VILNIAUS UNIVERSITETAS
FIZIKOS FAKULTETAS
DEPARTMENT OF THEORETICAL PHYSICS

Aleksejus Kononovičius

Modeling of Return in NASDAQ OMX Vilnius Stock Exchange

Research paper
Studijų programa - Theoretical Physics and Astrophysics
Kursas: I

Studentas	Aleksejus Kononovičius
Darbo vadovas	habil. dr. Vygintas Gontis
Katedros vedėjas	prof. habil. dr. Leonas Valkūnas

Vilnius, 2010

Contents

Foreword

Systems analyzed by social sciences tend to be rather complex. Their complexity is drawn not only from huge amounts seemingly unique parts (as huge amounts of data must be analyzed in order to scientifically understand system's behavior), but from tight interaction among them. Due to vital importance of interaction among system parts traditional, reductionist, approach seems to fail while attempting to scientifically describe and understand behavior exhibited macroscopically by complex system.

In the begging of XX century physicists also faced systems exhibiting complexity - Poincare solved three body problem [1], Ising proposed famous model for ferromagnetic interaction within lattice of magnetic spins [2], Lorenz shown that chaos can arise from deterministic equations [3]. Those are just few examples of tools developed by physicists for quantitative modeling and qualitative understanding of complex systems. In the last decade of XX century complexity physics, some may call it synergetics [4], seen expansion into non-traditional fields for physics - social research fields - becoming so-called physics of risk [5]. During this expansion some interdisciplinary sciences were born, on of them being econophysics [6, 7, 8] with econophysicists being physicists interested in complexity of financial markets.

Small scale time series of financial data exhibit interesting statistical properties, which in context of traditional economical theories can be considered anomalous. Even more interesting is the fact that many of these anomalous properties appear to be universal. Analysis of empirical stock trading data around the world have helped to establish a variety of statistical properties, which are called stylized facts [9, 10, 11]. It was also noticed that those statistical properties can be successfully defined with in framework of nonextensive statistical physics [12]. q-Gaussian distribution, nonextensive generalization of well known Gaussian distribution, seems to be perfect for description of distributions drawn from empirical realizations of various financial market quantities.

There are many stochastic models, i.e. ARCH models [13], which are defined through nonextensive statistical mechanics framework. Those models are able to reproduce stationary statistical properties of various financial market quantities, yet few of them are able to reproduce dynamical statistical properties of financial markets. Earlier we proposed model [14], which was also defined within nonextensive statistical mechanics framework, yet we used mathematical stochastic differential equation (SDE) framework [15], as it offers easier manipulation of stationary and dynamical statistical properties exhibited by model [16], to model return.

From previous discussion it is obvious that major stationary statistical property is q-Gaussian probability density function (further we abbreviate this term as PDF) of some observed quantities within financial markets (i.e. volatility measured as absolute return [14]). Major dynamical statistical property of various financial market quantities is their long-range memory. There are empirical evidences that trading activity, trading volume and volatility are random quantities exhibiting long-range memory [14, 17, 18, 19, 20]. Long-range memory is usually defined within time domain as power law autocorrelation

$$\rho(\Delta t) \sim \Delta t^{-\alpha}, \tag{1}$$

with $\alpha \to 0$, which can be translated in to frequency domain as as power law power spectral density (further we abbreviate this term as PSD)

$$S(f) \sim f^{-\beta}, \tag{2}$$

with $\beta \approx 1$. Thus in this work, as we did in [14], we will analyze long-range memory only through the frequency domain.

In this work we will:

- present previously proposed long-range memory stochastic model of return,
- compare statistical properties of trading activity and absolute return obtained from New York Stock Exchange (further abbreviated as NYSE) and NASDAQ OMX Vilnius Stock Exchange (further abbreviated as VSE) empirical data,
- adjust previously proposed long-range memory stochastic model of return in order to reproduce statistical properties of both financial markets.

1 Long-range memory stochastic model of return

Previously we, with co-authors, proposed stochastic model of return [14]. In this section we will discuss mathematics underlying that model. To prove model success we will compare model results with empirical statistical properties of return obtained from NYSE empirical data.

1.1 Stochastic model reproducing empirical return PDF and long-range memory

It is known that we can, in discrete time periods with discretization period of Δt, describe price changes as return

$$R_{t,\Delta t} = \ln\left[\frac{p(t)}{p(t-\Delta t)}\right], \tag{3}$$

here $p(t)$ is price function of time, which we can, for sake of simplicity, assume to be continuous and differentiable. By taking a limit of time period approaching zero, we can define continuous price changes via continuous return

$$r(t) = \lim_{\Delta t \to 0} \frac{R_{t,\Delta t}}{\Delta t}. \tag{4}$$

In proposed model of return we treat return as continuous stochastic long-range memory variable with pronounced non-extensiveness. This kind of variable, as we, with co-authors, have shown in [14], can be described by continuous stochastic differential equation

$$\mathrm{d}x = \left[\eta - \frac{\lambda}{2} - (x\epsilon^{\eta})^2\right] \frac{(1+x^2)^{\eta-1}}{(\epsilon\sqrt{1+x^2}+1)^2} x \mathrm{d}t_s + \frac{(1+x^2)^{\frac{\eta}{2}}}{\epsilon\sqrt{1+x^2}+1} \mathrm{d}W_s, \tag{5}$$

where x is continuous dimensionless return, which is function of dimensionless time (it is analogous to non-dimensionless $r(t)$ in Eq. (4)) t_s, and W_s is accordingly scaled Wiener process.

In [14] it is proposed to solve Eq. (5) numerically by using method of discretization with variable time step

$$h_k = \kappa^2 \frac{\left(\epsilon\sqrt{x_k^2+1}+1\right)^2}{(x_k^2+1)^{\eta-1}}, \tag{6}$$

where κ is model precision parameter, which should be less than 1. Using this numerical solution method Eq. (5) is transformed into set of difference equations for continuous dimensionless return

$$x_{k+1} = \kappa^2 \left[\eta - \frac{\lambda}{2} - (x_k\epsilon^{\eta})^2\right] x_k + \kappa\sqrt{1+x_k^2}\varepsilon_k, \tag{7}$$

where ε_k is k-th normal (with unit variance and mean zero) random variable, and dimensionless time

$$t_{s,k+1} = t_{s,k} + \kappa^2 \frac{\left(\epsilon\sqrt{x_k^2+1}+1\right)^2}{(x_k^2+1)^{\eta-1}}. \tag{8}$$

From Eq. (4) follows that by integrating solutions of Eq. (5), described by equations Eq. (7) and Eq. (8), we obtain discrete dimensionless return

$$X_{t_s,\tau} = \frac{\bar{R}_0}{\tau}\int_{t_s-\tau}^{t_s} x(T)\mathrm{d}T, \tag{9}$$

where τ is dimensionless discretization interval (corresponding to non-dimensionless Δt in Eq. (3)).

With rather high κ values it is convenient, due to possibility of h_k being larger τ, to additionally introduce constant time step (equal or less than τ) and in those extreme cases solve Eq. (5) using Euler-Maruyama method [21].

It can be analytically [15] and numerically (see Figure 1 (a)) shown that discrete dimensionless solutions of Eq. (5), described by Eq. (9), follow stationary PDF

$$\pi_r(X) = A_q\left(\frac{1}{1+X^2}\right)^{\frac{\lambda}{2}}, \tag{10}$$

where A_q is normalization constant. Eq. (10) can be rewritten to non-dimensionless form by assuming that dimensionless X is defined as

$$X_{t,\Delta t} = \frac{R_{t,\Delta t}}{R_0}, \tag{11}$$

where R_0 is characteristic scaling factor (in case of $\lambda \approx 4$ it equals standard deviation). By combining Eq. (10) and Eq. (11) we obtain

$$\pi_r(R_{t,\Delta t}) = B_q\left(\frac{R_0^2}{R_0^2+R_{t,\Delta t}^2}\right)^{\frac{\lambda}{2}}, \tag{12}$$

where B_q is normalization constant. Notice that Eq. (12) is actually q-Gaussian [14]. Thus we can, with some transformations

$$\lambda = \frac{2}{q-1}, \tag{13}$$

$$R_0 = \sigma_q\sqrt{\frac{3-q}{q-1}}, \tag{14}$$

rewrite Eq. (12) as actual q-Gaussian [12]

$$\pi_r(R_{t,\Delta t}) = C_q \exp_q\left[-\frac{R_{t,\Delta t}^2}{(3-q)\sigma_q^2}\right], \tag{15}$$

where q-exponential is defined as

$$\exp_q(x) = [1 + (1-q)x]^{\frac{1}{1-q}}. \tag{16}$$

PDF described by Eq. (10), Eq. (12) and Eq. (15) can be introduced through the variational principle applied to generalized entropy [13]

$$H_q = k\frac{1 - \int[\pi(X)]^q \mathrm{d}X}{1-q}. \tag{17}$$

PDF described by Eq. (10), Eq. (12) and Eq. (15) is known to fit empirical high-frequency return PDF [12, 14].

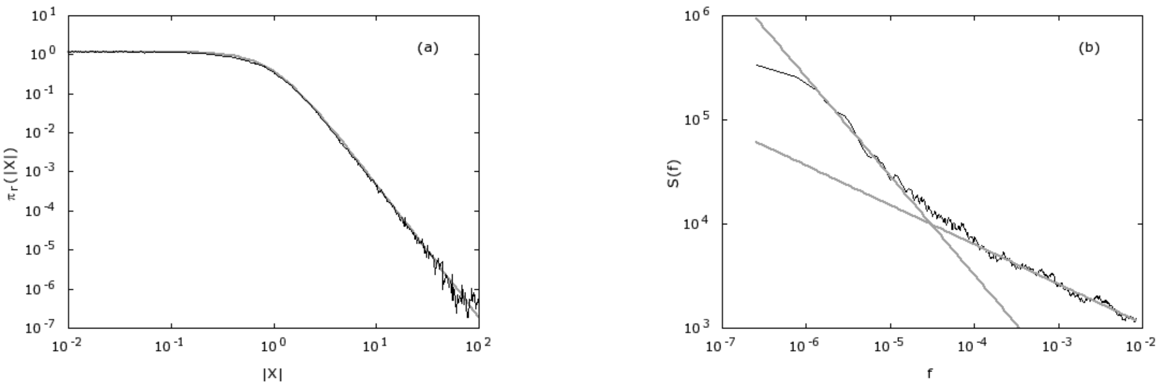

Figure 1: (a) numerically, through Eq. (9), calculated absolute return PDF (black thin line) approximated by q-Gaussian, Eq. (10), with $\lambda = 3.4$ (gray thick line), (b) numerically, through Eq. (9), calculated absolute return PSD (black thin line) approximated by two power law functions (gray thick lines) with $\beta_1 = 0.95$ and $\beta_2 = 0.38$. Used model parameter values: $\eta = 2.5$, $\lambda = 3.6$, $\tau = 10^{-4}$, $\epsilon = 0.01$.

It is known that financial markets exhibit long-range memory. Trivially this statistical property can be defined in frequency domain as spectral density inversely proportionate to frequency. Though actual empirical data from financial markets, such as NYSE and VSE (see Section 2), exhibit a little more sophisticated PSD - empirical PSD can be approximated by two power law functions with distinct power laws (β_i). This, more sophisticated, PSD is captured by [14] model (see Figure 1.

1.2 Empirical model adjustment

β values of model PSD (see Figure 1 (b)) are larger than those obtained from empirical data. In other words - model PSD and empirical PSD does not overlap, therefore further model adjustments are needed. In [14] model is improved by assuming that return process is double stochastic process composed of slowly diffusing long-range memory fluctuations and high amplitude rapid fluctuations modulating first one.

In that case empirical return $R_{t,\Delta t}$ can be mathematically formalized as

$$R_{t,\Delta t} = \xi \left\{ R_0[MA(R_{t,\Delta t}, T)], \lambda_2 \right\}, \tag{18}$$

where function MA is moving average of one-minute ($\Delta t = 60s$) trading return, $R_{t,\Delta t}$, within time window of one hour ($T = 3600s$) and $\xi \left\{ R_0, \lambda_2 \right\}$ is a q-Gaussian stochastic variable with R_0 dependent on modulating moving average of $R_{t,\Delta t}$, $MA(R_{t,\Delta t}, T)$, and empirically defined by λ_2. From analysis of NYSE empirical data [14] follows that $\lambda_2 = 5$ and R_0 can be written as linear function of $MA(R_{t,\Delta t}, T)$

$$R_0[MA(R_{t,\Delta t}), T] = R_{0b} + R_{0a}|MA(R_{t,\Delta t}, T)|, \tag{19}$$

with $R_{0b} = 1$, $R_{0a} = 2.5$. We introduced this decomposition as synthesis into our model by purposing to model $MA(R_{t,\Delta t}, T)$ as $X_{t,\tau}$ (as defined by Eq. (9)) with $\bar{R}_0 = 0.2$.

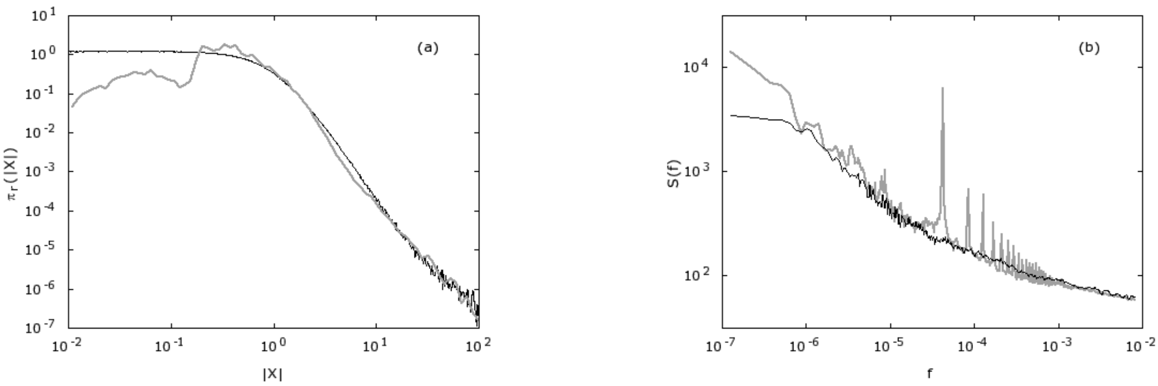

Figure 2: (a) numerically, through Eq. (18), calculated absolute return PDF (black thin line) and empirical return distribution (gray thick line), (b) numerically, through Eq. (18), calculated absolute return PSD (black thin line) and empirical return PSD (gray thick line). Used model parameter values: $\eta = 2.5$, $\lambda = 3.6$, $\tau = 10^{-4}/\sigma^2 = 60\,\mathrm{s}$, $\epsilon = 0.01$, $\lambda_2 = 5$, $\bar{R}_0 = 0.2$, $R_{0a} = 2.5$, $R_{0b} = 1$. Empirical statistical properties obtained by averaging corresponding statistical properties over 24 analyzed stocks from NYSE.

As we can see in Figure 2, adjusted model results fit NYSE empirical data. With negligible exceptions due to our consideration that money, and therefore return, is continuous variable (differences in PDF) and neglection of daily trading activity phenomena (differences in PSD).

2 Empirical analysis of NYSE and VSE empirical data

In this section we will analyze tick by tick trades of 4 stocks, APG1L, PTR1L, SRS1L, UKB1L, traded on VSE for 50 months since May, 2005. We will also extend and use results of our previous analysis [14, 20] of 24 stocks, ABT, ADM, BMY, C, CVX, DOW, FNM, GE, GM, HD, IBM, JNJ, JPM, KO, LLY, MMM, MO, MOT, MRK, SLE, PFE, T, WMT, XOM, traded on NYSE for 27 months from January, 2005.

2.1 Varying trading activity in analyzed financial markets

We can compare trading activity through statistical properties of inter-trade times (τ) and number of trades (N) per defined time window. Mathematically more sufficient description were proposed for inter-trade times [20], therefore we shall analyze statistical properties of τ.

Without defining or comparing precise distributions of τ in analyzed financial markets, we can calculate one of characteristic times - mean inter-trade time ($\bar{\tau}$). As we can see in Table 1 and Table 2 mean inter-trade times in analyzed financial markets vary greatly. This result ought to be expected as VSE is financial market of small country, at least in comparison with United States.

Table 1: Mean inter-trade times, $\bar{\tau}$, obtained from VSE empirical data

Stock	**$\bar{\tau}$, s**	**Stock**	**$\bar{\tau}$, s**
APG1L	337	PTR1L	565
SRS1L	381	UKB1L	164
		VSE mean	362

Table 2: Mean inter-trade times, $\bar{\tau}$, obtained from NYSE empirical data

Stock	**$\bar{\tau}$, s**	**Stock**	**$\bar{\tau}$, s**	**Stock**	**$\bar{\tau}$, s**	**Stock**	**$\bar{\tau}$, s**
ABT	4.09	ADM	4.22	BMY	3.27	C	1.79
CVX	2.34	DOW	3.9	FNM	5.4	GE	1.44
GM	2.34	HD	2.09	IBM	3.03	JNJ	2.64
JPM	2.41	KO	3.31	LLY	4.73	MMM	4.92
MO	3	MOT	1.66	MRK	2.47	PFE	1.24
SLE	6.58	T	2.34	WMT	1.84	XOM	1.44
						NYSE mean	3.02

Another characteristic time, τ_0, is drawn from mathematical description of τ distribution. While deriving formula for τ distribution it is assumed that each individual τ follows conditional exponential distribution

$$\varphi(\tau|n) = n \exp(-n\tau), \tag{20}$$

where n is average trading activity defined as rate of trades. From Eq. (20) follows that collective τ distribution in k-space can be defined as

$$\pi_\tau^{(k)}(\tau) = A \int_0^\infty \varphi(\tau|n)\pi_n(n)\mathrm{d}n, \tag{21}$$

where A is normalization constant, $\pi_n(n)$ is distribution of average trading activity, which, in case of single exponent power law model [22, 23], can be defined as

$$\pi_n(n) = B\left(\frac{n_0}{n}\right)^\lambda \exp\left[-\left(\frac{n_0}{n}\right)^m\right], \tag{22}$$

here B is normalization constant. Thus Eq. (21) can be rewritten as

$$\pi_\tau^{(k)}(\tau) = C\int_0^\infty x^{1-\lambda} \exp\left[-x^{-m} - x\left(\frac{\tau}{\tau_0}\right)\right] \mathrm{d}x, \tag{23}$$

with C being normalization constant. In case of parameter $m = 1$, PDF described by Eq. (23) can be expressed rather simply through Bessel function of second kind. While with $m > 1$ mathematical form of Eq. (23) is more complicated - hypergeometric functions are used.

Previously [20] there were evaluated τ_0 values for 24 NYSE stocks with Eq. (23) parameters set as follows: $m = 2$, $\lambda = 2.7$. While analyzing VSE empirical data we noticed that VSE stocks don't show sufficient fits with those parameters, therefore we tried other λ values to gain better fits. We found that with $\lambda = 1.95$ value τ distributions of VSE stocks are fitted sufficiently with τ_0 values given in Table 3.

Table 3: Empirically evaluated characteristic scaling times, τ_0, from VSE empirical data

Stock	τ_0, **s**	**Stock**	τ_0, **s**
APG1L	1000	PTR1L	1500
SRS1L	1150	UKB1L	390
		VSE mean	1010

In order to have comparison material, we estimated τ_0 for NYSE stocks with new Eq. (23) parameter set ($m = 2$, $\lambda = 1.95$). Actual τ_0 values are given in Table 4.

Table 4: Empirically evaluated characteristic scaling times, τ_0, from NYSE empirical data

Stock	τ_0, **s**	**Stock**	τ_0, **s**	**Stock**	τ_0, **s**	**Stock**	τ_0, **s**
ABT	8.5	ADM	11	BMY	9.5	C	4
CVX	5	DOW	8.5	FNM	13.5	GE	2.6
GM	7.6	HD	5.1	IBM	6.5	JNJ	5.25
JPM	5.25	KO	7.5	LLY	10	MMM	10
MO	6.5	MOT	5	MRK	6.25	PFE	3
SLE	16	T	10	WMT	4.25	XOM	3
						NYSE mean	7.2

Notice that ratio of $\bar{\tau}$ and τ_0 in different stocks of different financial markets is approximately the same (0.41 for NYSE and 0.36 for VSE). Therefore qualitatively trading in NYSE and VSE is the same, at least from distribution point of view. Quantitatively NYSE and VSE differs only in market size - it's natural to expect larger amount of traders to trade stocks of bigger companies more intensively than smaller amount of traders making trades with stock of smaller companies.

From presented numerical data one significant, in terms of trading activity, financial market statistical property is not visible. Though we can visualize it by ploting τ distributions (black thin lines in Figure 3) and their approximations (gray thick lines in Figure 3).

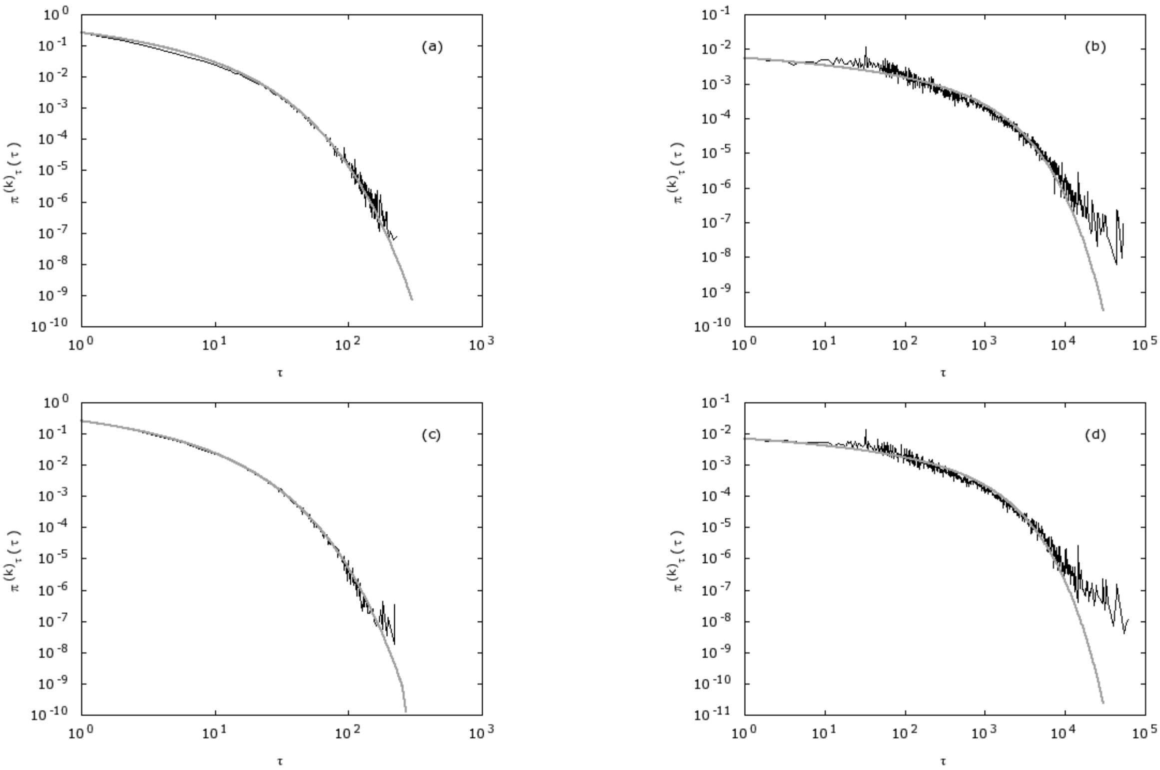

Figure 3: Empirical τ distributions (black thin lines) and proposed approximations (gray thick lines) of corresponding stocks - (a) LLY, (b) PTR1L, (c) DOW, (d) SRS1L. LLY and DOW are stocks traded on NYSE. PTR1L and SRS1L are stocks traded on VSE.

Note significantly fatter, not fitted by Eq. (23), tails of VSE empirical distributions (Figure 3 (b) and (d)) and some smaller other deviations are visible in VSE figures, though they are not present in empirical NYSE distributions (Figure 3 (a) and (c)), which are fitted almost perfectly. This happens due to fact that formula for τ distribution was derived by assuming τ_0 value being constant for whole realization to be fitted, which is true for stable mature financial markets, such as NYSE, but might not necessarily be true in case of maturing financial markets, such as VSE.

We can test this assumption by cutting any VSE stock realization into few periods. Most comfortably it can be done with SRS1L realization as it has sufficient number of trades and we know two dates, which are comfortable enough, then liquidity "shocks" occurred. We cut along those dates and obtain empirical distributions of τ, which we can afterward fit by Eq. (23) (see Figure 4).

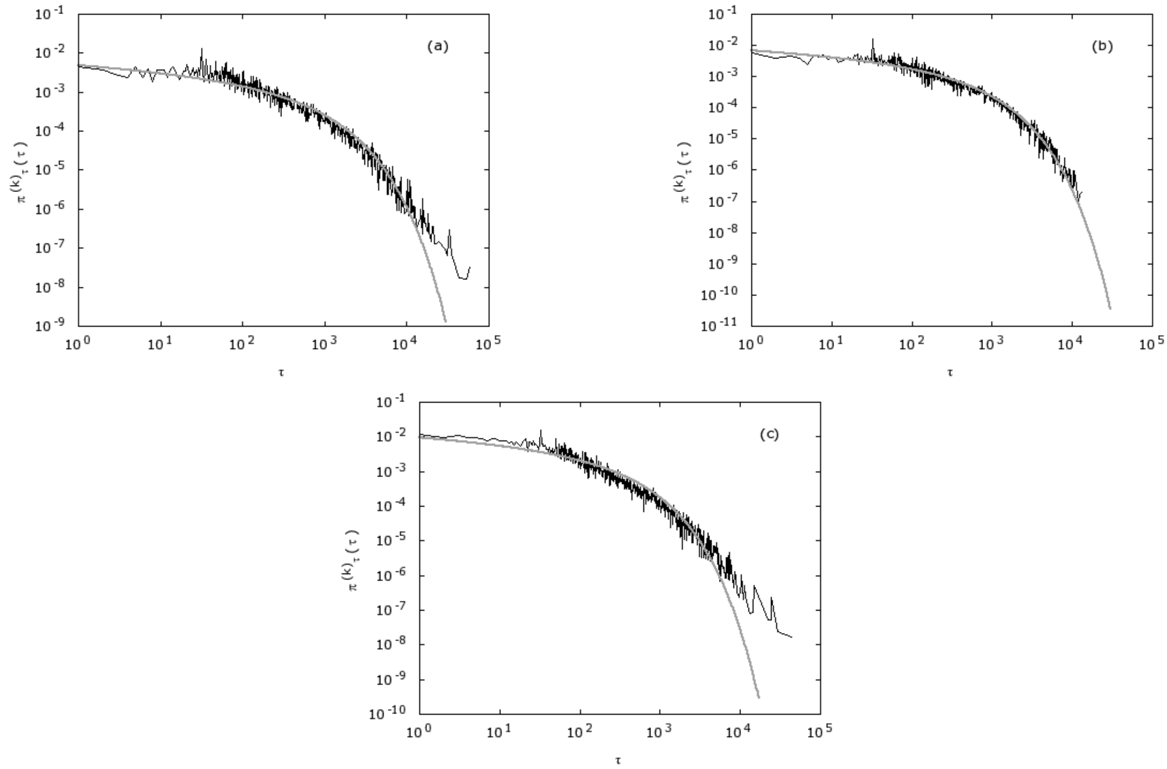

Figure 4: Empirical τ distributions (black thin lines) and proposed their approximations (gray thick lines) for different time periods in SRS1L realization ((a) from 2005-05-30 to 2007-01-04, (b) from 2007-01-18 to 2008-09-19, (c) from 2008-09-22 to 2009-07-13).

Table 5: Estimated τ_0 values for SRS1L realization in different time windows

Time window	τ_0, s	**Amount of trades**
from 2005-05-30 to 2007-01-04	1800	10002
from 2007-01-18 to 2008-09-19	1200	16046
from 2008-09-22 to 2009-07-13	800	16632
from 2005-05-30 to 2009-07-13	1150	42680

From Table 5 it is evident that τ_0 of SRS1L decreased with time, therefore aggregate τ distribution of SRS1L just represents it's mean behavior with fatter tails originating from varying τ_0. Note that only in time window from 2007-01-18 to 2008-09-19 theoretical distribution fitted empirical distribution perfectly (see Figure 4 (b)), while other time windows were fitted with same inconsistencies as aggregate distribution. From these results we can draw a conclusion that VSE is developing market - stock liquidity is decreasing rapidly with large "shocks" and fluctuating between those "shocks". NYSE, in contrary, is already developed financial market - stocks are fitted almost perfectly, therefore stock liquidity in NYSE is almost constant.

We can summarize this subsection's discussion with two points:

- VSE is smaller market than NYSE, therefore evident quantitative differences occur,

- qualitative differences are also present, though almost negligible, because VSE is still maturing market.

2.2 Stock return statistical properties in analyzed financial markets

Statistical properties of absolute return are usually analyzed and modeled at scales near 60 seconds [14]. Absolute return distributions of large markets, such as NYSE, at this scale are perfectly consistent, because mean inter-trade time (in case of NYSE it equals 3.02 s) in that case is significantly smaller than return discretization period. Different situation is observed in small markets, such as VSE (mean inter-trade time equals 362 s and is significantly larger than discretization period). This difference in trading activity causes difference in empirical return distribution - zero return probabilities peak above any other probability in VSE (see Table 6).

Table 6: One minute zero return value probability comparison

Stock	Zero return value probability
APG1L	95.08%
PTR1L	96.86%
SRS1L	96.56%
UKB1L	92.1%
VSE mean	95.15%
NYSE mean	23.71%

In order to model new, smaller, market with proposed model, we should eliminate low liquidity effects or introduce them into proposed model. Solution to his problem seems trivial - we should increase return discretization period from 1 minute to 2 hours or so. Yet we can't do that as in that case realization of 50 months would effectively shorten to realization of 14 days, which is obviously too short to analyze. Another problem of this solution lies within non-stationarity of statistical properties exhibited by return in financial markets [24]. Therefore large increase of discretization period might distort interesting statistical properties observed at scale of minutes. Though small increase of scale might do good.

Alternatively we can ignore zero return values, as nothing price-wise happens during those periods, then calculating return PDF, though we can't ignore them then we calculate PSD. In this case we also somewhat shorten realization length in terms of PDF, yet we are not threaten to have to deal with effects of non-stationarity. This solution seems a little rigid, yet it works perfectly as with eliminated zero probability VSE empirical return PDF overlaps with NYSE empirical return PDF (see Figure 5 (a)). In order to yield greater precision empirical PDFs more complex algorithms filtering zero return values (such as ignore time periods with no deals or ignore time periods with no price changes) might be implemented.

As we can see in Figure 5 (b) despite low financial market liquidity PSD of VSE and NYSE almost overlap. Difference is clearly seen only for higher frequencies and is caused by lower

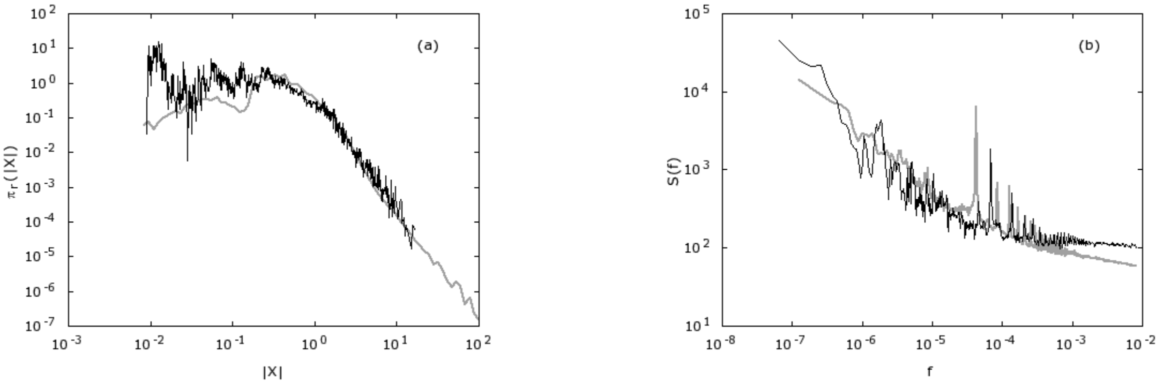

Figure 5: (a) mean VSE empirical PDF of absolute dimensionless one minute return, zero values excluded, (black thin line) overlapping with mean NYSE empirical PDF of absolute dimensionless one minute return (gray thick line). (b) mean VSE empirical PSD of absolute dimensionless one minute return (black thin line) overlapping with mean NYSE empirical PSD of absolute dimensionless one minute return (gray thick line). Visualized statistical properties obtained by averaging corresponding statistical properties of all analyzed stocks from corresponding financial market.

VSE market liquidity (white noise in PSD) and by different trading session length in financial markets (non-overlapping PSD maxima's).

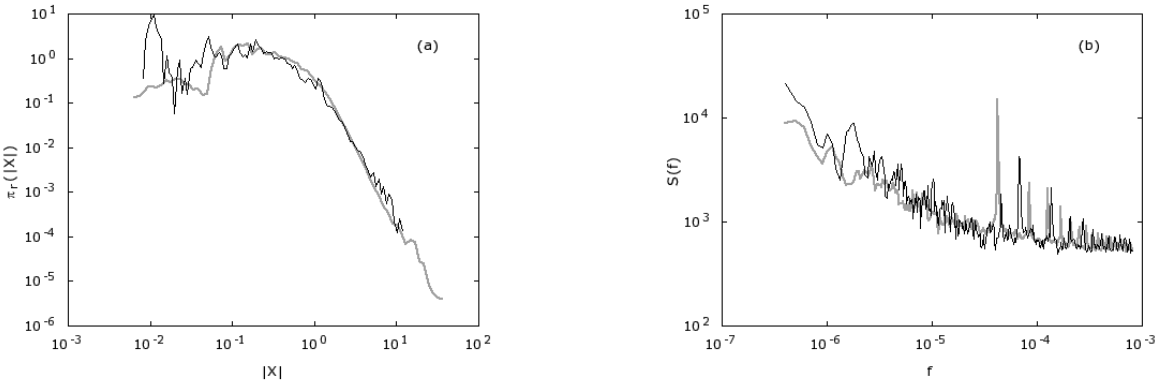

Figure 6: (a) mean VSE empirical PDF of absolute dimensionless ten minutes return, zero values excluded, (black thin line) overlapping with mean NYSE empirical PDF of absolute dimensionless ten minutes return (gray thick line). (b) mean VSE empirical PSD of absolute dimensionless ten minutes return (black thin line) overlapping with mean NYSE empirical PSD of absolute dimensionless ten minutes return (gray thick line). Visualized statistical properties obtained by averaging corresponding statistical properties of all analyzed stocks from corresponding financial market.

As we can see in Figure 6 (b) PSDs of VSE and NYSE overlap even better at larger time scale (600 seconds), which actually proves our previous assumption that differences between VSE and NYSE return statistical properties are caused by differing market liquidity. Though by increasing return integration window ten times we still do not solve zero return probability problem, which has to be solved as previously, by ignoring those values, in order to obtain comparable PDFs (which are present at Figure 6 (a)).

This subsection, as previous, can be summarized with two points:

- only difference between VSE and NYSE, in case of return statistical properties, is differing market liquidity,
- which can be easily eliminated by ignoring zero return values (improves PDF overlapping) or increasing integration window (improves PSD overlapping).

2.3 Time domain correlations between financial market indexes

In previous subsections we have shown similarities of statistical laws governing NYSE and VSE. From presented results we can easily assume that both markets are governed by same "mechanic", which could be interpreted as money flux. Recent years financial markets' trends suggest that this money flux is also a global phenomenon (see Figure 7), possibly due to economical globalization.

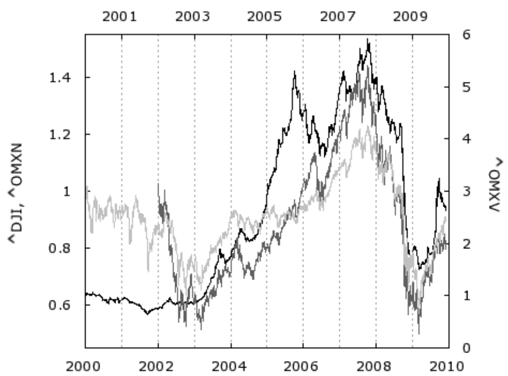

Figure 7: Evolution of ˆDJI (light gray line), ˆOMXN (dark gray line) and ˆOMXV (black thin line) financial market indexes. Indexes normalized with their value on 2000-01-01 (or 2002-01-01 in case of ˆOMXN) as a reference point.

In Figure 7 we see that three different, Nordic (represented by OMX Nordic 40 (abbreviated as ˆOMXN)), USA (represented by Dow Jones Industrial Average (abbreviated as ˆDJI)), Lithuanian (represented by OMX Vilnius (abbreviated as ˆOMXV)), financial markets follow similar trends within current decade. All three markets were caught in financial bubble inflating since 2003 until 2007, and therefore are now caught in it's explosion.

Quantitative analysis of correlations, Figure 8, reveal that despite general agreement between indexes in some time periods indexes do not correlate. Those disagreements can be justified by local specifics as developing markets depend less on global financial market than more stable markets. Stable markets react mostly to world-wide events, while developing markets have unused inner potential, which can give some independence from global trends as they can grow from within. Though overly rapid growth might lead to financial market overheat - stock prices fall as they become significantly overestimated.

Lithuanian economics, and financial market, was growing rapidly, due to development of inner economical potential facilitated by world-wide financial bubble, from 2003 until 2006. In 2006

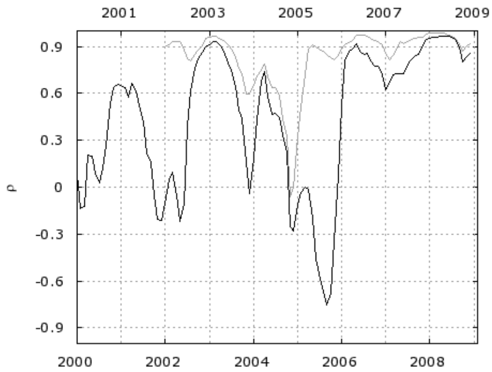

Figure 8: Evolution of yearly correlation between ˆDJI and ˆOMXV (black line), and between ˆDJI and ˆOMXN (gray line). Abscissa value corresponds to first date in analyzed, year wide, time window.

Lithuanian financial market index fell possibly due to financial overheat. Similar dynamics can be observed in evolution of Nordic financial market index. Though changes are lesser due to Nordic economics being better developed, and therefore more dependent from global trend, than Lithuanian.

3 Adjustment of long-range memory stochastic model of return

Numerical solutions of Eq. (5) tend to become unstable with τ value significantly differing from 10^{-4} value, due to diffusion limiter $(x\epsilon^{\eta})^2$ becoming inadequate. Therefore it would be convenient to change model SDE Eq. (5) by introducing new parameter, x_{max}, responsible for limiting diffusion area. Thus we rewrite Eq. (5) as

$$\mathrm{d}x = \left[\eta - \frac{\lambda}{2} - \left(\frac{x}{x_{max}}\right)^2\right] \frac{(1+x^2)^{\eta-1}}{(\epsilon\sqrt{1+x^2}+1)^2} x \mathrm{d}t_s + \frac{(1+x^2)^{\frac{\eta}{2}}}{\epsilon\sqrt{1+x^2}+1} \mathrm{d}W_s. \tag{24}$$

Eq. (24) can be numerically solved using same methods, that were used for Eq. (5) - variable time step method and Euler-Maruyama method [21] for possible extreme cases (variable time step, h_k, being larger than return discretization interval, τ).

With greater ease we fit empirical statistical properties of VSE (see Figure 9)and NYSE (see Figure 10) using model, with different parameter sets, results. Though from similarity of empirical PSD (see Figure 5) follows that empirical statistical properties can be fitted using same model results - qualitatively best parameter set for this is NYSE parameter set as it clearly reproduces sophisticated long-range memory statistical property.

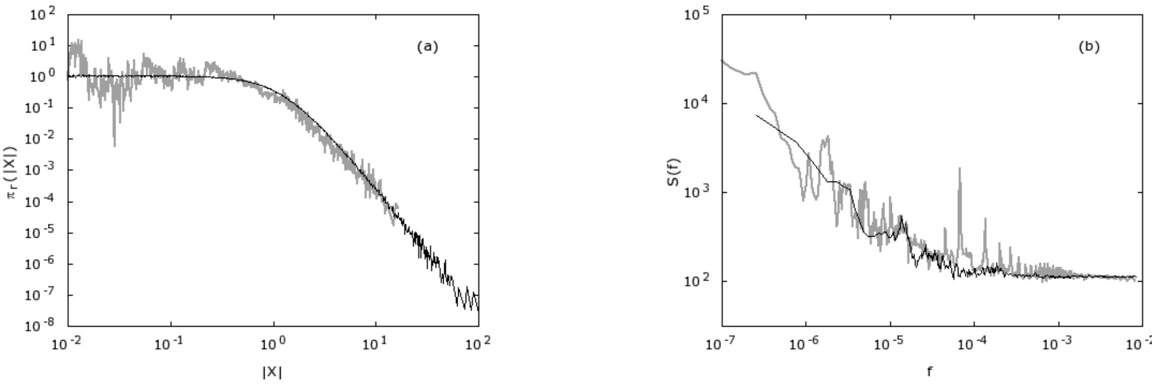

Figure 9: (a) VSE empirical absolute one minute return, zero values excluded, PDF (gray thick line) and PDF reproduced by adjusted model (black thin line), (b) VSE empirical absolute one minute return PSD (gray thick line) and PSD reproduced by adjusted model (black thin line). Shown model statistical properties obtained by averaging over 100 realizations, while VSE statistical properties obtained by averaging statistical properties of 4 analyzed VSE stocks. Model parameters are set as follows: $\lambda = 3.6$, $\eta = 2.5$, $x_{max} = 100$, $\epsilon = 0.008$, $\tau = 2 \cdot 10^{-6}/\sigma^2 = 60s$, $\lambda_2 = 4.25$, $R_{0b} = 1$, $R_{0a} = 2$, $\bar{R}_0 = 0.2$.

As we can see in Figure 11 model parameter set coincide with parameter set used to model NYSE one minute return, with exception of R_{0b}. This parameter value decreases as time scale grows larger due to bigger amounts of individual trades are being averaged, therefore smoothing momentum bursts. We strengthen this assumption by fitting empirical statistical properties at thirty minutes time scale with same parameter set, but R_{0b}, which decreases once again (see Figure 12).

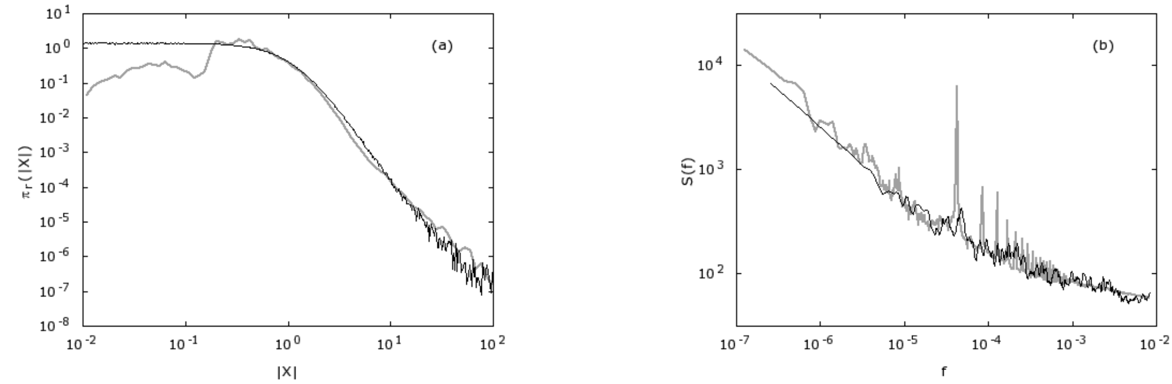

Figure 10: (a) NYSE empirical absolute one minute return PDF (gray thick line) and PDF reproduced by adjusted model (black thin line), (b) NYSE empirical absolute one minute return PSD (gray thick line) and PSD reproduced by adjusted model (black thin line). Shown model statistical properties obtained by averaging over 100 realizations, while NYSE statistical properties obtained by averaging statistical properties of 24 analyzed NYSE stocks. Model parameters are set as follows: $\lambda = 3.6$, $\eta = 2.5$, $x_{max} = 10^4$, $\epsilon = 0.017$, $\tau = 2 \cdot 10^{-5}/\sigma^2 = 60s$, $\lambda_2 = 5.0$, $R_{0b} = 1$, $R_{0a} = 2$, $\bar{R}_0 = 0.2$.

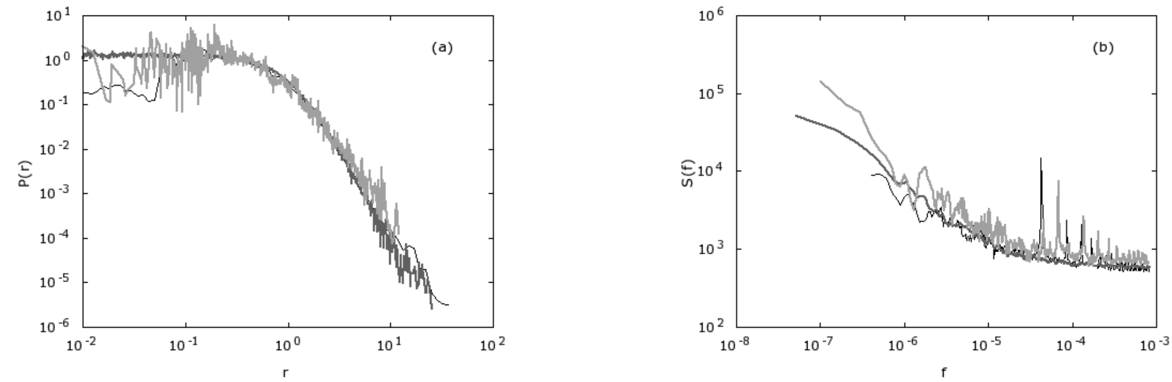

Figure 11: (a) NYSE empirical absolute ten minutes return PDF (black thin line), VSE empirical absolute ten minutes return, zero values excluded, PDF (black thin line) and PDF reproduced by adjusted model (dark gray thick line line), (b) NYSE empirical absolute ten minutes return PSD (black thin line), VSE empirical absolute ten minutes return PSD (light gray thick line) and PSD reproduced by adjusted model (dark gray thick line line). Shown model statistical properties obtained by averaging over 100 realizations, while NYSE statistical properties obtained by averaging statistical properties of 24 analyzed NYSE stocks and VSE statistical properties obtained by averaging statistical properties of 4 analyzed VSE stocks. Model parameters are set as follows: $\lambda = 3.6$, $\eta = 2.5$, $x_{max} = 10^4$, $\epsilon = 0.017$, $\tau = 2 \cdot 10^{-4}/\sigma^2 = 600s$, $\lambda_2 = 5.0$, $R_{0b} = 1$, $R_{0a} = 1.8$, $\bar{R}_0 = 0.2$.

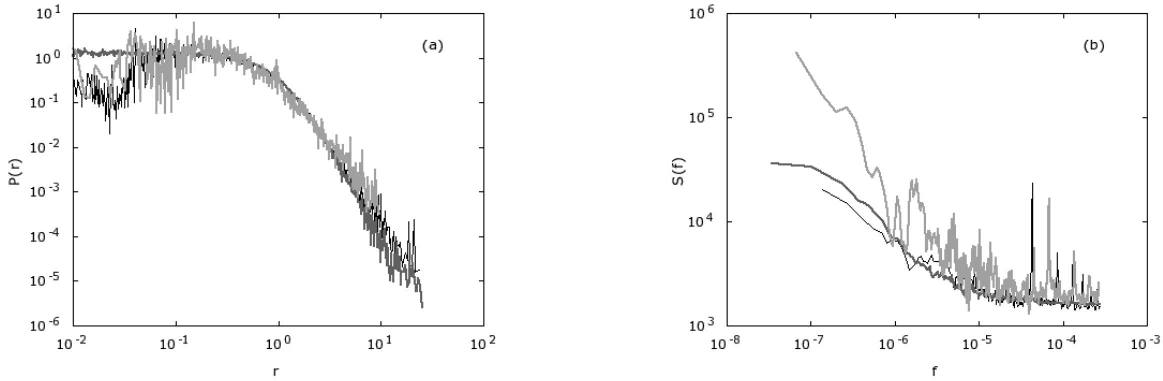

Figure 12: (a) NYSE empirical absolute thirty minutes return PDF (black thin line), VSE empirical absolute thirty minutes return, zero values excluded, PDF (black thin line) and PDF reproduced by adjusted model (dark gray thick line line), (b) NYSE empirical absolute thirty minutes return PSD (black thin line), VSE empirical absolute thirty minutes return PSD (light gray thick line) and PSD reproduced by adjusted model (dark gray thick line line). Shown model statistical properties obtained by averaging over 100 realizations, while NYSE statistical properties obtained by averaging statistical properties of 24 analyzed NYSE stocks and VSE statistical properties obtained by averaging statistical properties of 4 analyzed VSE stocks. Model parameters are set as follows: $\lambda = 3.6$, $\eta = 2.5$, $x_{max} = 10^4$, $\epsilon = 0.017$, $\tau = 6 \cdot 10^{-4}/\sigma^2 = 1800s$, $\lambda_2 = 5.0$, $R_{0b} = 1$, $R_{0a} = 1.5$, $\bar{R}_0 = 0.2$.

4 Paper summary and conclusion

- In Section 1 we discussed previously, with co-authors, proposed model of return in financial markets ([14]). We, in this work, have shown that it reproduces two interesting properties - power law PDF, which originates from non-extensiveness of financial markets, and sophisticated long-range memory as double power law PSD.
- In Section 2 we have analyzed empirical trade by trade data from VSE and NYSE:
  - While comparing statistical properties of trading activity in different financial markets we discovered only obvious differences - market size and market development level. Difference in market size underlies differing financial market stock liquidity (see Table 1, Table 2, Table 3 and Table 4), while non-constant financial market stock liquidity (see Table 5) is a clear sign that market is young and developing.
  - Differences in statistical properties of return originate from aforementioned differences in trading activity - both PDF and PSD are distorted by overwhelmingly frequent zero return values (see Table 6). We also have found that if we ignore those differences (improves PDF overlapping) or increase return discretization times (improves PSD overlapping), we obtain very similar statistical properties from both markets (see Figure 5 and Figure 6).
  - We noticed that financial market indexes agree most of the time with exceptions due to local specifics (see Figure 8). This suggests that underlying process is global.
- We studied how [14] model SDE and it's numerical solutions behave with different return discretization times. In Section 3 we came to conclusion that model should be adjusted by introducing parameter responsible for limiting return diffusion area. Without this adjustment numerical solutions of model SDE might have not been stable enough with differing discretization times.
- We proposed new adjusted [14] model parameter sets which enable reproduction of statistical properties as obtained from VSE (see Figure 9) and NYSE (see Figure 10) empirical data.
- Obtained empirical similarities and model universality (see Figure 11) suggest that different financial markets are globally interconnected and thus driven by mechanic, which can be interpreted as global money flux.

Acknowledgment

We would like to express gratitude towards NASDAQ OMX Vilnius Stock Exchange, which provided empirical trade by trade data for our research. And also for organizing contest for best student scientific work on Lithuanian financial markets, which encouraged us to develop our work in this direction.

References

1. J. H. Poincare. *On the problem of three bodies and the equations of dynamics. Divergence series of Mr. Lindstedt.* Acta Mathematica 13, 1890, p. 1-270.

2. E. Ising. *Contribution to the theory of ferromagnetism.* Zeitschrift fur Physik 31, 1925, p. 253-258.

3. E. N. Lorenz. *Deterministic non-periodic flow.* Journal of the Atmospheric Sciences 20, 1963, p. 130-141.

4. H. Haken. *Synergetics, an Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology.* Springer-Verlag, New York, 1983.

5. B. Kaulakys, V. Gontis. *Rizikos fizika.* URI: http://mokslasplius.lt/rizikos-fizika/node/44.

6. J. P. Bouchaud. *The (unfortunate) complexity of the economy.* Physics World 4, 2009, p. 28-32. arXiv: 0904.0805v1 [q-fin.GN].

7. Y. Wang, J. Wu, Z. Di. *Physics of econophysics.* 2004. arXiv: cond-mat/0401025v1 [cond-mat.soft].

8. V. M. Yakovenko. *Econophysics, statistical mechanics approach to.* Encyclopedia of complexity and system science, 2008. arXiv: 0709.3662v4 [q-fin.ST].

9. R. N. Mantegna, H. E. Stanley. *Scaling behaviour in the dynamics of an economic index.* Nature 376, 1995, p. 46-49.

10. R. F. Engle. *The Econometrics of Ultra High Frequency Data.* Econometrica 68, 2000, p. 1-22.

11. J. P. Bouchaud, M. Potters. *Theory of financial risks and derivative pricing.* Cambridge University Press, New York, 2004.

12. Cf. Gell-Mann, C. Tsallis. *Nonextensive Entropy - Interdisciplinary Applications.* Oxford University Press, New York, 2004.

13. S. M. Duarte Queiros, C. Anteneodo, C. Tsallis. *Power-law distributions in economics: A nonextensive statistical approach.* arXiv: physics/0503024v1 [physics.soc-ph].

14. V. Gontis, J. Ruseckas, A. Kononovičius. *Long-range memory stochastic model of the return in financial markets.* Physica A 389, 2010, p. 100-106. arXiv: 0901.0903v3 [q-fin.ST].

15. C. W. Gardiner. *Handbook of stochastic methods.* Springer, Berlin, 1997.

16. C. Anteneodo, R. Riera. *Additive-multiplicative stochastic models of financial mean-reverting processes.* Physical Review E 72, 2005, p. 026106. arXiv: physics/0502119v1 [physics.soc-ph].

17. R. F. Engle, A. J. Paton. *What good is a volatility Model?* Quantitative Finance 1, 2001, p. 237-245.

18. V. Plerou, P. Gopikrishnan, X. Gabaix, L. A. Nunes Amaral, H. E. Stanley. *Price fluctuations, market activity and trading volume*. Quantitative Finance 1, 2001, p. 262-269.

19. X. Gabaix, P. Gopikrishnan, V. Plerou, H. E. Stanley. *A theory of power laws in financial markets fluctuations*. Nature 423, 2003, p. 267.

20. V. Gontis, B. Kaulakys, J. Ruseckas. *Trading activity as driven Poisson process: comparison with empirical data*. Physica A 387, 2008, p. 3891-3896. arXiv: 0710.1439v1 [q-fin.TR].

21. P. E. Kloeden, E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1999.

22. V. Gontis, B. Kaulakys. *Long-range memory model of trading activity and volatility*. Journal of Statistical Mechanics, P10016, 2006. arXiv: physics/0606115v1 [physics.soc-ph]

23. V. Gontis, B. Kaulakys. *Modeling long-range memory trading activity by stochastic differential equations*. Physica A 382, 2007, p. 114-120. arXiv: physics/0608036v1 [physics.soc-ph].

24. V. Plerou, P. Gopikrishnan, L. A. Amaral, M. Meyer, H. E. Stanley. *Scaling of the distribution of price fluctuations of individual companies*. Physical Review E 60, 1999, p. 6519-6529. arXiv: cond-mat/9907161v1 [cond-mat.stat-mech].

Aleksejus Kononovičius

Akcijų grąžos NASDAQ OMX Vilnius vertybinių popierių biržoje modeliavimas

Santrauka

Šiame darbe mes aptarėme anksčiau mūsų, su bendraautoriais, pasiūlytą grąžos modelį [14], o taip pat pasiūlėme šio modelio papildymą, kurio dėka tapo paprasčiau atkurti Vilniaus vertybinių popierių biržos empiriškai nustatytas statistines savybes. Papildytam modeliui pasiūlėme parametrų rinkinius su kuriais galima atkurti Vilniaus ir Niujorko vertybinių popierių biržos statistines savybes. Taip pat pastebėjome, kad didinant laiko mastelį skirtintų vertybinių popierių biržų statistinės savybės panašėja ir jas įmanoma modeliuoti naudojant parametrų rinkinį praktiškai identišką parametrų rinkiniui naudotam atkurti NYSE statistines savybes vienos minutės mastelyje.

Šiame darbe taip pat atlikome šių dviejų biržų statistinių savybių palyginimą. Palyginimo metu nustatėme, kad nepaisant kiekybinio skirtumo (Niujorko vertybinių popierių birža yra daug likvidesnė nei Vilniaus vertybinių popierių birža) kokybiškai prekyba skirtingose biržose yra panaši. Vienintelis kokybinis skirtumas matomas lyginant biržų empirinius duomenis yra prekybos aktyvumo charakteringų laikų kitimas stebimas Vilniaus vertybinių popierių biržoje, o Niujorko vertybinių popierių biržoje charakteringi laikai beveik nekinta. Šis skirtumas gali būti paaiškintas labai paprastai - Vilniaus vertybinių popierių birža, priešingai nei Niujorko, yra jauna finansų rinka atstovaujanti besivystančiai ekonomikai, kuriai augant ir stiprėjant didėja jos įmonių akcijų likvidumas (tuo pačiu kinta ir charakteringi prekybos aktyvumo laikai).

Iš pademonstruotų rezultatų - papildyto grąžos modelio universalumo ir empiriškai nustatytų statistinių savybių panašumo - galima spręsti apie tai, kad finansų rinkos yra tarpusavyje susijusios ir paklūsta tam pačiam mechanizmui, kurį galima būtų įvardinti kaip pasaulinį pinigų srautą. To priežastimi yra kitų rinkų glaudūs ryšiai arba kitaip tariant ekonomikos globalizacija. Ryšys tarp finansų rinkų yra dar labiau sustiprėjęs pasaulinei ekonomikai susidūrus su bendra problema - pasauline finansų krize.