory estimation methods were applied: GPH log-periodogram regression and local Whittle estimation together with their modifications: bias reduced log-periodogram regression (Andrews – Guggenberger, 2003), pooled log-periodogram regression (Shimotsu – Philips, 2002a), exact local Whittle estimation (Shimotsu – Philips, 2002b) and local polynomial Whittle estimation (Andrews – Sun, 2002). For fixed m all the methods gave similar results but when the bandwidth m changed, long-memory estimates differed slightly: the smaller m, the higher estimates values. Despite these differences in values, the same conclusions can be drawn for different m because relationships of long-memory estimates across subsamples remained unchanged. Due to the lack of space only results of local Whittle estimation are presented in *Table 2*. The other results can be supplied to the reader upon request.

It is no surprise that univariate long-memory estimators for two different measures of volatility, i.e. absolute and squared returns, are quite similar. Moreover, in all but two presented cases volatility long-memory estimates are in the stationary region, i.e. $-0.5 < d < 0.5$. This is in line with previous results in the literature ((Ding et al., 1993), (Bollerslev – Mikkelsen, 1996), (Baillie et al., 1996), (Lobato – Savin, 1998)).

When we focus our attention on the behavior of long memory of volatility series in different subperiods, it turns out that for majority of samples estimates increase, i.e. in most cases value of memory parameter increases from one subperiod to another (it is particularly visible between the second and third subperiods) and reaches its maxima in the most recent subperiod: in 20 cases of absolute returns and 15 cases of squared returns.

In contrast to volatility long-memory estimates, a large number of long-memory estimates of log-volume lie in the nonstationary region $d > 0.5$. This is again in line with results from the literature (Gallant et al., 1992), (Andersen, 1996), (Bollerslev – Jubinski, 1999). The nonstationarity of log-volume series is particularly visible among estimates of memory parameter d in the whole period (13 cases). On the other hand, there is no visible increasing or decreasing trend in the behavior of d across time, however there are many more cases when long memory decreases. Moreover, it should be noted here that the value of memory parameters in the whole period is superior to d in subperiods for the majority of stocks.

Based on the asymptotic normality of the considered long-memory estimates, the standard cases $I(0)$ and $I(1)$ were tested. The null hypothesis about week dependence, i.e. $I(0)$ was rejected in the great majority of samples, whereas the $I(1)$ hypothesis about the unit root was rejected in all cases.

As a next step, we examined the long memory of bivariate series consisting of different measures of volatility and log-volume. As it was pointed out in Section 3, the GPH bivariate estimates are identical to the univariate estimates. Thus only multivariate modifications of local Whittle procedures were used. The bivariate memory parameter of volatility-volume series was computed via the two-step procedure. In literature, some authors – e.g. Bollerslev and Jubinski (1999), Lobato and Velasco (2000) – raise the question about the existence of common long memory of volatility and log-volume. In order to answer this question we performed tests of equality of long-memory parameters of volatility and log-volume. We use bivariate local Whittle estimates as well as GPH estimates. When the null hypothesis that returns volatility and trading volume share the same long-memory parameter could not be rejected, a common memory was estimated. As an example, *Tables 3* and *4* summarise the results of these estimations for volume-squared returns pairs. The Wald test described in Section 3 rejects the null hypothesis of a common long memory of squared returns and log-volume in the whole period for 17 cases of GPH bivariate estimates and for 16 cases of LW bivariate estimates. On the other hand, in the most recent period it rejects the null hypothesis only in three and two cases respectively. There are similar results regarding the long-run relationship between absolute returns and log-volume. The results of testing the existence of the common long-memory parameter suggest that generally it can be assumed that log-volume and returns volatility have the same long memory, particularly in the most recent period. Therefore, a new question arises: do they move together in a long time horizon? To answer this question, fractional cointegration between log-volume and volatility series should be examined by means of estimation of squared coherency and parameter β, described in Subsection 3.4. Typically, the estimated squared coherency for either of the considered pairs, i.e. log-volume and absolute returns or log-volume and squared returns, is in the range 0.1–0.3, and only in two cases it is greater than 0.4. However, it is too far from 1 to assume trading volume and returns volatility are fractionally cointegrated. A similar conclusion can be drawn from examining β estimates and the long-memory parameter of residuals in a frequency domain least squares estimation. As it was mentioned

TABLE 3 GPH Estimates of Long Memory Parameter $\boldsymbol{d} = (d_1, d_2)^T$ of Squared Returns (upper row) and Log-volume (lower row). (If there is no rejection of the null hypothesis: $d_1 = d_2$ at 0.05 significance level, the common long memory parameter is displayed (in bold).)

	Squared Returns							
	01.94–11.05		01.94–12.97		01.98–12.01		01.02–11.05	
Allianz	0.40		0.24	**0.24**	0.18		0.58	**0.55**
	0.65		0.23		0.47		0.53	
BASF	0.28		0.18	**0.23**	0.24		0.36	**0.34**
	0.43		0.31		0.44		0.32	
Bayer	0.21		0.24	**0.30**	0.14		0.09	
	0.46		0.36		0.46		0.35	
BMW	0.41	**0.41**	0.35	**0.34**	0.36	**0.38**	0.38	**0.35**
	0.41		0.33		0.39		0.31	
CommB	0.40	**0.42**	0.29	**0.30**	0.40		0.40	**0.41**
	0.43		0.34		0.18		0.43	
Daimler	0.36		0.18		0.36	**0.42**	0.49	**0.42**
	0.50		0.43		0.50		0.35	
DBank	0.39	**0.37**	0.24	**0.28**	0.26		0.49	
	0.33		0.35		0.45		0.29	
DTelekom	0.45		0.23	**0.29**	0.24		0.51	**0.40**
	0.65		0.34		0.66		0.31	
Henkel	0.29		0.30		0.20	**0.27**	0.23	**0.21**
	0.57		0.50		0.31		0.20	
HVB	0.47		0.20		0.51	**0.57**	0.55	**0.51**
	0.63		0.36		0.63		0.49	
Linde	0.34		0.35	**0.38**	0.27	**0.31**	0.26	**0.17**
	0.63		0.44		0.38		0.10	
Lufthansa	0.26		0.14		0.50		0.25	**0.28**
	0.58		0.41		0.34		0.30	
MAN	0.38		0.20	**0.28**	0.21	**0.24**	0.44	**0.43**
	0.57		0.36		0.30		0.40	
Metro	0.30		0.15	**0.25**	0.15		0.25	**0.25**
	0.55		0.31		0.36		0.24	
MuRe	0.45		0.22	**0.26**	0.29		0.40	**0.49**
	0.68		0.30		0.52		0.55	
RWE	0.38		0.22	**0.23**	0.25		0.48	**0.41**
	0.54		0.23		0.57		0.32	
SAP	0.24		-0.03		0.25		0.37	**0.35**
	0.66		0.39		0.67		0.32	
Schering	0.18		0.26	**0.34**	0.22	**0.23**	0.22	**0.25**
	0.41		0.40		0.24		0.29	
Siemens	0.37	**0.37**	0.29	**0.29**	0.23		0.37	**0.33**
	0.36		0.28		0.46		0.27	
Thyssen	0.19		0.21	**0.22**	0.07	**0.11**	0.19	**0.25**
	0.42		0.23		0.17		0.33	
TUI	0.39		0.22		0.37	**0.35**	0.28	
	0.67		0.49		0.30		0.61	
VW	0.38	**0.35**	0.22	**0.24**	0.46	**0.39**	0.44	**0.39**
	0.31		0.27		0.36		0.35	

TABLE 4 Multivariate Whittle Estimates of the Long Memory Parameter $\boldsymbol{d} = (d_1, d_2)^T$ of Squared Returns (upper row) and Log-volume (lower row). (If there is no rejection of the null hypothesis: $d_1 = d_2$ at 0.05 significance level, the common long memory parameter is displayed (in bold).)

	Squared Returns							
	01.94–11.05		01.94–12.97		01.98–12.01		01.02–11.05	
Allianz	0.38		0.29	**0.28**	0.23		0.52	**0.53**
	0.66		0.28		0.49		0.54	
BASF	0.26		0.18		0.21		0.35	**0.33**
	0.47		0.35		0.38		0.29	
Bayer	0.19		0.26	**0.31**	0.10		0.13	**0.20**
	0.46		0.37		0.41		0.27	
BMW	0.41	**0.44**	0.34	**0.36**	0.32	**0.36**	0.42	**0.38**
	0.45		0.39		0.39		0.28	
CommB	0.38		0.27	**0.31**	0.44		0.37	**0.38**
	0.49		0.36		0.30		0.40	
Daimler	0.35		0.24		0.35	**0.41**	0.44	**0.41**
	0.54		0.49		0.45		0.34	
DBank	0.37	**0.40**	0.27		0.33	**0.36**	0.42	**0.38**
	0.42		0.41		0.37		0.28	
DTelekom	0.42		0.29	**0.31**	0.26		0.43	
	0.65		0.34		0.62		0.28	
Henkel	0.32		0.30		0.27	**0.33**	0.26	**0.25**
	0.54		0.48		0.39		0.23	
HVB	0.45		0.19		0.46	**0.51**	0.48	**0.46**
	0.65		0.38		0.53		0.43	
Linde	0.33		0.34	**0.37**	0.26	**0.32**	0.28	**0.25**
	0.58		0.40		0.38		0.20	
Lufthansa	0.34		0.23		0.50		0.30	**0.32**
	0.58		0.43		0.35		0.34	
MAN	0.38		0.16		0.22	**0.25**	0.41	**0.43**
	0.55		0.34		0.27		0.45	
Metro	0.32		0.20		0.21		0.28	**0.26**
	0.56		0.39		0.39		0.22	
MuRe	0.43		0.20	**0.22**	0.28		0.41	**0.46**
	0.67		0.23		0.53		0.50	
RWE	0.36		0.25	**0.24**	0.25		0.38	**0.37**
	0.57		0.22		0.48		0.36	
SAP	0.26		0,00		0.23		0.28	**0.28**
	0.66		0.36		0.66		0.27	
Schering	0.24		0.22		0.21	**0.23**	0.22	**0.25**
	0.41		0.36		0.25		0.28	
Siemens	0.39	**0.41**	0.34	**0.35**	0.22		0.38	**0.36**
	0.43		0.35		0.42		0.31	
Thyssen	0.25	**0.36**	0.16	**0.20**	0.13	**0.20**	0.30	**0.33**
	0.42		0.24		0.25		0.36	
TUI	0.36	**0.53**	0.22		0.34	**0.30**	0.32	
	0.61		0.41		0.27		0.58	
VW	0.44	**0.40**	0.23	**0.26**	0.42	**0.38**	0.40	**0.40**
	0.37		0.30		0.33		0.39	

in Subsection 3.4, the memory of residuals must be smaller than the common long memory of the series under consideration (i.e. log-volume and volatility). The above condition is fulfilled only in a minority of cases. Moreover, if it is fulfilled, the difference between the common long memory of volume – volatility and residuals is insignificant. These findings are in line with those of Lobato and Velasco (2000) and indicate that even though returns volatility and trading-volume series might share a common long-memory parameter, they do not move together.

6. Conclusions

We tested by means of several methods for stochastic long memory in the stock data of German companies included in DAX index. The subject of our investigations were trading volume and volatility of returns (approximated by absolute returns and alternatively by squared returns). We established that for the equities listed in the DAX index the log-volume (the logarithm of trading volume) and returns volatility exhibit long memory. Moreover these two series have the same long-memory parameters for most of the equities. This common long memory of both series is especially strongly pronounced in the latest data. On the other hand, there is no evidence that log-volume and volatility share the same long memory component.

The presence of long memory represents nonlinearity in the mean of the process. This suggests a possibility for constructing nonlinear econometric models which could be applied to forecasting, especially over long forecasting horizons. According to our experience the known methods of long-memory parameter estimation lead to estimators whose values are very close together. One important question which arises here concerns the source of long memory in the series. The existence of long memory in the investigated series may reflect the statistical properties of fundamental factors underlying their behavior or qualitative changes which take place on stock markets. According to empirical investigations, the growing share of stocks by institutional investors is accompanied by an increasing autocorrelation in returns and trading volume data. On the other hand, long memory is related to autocorrelation. Thus, in our opinion the increasing presence of long memory in the latest German trading-volume data may be caused by the growing share of equities by institutional investors. This is in line with results presented in the contribution by Gurgul and Majdosz (2006) for the Polish capital market.

The increasing level of long memory in stock data also suggests the possibility of risk-level reduction in response to increasing activities of institutional investors on the German capital market. This finding supports particularly the hypothesis concerning the stabilizing impact of pension reform on the German capital market and corresponds to the empirical evidence presented in the literature for other markets.

REFERENCES

AGIAKLOGLOU, C. – NEWBOLD, P. – WOHAR, M. (1993): Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis,* vol. 14, 1993, pp. 235–246.

ANDERSEN, T. G. (1996): Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance*, vol. 51, 1996, pp. 169–204.

ANDREWS, D. W. K. – GUGGENBERGER, P. (2003): A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica,* vol. 71, 2003, pp. 675–712.

ANDREWS, D. W. K. – SUN, Y. (2004): Adaptive local polynomial Whittle estimation of long--range dependence. *Econometrica,* vol. 72, 2004, pp. 569–614.

AVOUYI-DOVI, S. – JONDEAU, E. (2000): International transmission and volume effects in G5 stock market returns and volatility. *BIS Conference Papers*, 2000, no. 8, pp. 159–174.

BAILLIE, R. T. (1996): Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, vol. 73, 1996, pp. 5–59.

BAILLIE, R. T. – BOLLERSLEV, T. – MIKKELSEN, H. O. (1996): Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics,* vol. 74, 1996, pp. 3–30.

BARKOULAS, J. T. – BAUM, C. F. (1996): Long term dependence in stock returns. *Economics Letters*, vol. 53, 1996, pp. 253–259.

BARKOULAS, J. T. – LABYS, W. C. – ONOCHIE, J. (1997): Fractional dynamics in international commodity prices. *Journal of Futures Markets*, vol. 17, 1997, pp. 161–189.

BERAN, J. A. (1992): Statistical methods for data with long-range dependence. *Statistical Science*, vol. 7, 1992, pp. 404–427.

BERAN, J. A. (1994): *Statistics for Long-Memory Processes*. Chapman and Hall, 1994.

BESSEMBINDER, H. – SEGUIN, P. J. (1993): Price volatility, trading volume and market depth: evidence from futures markets. *Journal of Financial and Quantitative Analysis*, vol. 28, 1993, pp. 21–39.

BEVERIDGE, S. – OICKLE, C. (1997): Long memory in the Canadian stock market. *Applied Financial Economics*, 1997, pp. 667–672.

BLUME, L. – EASLEY, D. – O'HARA, M. (1994): Market statistics and technical analysis: The role of volume. *Journal of Finance*, vol. 49, 1994, pp. 153–181.

BOLLERSLEV, T. – JUBINSKI, D. (1999): Equity trading volume volatility: latent information arrivals and common long-run dependencies. *Journal of Business & Economic Statistics,* vol. 17, 1999, pp. 9–21.

BOLLERSLEV, T. – MIKKELSEN, H. O. (1996): Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, vol. 73, 1996, pp. 151–184.

BRAILSFORD, T. J. (1996): The empirical relationship between trading volume, returns and volatility. *Accounting and Finance*, vol. 35, 1996, pp. 89–111.

BROCK, W. A. – LEBARON, B. D. (1996): A dynamic structural model for stock return volatility and trading volume. The *Review of Economics and Statistics*, vol. 78, 1996, pp. 94–110.

CAMPBELL, J. Y. – GROSSMANN, S. J. – WANG, J. (1993): Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics*, vol. 108, 1993, pp. 905–939.

CHEUNG, Y. W. – LAI, K. – LAI, M. (1993): Are there long cycles in foreign stock returns? *Journal of International Financial Markets, Institutions and Money*, vol. 3, 1993, pp. 33–47.

CHEUNG, Y. W. – LAI, K. S. (1993): Do gold market returns have long memory? *Financial Review*, vol. 28, 1993, pp. 181–202.

CHEUNG, Y. W. – LAI, K. S. (1995): A search for long memory in international stock market returns. *Journal of International Money and Finance*, vol. 14, 1995, pp. 597–615.

CHORDIA, T. – SUBRAHMANYAM, A. – ANSHUMAN, V. R. (2001): Trading activity and expected stock returns. *Journal of Financial Economics,* vol. 59, 2001, pp. 3–32.

CHORDIA, T. – SWAMINATHAN, B. (2000): Trading volume and cross-autocorrelations in stock returns. *Journal of Finance*, vol. 55, 2000, pp. 913–935.

CLARK, P. K. (1973): A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, vol. 41, 1973, pp. 135–155.

CONNOLLY, R. – STIVERS, C. (2003): Momentum and reversals in equity-index returns during periods of abnormal turnover and return dispersion. *Journal of Finance*, vol. 58, 2003, pp. 1521–1555.

COPELAND, T. (1976): A model of asset trading under the assumption of sequential information arrival. *Journal of Finance*, vol. 31, 1976, pp. 135–155.

COX, D. R. – REID, N. (1987): Parameter orthogonality and approximate conditional inference (with discussion). *Journal of the Royal Statistical Society Series B*, vol. 49, 1987, pp. 1–39.

CRATO, N. (1994): Some international evidence regarding the stochastic memory of stock returns. *Applied Financial Economics*, vol. 4, 1994, pp. 33–39.

DARRAT, A. F. – RAHMAN, S. – ZHONG, M. (2003): Intraday trading volume and return volatility of the DJIA stocks: A note. *Journal of Banking and Finance*, vol. 27, 2003, pp. 2035–2043.

DING, Z. – GRANGER, C. W. J. – ENGLE, R. F. (1993): A long memory property of stock market returns and a new model. *Journal of Empirical Finance,* vol. 1, 1993, pp. 83–106.

DOORNIK, J. A. – OOMS, M. (2003): Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics and Data Analysis*, vol. 42, 2003, pp. 333–348.

FAMA, E. F. (1970): Efficient capital markets: A review of theory and empirical work. *Journal of Finance,* vol. 25, 1970, pp. 383–417.

FANG, H. – LAI, K. – LAI, M. (1994): Fractal structure in currency futures prices. *Journal of Futures Markets*, vol. 14, 1994, pp. 169–181.

GALLANT, A. R. – ROSSI, P. E. – TAUCHEN, G. E. (1992): Stock prices and volume. *Review of Financial Studies,* vol. 5, 1992, pp. 199–242.

GALLO, G. M. – PACINI, B. (2000): The effects of trading activity on market volatility. The *European Journal of Finance*, vol. 6, 2000, pp. 163–175.

GEWEKE, J. – PORTER-HUDAK, S. (1983): The estimation and application of long memory time series models. *Journal of Time Series Analysis*, vol. 4, 1983, pp. 221–238.

GRANGER, C. W. J. – JOYEUX, R. (1980): An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, vol. 1, 1980, pp. 15–29.

GRANGER, C. W. J. – ZHUANXIN, D. (1996): Varieties of long memory models. *Journal of Econometrics*, vol. 73, 1996, pp. 61–67.

GREENE, M. T. – FIELITZ, B. D. (1977): Long-term dependence in common stock returns. *Journal of Financial Economics*, vol. 5, 1977, pp. 339–349.

GURGUL, H. – MAJDOSZ, P. (2006): The impact of institutional investors on risk and stock return autocorrelation in the context of the Polish pension reform. *Badania Operacyjne i Decyzje*, forthcoming.

HELMS, B. P. – KAEN, F. R. – ROSENMAN, R. E. (1984): Memory in commodity futures contracts. *Journal of Futures Market*, vol. 4, 1984, pp. 559–567.

HENRY, M. – ROBINSON, P. M. (1996): Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In: Robinson, P. M. – Rosenblatt, M. (eds.): *Athens Conference on Applied Probability and Time Series Analysis*, Volume II: *Time Series Analysis, In Memory of E. J. Hannan.* New York, Springer, pp. 220–232.

HIEMSTRA, C. – JONES, J. D. (1994): Testing for linear and nonlinear Granger causality in the stock price – volume relation. *Journal of Finance,* vol. 49, 1994, pp. 1639–1664.

HOSKING, J. R. M. (1981): Fractional differencing. *Biometrika*, vol. 68, 1981, pp. 165–176.

HURST, H. R. (1951): Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, vol. 1, 1951, pp. 519–543.

HURVICH, C. M. – DEO, R. S. – BRODSKY, J. (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 Series Analysis,* vol. 19, 1998, pp. 19–46.

KARPOFF, J. M. (1987): The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, vol. 22, 1987, pp. 109–126.

KOOP, G. – LEY, E. – OSIEWALSKI, J. – STEEL, M. F. J. (1997): Bayesian analysis of long memory and persistence using ARFIMA models. *Journal of Econometrics*, vol. 76, 1997, pp. 149–169.

KÜNSCH, H. R. (1987): Statistical aspects of self-similar processes. In: Prokhorov, Y. – Sazanov, V. V. (eds.): *Proceedings of the First World Congress of the Bernoulli Society.* Utrecht, VNU Science Press, 1987, pp. 67–74.

LAMOUREUX, C. G. – LASTRAPES, W. D. (1990): Heteroscedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, vol. 45, 1990, pp. 221–229.

LAMOUREUX, C. G. – LASTRAPES, W. D. (1994): Endogenous trading volume and momentum in stock-return volatility. *Journal of Business and Economic Statistics*, vol. 12, 1994, pp. 253–260.

LEE, B. S. – RUI, O. M. (2002): The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence. *Journal of Banking and Finance*, vol. 26, 2002, pp. 51–78.

LO, A. W. (1991): Long-term memory in stock market prices. *Econometrica*, vol. 59, 1991, pp. 1279–1313.

LOBATO, I. N. (1999): A semiparametric two-step estimator in a multivariate long memory model. *Journal of Econometrics,* vol. 90, 1999, pp. 129–153.

LOBATO, I. N. – SAVIN, N. E. (1998): Real and spurious long-memory properties of stock-market data (with discussion). *Journal of Business & Economic Statistics,* vol. 16, 1998, pp. 261–283.

LOBATO, I. N. – VELASCO, C. (2000): Long memory in stock-market trading volume. *Journal of Business & Economic Statistics*, vol. 18, 2000, no. 4, pp. 410–427.

MANDELBROT, B. B. (1971): When can a price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. *Review of Economics and Statistics*, vol. 53, 1971, pp. 225–236.

MCKENZIE, M. D. – FAFF, R. W. (2003): The determinants of conditional autocorrelation in stock returns. *Journal of Financial Research*, vol. 26, 2003, pp. 259–274.

OMRAN, M. F. – MCKENZIE, E. (2000): Heteroscedasticity in stock returns data revisited: Volume versus GARCH effects. *Applied Financial Economics*, vol. 10, 2000, pp. 553–560.

PETERS, E. E. (1994): *Fractal market analysis: applying chaos theory to investment and economics.* J. Wiley & Sons, 1994.

PHILLIPS, P. C. B. – SHIMOTSU, K. (2004): Local Whittle estimation in nonstationary and unit root cases. *Annals of Statistics,* vol. 34, 2004, no. 2, pp. 656–692.

ROBINSON, P. (1995a): Log-periodogram regression of time series with long range dependence. *Annals of Statistics,* vol. 23, 1995, pp. 1048–1072.

ROBINSON, P. M. (1995b): Gaussian semiparametric estimation of long range dependence. *Annals of Statistics,* vol. 23, 1995, pp. 1630–1661.

ROBINSON, P. M. – MARINUCCI, D. (2001): Narrow-band analysis of nonstationary processes. *Annals of Statistics,* vol. 29, 2001, no. 4, pp. 947–986.

ROBINSON, P. M. – YAJIMA, Y. (2002): Determination of cointegrating rank in fractional systems. *Journal of Econometrics,* vol. 106, 2002, no. 2, pp. 217–241.

SHIMOTSU, K. – PHILLIPS, P. C. B. (2002a): Pooled log periodogram regression. *Journal of Time Series Analysis,* vol. 23, 2002, pp. 57–93.

SHIMOTSU, K. – PHILLIPS, P. C. B. (2002b): Exact local Whittle estimation of fractional integration. *Cowles Foundation Discussion Paper,* no. 1367.

SOWELL, F. B. (1992): Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics,* vol. 53, 1992, pp. 165–188.

SUOMINEN, M. (2001): Trading volume and information revelation in stock markets. *Journal of Financial and Quantitative Analysis*, vol. 36, 2001, pp. 545–565.

TANAKA, K. (1999): The nonstationary fractional unit root. *Econometric Theory,* vol. 15, 1999, pp. 549–582.

VELASCO, C. (1999a): Non-stationary log-periodogram regression. *Journal of Econometrics,* vol. 91, 1999, pp. 325–371.

VELASCO, C. (1999b): Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis,* vol. 20, 1999, pp. 87–127.

SUMMARY

JEL Classification: G14, G35
Keywords: DAX 30; trading volume; univariate and bivariate long memory

Long Memory on the German Stock Exchange

Henryk GURGUL – Faculty of Management, University of Science and Technology, Krakow – corresponding author (h_gurgul@o2.pl)
Tomasz WÓJTOWICZ – Faculty of Management, University of Science and Technology, Krakow

In this study, the contributors present the results of their investigations into the long-memory properties of trading volume and the volatility of stock returns (given by absolute returns and alternatively by square returns). Their database is daily stock data of German companies in the DAX segment of the German Stock Exchange. The purpose of these investigations is the calculation of memory parameters and to determine whether there exists the same degree of long memory for trading-volume and return-volatility data. Calculations are performed on daily results from January 1994 to November 2005 and in three sub-periods: January 1994 to December 1997, January 1998 to December 2001, and January 2002 to November 2005.