Master of Advanced Studies in Finance

Identification of bubble phases from a CEV-type model

Anastasia Filimon

Master Thesis submitted to

ETH ZURICH and UNIVERSITY OF ZURICH

Supervisor in ETH Zürich

Prof. Didier Sornette

and

Supervisors in LGT Capital Management

Dr. Walter Pfaff

Dr. Magnus Pirovino

Roman Babenko

April 2010

Declaration

I hereby declare that this thesis was performed and written on my own and that references and resources used within this work have been explicitly indicated.

I am aware that making a false declaration may have serious consequences.

Zürich, 15/04/2010 _____

(Signature)

Acknowledgement

This paper was sponsored by LGT Capital Management and conducted as a part of the internship in the Asset Allocation team in Pfäffikon, Switzerland.

I would like to thank the supervisors of my master thesis in the academy, Prof. Didier Sornette for his motivating way of supervising and his availability, Dr. Stefan Reimann for the discussions and guidance throughout the work. Furthermore, I would like to thank my supervisors in the industry, Dr. Walter Pfaff and Dr. Magnus Pirovino for their numerous valuable comments, and Roman Babenko for his helpful input and patience, and the whole Asset Allocation and LGT and Science teams for their support through my time in the company.

And last but not least, many thanks to my mother and my best friends for their support and encouragements.

Contents

Chapter 1

Introduction

Recently, the analysis of bubbles and crashes in financial markets is attracting much attention as they have appeared in almost every market, and sometimes their impact on the society might be compared with that of natural disasters.

Following Case and Shiller [6], the term "bubble" refers to a situation in which excessive public expectations of future price increases cause prices to be temporarily elevated. In other words, bubble is defined as a deviation of the stock price from its fundamental value [5], though such definition lacks practical use. As mentioned by the former Federal Reserve Chairman A.Greenspan - *"despite our suspicions, it was very difficult to definitively identify a bubble until after the fact–that is, when its bursting confirmed its existence"* [9].

Watanabe et al. [19], [20] introduced the following mathematical definition of a financial bubble:

$$X_t - X_{t-1} = (\omega_1(t;T) - 1)(X_{t-1} - X_0(t;T)) + W_t, \tag{1.1}$$

where X_t is a price at time t, W_t is the residual noise term, parameters $\omega_1(t;T)$ and $X_0(t;T)$ are uniquely determined from the past T data points minimizing the root-mean-square of W_t. This formula describes three different kinds of behavior:

1. if $\omega_1(t;T) > 1$, the price is either exponentially increasing or decreasing and $X_0(t;T)$ gives the base line of the exponential divergence. Such behavior is defined as a bubble,
2. if $\omega_1(t;T) = 1$, the price follows a random walk,
3. if $\omega_1(t;T) < 1$, the price is convergent to $X_0(t;T)$.

In this definition the price motion obeys the exponential form for a given time scale from t-T to t:

$$X_t = (X_{t-T} - X_0(t;T))e^{T \ln \omega_1(t;T)} + X_0(t;T)) + \sum_{i=0}^{T-1} \omega_1(t;T)^i W_{t-i} \qquad (1.2)$$

The equation 1.2 characterizes the exponential growth, when the exponential function is regarded as a constant; the double exponential growth and power law growth with a critical time, when the exponent changes exponentially with time. These characterizations are known from the literature to be good approximations of the price increases related to financial bubbles. For example, Watanabe et al. [19], [21] considered exponential fitting for stock market data with regard to bubbles, Sornette et al. [15], [16], [17] showed that hyperinflations and crashes behaviors can be characterized by a power low singularity culminating at a critical time.

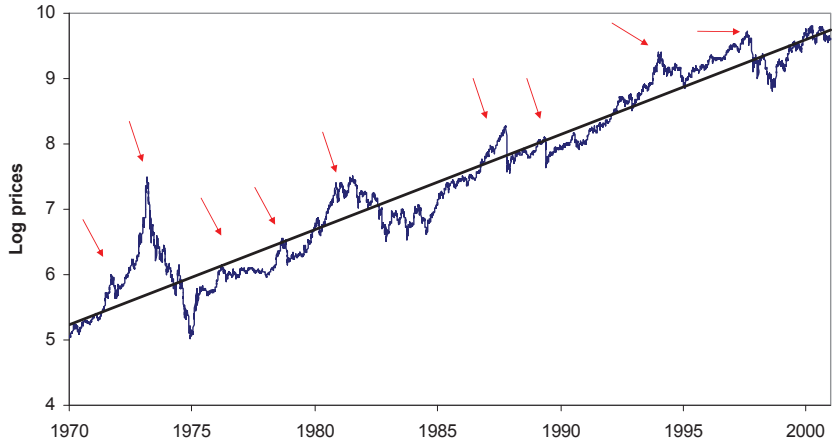


Figure 1.1: Trajectory of the Hang Seng index from 1970 to 2000. Log-prices are plotted against the time, black line corresponds to exponential growth with a constant growth rate. The red arrows point to the crashes - decline in the price of more than 15% in less than three weeks.

The definition 1.2 was shown to be powerful for the detection of bubbles and for the identification of the start of a bubble even before the burst, but not for the prediction of the end of a bubble. Sornette et al. [15]-[17] provided a mathematical description of the end of the bubbles as a faster-than-exponential growth singularity at a critical time t_c. An example of this definition is on the Figure 1.1, where the logarithm price of the Hang Seng

index from 1970 to 2000 is plotted. As it is seen, the price is not growing exponentially with a constant growth rate, but exhibits periods of faster-than-exponential growth with critical times (red arrows on the graph). This example suggests that faster-than-exponential price growth is followed by a crash.

Further, Ramirez et al. [1] showed that the peso depreciation growth during the Mexican financial crises of 1990s was greater than an exponential and those growth rates were compatible with a spontaneous singularity occurred at a critical time.

The Financial Instability Hypothesis provides an economic rational for periodically occurring instabilities caused by a feedback due to which the system's reactions to a movement of the economy amplify the movement [13]. The so called positive price-to-price feedback (i.e. conditioned on the observation that the market has recently moved up (down), this makes it more probable to keep it moving up (down), so that a large cumulative move ensues [18]) leads to faster-than-exponential growth regimes in the prices, i.e. periods of instabilities. Among the mechanisms that may lead to positive feedbacks on prices are:

1. hedging derivatives,
2. insurance portfolios,
3. "herd" behavior and "crowd" effect.

In the present work we will try to identify the periods of faster-than-exponential growth from empirical data. The analyzed methods are based on the assumption that the price follows the CEV process. The model is described in Chapter 2, Chapter 3 provides detailed description of the methods, as well as the suggestion of the model modification and model diagnostics. In Chapter 4 and 5 the proposed methods are tested on the stochastic benchmarks and on the NASDAQ daily prices, respectively.

Chapter 2

The Model

The current study is based on the assumption that the price follows constant elasticity of variance model

$$dX_t = \mu X_t dt + \sigma X_t^m dW_t \tag{2.1}$$

where W_t is a Brownian motion.

This is the so called "risk-driven" model, where the traders may increase the instability of the market at certain times by changing their opinions suddenly on a large scale.

The model 2.1 was proposed by Cox and Ross [8] as an alternative to the Black-Scholes model. Under this model the instantaneous volatility of the percentage price change is equal to σX_t^{m-1}. Traditional log-normal Black-Scholes model corresponds to the limiting case $m = 1$, when the volatility is not a function of the stock price itself.

The model with $m > 1$ admits positive feedback since in this case the stochastic growth rate of the process increases with X_t and this case is particularly interesting for us to detect periods of faster-than-exponential growth. While $m < 1$ causes the inverse relationship between the price level and the volatility. The model describing such relation might look inappropriate in the real world, nonetheless, the studies suggest that stock return volatility is asymmetric. For example [2], an increase in the stock return volatility is lower when the stock price increases following "good news" announcement than when the stock price declines corresponding to "bad news" announce. One of the possible explanation of this asymmetry in stock return volatility is the "leverage effect". Decline in the company's stock price causes more rapid fall of the market value of its equity than the market value of its debt. This

results in the increase of company's financial leverage, hence an increase in the volatility of equity [11],[7]. This similar effect could be observed even in the absence of debt. As every company faces some fixed cost which have to be met irrespective of the income, a decline in the income decreases the value of the company and increases its riskiness. These arguments of operating and financial leverage represent grounded explanation for the inverse relationship between the variance and stock price level to model heteroscedasticity in stocks returns.

Among the other evidences for the model 2.1, the results of Macbeth and Merville [12]. They tested the Cox call option valuation model for constant elasticity of variance diffusion processes against the Black-Scholes call option valuation model. The outcome was that stock prices appeared to be generated by constant elasticity of variance diffusion processes and the Cox valuation model fit market prices of call options significantly better than the Black-Scholes model. Thus, the model has an important implications in finance.

Chapter 3

Estimation methods

3.1 Macbeth-Merville method

From the constant elasticity of variance diffusion process equation

$$dX_t = \mu X_t dt + \sigma X_t^m dW_t \tag{3.1}$$

it follows that

$$\frac{(dX_t - \mu X_t dt)^2}{\sigma^2 X_t^{2m} dt} \equiv u_t \sim \chi^2(1) \tag{3.2}$$

Applying natural logarithm to both sides of 3.2 gives a linear regression equation to obtain consistent estimates of parameter m:

$$\ln[(dX_t - \mu X_t dt)^2] = \ln[\sigma^2] + \ln[X_t^{2m}] + \ln[dt] + \ln[u_t] \tag{3.3}$$

$$\ln[(dX_t - \mu X_t dt)^2] - \ln[dt] = 2\ln[\sigma] + 2m\ln[X_t] + \ln[u_t] \tag{3.4}$$

In other words, we have a linear model of the form:

$$y_t = \alpha + \beta x_t + \epsilon_t, \tag{3.5}$$

where

$$y_t = \ln[(dX_t - \mu X_t dt)^2] - \ln[dt], \tag{3.6}$$

$$x_t = \ln[X_t], \tag{3.7}$$

$$\epsilon_t = \ln[u_t] - E[\ln[u_t]] \tag{3.8}$$

$$E[\widehat{\alpha}] = 2\ln[\sigma] + E[\ln[u_t]], \tag{3.9}$$

$$E[\widehat{\beta}] = 2m \tag{3.10}$$

3.2 Entropy method

In information theory, entropy is a measure of the uncertainty associated with a random variable.

Definition 1. *The Shannon entropy H of a discrete random variable X taking possible values* $\{x_1,...,x_n\}$ *with a probability mass function p is defined as*

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p\,(x_i) \tag{3.11}$$

The concept was introduced by C. E. Shannon in his paper "A Mathematical Theory of Communication", 1948. The Shannon entropy is restricted to random variables taking discrete values. The extension of the Shannon entropy to the domain of real numbers is usually referred to as the continuous entropy, or differential entropy.

Definition 2. *The continuous entropy h of a continuous random variable X with a probability density function f whose support is a set R is definied as*

$$h(X) = -\int_R f(x) \log(f(x))dx. \tag{3.12}$$

Since probability density functions can take values greater than 1, continuous entropy h might take negative values as opposed to the Shannon entropy H.

The Renyi entropy, a generalisation of the Shannon entropy, is one of a family of functionals for quantifying the diversity, uncertainty or randomness of a system. The concept was introduced by A. Renyi.

Definition 3. *The Renyi entropy of order q, where* $q \geq 0$*, of a discrete random variable X taking possible values* $\{x_1,...,x_n\}$ *with a probability mass function p is defined as*

$$H_q(X) = \frac{1}{1-q} \log(\sum_{i=1}^{n} p^q(x_i)) \tag{3.13}$$

The trivial case $q = 1$ gives the Shannon entropy.

From the constant elastisity of variance model, the disribution of log-returns conditioned on the price is

$$f(r|x) = \beta\phi(\beta(r - (\mu - \frac{1}{2\beta^2}))), \tag{3.14}$$

where $\beta = \frac{1}{\sigma x^{m-1}}$ (see Appendix A.1 for the derivation).

Applying the Renyi entropy of the log-returns variables conditioned on the prices $(r|x)$ [14], we have

$$H_q(r|x) = \frac{1}{1-q} \log \int f^q(r|x)dr = -\ln \beta + H_q(\phi) \tag{3.15}$$

Taking the difference of the Renyi entropies for any two different price levels brings the following equation which allows estimation of parameter m from the real data:

$$H_q(r|x) - H_q(r|x') = (m_q(x, x') - 1)(\ln x - \ln x') \tag{3.16}$$

The estimation procedure from the above equation is the following:

1. start with the time series of prices (X_t) and the corresponding time series of log-returns (R_t),
2. the interval $[\min(R_t), \max(R_t)]$ is divided into M equal bins,
3. for every time series of prices X_k corresponding to window T_k, build an equidistant N-points grid of the interval $[\min(X_k), \max(X_k)]$ of the form $\min(X_k) = x_k^1, ..., x_k^n, ..., x_k^N = \max(X_k)$,
4. this grid is used to construct the data to estimate m for the linear regression model 3.16, i.e. $N(N-1)/2$ unique pairs of the form (x_k^m, x_k^n), where $m < n$,
5. for the simplicity the Shannon entropy, a trivial case of the Renyi entropy, is used in the linear regression equation 3.16.

The left-hand side of 3.16 requires the knowledge of density of log-returns distribution conditioned on the price level $f(r|x)$. One possibility to approximate this density is to use the empirical distribution. In this case $f(r_i|x)$ is the empirical probability of the returns in the $i-th$ bin for the given price level x and the Shannon entropy is calculated according to the formula:

$$H(r|x) = -\sum_{i=1}^{M} f(r_i|x) \log f(r_i|x)) \tag{3.17}$$

.

Another possibility to estimate the conditional return distribution is using the formula 3.14 derived from the CEV model. Though the distribution in this case has a closed form, it is a function of parameter m which is unknown upfront. Therefore, in this case only the assumption about the functional form (i.e. normal distribution) of the log-returns distribution is made. Under this assumption the differential entropy has the following form (see Appendix A.2 for the derivation):

$$H(r|x) = \ln(\sqrt{2\pi e}\sigma), \tag{3.18}$$

depending only on the standard deviation, which can be estimated empirically.

3.3 "Best fit" method

This method represents a search for the value of m that would yield a sample of

$$dW_t = \frac{dX_t - \mu X_t dt}{\sigma X_t^m} \tag{3.19}$$

values which fits the normal distribution best.

The Pearson's chi-square test for normality (see Appendix A.4.5 for the description of the test) is used and m is selected on an interval in a such way, that p-value of the test is ≥ 0.05 and is the highest among others.

3.4 Modified model

Instead of initial model 2.1, we can consider a modified model, where drift is not constant but is a function of time:

$$dX_t = \mu(t)X_t dt + \sigma X_t^m dW_t \tag{3.20}$$

where W_t is a Brownian motion.

So, instead of using constant return for the whole time window

$$\mu = ln(X_T \setminus X_0), \tag{3.21}$$

the log-price process is described by power law function of time, i.e.

$$\ln X_t = at^{\alpha} + b, \tag{3.22}$$

where unknown parameters are ordinary least squares estimates.

An example of such approximation of log-price process is shown on the Figure 3.1. And in this case the drift is defined by

$$\mu(t) = ln(X_t \setminus X_{t-1}) = a(t^{\alpha} - (t-1)^{\alpha}). \tag{3.23}$$

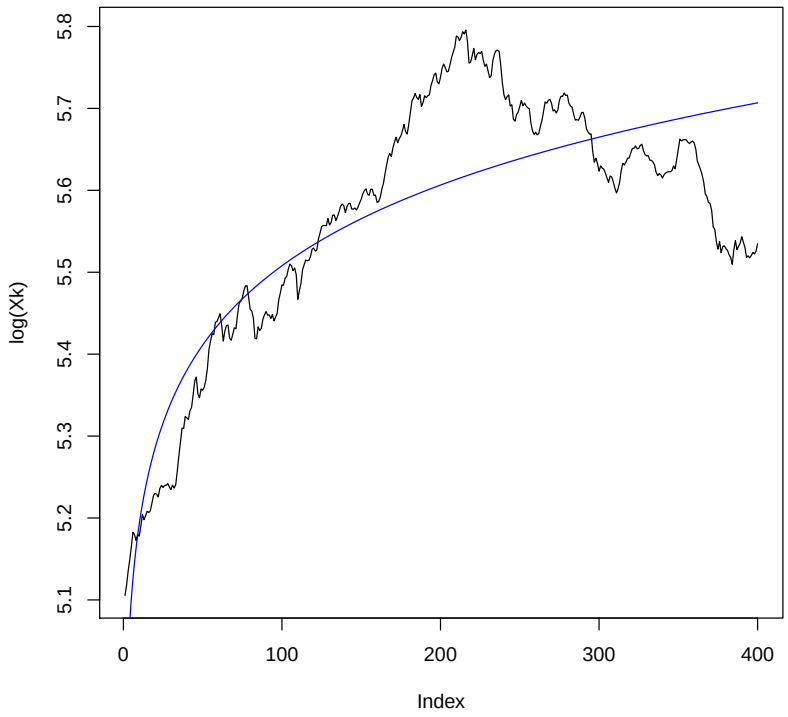


Figure 3.1: Approximation of log-price process by power-law function.

3.5 Model validity

After the parameter m is estimated the correctness of CEV model 2.1 should be checked. For this reason for the estimated m, the distribution of variables

$$dW_t = \frac{dX_t - \mu X_t dt}{\sigma X_t^m} \tag{3.24}$$

should be checked for normality.

The following test are used in the current study (see Appendix A.3 for the detailed description of these tests):

1. Anderson-Darling test,
2. Jarque-Bera test,
3. Shapiro-Wilk test,
4. Cramer von Mises test,
5. Pearson's chi-square test.

Both methods for estimation parameter m use linear regression model. The estimation from the linear regression model depends on several assumptions which need to be checked using regression diagnostics. Diagnostic techniques can be numerical, which are narrower in scope, but require no intuition, or graphical, which are more flexible but harder to interpret. In the present work numerical techniques are used.

The following residual assumptions need to be checked (see Appendix A.4 for the detailed description):

1. Constant variance,
2. Normality,
3. Serial correlation.

Another possibility is to use BDS test, first devised by W.A. Brock, W. Dechert and J. Scheinkman in 1987. BDS test is a powerful tool for detecting serial dependence in time series.

The test is defined as (see Appendix A.5 for the description):

H_0: $\left\{X_i, i = \overline{1,n}\right\}$ are i.i.d.,

H_a: $\left\{X_i, i = \overline{1,n}\right\}$ are not i.i.d..
and can be applied to:

1. dW_t, if the assumption of normality is changed for i.i.d.,

2. the residuals in the linear regression models 3.5, 3.16.

In particular, when applied to the residuals from a fitted linear time series model, the BDS test can be used to detect remaining dependence and the presence of omitted nonlinear structure. If the null hypothesis cannot be rejected, then the original linear model cannot be rejected; if the null hypothesis is rejected, the fitted linear model is mis-specified, and in this sense, it can also be treated as a test for nonlinearity.

Chapter 4

Accuracy of the estimation methods

The described algorithms for estimation parameter m are tested against stochastic benchmarks, when the parameter m is given by construction:

1. Construction of new price trail using reshuffled log-returns,
2. Simulation from the CEV model.

4.1 Construction of new price trail using reshuffled log-returns

An artificial price trail can be constructed using random reshuffling of the real log-returns. The main characteristic of the new trail - it is purely random as by construction any trend in the prices is destroyed. Therefore, in general, we expect the new price trail follows the standard log-normal Black-Scholes diffusion process (CEV, $m = 1$) with the relaxed assumption on dW_t, which are not normally distributed.

To construct the new price trail using reshuffled log-returns the following procedure is implemented:

1. Given the existing price trail $(X_t)_{t\in T}$ and the corresponding trail of log-returns is $(R_t)_{t\in T}$, where $R_t = \ln(\frac{X_t}{X_{t-1}})$,

2. the trail of log-returns $(R_t)_{t\in T}$ is then randomly permuted to obtain a new trail $(\widetilde{R}_t)_{t\in T}$, which is used to build a new price process according to the rule:

$$\widetilde{X}_t = e^{\widetilde{R}_t}\widetilde{X}_{t-1}, \tag{4.1}$$

where $\widetilde{X}_0 = X_0$.

The methods provide the following results:

Macbeth-Merville method							
min	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**	**BDS**
0.95	0.99	1.00	1.01	1.05	1.00	0.02	95%
Macbeth-Merville method, modified model							
0.93	0.98	0.99	1.01	1.05	0.99	0.02	96%
"Best fit" method							
0.85	0.95	1.05	1.05	1.15	1.02	0.07	
"Best fit" method, modified model							
0.85	0.95	1.05	1.05	1.15	1.02	0.07	
Entropy method, M=20, N=20							
0.81	1.10	1.21	1.33	1.57	1.21	0.17	0%
M=20, N=20, semi-empirical distr.							
0.46	0.95	1.09	1.23	1.64	1.09	0.22	2%
M=20, N=50							
0.74	0.98	1.09	1.16	1.42	1.08	0.14	0%
M=20, N=50, semi-empirical distr.							
0.70	0.97	1.07	1.17	1.31	1.06	0.15	0%
M=20, N=100							
0.67	0.94	1.04	1.12	1.43	1.03	0.14	0%
M=20, N=100, semi-empirical distr.							
0.77	0.96	1.03	1.09	1.27	1.03	0.10	0%
M=40, N=50							
0.71	0.96	1.12	1.28	1.54	1.12	0.20	0%
M=40, N=50, semi-empirical distr.							
0.73	0.95	1.05	1.13	1.37	1.05	0.13	0%

Table 4.1: Reshuffled returns, NASDAQ daily series for the period 1980-2010, number of simulations 1000.

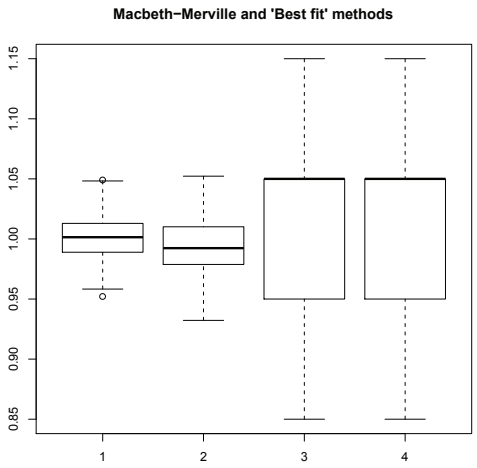


Figure 4.1: Macbeth-Merville and "Best fit" methods, reshuffled returns, box-plots of results. First and second box-plots from the left are for Macbeth-Merville method applied to initial and modified models, respectively. Third and fourth box-plots from the left are for "Best fit" method applied to initial and modified models, respectively. Reshuffling is done for NASDAQ daily series for the period 1980-2010, number of simulations 1000.

Macbeth-Merville method, applied to both the initial and modified models, as it is seen from the Table 4.1, provides better estimations than both "Best fit" and Entropy methods, though "Best fit" provides better estimates than Entropy method. It should be noted that the 50% and 75% quantiles for the "Best fit" model coincide. This is due to the fact that m is chosen with accuracy up to 0.05. The linear model 3.5 in Macbeth-Merville method is rejected only in 4-5% cases according to BDS test, while for Entropy method there is a non-linear relation in the model 3.16. Analyzing Entropy method itself, based on these estimation results, it is difficult to conclude if the model with empirical or with semi-empirical distribution of log-returns provides better results. Nonetheless, as it is seen, estimation using semi-empirical distribution with the values of $M = 20$, $N = 100$ provides more accurate results. For the convenience, these results are presented in the form of box-plots on the Figures 4.1 and 4.2.

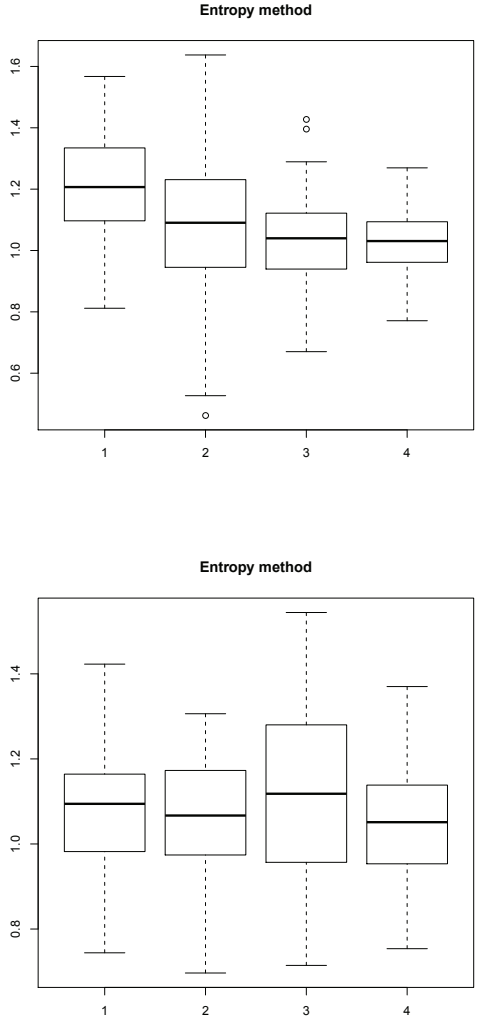
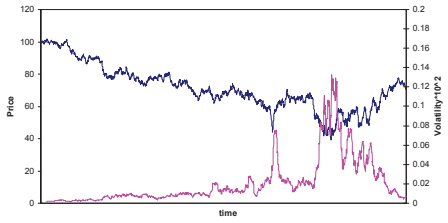


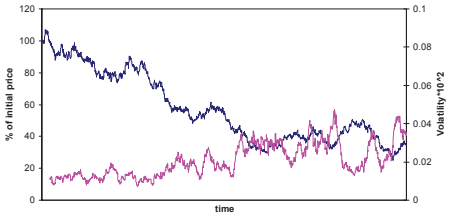
Figure 4.2: Upper figure, first and second box-plots from the left are for the values $M = 20$, $N = 20$ using empirical and semi-empirical distribution of log-returns, respectively. Upper figure, third and fourth box-plots from the left are for the values $M = 20$, $N = 100$ using empirical and semi-empirical distribution of log-returns, respectively. Lower figure, first and second box-plots from the left are for the values $M = 20$, $N = 50$ using empirical and semi-empirical distribution of log-returns, respectively. Lower figure, third and fourth box-plots from the left are for the values $M = 40$, $N = 50$ using empirical and semi-empirical distribution of log-returns, respectively. Reshuffling is done for NASDAQ daily series for the period 1980-2010, number of simulations 1000.

4.2 Simulation from the CEV model

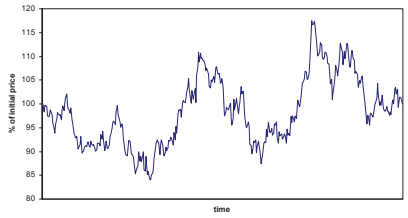
Another benchmark for testing the proposed algorithms is the CEV model 2.1 itself. Selected simulated paths for different values of m, are shown on the Figure 4.3. The paths simulated from the model with $m > 1$ exhibit periods of faster-than-exponential growth with singularities.



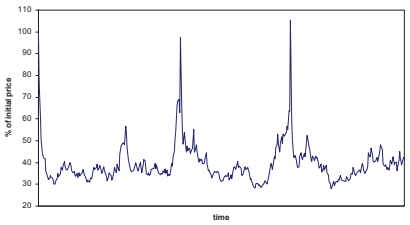
(a) m=-1.5



(b) m=0.5



(c) m=1.5



(d) m=3

Figure 4.3: Simulation results from the CEV model, for values of m=-1.5, 0.5, 1.5, 3. On the figures a) and b) volatilities are shown with pink, so that the negative relation between the price and the volatility is seen.

Macbeth-Merville method applied to the initial model 2.1 provides the following results:

Macbeth-Merville method									
window size	**true m**	**min**	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**	**BDS**
200	0.5	-1.15	0.25	0.49	0.84	2.07	0.49	0.53	93.6%
250	0.5	-1.18	0.25	0.46	0.70	1.98	0.46	0.41	91.4%
300	0.5	-1.49	0.31	0.49	0.67	1.80	0.47	0.35	92.6%
350	0.5	-0.72	0.32	0.50	0.66	1.65	0.48	0.30	93.8%
400	0.5	-0.72	0.34	0.51	0.63	1.88	0.48	0.28	93%
200	0.7	-0.31	0.52	0.70	0.85	1.58	0.68	0.31	93.6%
250	0.7	-0.31	0.55	0.68	0.80	1.61	0.67	0.23	91.8%
300	0.7	-0.52	0.59	0.69	0.79	1.44	0.68	0.20	93.2%
350	0.7	-0.08	0.61	0.70	0.78	1.30	0.69	0.17	93.6%
400	0.7	0.02	0.61	0.70	0.77	1.44	0.69	0.16	93.4%
200	1	0.52	0.91	0.98	1.05	1.39	0.98	0.13	94.4%
250	1	0.58	0.93	0.99	1.05	1.31	0.99	0.10	91%
300	1	0.66	0.95	0.99	1.04	1.29	0.99	0.08	92.4%
350	1	0.71	0.96	1.00	1.03	1.30	0.99	0.07	92.8%
400	1	0.78	0.96	1.00	1.02	1.20	0.99	0.06	92.6%
200	1.3	0.95	1.24	1.29	1.33	2.23	1.29	0.1	56.2%
250	1.3	1.04	1.24	1.30	1.33	2.23	1.29	0.1	57.8%
300	1.3	1.07	1.25	1.29	1.33	2.23	1.30	0.09	57.8%
350	1.3	1.12	1.25	1.29	1.32	2.23	1.30	0.08	59.4%
400	1.3	1.13	1.26	1.29	1.32	2.23	1.30	0.08	59.8%
200	1.5	0.91	1.40	1.48	1.57	2.01	1.48	0.14	90.6%
250	1.5	0.94	1.40	1.49	1.57	1.98	1.49	0.13	92.6%
300	1.5	1.13	1.43	1.49	1.55	1.93	1.49	0.10	93.6%
350	1.5	1.20	1.43	1.50	1.54	1.82	1.49	0.09	93.6%
400	1.5	1.23	1.44	1.50	1.54	1.87	1.49	0.08	94.8%
200	1.7	1.14	1.58	1.67	1.77	2.90	1.68	0.17	89%
250	1.7	0.84	1.58	1.68	1.77	2.90	1.68	0.15	90.2%
300	1.7	0.84	1.61	1.69	1.75	2.90	1.68	0.14	91.6%
350	1.7	0.84	1.62	1.69	1.75	2.90	1.69	0.12	93%
400	1.7	0.84	1.63	1.69	1.75	2.90	1.69	0.12	93%

Table 4.2: Macbeth-Merville method, initial model, distributions of estimations as a result of simulations from the CEV model for different m and window size T_k, number of simulations 1000.

As it is seen from the tables 4.2, 4.3 and 4.4, the estimations are less volatile the longer the time window. Comparing Macbeth-Merville method applied to different models, the distribution of the estimations for the modified model is less volatile than of the initial model with better results for the means and medians. "Best fit" method provides more volatile results with underestimated means and medians.

Macbeth-Merville method applied to the modified model 3.20 provides the following results:

Macbeth-Merville method, modified model									
window size	**true m**	**min**	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**	**BDS**
200	0.5	-0.13	0.40	0.50	0.58	1.10	0.49	0.16	92.6%
250	0.5	-0.06	0.42	0.50	0.57	0.91	0.49	0.14	91.6%
300	0.5	0.03	0.43	0.50	0.57	0.91	0.49	0.12	91%
350	0.5	0.12	0.44	0.50	0.57	0.91	0.50	0.11	91%
400	0.5	0.13	0.44	0.50	0.57	0.91	0.50	0.10	91.6%
200	0.7	0.14	0.62	0.69	0.76	1.11	0.68	0.14	91.8%
250	0.7	0.21	0.64	0.70	0.76	1.10	0.69	0.11	90.8%
300	0.7	0.35	0.65	0.70	0.75	0.98	0.70	0.09	92.4%
350	0.7	0.32	0.66	0.70	0.74	1.02	0.70	0.08	93%
400	0.7	0.41	0.66	0.70	0.74	0.95	0.70	0.07	93.4%
200	1	0.38	0.92	0.99	1.05	1.34	0.98	0.13	90.4%
250	1	0.61	0.94	1.00	1.05	1.31	1.00	0.10	90.8%
300	1	0.71	0.95	1.00	1.04	1.29	1.00	0.08	91.2%
350	1	0.77	0.97	1.00	1.03	1.31	1.00	0.07	93.2%
400	1	0.79	0.97	1.00	1.03	1.21	1.00	0.06	94.6%
200	1.3	0.89	1.20	1.29	1.37	1.68	1.29	0.13	90.6%
250	1.3	1.00	1.23	1.29	1.37	1.61	1.30	0.11	91.4%
300	1.3	1.02	1.24	1.30	1.35	1.76	1.30	0.09	91.8%
350	1.3	1.03	1.25	1.30	1.34	1.64	1.30	0.08	93.4%
400	1.3	1.08	1.25	1.30	1.34	1.56	1.30	0.07	93.4%
200	1.5	1.04	1.40	1.49	1.59	1.92	1.49	0.14	90.8%
250	1.5	1.11	1.42	1.50	1.57	1.86	1.50	0.12	90.8%
300	1.5	1.19	1.43	1.50	1.56	2.01	1.50	0.11	90.6%
350	1.5	1.22	1.45	1.50	1.55	1.85	1.50	0.09	92.8%
400	1.5	1.26	1.45	1.50	1.55	1.84	1.50	0.08	92.2%
200	1.7	1.22	1.59	1.69	1.78	2.55	1.69	0.15	90%
250	1.7	1.30	1.61	1.70	1.78	2.56	1.70	0.14	90.8%
300	1.7	1.34	1.63	1.70	1.77	2.56	1.70	0.13	90%
350	1.7	1.39	1.63	1.70	1.76	2.56	1.70	0.11	92.2%
400	1.7	1.41	1.64	1.70	1.76	2.56	1.70	0.11	91.8%

Table 4.3: Macbeth-Merville method, modified model, distributions of estimations as a result of simulations from the CEV model for different m and window size T_k, number of simulations 1000.

"Best fit" method applied to the initial model 2.1 provides the following results:

"Best fit" method								
window size	**true m**	**min**	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**
200	0.5	-0.95	0.2	0.55	0.81	1.45	0.47	0.49
250	0.5	-0.60	0.25	0.45	0.70	1.65	0.43	0.45
300	0.5	-0.95	0.10	0.40	0.66	1.10	0.35	0.46
350	0.5	-0.75	0.15	0.40	0.70	1.15	0.39	0.39
400	0.5	-0.65	0.19	0.45	0.70	1.10	0.39	0.37
200	0.7	-0.85	0.40	0.60	0.80	1.85	0.62	0.39
250	0.7	-0.40	0.34	0.65	0.85	1.45	0.59	0.35
300	0.7	-0.65	0.35	0.65	0.80	1.35	0.58	0.33
350	0.7	-0.05	0.50	0.60	0.71	1.25	0.58	0.25
400	0.7	-0.25	0.55	0.70	0.81	1.10	0.68	0.21
200	1	-0.50	0.64	0.85	1.11	1.95	0.86	0.40
250	1	-0.30	0.74	0.85	1.00	1.70	0.85	0.32
300	1	0.20	0.75	0.85	1.00	1.45	0.86	0.24
350	1	0.25	0.75	0.90	1.05	1.35	0.89	0.22
400	1	0.25	0.85	0.95	1.05	1.40	0.93	0.20
200	1.3	0.55	0.85	1.05	1.41	2.45	1.14	0.38
250	1.3	0.45	0.90	1.10	1.35	2.30	1.12	0.33
300	1.3	0.50	0.95	1.10	1.30	1.90	1.12	0.26
350	1.3	0.55	1.04	1.15	1.40	1.90	1.20	0.25
400	1.3	0.45	1.05	1.20	1.36	1.80	1.20	0.24
200	1.5	0.45	1.04	1.30	1.70	2.50	1.37	0.41
250	1.5	0.55	1.05	1.30	1.55	2.30	1.31	0.35
300	1.5	0.50	1.10	1.30	1.46	2.10	1.30	0.30
350	1.5	0.70	1.20	1.35	1.55	2.15	1.35	0.27
400	1.5	0.60	1.30	1.45	1.65	2.10	1.43	0.28
200	1.7	0.30	1.24	1.50	1.90	2.65	1.57	0.49
250	1.7	0.50	1.20	1.50	1.75	3.30	1.50	0.41
300	1.7	0.65	1.25	1.48	1.75	2.40	1.51	0.33
350	1.7	0.80	1.35	1.50	1.70	2.50	1.52	0.30
400	1.7	0.70	1.45	1.65	1.81	2.45	1.63	0.33

Table 4.4: "Best fit" method, distributions of estimations as a result of simulations from the CEV model for different m and window size T_k, number of simulations 1000. The value of m is selected from the interval [-4,4] with accuracy up to 0.05.

Entropy method using empirical distribution for log-returns provides the following results:

Entropy method								
window size	**true m**	**min**	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**
200	0.5	-1.43	0.44	1.15	1.95	5.34	1.25	1.28
250	0.5	-1.53	0.61	1.21	1.88	5.68	1.33	1.21
300	0.5	-2.75	0.63	1.22	1.70	4.34	1.23	1.03
350	0.5	-2.09	1.63	1.15	1.55	4.66	1.15	1.01
400	0.5	-0.44	0.53	1.02	1.71	2.88	1.14	0.82
200	0.7	-0.41	0.66	1.14	1.54	2.91	1.11	0.70
250	0.7	-0.24	0.79	1.17	1.58	2.88	1.19	0.63
300	0.7	-0.28	0.70	1.09	1.70	2.71	1.19	0.76
350	0.7	-0.31	0.71	1.21	1.66	2.11	1.14	0.65
400	0.7	0.31	0.83	1.09	1.73	2.13	1.24	0.61
200	1	0.01	0.65	0.96	1.27	2.44	0.98	0.42
250	1	0.10	0.75	1.06	1.26	2.10	1.01	0.35
300	1	0.10	0.86	1.07	1.25	1.82	1.04	0.34
350	1	0.20	0.86	1.08	1.23	1.56	1.03	0.29
400	1	0.49	0.94	1.08	1.25	1.65	1.07	0.24
200	1.3	0.12	0.49	0.83	0.91	1.03	0.69	0.35
250	1.3	-0.19	0.64	0.96	1.20	1.92	0.91	0.40
300	1.3	-0.24	0.87	1.26	1.65	3.39	1.28	0.59
350	1.3	-0.36	0.67	1.07	1.22	1.88	0.95	0.40
400	1.3	0.19	0.78	1.03	1.22	1.73	0.99	0.31
200	1.5	-0.33	0.53	0.92	1.21	2.34	0.88	0.53
250	1.5	-0.69	0.95	1.36	1.88	4.27	1.48	0.88
300	1.5	-0.25	0.97	1.34	1.80	5.17	1.43	0.81
350	1.5	-0.40	1.06	1.45	1.77	4.88	1.47	0.71
400	1.5	0.33	1.11	1.49	1.71	3.28	1.45	0.51
200	1.7	-1.65	0.86	1.38	2.17	5.87	1.63	1.31
250	1.7	-0.68	1.01	1.43	2.21	5.84	1.63	1.11
300	1.7	-0.56	1.04	1.49	2.01	5.80	1.59	1.00
350	1.7	-0.78	1.10	1.57	1.92	5.14	1.64	0.95
400	1.7	0.32	1.27	1.62	1.87	4.72	1.60	0.62

Table 4.5: Entropy method, empirical log-returns distribution, distributions of estimations as a result of simulations from CEV model for different m, $M = 20$, $N = 100$.

Estimations with Entropy method using empirical distribution of log-returns provide highly volatile results as it is seen in the Table 4.5. Using semi-empirical distribution instead, provides comparably better results 4.6, but they are still more volatile then those achieved with other methods. Using BDS test the linearity in the model 3.16 is rejected for the simulation results.

Entropy method using semi-empirical distribution for log-returns provides the following results:

Entropy method, semi-empirical distribution								
window size	**true m**	**min**	**25% q.**	**50% q.**	**75% q.**	**max**	**mean**	**sd**
200	0.5	-0.74	0.26	0.49	0.75	1.54	0.51	0.42
250	0.5	-0.80	0.31	0.47	0.71	1.41	0.47	0.35
300	0.5	-0.49	0.30	0.48	0.63	1.50	0.47	0.32
350	0.5	-0.62	0.34	0.50	0.62	1.46	0.47	0.30
400	0.5	-0.06	0.36	0.48	0.74	1.15	0.53	0.27
200	0.7	-0.05	0.58	0.71	0.89	1.44	0.72	0.24
250	0.7	-0.06	0.60	0.70	0.82	1.10	0.70	0.20
300	0.7	0.43	0.57	0.74	0.79	1.20	0.71	0.19
350	0.7	0.43	0.58	0.74	0.81	0.99	0.72	0.16
400	0.7	0.54	0.61	0.67	0.86	1.00	0.73	0.16
200	1	0.73	0.95	1.02	1.10	1.32	1.03	0.12
250	1	0.77	0.96	1.02	1.07	1.32	1.02	0.11
300	1	0.76	0.97	1.01	1.07	1.32	1.02	0.09
350	1	0.78	0.96	1.01	1.07	1.39	1.02	0.08
400	1	0.87	0.97	1.01	1.06	1.24	1.02	0.07
200	1.3	1.25	1.28	1.33	1.55	1.62	1.40	0.17
250	1.3	1.04	1.25	1.33	1.40	1.82	1.33	0.13
300	1.3	1.10	1.25	1.31	1.39	1.83	1.32	0.11
350	1.3	1.05	1.25	1.31	1.39	1.83	1.32	0.11
400	1.3	1.14	1.25	1.33	1.39	1.74	1.32	0.10
200	1.5	0.05	1.39	1.51	1.60	1.99	1.50	0.31
250	1.5	0.32	1.34	1.48	1.60	2.46	1.48	0.29
300	1.5	0.98	1.41	1.49	1.62	2.27	1.52	0.27
350	1.5	1.01	1.40	1.50	1.63	2.08	1.52	0.18
400	1.5	1.16	1.42	1.52	1.61	2.21	1.53	0.16
200	1.7	0.26	1.53	1.67	1.81	3.34	1.70	0.38
250	1.7	0.77	1.53	1.69	1.84	2.67	1.69	0.30
300	1.7	0.87	1.61	1.69	1.82	2.36	1.72	0.23
350	1.7	1.21	1.61	1.70	1.81	2.23	1.72	0.17
400	1.7	1.33	1.63	1.71	1.82	2.50	1.73	0.17

Table 4.6: Entropy method, semi-empirical log-returns distribution, distributions of estimations as a result of simulations from CEV model for different m, $M = 20$, $N = 100$.

Chapter 5

Results

5.1 Macbeth-Merville method

In the present work NASDAQ index daily prices for the period 1980-2010 are analyzed. Based on the results of the previous section, Macbeth-Merville method is applied to the modified model with the length of moving time window $T_k = 400$ days and 20 days shift.

On the first graph of the Figure 5.1 the estimated values of m are presented, while on the second graph of the same figure there are the corresponding volatilities of log-returns. It is seen, that the periods where estimated $m < 1$ are those where there is an inverse relationship between the price and the volatility. For example, a decline of the price in October 1987 increased the volatility of log-returns and the consequent price growth decreased it. Another example, in the period 2005-2007 the estimated $m < 1$: while the price is growing the volatility is declining. Figure 5.2 provides zoom for the period 1987-1993, and it is seen that estimated m shows a change of regime.

Regression diagnostics for Macbeth-Merville method shows that the equation 3.5 can be used as a linear regression equation. And though the expected value of the term $\ln[\chi^2(1)]$ is not equal to zero, it sums to the estimator of an intercept coefficient, as shown in the equation 3.9. Residuals are shown to be uncorrelated and homoscedastic, but not normally distributed. With the latter, least squares estimates are still the best linear unbiased estimates.

As with the artificial time series in the previous section, it is interesting to analyze the sensitivity of the estimation to the amount of data available. For this purpose the estimation is done for different sizes of time windows T_k. As it is seen from the Table 5.1, the values of m, estimated for the time

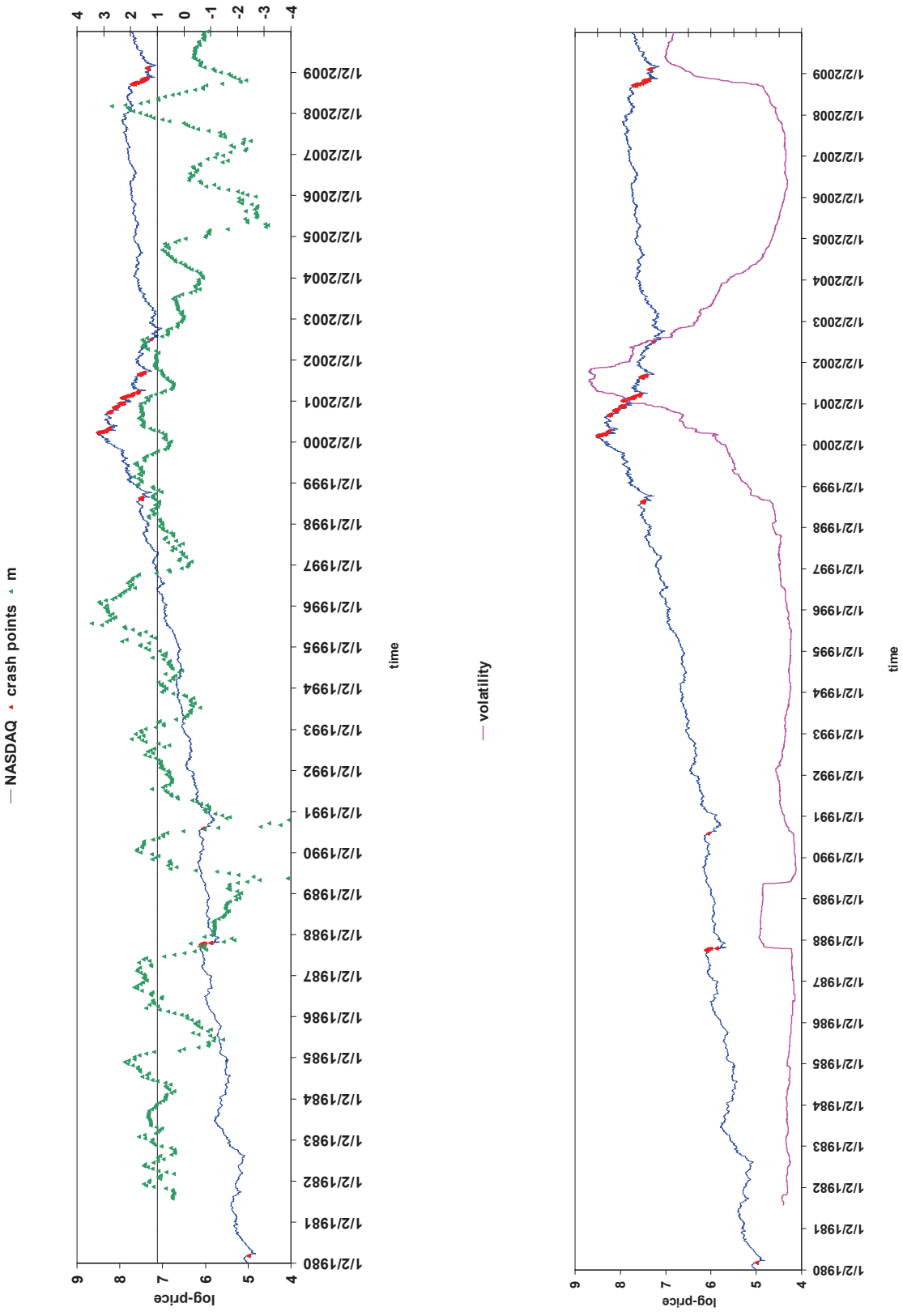
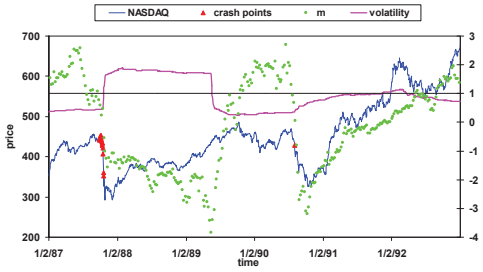


Figure 5.1: Macbeth-Merville method, estimated m for the NASDAQ time series, for the period 1980-2010, $T_k = 400$. Marked by red are "crash points", followed by a decline in price of more than 15% in less than three weeks. On the second graph volatilities of log-returns in the moving windows are shown.



(a) 1987-1993

Figure 5.2: Macbeth-Merville method, estimated m for the NASDAQ time series, for the period 1987-1993, $T_k = 400$, shift is 5 days. The volatility is re-scaled to fit the m axes. The red points are so called "crash points", which are followed by a decline in price of more than 15% in less than three weeks.

windows of similar lengths are highly correlated. Also from the Figure 5.7, the less the length of the time window the fatter the tails of the distribution of m. This happens because in the regression model the less data there is, the more influence the extreme points have, therefore, they pull the repression line, increasing the slope in the absolute value.

	T_k=200	T_k=250	T_k=300	T_k=350	T_k=400	T_k=800	sd
T_k=200	1	0.77	0.64	0.47	0.4	0.28	2.22
T_k=250	0.77	1	0.85	0.69	0.56	0.4	1.76
T_k=300	0.64	0.85	1	0.88	0.74	0.46	1.58
T_k=350	0.47	0.69	0.88	1	0.9	0.54	1.4
T_k=400	0.4	0.56	0.74	0.9	1	0.59	1.28
T_k=800	0.28	0.4	0.46	0.54	0.59	1	1.3

Table 5.1: Macbeth-Merville method, correlation coefficients and standard deviations of m series estimated for T_k=200, 250, 300, 350, 400, 800.

Based on the results from the Figures 5.3, 5.4, 5.5 and 5.6, the regions where the initial model 2.1 is shown to be correct coincide, but they are getting narrower for the longer time window T_k. So, for time windows of the shorter length the CEV model is a good descriptor of the price process. The same figures also present the power α from the equation 3.22 which

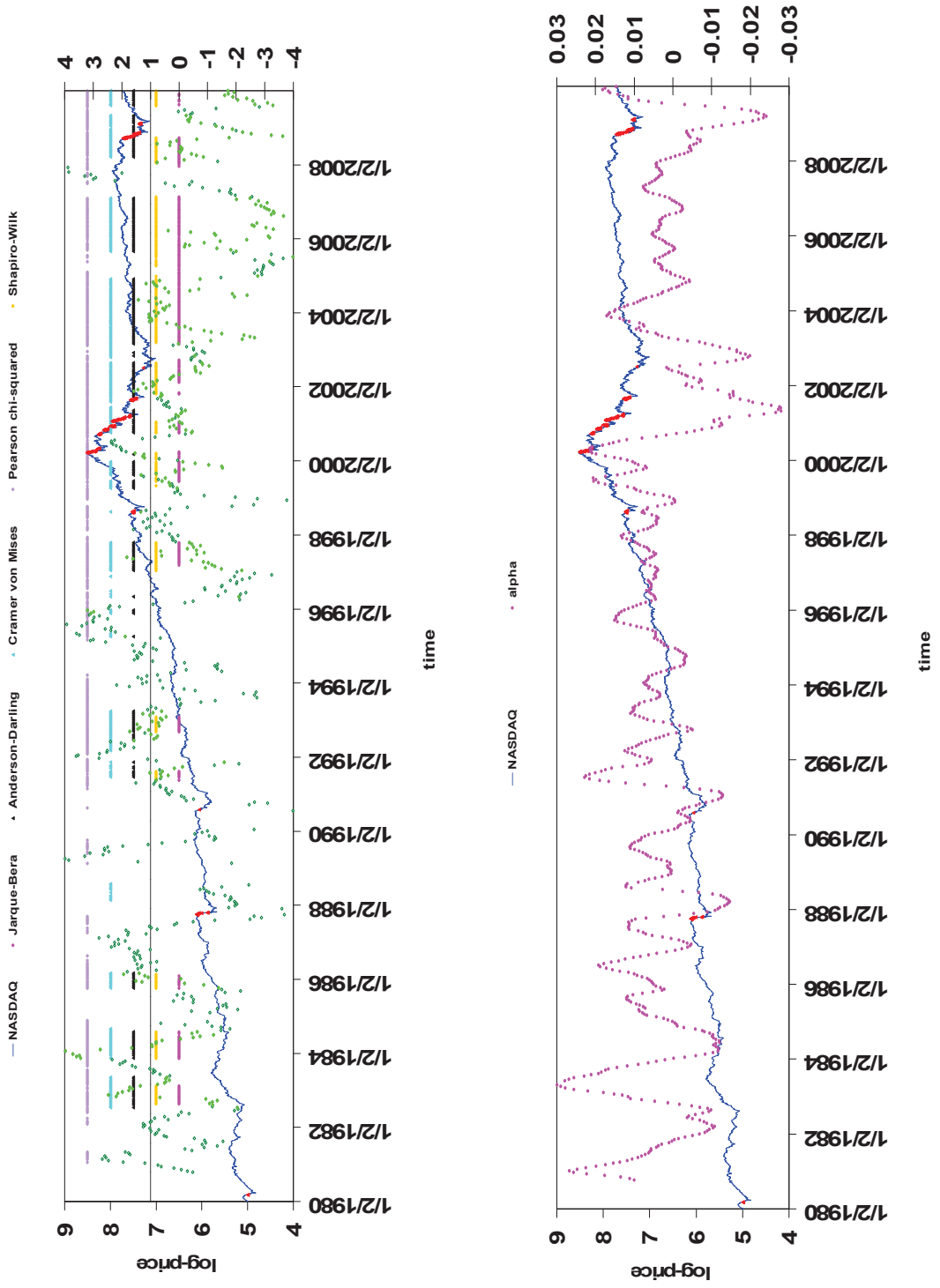


Figure 5.3: Macbeth-Merville method, estimated m for the NASDAQ time series, for the period 1980-2010, T_k=200. The colored markers are the regions where for the estimated m the null hypothesis about the normality of dW_t is not rejected. The estimated m marked with light green are those for which at least three normality tests agree. The right graph shows estimated power α for the modified model.

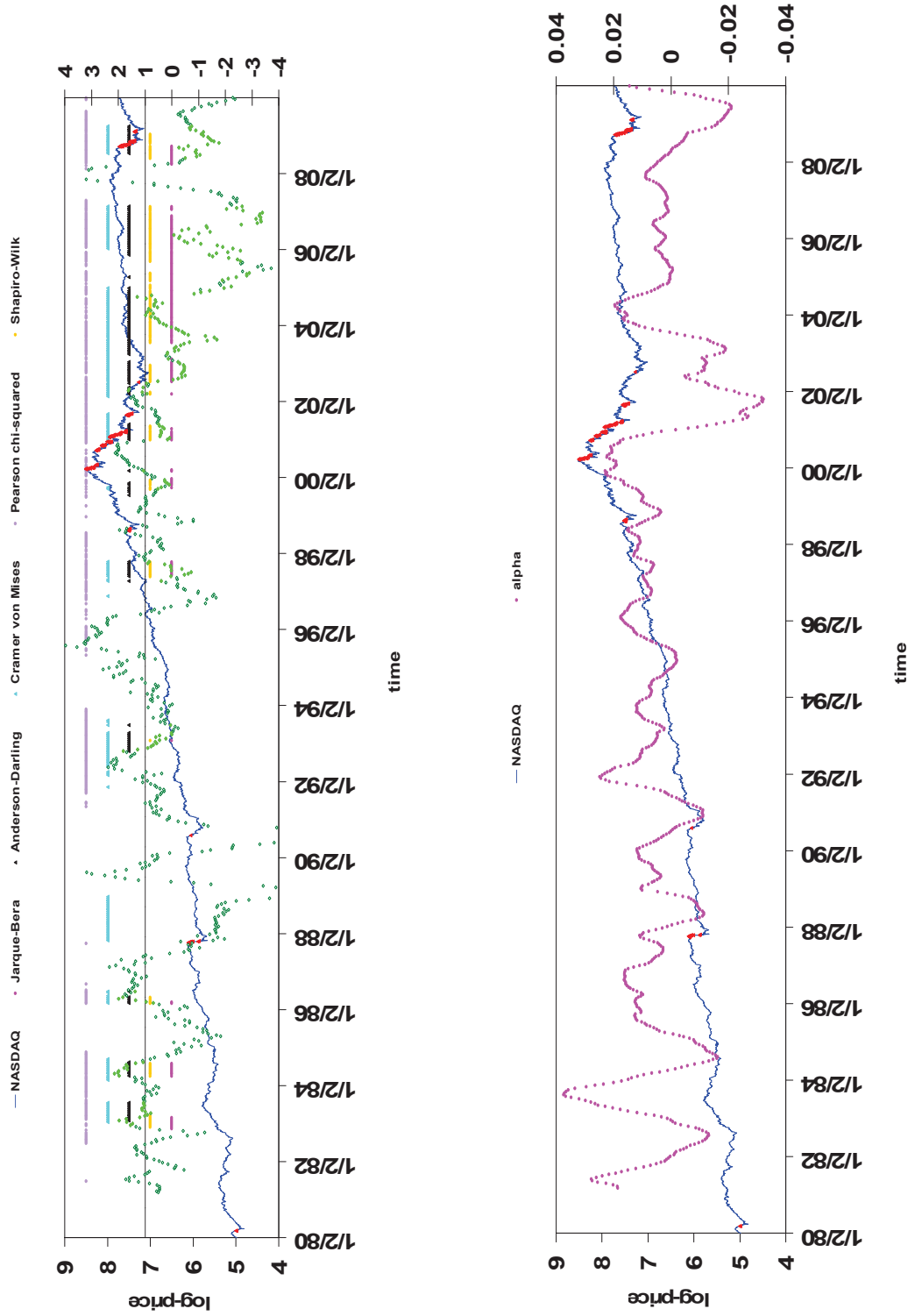


Figure 5.4: Macbeth-Merville method, estimated m for the NASDAQ time series, for the period 1980-2010, T_k=300. The colored markers are the regions where for the estimated m the null hypothesis about the normality of dW_t is not rejected. The estimated m marked with light green are those for which at least three normality tests agree. The right graph shows estimated power α for the modified model.

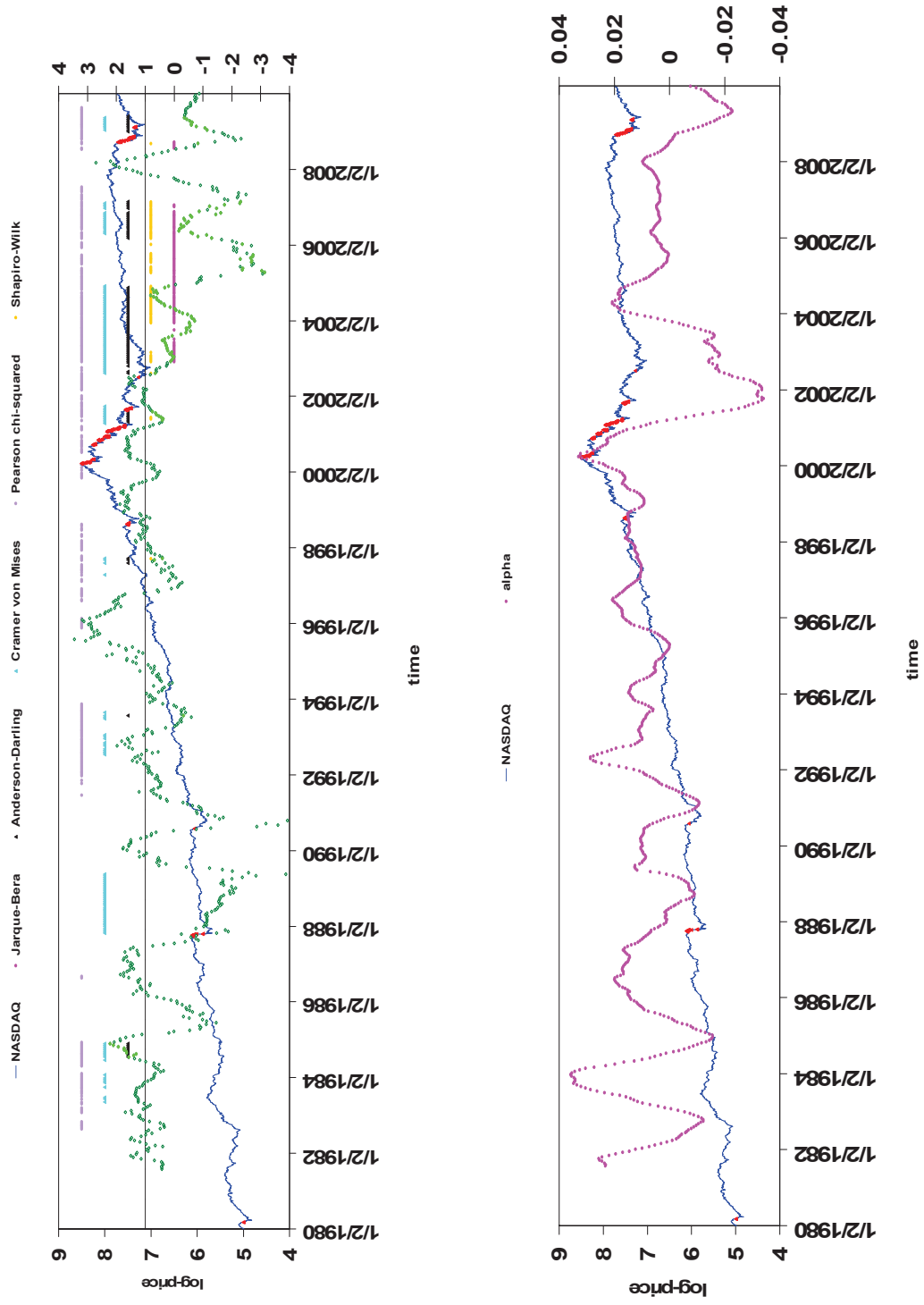


Figure 5.5: Macbeth-Merville method, estimated m for the NASDAQ time series, for the period 1980-2010, T_k=400. The colored markers are the regions where for the estimated m the null hypothesis about the normality of dW_t is not rejected. The estimated m marked with light green are those for which at least three normality tests agree. The right graph shows estimated power α for the modified model.

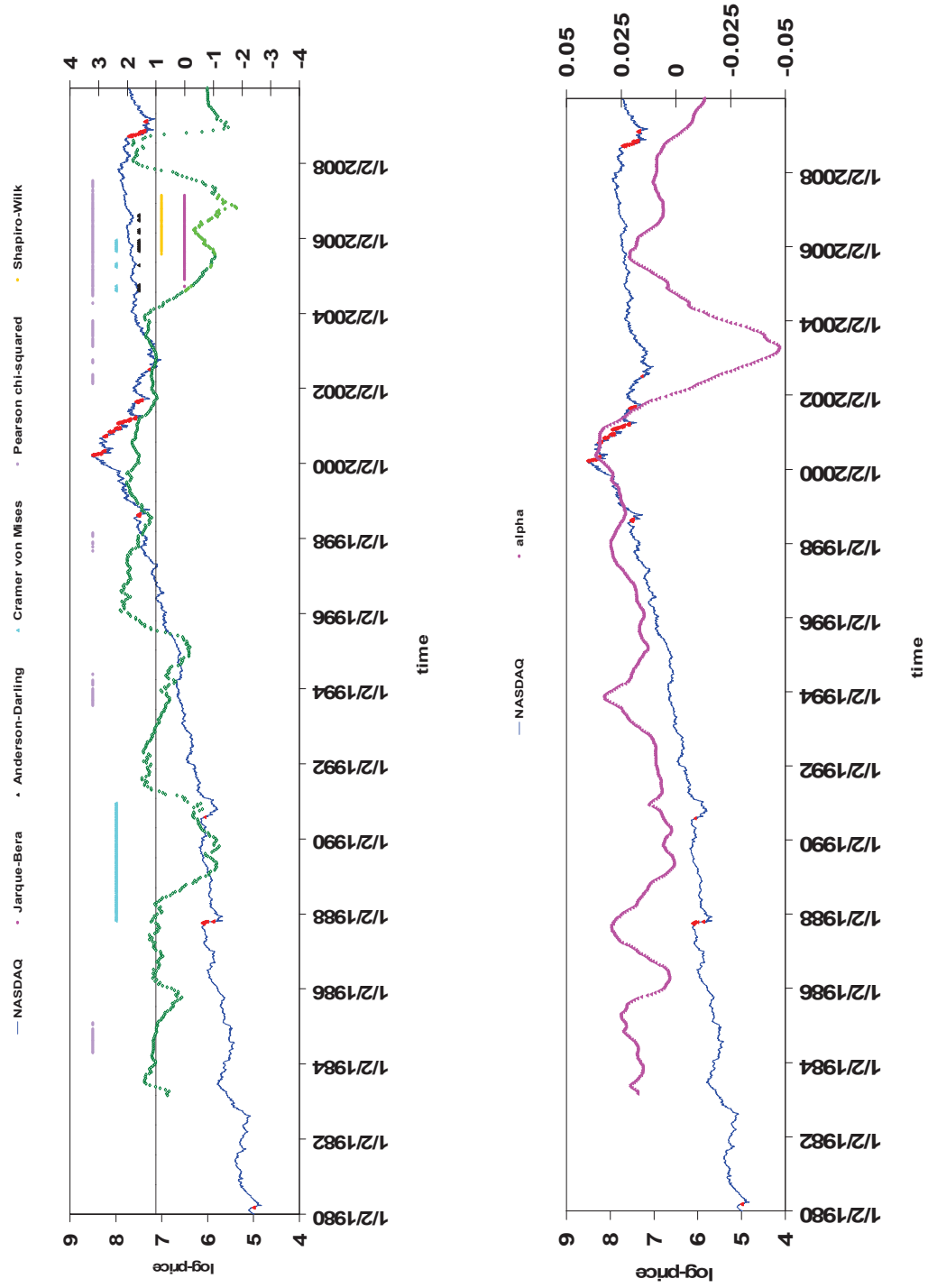


Figure 5.6: Macbeth-Merville method and "Best fit" methods, estimated m for the NASDAQ time series, for the period 1980-2010, T_k=800. The colored markers are the regions where for the estimated m the null hypothesis about the normality of dW_t is not rejected. The estimated m marked with light green are those for which at least three normality tests agree. The right graph shows estimated power α for the modified model.

has an important implication in relation to the finite-time singularity in the log-price process. As it is seen from the Figure 5.8, the longer the time window T_k the more negatively skewed the distribution of α.

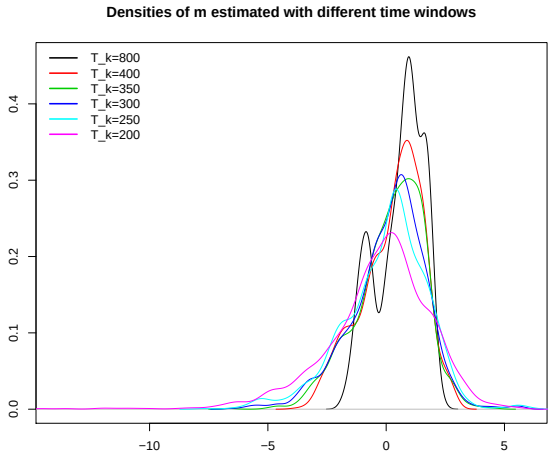


Figure 5.7: Macbeth-Merville method, distributions of m for different sizes of time window.

Applying BDS test to dW_t widens those regions where at least three normality tests agree. The results of BDS test, applied to the residuals from a fitted linear model, show that, in general, for different sizes of time window T_k, the linear model is not rejected. That means that for NASDAQ daily time series the linear model 3.5 in Macbeth-Merville method for the estimation of m is reasonable. These results are compared to those obtained under assumption that the price follows Black-Scholes model (Table 5.2).

On the Figure 5.9 the scatter plots for Macbeth-Merville method where estimated m is plotted against the volatility of log-returns in the following 20 days, are shown. Also linear regression lines are added - the trend suggests that for the higher m the volatility of log-returns in the following 20 days is expected to increase.

	T_k	initial model	modified model	B-S	B-S, modified
1	200	86.70%	89%	84.80%	87.50%
2	250	87.30%	88.50%	85.10%	87%
3	300	86.93%	88.86%	85.97%	87.90%
4	350	87.81%	89.20%	86.84%	87.12%
5	400	87.59%	87.58%	85.91%	86.33%
6	800	83.90%	78.58%	80.95%	76.96%

Table 5.2: Comparison of non-linearity diagnostic in 3.5 for Macbeth-Merville method for initial and modified models and Black-Scholes standard model with constant and non-constant drifts for different sizes of time window T_k. Table represents the percent of windows in the whole time series where the linear model 3.5 is not rejected by BDS test.

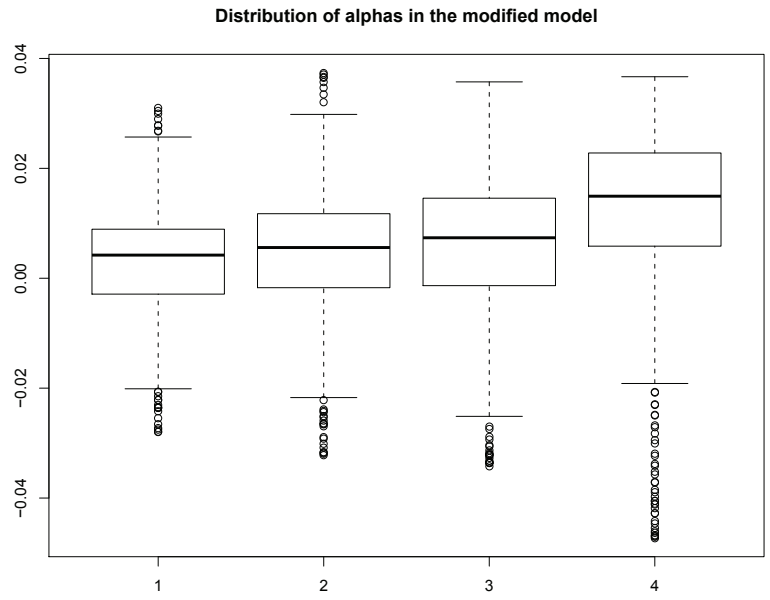
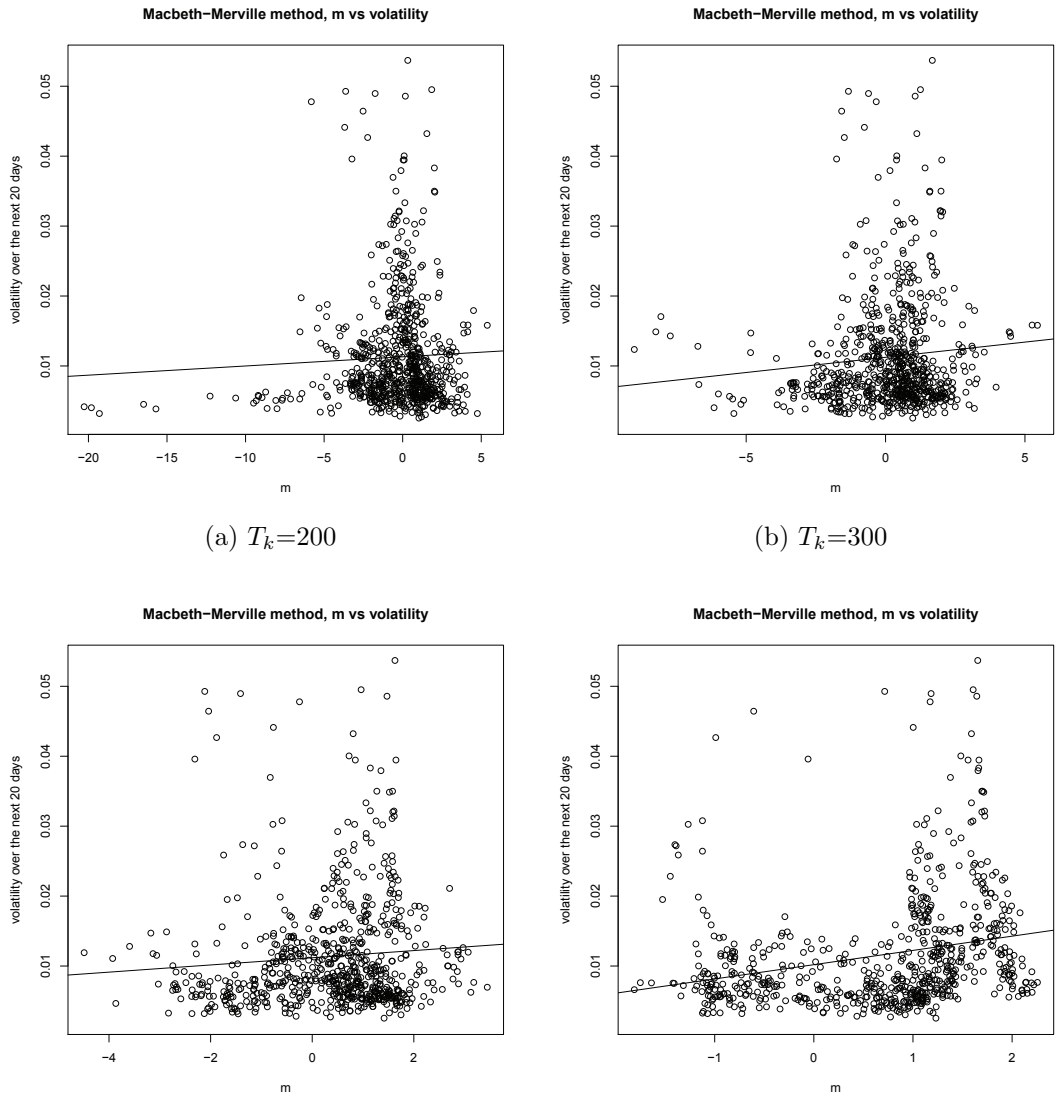


Figure 5.8: Distribution of α in the modified model for different sizes of time window T_k=200, 300, 400, 800 from the left to the right, respectively.



(a) T_k=200

(b) T_k=300

(c) T_k=400

(d) T_k=800

Figure 5.9: Macbeth-Merville method, scatter plots m vs volatility over the next 20 days for different sizes of time window T_k=200, 300, 400, 800.

5.2 Entropy method

The estimation of the parameter m using Entropy method depends on the values of the parameters N and M. Parameter N is responsible for the construction of the grid on the interval $[\min(X_k), \max(X_k)]$ and it influences how many data points there will be for the linear regression estimation 3.16. Therefore, it should be chosen in a such way that, from one side, there is enough data for the estimation from the regression model and the information is not lost, i.e. the log-returns between two consequent price levels. And, from the other side, the data is not redundant and we do not loose calculation capacity, i.e. if the grid is too dense, the entropies for the consequent price levels will coincide. Parameter M is dividing the interval $[\min(R_t), \max(R_t)]$ into bins to approximate the distribution of log-returns and also involves a trade-off between smaller and higher values.

On the Figure 5.10, the estimated m series for the different values of N are shown, the time window $T_k = 400$. As it is seen, the m series follow the same trend, though for the smaller values of N estimated m is higher. The same result is shown in the previous section on the Figure 4.2 and Table 4.1. Based on the results of the previous section the following values of the parameters are chosen for the further analysis $M = 20$, $N = 100$.

	T_k=200	T_k=250	T_k=300	T_k=350	T_k=400	T_k=800	sd
T_k=200	1	0.51	0.33	0.31	0.31	0.15	2.41
T_k=250	0.51	1	0.58	0.39	0.38	0.1	1.88
T_k=300	0.33	0.58	1	0.61	0.48	0.1	1.57
T_k=350	0.31	0.39	0.61	1	0.68	0.16	1.4
T_k=400	0.31	0.38	0.48	0.9	1	0.14	1.33
T_k=800	0.15	0.01	0.1	0.54	0.14	1	1.06

Table 5.3: Entropy method, semi-empirical log-returns distribution, correlation coefficients of m series estimated for T_k=200, 250, 300, 350, 400, 800, $M = 20$, $N = 100$.

Regression diagnostics for the Entropy method shows that the residuals are correlated and not normally distributed. Also, repeating the results of the previous section, BDS test rejects the linear model 3.16.

Figures 5.12 and 5.13 allow to compare the estimation results for Macbeth-Merville method applied to the modified model and Entropy method

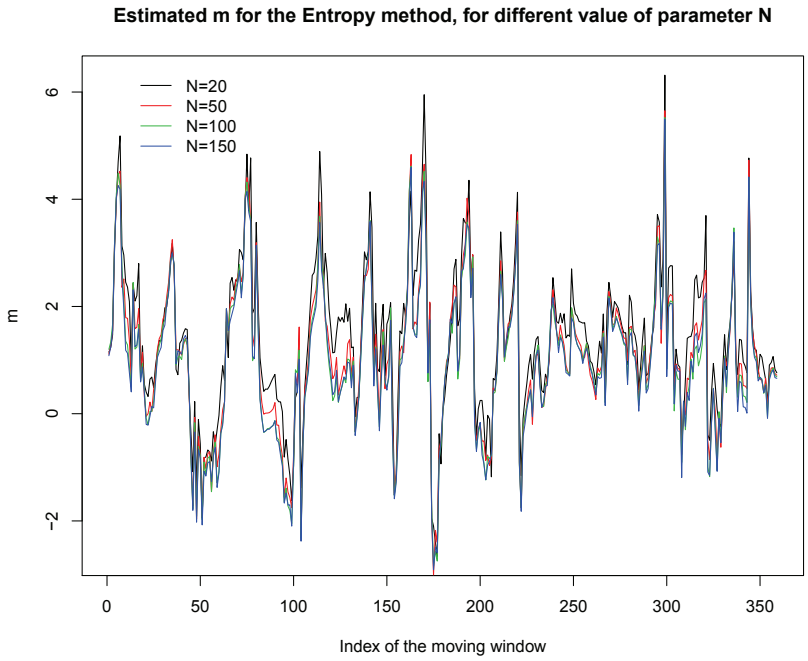


Figure 5.10: Entropy method, estimated m for the NASDAQ time series, for the period 1980-2010, T_k=400, $N = 20, 50, 100, 150$.

using both semi-empirical and empirical log-returns distributions. For example, CEV model is again shown to be a good descriptor of the price process for the time windows of shorter length. Also, it is seen that estimation using Entropy method with the semi-empirical log-returns distribution provides results closer to those obtained by Macbeth-Merville method, than with the empirical distribution. This fact is well observed on the Figure 5.13, where the length of the time window is $T_k = 800$.

In comparison to Macbeth-Merville method, as it is seen from the Table 5.3, m series estimated by Entropy method with semi-empirical distribution for log-returns for the windows of different lengths are much less correlated.

On the Figure 5.14 the scatter plots for Entropy method where estimated m is plotted against the volatility of log-returns in the following 20 days, are shown. Also linear regression lines are added - the trend suggests that for the higher m the volatility of log-returns in the following 20 days is expected to increase.

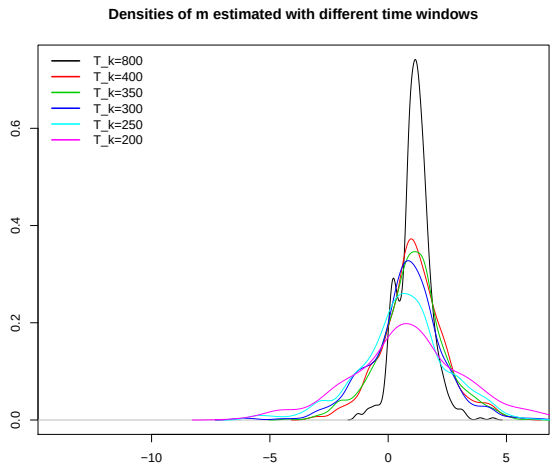
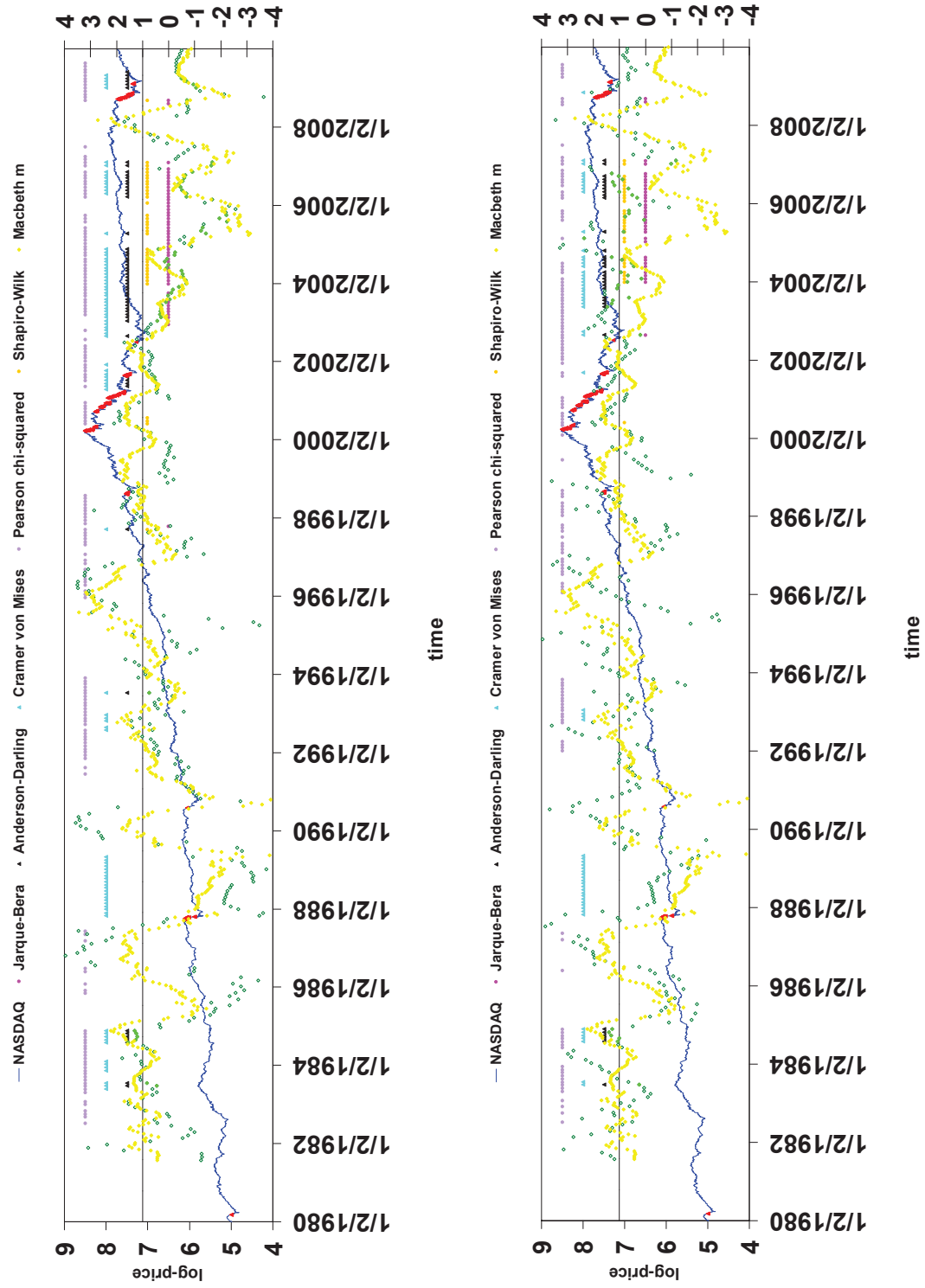


Figure 5.11: Entropy method, distributions of m for different sizes of time window.

5.3 "Best fit" method

Using this method the search of m is done on the interval $[-4, 4]$ with accuracy up to 0.05. We search for m which will produce the highest p-value for the Pearson's chi-square normality test applied to the dW_t variables. The corresponding m is accepted only if the corresponding p-value is higher than 0.05. Therefore, this method is unable to find m for some of the time windows. An example of this estimation is shown on the Figure 5.15 in comparison to Macbeth-Merville method. It seems that the algorithm provides estimations for m only in those time windows where the model with estimations by Macbeth-Merville method passes Pearson's chi-square test for normality. Though two time series for m seem to follow the same trend in those regions, they don't completely coincide.

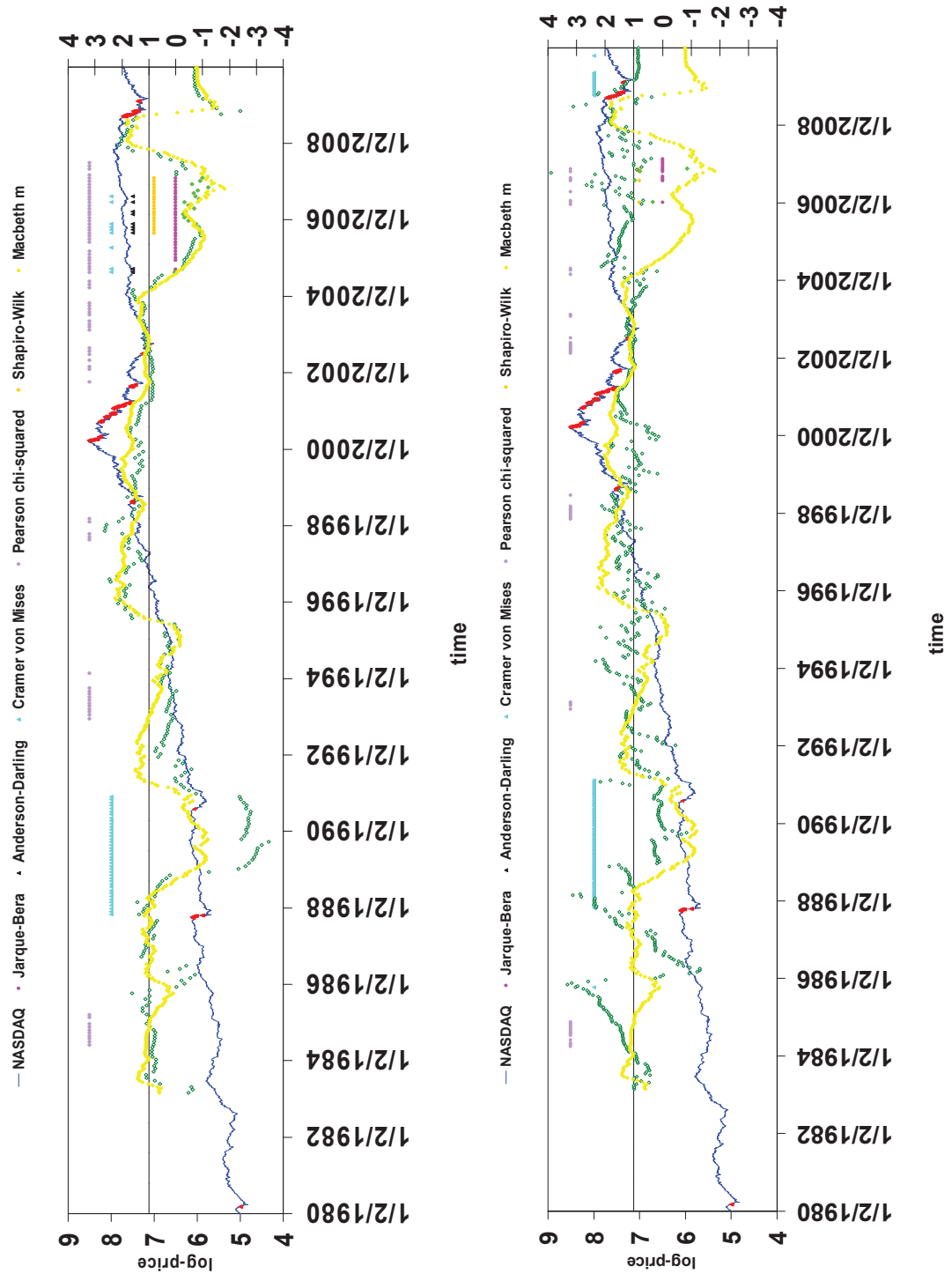
On the Figure 5.17 the scatter plots for 'Best fit' method where estimated m is plotted against the volatility of log-returns in the following 20 days, are shown, linear regression lines are also added. These scatter plots cannot be used to make enough reasonable conclusions about the dependence between an estimated m and volatility in the following period. The reason for this - the series of estimations consist only of those m for which Pearson's chi-square test did not reject normality for dW_t, i.e. the series are not complete and are based on the strong normality assumption.



(a) Semi-empirical log-returns distribution

(b) Empirical log-returns distribution

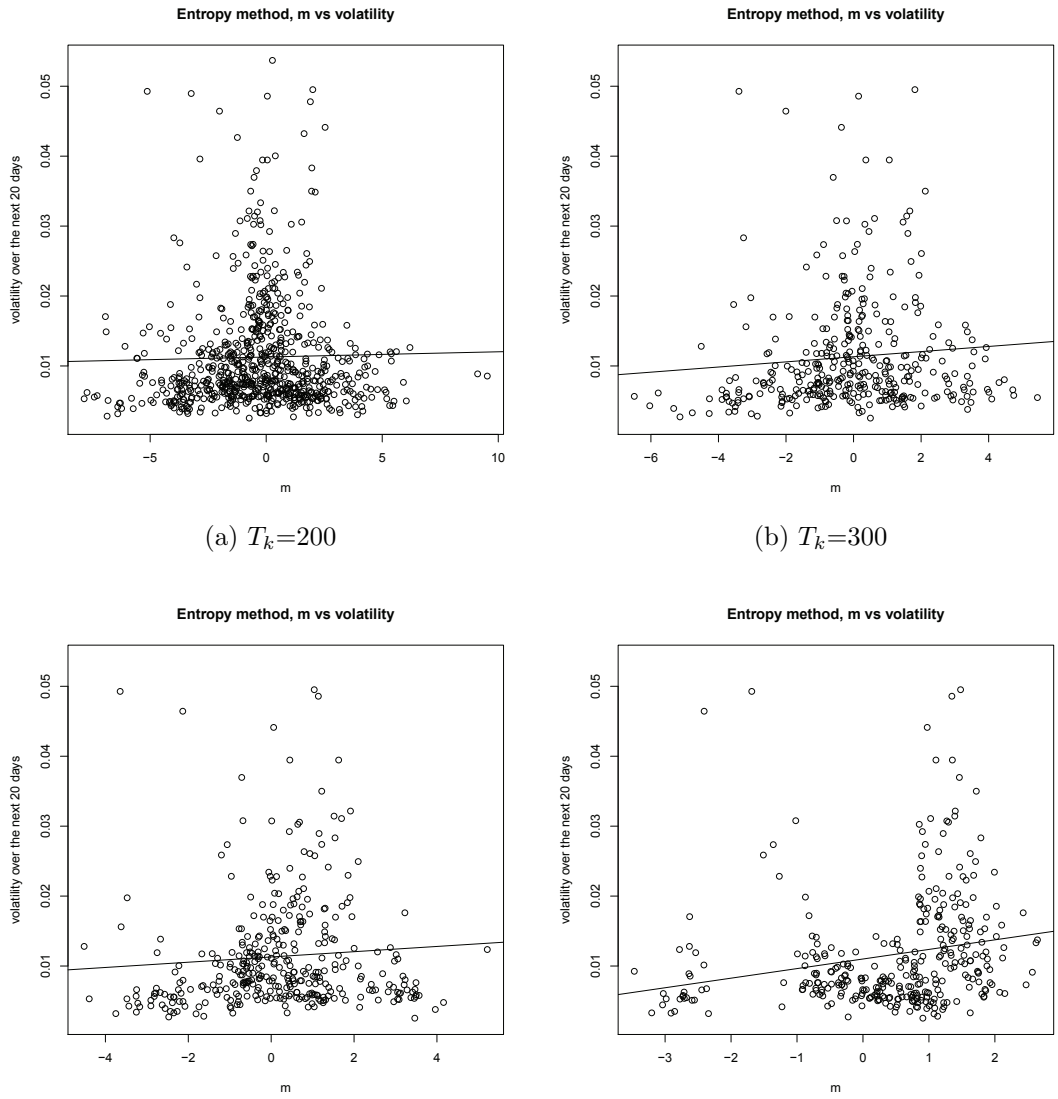
Figure 5.12: Comparison of results for Macbeth-Merville and Entropy methods, T_k=400, $M = 20$, $N = 100$. Left and right graphs show the estimation results using the semi-empirical and empirical log-returns distributions, respectively. The results are compared with those obtained by Macbeth-Merville method applied to the modified model with the same length of time window T_k.



(a) Semi-empirical log-returns distribution

(b) Empirical log-returns distribution

Figure 5.13: Comparison of results for Macbeth-Merville and Entropy methods, T_k=800, $M = 20$, $N = 100$. Left and right graphs show the estimation results using the semi-empirical and empirical log-returns distributions, respectively. The results are compared with those obtained by Macbeth-Merville method applied to the modified model with the same length of time window T_k.



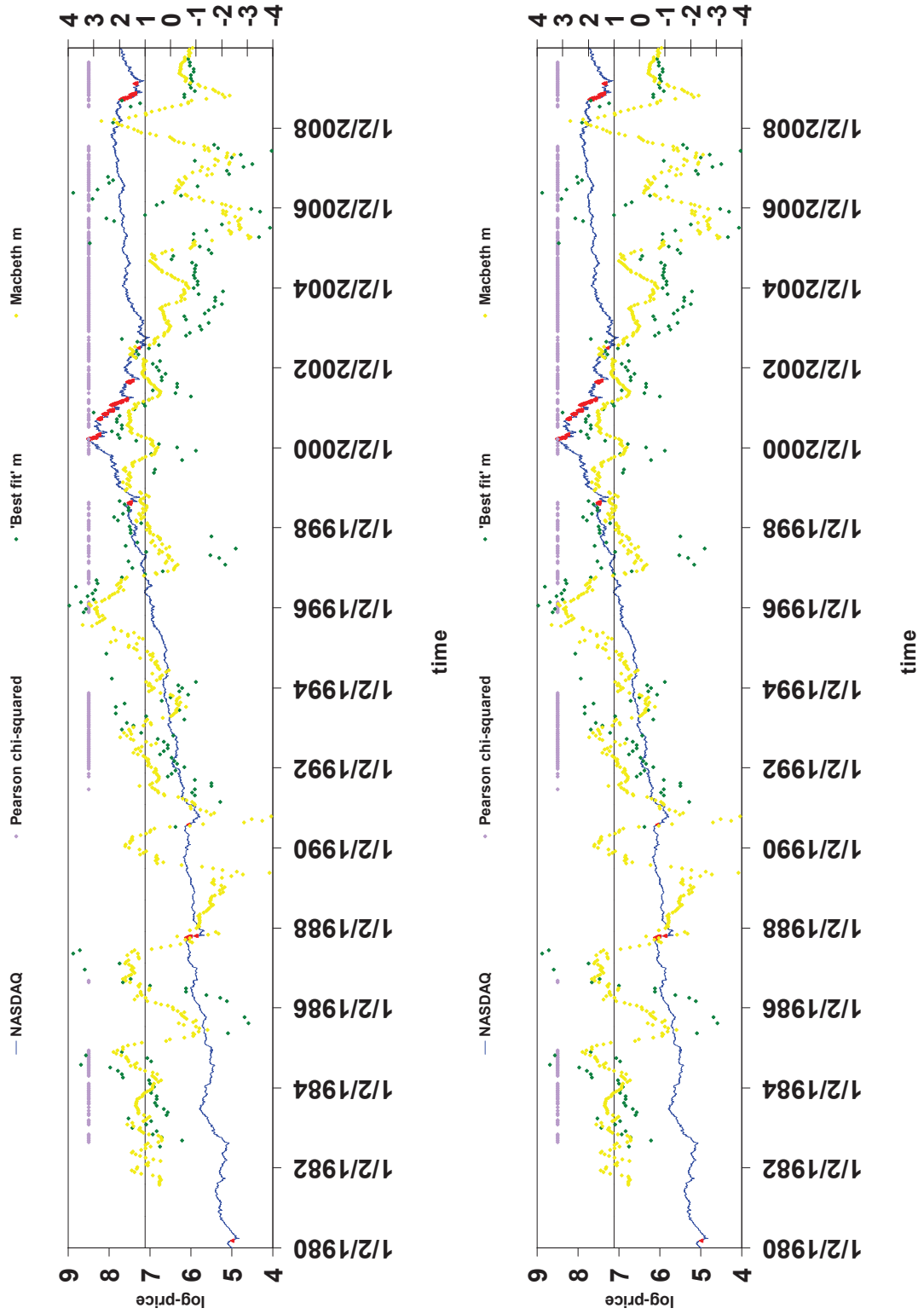
(a) T_k=200

(b) T_k=300

(c) T_k=400

(d) T_k=800

Figure 5.14: Entropy method, scatter plots m vs volatility over the next 20 days for different sizes of time window T_k=200, 300, 400, 800.



(a) T_k=400 (b) T_k=800

Figure 5.15: Comparison of results for "Best fit" and Macbeth-Merville methods, T_k=400, 800. Macbeth-Merville and 'Best fit' estimates are marked with yellow and green, respectively. WIth lilac those windows where Macbeth-Merville estimates pass Pearson's chi-suqare normality test are marked.

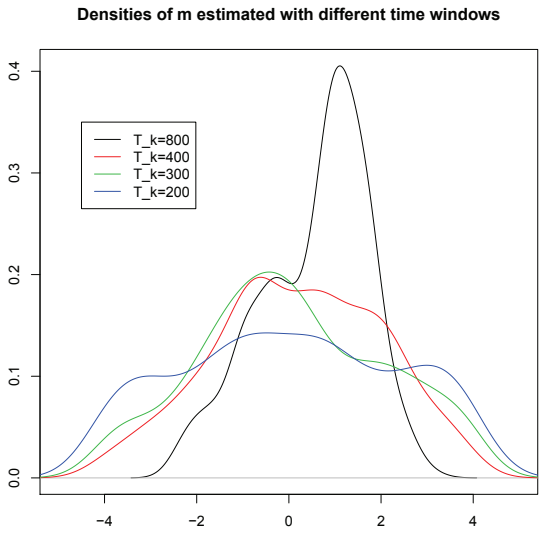
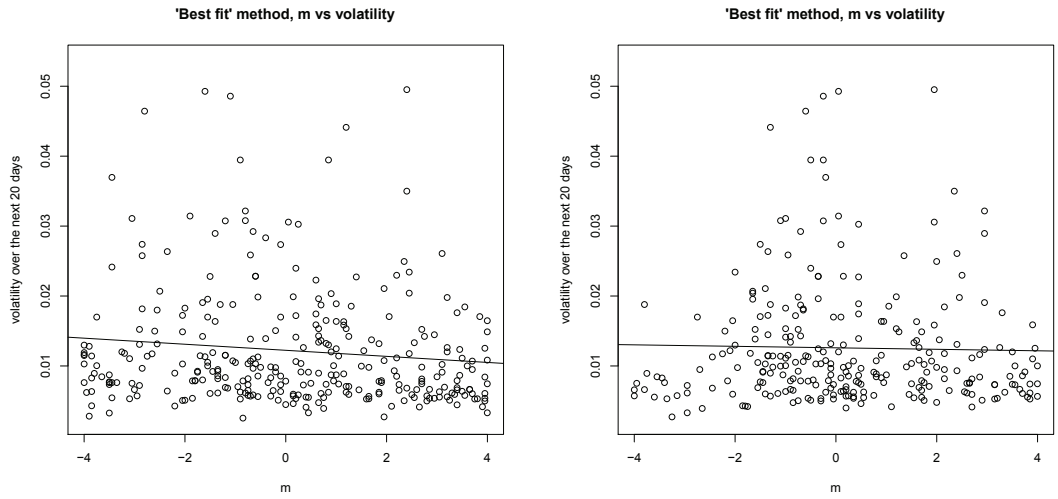
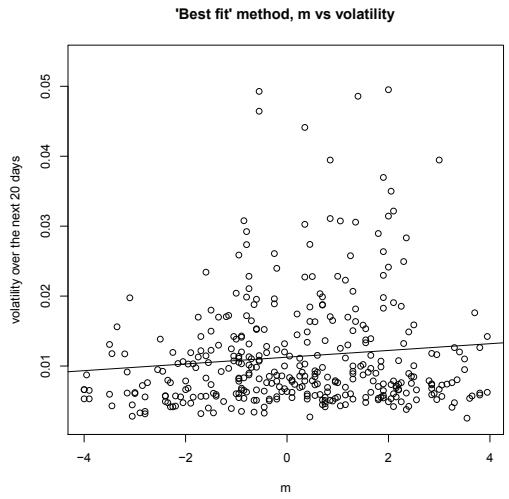


Figure 5.16: "Best fit", distributions of m for different sizes of time window.

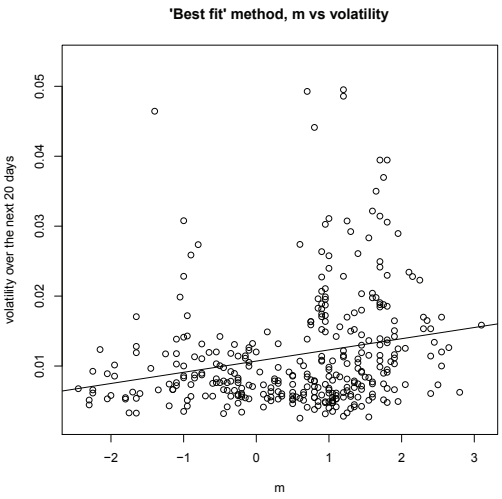


(a) T_k=200

(b) T_k=300



(c) T_k=400



(d) T_k=800

Figure 5.17: "Best fit" method, scatter plots m vs volatility over the next 20 days for different sizes of time window T_k=200, 300, 400, 800.

Chapter 6

Conclusions

In the current research we were trying to detect the periods in prices where the growth is faster-than-exponential. These periods according to studies, correspond to bubble periods. Using the hypothesis that bubble periods are those where positive feedback in prices exist, we analyzed different methods for estimation the co-operativity parameter, based on the assumption that price process follows the CEV model.

As it is shown, Macbeth-Merville method provides quite accurate results and linear model in the main equation is indicated be statistically probable. While Entropy method provides volatile results and the linear model in the main equation of the method is doubtful and might need further investigation. And the last, 'Best fit' method provides comparatively accurate results, though it is fully based on the assumption of normality which is a very strong one. All three methods are shown to perform better when more data is available.

The topic of the present research has an important implication for the investment industry professionals as detecting bubbles may prevent asset managers from investments in late bubble cycles. But though in the current study the mathematical detection of bubble periods shows some indications, the results are not clear, and, hence, needs to be further investigated.

Appendix A

A.1 Log-returns distribution

$$dX_t = \mu X_t dt + \sigma X_t^m dW_t \tag{A.1}$$

To find the distribution of log-returns, we apply Ito's lemma to the function $f = lnX$:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}(dW_t)^2 \tag{A.2}$$

$$\frac{\partial f}{\partial X} = \frac{1}{X}, \frac{\partial^2 f}{\partial X^2} = -\frac{1}{X^2}, \frac{\partial f}{\partial t} = 0, \tag{A.3}$$

$$\begin{aligned} df &= 0dt + \frac{1}{X_t}dX_t + \frac{1}{2}(-\frac{1}{X_t^2})\sigma^2 X_t^{2m}dt \\ &= \frac{1}{X_t}(\mu X_t dt + \sigma X_t^m dW_t) - \frac{1}{2X_t^2}\sigma^2 X_t^{2m}dt \\ &= (\mu - \frac{1}{2}\sigma^2 X_t^{2(m-1)})dt + \sigma X_t^{m-1}dW_t, \end{aligned} \tag{A.4}$$

therefore,

$$\ln \frac{X_t}{X_0} \sim \phi\left[(\mu - \frac{1}{2}\sigma^2 X_t^{2(m-1)})t, \sigma X_t^{m-1}\sqrt{t}\right]$$

where X_t denotes the stock price at time t, X_0 denotes the price at time 0, and $\phi(\mu, \sigma)$ denotes a normal distribution with mean μ and standard deviation σ.

Applying "standardizing" transformation, we have:

$$f(r|x) = \frac{1}{\sigma x^{m-1}}\phi(\frac{r - (\mu - \frac{1}{2}\sigma^2 x^{2(m-1)})}{\sigma x^{m-1}})$$

A.2 Entropy for Normal distribution

Let X be a random variable with a probability density function f whose support is a set X. The differential entropy $H(X)$ or $H(f)$ is defined as

$$H(X) = -\int_X f(x)\ln(f(x))dx$$

Differential form of the entropy for the normal distribution have the following closed form:

$$\begin{aligned} H(\phi(x)) &= -\int_{-\infty}^{\infty} \phi(x)\ln(\phi(x))dx \\ &= -\int_{-\infty}^{\infty} \phi(x)(\ln(\frac{1}{\sqrt{2\pi}\sigma}) - \frac{(x-\mu)^2}{2\sigma^2})dx \\ &= \ln(\sqrt{2\pi}\sigma) + \frac{\sigma}{2}\int_{-\infty}^{\infty}\left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2)d\left(\frac{x-\mu}{\sigma}\right) \\ &= \ln(\sqrt{2\pi}\sigma) + \frac{1}{2} = \ln(\sqrt{2\pi e}\sigma) \end{aligned} \tag{A.5}$$

A.3 Linear Regression Diagnostics

a Constant variance

In this case we need to check whether the variance in the residuals is related to the fitted values [10]. Constant variance assumption holds if on the plot of residuals against fitted values residuals are symmetric vertically around zero. Numerically, this can be checked applying linear regression model to residuals against the fitted values and to absolute values of residuals against the fitted values: The estimated model we are expecting to see is

$$residuals = a_1 + b_1 * \textit{fitted values}, \tag{A.6}$$

$$|residuals| = a_2 + b_2 * \textit{fitted values}, \tag{A.7}$$

where b_1 and b_2 are not significantly different from zero.

b Normality

Graphically the residuals can be assessed for normality using a Q-Q-plot, which compares the residuals to normal observations. In the present work Pearson's phi-squared test is used as a test for normality (see Appendix A.2.5):

H_0: residuals are normally distributed,

H_a: residuals are not normally distributed.

For p-values of the test smaller than 0.05 the null hypothesis is rejected. When the errors are not normal, least squares estimates may not be optimal. They will still be best linear unbiased estimates, but other robust estimators may be more effective.

c Serial correlation

Another assumption of linear model is that the errors are uncorrelated. Graphically this can be checked plotting the residuals against the time. To check this assumption numerically we use Durbin-Watson test with the following statistic:

$$DW = \sum_{i=2}^{n} (\widehat{\epsilon}_i - \widehat{\epsilon}_{i-1})^2 / (\sum_{i=1}^{n} \widehat{\epsilon}_i^2) \tag{A.8}$$

The Durbin-Watson statistic has a range from 0 to 4 with a midpoint of 2 and has the following interpretation:

1. if $DW < 1$, the errors are positively correlated: an increase in one period follows an increase in the previous period,
2. if $DW = 2$, there is no autocorrelation in the errors,
3. if $DW > 3$, the errors are negatively correlated: an increase in one period follows an decrease in the previous period.

Therefore, we are interested in the values of DW on the interval [1,3].

A.4 Normality tests

Normality tests are used to determine how likely an underlying random variables are to be normally distributed.

The tests are defined as
H_0: $\left\{X_i, i = \overline{1,n}\right\}$ follow a normal distribution
H_a: $\left\{X_i, i = \overline{1,n}\right\}$ do not follow the normal distribution.

A.4.1 Anderson-Darling test for normality

It is a modification of Kolmogorov-Smirnov test, which gives more weight to the tails than does the latter test. Makes use of the specific distribution (in our case normal distribution) in calculating critical values.

The testing procedure is the following:

1. the tested data $\left\{X_i, i = \overline{1,n}\right\}$ is sorted in increasing order,
2. the sample mean and standard deviation are calculated

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ s = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \tag{A.9}$$

3. the values X_i are standardized to create new values

$$Y_i = \frac{X_i - \overline{X}}{s},\ i = \overline{1,n}, \tag{A.10}$$

4. the test statistic is calculated as

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)(\ln(\Phi(Y_i)) + \ln(1 - \Phi(Y_{n+1-i}))), \tag{A.11}$$

where Φ is the cumulative normal distribution function.

This test is a one-sided and the hypothesis that the distribution is of a specific form is rejected if the test statistic, A, is greater than the critical value.

A.4.2 Jarque-Bera test for normality

This test for normality is based on the sample kurtosis and skewness coefficients. The Jarque-Bera test statistic is defined as

$$JB = \frac{n}{6}(S^2 + \frac{1}{4}K^2), \tag{A.12}$$

where n is the number of observations and S is the sample skewness, and K is the sample kurtosis:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^3}{(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2)^{3/2}} \tag{A.13}$$

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^4}{(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2)^2} - 3 \tag{A.14}$$

For large sample sizes the Jarque-Beta statistic has an asymptotic chi-square distribution with two degrees of freedom. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero, since samples from a normal distribution have an expected skewness zero and an expected excess kurtosis zero (which is the same as a kurtosis of 3).

A.4.3 Shapiro-Wilk test for normality

The testing procedure is the following:

1. the tested data $\left\{X_i, i = \overline{1,n}\right\}$ is sorted in increasing order,
2. the sample mean is calculated

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \tag{A.15}$$

3. The test statistic is calculated as

$$W = \frac{\left(\sum_{i=1}^{n} a_i X_i\right)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{A.16}$$

where a_i are constants generated from the means, variances and covariances of the order statistics of a sample of size n from a normal distribution

$$(a_1, ..., a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}},\ m = (m_1, ..., m_n)^T$$

are the expected values of the order statistics of independent and identically-distributed random variables sampled from the standard normal distribution, and V is the covariance matrix of those order statistics.

Shapiro-Wilk test is closely tied to the normal probability plot, since it is based on the correlation between the normal quantiles and the sample quantiles. The correlation measures how close the normal plot is to being a straight line.

A.4.4 Cramer von Mises test for normality

In statistics the Cramer-von-Mises criterion is a form of minimum distance estimation used for judging the goodness of fit of a probability distribution $F*$ (in our case Φ) compared to a given empirical distribution function F_n. It is defined as

$$w^2 = \int_{-\infty}^{\infty} [F_n(x) - \Phi(x)]^2 \, d\Phi(x) \tag{A.17}$$

The testing procedure is the following:

1. the tested data $\left\{X_i, i = \overline{1, n}\right\}$ is sorted in increasing order,
2. the test statistic is calculated as

$$T = nw^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i - 1}{2n} - \Phi(X_i) \right]^2 \tag{A.18}$$

If this value is larger than the critical value we can reject the null hypothesis.

A.4.5 Pearson's chi-square test for normality

The Pearson's chi-square goodness-of-fit test is applied to binned data (i.e., data put into classes). This is not a restriction since for non-binned data it is possible to calculate a histogram or frequency table before generating the chi-square test. However, the value of the chi-square test statistic is dependent on how the data is binned. Another disadvantage of the chi-square test is that it requires a sufficient sample size in order for the chi-square approximation to be valid.

The testing procedure is the following:

1. the data is divided into k bins,
2. the test statistic is defined as

$$\chi^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i, \tag{A.19}$$

where O_i is the observed frequency for the bin i, and E_i is the expected frequency for bin i. The expected frequency is calculated by

$$E_i = n(\Phi(Y_u) - \Phi(Y_l)), \tag{A.20}$$

where Y_u is the upper limit for class i, Y_l is the lower limit for class i, and n is the sample size.

The chi-square statistic can then be used to calculate a p-value by comparing the value of the statistic to a chi-square distribution.

A.5 BDS test

The computation of BDS test follows the below procedure:

1. given is an N observations time series, which are the first difference of the natural logarithms of raw data:

$$\{x_i\} = [x_1, x_2, ..., x_N]$$

2. select value of m (embedding dimension), embed the time series into m-dimensional vectors, by taking each m successive points in the series. This way we convert the series of scalars into a series of vectors with overlapping entries:

$$\begin{aligned} x_1^m &= (x_1, x_2, ..., x_m), \\ x_2^m &= (x_2, x_3, ..., x_{m+1}), \\ &... \\ x_{N-m}^m &= (x_{N-m}, x_{N-m+1}, ..., x_N), \end{aligned}$$

3. compute the correlation integral, which is a measure of the frequency with which temporal patterns are repeated in the data. This is done by adding the number of pairs of points (i, j), where $1 \leq i \leq N$ and $1 \leq j \leq N$, in the m-dimensional space which are "close" in the sense that the points are within a radius or tolerance ϵ of each other:

$$C_{\epsilon,m} = \frac{1}{N_m(N_m - 1)} \sum_{i \neq j} I_{i,j;\epsilon}, \tag{A.21}$$

where $I_{i,j;\epsilon} = 1$, if $||x_i^m - x_j^m|| \leq \epsilon$ and 0, otherwise,

4. Brock, Dechert and Scheinkman [3] showed that if the time series is i.i.d.

$$C_{\epsilon,m} \approx [C_{\epsilon,1}]^m \tag{A.22}$$

the quantity $C_{\epsilon,m} - [C_{\epsilon,1}]^m$ has an asymptotic normal distribution with zero mean and a variance $V_{\epsilon,m}$ defined by:

$$V_{\epsilon,m} = 4[K^m + 2\sum_{j=1}^{m-1} K^{m-j} C_\epsilon^{2j} + (m-1)^2 C_\epsilon^{2m} - m^2 K C_\epsilon^{2m-2}], \tag{A.23}$$

where $K = \frac{6}{N_m(N_m-1)(N_m-2)} \sum_{i<j<N} (\frac{I_{i,j;\epsilon} I_{j,N;\epsilon} + I_{i,N;\epsilon} I_{N,j;\epsilon} + I_{i,j;\epsilon} I_{i,N;\epsilon}}{3})$

5. The BDS test statistic is calculated as:

$$BDS_{\epsilon,m} = \frac{\sqrt{N}[C_{\epsilon,m} - [C_{\epsilon,1}]^m]}{\sqrt{V_{\epsilon,m}}}, \tag{A.24}$$

APPENDIX A.

BDS test is a two-tailed test, the null hypothesis is rejected if the test statistic is greater than or less than the critical values (e.g. if $\alpha = 0.05$, the critical value $= \pm 1.96$).

Bibliography

[1] J. Alvarez-Ramirez, C. Ibarra-Valdez, "Finite-time singularities in the dynamics of Mexican financial crises", Physica A331, 253-268, 2004.

[2] F. Black, "Studies of stock price volatility changes", Proceedings of the American Statistical Association Annual Meetings, Business and Economics Section, Washington DC (177-181), 1976.

[3] W. A. Brock, W. Dechert, J. Scheinkman, "A test for independence based on the correlation dimension", 1987.

[4] W. A. Brock, C. L. Sayers, "Is the business cycle characterized by deterministic chaos?" Journal of Monetary Economics, 22, pp. 71 90, 1988.

[5] J. Y. Campbell, A. W. Lo, A.C. MacKinlay, "The Econometrics of Financial Markets", 1997.

[6] K. E. Case, R. J. Shiller, "Is There a "Bubble" in the Housing Market?", 2004.

[7] A. Christie, "The stochastic behavior of stock return variances: Value, leverage and interest rate effects", Journal of Financial Economics 10 (407-432), 1982.

[8] J. C. Cox and S. A. Ross, "The Valuation of Options for Alternative Stochastic Processes", Journal of Financial Economics, 3, 145-166, 1976.

[9] "Economic volatility" At a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, August 30, 2002.

[10] J.J. Faraway, "Linear Models with R", 2005.

[11] G. A. Karolyi "Why Stock Return Volatility Really Matters", 2001.

[12] J.D. Macbeth, L.J. Merville "Tests of the Black-Scholes and Cox Call Option Valuation Models", Journal of Finance, Volume 35, Issue 2, 285-301, 1979.

[13] H. P. Minsky, "The Financial Instability Hypothesis", Working Paper No. 74, 1992.

[14] S. Reimann, R. Woodard, D. Sornette, "Detecting bubble phases in empirical data: positive feedback in a CIR-CEV framework", unpublished paper, 2010.

[15] D. Sornette, A. Johansen, "Finite-time singularity in the dynamics of the world population, economic and financial indices", Physica A294, 465-502, 2001.

[16] D. Sornette, W.X. Zhou, "Is there a real-estate bubble in the US?", Physica A361, 297-308, 2006.

[17] D. Sornette, W.X. Zhou, H. Takayasu, "Finite-time signature of hyperinflation", Physica A325, 492-506, 2003.

[18] D. Sornette, "Why stock markets crash", Princeton University Press, 2003.

[19] K. Watanabe, H. Takayasu, M. Takayasu, "A mathematical definition of the financial bubbles and crashes", Physica A383, 120-124, 2007.

[20] K. Watanabe, H. Takayasu, M. Takayasu, "Observation of two types of behaviors of financial bubbles and the related higher-order potential forces", Progress of Theoretical Physics Supplement No.179, 2009.

[21] K. Watanabe, H. Takayasu, M. Takayasu, "Extracting the exponential behaviors in the market data", Physica A382, 336-339, 2007.

List of Figures

List of Tables