Modeling the Behavior of Prague Stock Exchange Index (PX-50)

Martina Horníková

Introduction

This paper examines behavior of the Prague Stock-Exchange Index, PX-50, which includes 50 leading Czech companies. We will see that this index exhibits typical econometric properties of financial time series, in which case the estimation is usually made with the use of ARCH models. The data suggests that the best fitting model that should be used in our case is the GARCH(1,1) model.

Modeling PX-50 Index

The Prague Stock Exchange introduced its official index PX-50 on the occasion of the first anniversary of the opening of its trading. A standard method of calculation has been chosen for the PX-50 Index in accordance with the IFC (International Finance Corporation) methodology recommended for the creation of indices in emerging markets. Based on relevant analyses, it was decided to create a base composed of 50 issues. At the present time, the actual number of the basic issues is varying. However, in accordance with the Principles of Updating the Base of the Index PX-50, which was approved in December 2001, the base cannot consist of more than 50 issues. The Index base incorporates no issues coming from the Sector No. 18 (investment funds) nor does it issues of those holding companies that have emerged from transformation of investment funds, because their price quotations already reflect the price movements occurred in the basic issues. Selected as the starting exchange day (a benchmark date) for the Index PX-50 was April 5, 1994 and its opening value was fixed at 1 000 points.

Autoregressive Conditional Heteroskedasticity (ARCH) models are specifically designed to model and forecast conditional variances. It has proven to be useful in studying a variety of macroeconomic phenomena. ARCH models were introduced and developed by Engle (1982).

Engle (1982,1983) and Cragg (1982) have found evidence that for some phenomena the disturbance variance in time-series models is less stable than it is usually assumed. More recent studies of financial markets suggest that this phenomenon is very common. This particular specification of heteroskedasticity was motivated by the observation that in many financial time series, the magnitude of residuals appeared to be related to the magnitude of recent residuals. ARCH in itself does not invalidate standard LS inference. However, ignoring ARCH effects may result in loss of efficiency.

Data description

The Prague Stock-Exchange Index, PX-50, is an index composed of 50 leading Czech companies. It includes companies of the highest market capitalization. The data was extracted from several internet sources, mostly from www.finance.yahoo.com, www.pse.cz, www.akcie.cz.

The range of data is from February 16, 1994 till November 25, 2002 and includes daily data. The overall number of observations is 2187, which is a highly representative sample. The entire estimation procedure is performed with the use of Eviews 4.1 software.

Description of variables used in the following models:

PX50, DAX, DJI, FTSE, NASDAQ	the original *daily closing* indices
CPX50, CDAX, CDJI, CFTSE, CNASDAQ	corrected indices; any missing (due to national holidays) is replaced by the previous day's closing value
RPX50, RDAX, RDJI, RFTSE, RNASDAQ	returns on various indices, defined as change in logs: $RPX50_t = \log(RPX50_t) - \log(RPX50_{t-1})$
DAY, MONTH, YEAR	The date of a particular observation

All of these data are for the time interval from February 16, 1994 till November 25, 2002.

Events:

- Argentinean and Mexico crises, 1994,
- Asian crisis, August 1997,
- Russian crisis, August 1998,
- September 11, 2001,

Modeling

First we have a look at the graph of the Index PX-50, Figure 1, and at its returns, Figure 2. Mainly from the secondly mentioned one, we can deduct some important aspects of the PX-50 time series development. The returns are corresponding to important word events (mentioned in Data description section of the paper) during the history. I will comment on them later on in the paper.

In the histogram, Figure 3 we can see negative skewness (approximately -0.33) of the PX-50 returns, which indicates some large negative shocks during the period from February 16, 1994 till November 25, 2002. The kurtosis value is approx. 6.165 (which is well above 3, which is, according to the theoretical sources, the appropriate for the normal distribution). We can say that the kurtosis is indicating fat tails relative to normal distribution.

Nevertheless, the time series of PX-50 returns is stationary, which is proven by ADF-test in Table 1, from where we can see that the null hypothesis of a unit root can be rejected even at 1% level of significance. It means that the series is stationary.

From the correlogram of PX-50 returns, Table 2, we can see that we can surely reject the null hypothesis of no autocorrelation for the whole sample period, from 1994 till 2002. It leads us to the result that the Czech stock exchange is inefficient.

From this correlogram we can also see that ARMA(2,2) process could be appropriate to fit. When we estimate an ARMA(2,2) equation for PX-50 returns and estimate by OLS method, we can see that there is no remaining serial autocorrelation in residuals (proven by Breusch-Godfrey Serial Correlation LM Test in Table 3), but according to the correlogram there is still some remaining autocorrelation in squared residuals, proven by significant ARCH LM Test (H0 of no autocorrelation is rejected at 5% level of significance). This test suggests to fit an ARCH process to our equation. We run GARCH(1,1) on this equation that gives us significant results, Table 4. When we sum up the arch (0.115045) and garch values (0.878693) we get 0.993738, which is very close to 1. That means that the shock is very persistent (is dying off very slowly). Nevertheless, we cannot find unit roots in our model, which means that the model is stationary.

If we run the usual tests now, we can see that there's no more remaining autocorrelation present in the model, according to both the correlograms and insignificant ARCH LM Test. Unfortunately, according to the histogram-normality test, Figure 4, normality wasn't achieved. Nevertheless, it is quite usual characteristics of financial time series.

If we try to include returns of some foreign indices into the estimated equation (Dow Jones, NASDAQ, FTSE or DAX), we get worse results in comparison with those, which we got with the use of previously mentioned ARMA(2,2) model only. Also the forecast, Figure 5, is much better in the case of ARMA(2,2) model.

In the forecast, the bias proportion tells us how far the mean of the forecast is from the mean of the actual series and the variance proportion tells us how far the variation of the forecast is from the variation of the actual series. Apparently, these should be as small as possible. Our results (bias proportion of 0.019143 and variance proportion of 0.328020) are very good in this respect. On the other hand, the covariance proportion measures the remaining unsystematic forecasting errors and this one should be as high as possible. Our result of 0.662037 is reasonably high. All these factors lead to the conclusion of a good forecast.

In the forecast we can see that both the bias proportion and the variance proportion are very small and the covariance proportion is high so we can claim the validity and good quality of the forecast.

When we look at the graph of PX-50 returns, Figure 1, we can observe impact of some important events in the history on the enormously high negative or positive values of the returns. These were some significant changes in the time series of the returns in the late August 1998. The index was influenced first by the Asian crisis (August 1997) and later on by the Russian crisis (in the second half of August 1998). The second mention event had even bigger impact on the financial market than the first one, since there is a huge trade between the Czech Republic and Russia. Since September 1998, the Czech stock-exchange market started to be much more volatile. Other important events that can be read from the graph of PX-50 returns is the Mexico and Argentina crisis in spring 1994, which resulted into a big decline in returns just at the very beginning of theCzech stock-exchange, from March till July 1994. During that period we can see mostly negative signs in returns. The next very significant event that influenced the Czech (and the whole world) market was the terrorist attack on September 11, 2001 in the USA. Afterwards the American Stock Exchange was closed for several days and there was a high uncertainty among the investors. The Czech financial market reacted on September 18, 2001 with a huge decrease to approximately – 7,53% (the second lowest return in the history of the Czech stock exchange).

Responding to the mentioned events, we divide the sample into two sub-samples, the first one for observations before the Russian crisis, i.e. 1-1133 (from February 1994 till August 27, 1998) and the second one for the rest of the sample, 1134-2187 (after September 1998 till nowadays). We test for validity (robustness) of the ARMA(2,2) model. In the first sub-sample we can obtain significant results using the ARMA(2,2) model and running GARCH(1,1) but, unfortunately, in case of the second sub-sample we cannot say the same. It means that ARMA(2,2) model is not robust and that we have to find out another estimating model for the second part of the sample period.

The first sub-sample:

Even for the first sub-sample we can find a better model than ARMA(2,2) (because when we test by for omitted variables in ARMA(2,2), we can see that RFTSE should be included in the model). We develop another model that suits the situation better. From the results we got by estimating with the use of OLS method, we can see that there is autocorrelation in residuals squared and also the ARCH LM test (a Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) in the residuals) is significant. The null hypothesis of no ARCH must be rejected. After running GARCH(1,1) we get results as in Table 5. According to the correlogram of residuals and residuals squared (not shown in the paper), there is no remaining autocorrelation and according to the ARCH LM test H0 of no ARCH cannot be rejected at 5% level of significance. The sum of ARCH and GARCH values is approximately 0.983 and because it is less than 1 the conditional variance converges to the unconditional variance, which is also a good characteristics of the results. Nevertheless, it's very close to 1, which concludes persistency of a shock. Testing for normality by histogram-normality test, Table 6, the Jarque-Bera statistics is significant, which implies that the standardized residuals aren't even in the first sub-sample normally distributed. There is still negative skewness, indicating some large negative shocks and high kurtosis (approx.7.77) and fat tails.

The second sub-sample (after the Russian crisis):

ARMA(2,2) model doesn't work for this period (both AR(1) and MA(1) are insignificant). We can claim the same also from the correlogram of the second sub-sample, Table 7. H0 hypothesis of the augmented DF-unit root test (of a unit root presence) has to be rejected (Table 8) and it means that there is no unit root and the time series is therefore still stationary. In the histogram, Table 9, we can see that even for the second period the normality has to be rejected and the time series is still skewed to the left and kurtosis is still very high (approx. 9.37) which again implies the fat tails. OLS model based on foreign indices appeared to be the best one, as shown in Table 10. In the OLS model there is no autocorrelation in residuals confirmed but autocorrelation in residuals squared, ARCH present (H0 for no ARCH cannot be rejected at 5% level of significance, according to ARCH LM Test). After running GARCH(1,1), Table 11, we get rid of autocorrelation, ARCH LM Test is insignificant. The sum of ARCH and GARCH values is 0.935467, which is quite close to 1. That again means persistency of shocks. No unit roots are present, which indicates stationarity. Again, normality is not achieved but the results from the histogram-normality test are quite good, Figure 7. The skewness is very small (only 0.06 approx.) and the kurtosis is only around 4.35. Also the forecasted values, in Table 12, seem to be reasonable. Both bias and variance proportions are small and covariance proportion is quite high. Theil inequality coefficient is around 0.75, in which case we can speak about a good fit of the model.

We can see that normality wasn't achieved in any of our models. The negative skewness is due to asymmetry in the modeled data. With financial data, the asymmetry is quite usual, because of the fact that negative shocks cause higher volatility in the near future than positive shocks. In finance this effect is called the leverage effect.

Conclusion

This paper examined the behavior of the Prague Stock-Exchange Index, PX-50; the index, which represents stocks of 50, Czech companies with the highest market capitalization. The values are closing-time data of the index. The index exhibits traditional econometric properties of financial time series – no achievement of normality, clusters, negative skewness, large kurtosis (presence of fat tails), and autocorrelation. Therefore we applied Generalized ARCH model to model the behavior of the PX-50 Index. Final conclusion is that the best fitting model is GARCH(1,1) where the Coefficient Covariance is adjusted to Heteroskedasticity Consistent Covariance (according to Bollerslev - Wooldridge). Using the mentioned method we obtained very good results and also a forecast that seems to be reasonable for the future development of the Prague Stock-Exchange Index PX-50.

References

1. Gujarati, D. N. *Essentials of econometrics*, New York : McGraw-Hill, 1992,
2. Greene, W. H. *Econometric Analysis*, Prentice Hall, 2000,
3. Maddala, G. S. *Introduction to Econometrics*; Wiley, 3rd edition, 2001,
4. Web sources: Yahoo Finance Database: www.finance.yahoo.com,
   Prague Stock-Exchange: www.pse.cz,
   www.akcie.cz,
   www.patria.cz.

Appendix:

Figure 1: PX-50 Index

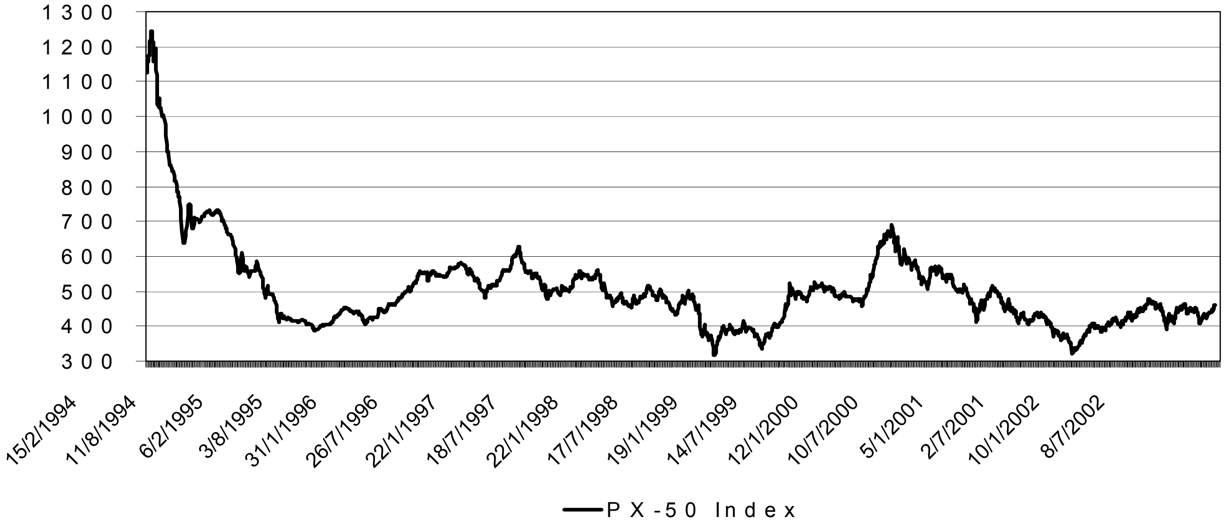


Figure 2: Returns of PX-50 Index

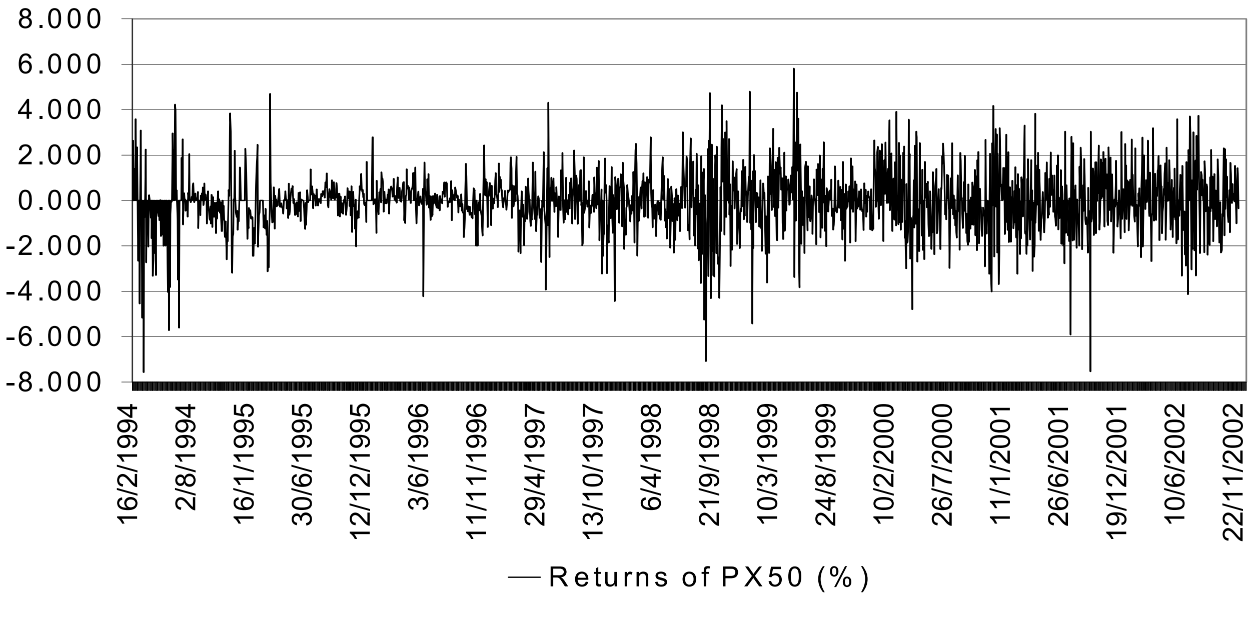
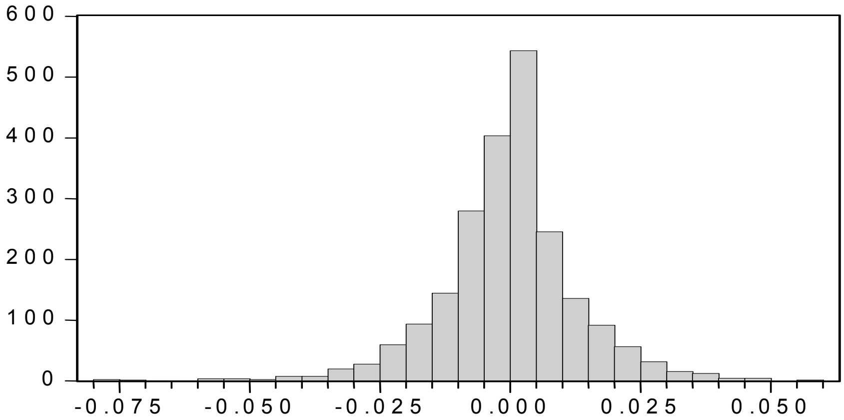


Figure 3: Histogram of PX-50 returns



Series: RPX50	
Sample 2 2187	
Observations 2186	
Mean	-0.000409
Median	0.000000
Maximum	0.058200
Minimum	-0.075664
Std. Dev.	0.012964
Skewness	-0.333217
Kurtosis	6.164944
Jarque-Bera	952.8231
Probability	0.000000

Table 1: Augmented Dickey-Fuller unit root test on PX-50 returns

Null Hypothesis: RPX50 has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic based on SIC, MAXLAG=25)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-28.60086	0.0000
Test critical values:	1% level	-3.433146	
	5% level	-2.862661	
	10% level	-2.567413	

Table 2: Correlogram for PX-50 returns

Sample: 1 2187
Included observations: 2186

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>|* |</td><td>|* |</td><td>1</td><td>0.128</td><td>0.128</td><td>36.044</td><td>0.000</td></tr>
<tr><td>|* |</td><td>|* |</td><td>2</td><td>0.092</td><td>0.076</td><td>54.381</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>3</td><td>0.060</td><td>0.040</td><td>62.149</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>4</td><td>0.019</td><td>0.000</td><td>62.932</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>5</td><td>-0.036</td><td>-0.047</td><td>65.771</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>6</td><td>-0.008</td><td>-0.003</td><td>65.925</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>7</td><td>0.014</td><td>0.022</td><td>66.354</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>8</td><td>0.012</td><td>0.013</td><td>66.652</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>9</td><td>0.015</td><td>0.011</td><td>67.115</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>10</td><td>0.046</td><td>0.039</td><td>71.796</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>11</td><td>0.030</td><td>0.016</td><td>73.712</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>12</td><td>0.019</td><td>0.008</td><td>74.546</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>13</td><td>0.051</td><td>0.042</td><td>80.330</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>14</td><td>0.001</td><td>-0.015</td><td>80.331</td><td>0.000</td></tr>
<tr><td>| |</td><td>| |</td><td>15</td><td>0.018</td><td>0.015</td><td>81.083</td><td>0.000</td></tr>
</table>

Table 3: Breusch-Godfrey Serial Corr. LM Test, ARMA(2,2) on PX-50 returns

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.523341	Probability	0.666252
Obs*R-squared	1.574241	Probability	0.665245

Table 4: GARCH(1,1) for ARMA(2,2) process

Dependent Variable: RPX50
Method: ML – ARCH (Marquardt)
Sample(adjusted): 4 2187
Included observations: 2184 after adjusting endpoints
Convergence achieved after 48 iterations
Bollerslev-Wooldrige robust standard errors & covariance
MA backcast: 2 3, Variance backcast: ON

	Coefficient	Std. Error	z-Statistic	Prob.
C	4.22E-06	3.20E-06	1.318167	0.1874
RPX50(-1)	1.373151	0.060908	22.54451	0.0000
RPX50(-2)	-0.383539	0.061462	-6.240266	0.0000
MA(1)	-1.173107	0.058542	-20.03889	0.0000
MA(2)	0.189458	0.059965	3.159452	0.0016
	Variance Equation			
C	2.13E-06	8.83E-07	2.409647	0.0160
ARCH(1)	0.115045	0.014936	7.702539	0.0000
GARCH(1)	0.878693	0.014588	60.23554	0.0000

R-squared	0.014009	Mean dependent var	-0.000422
Adjusted R-squared	0.010837	S.D. dependent var	0.012957
S.E. of regression	0.012887	Akaike info criterion	-6.129119
Sum squared resid	0.361356	Schwarz criterion	-6.108280
Log likelihood	6700.998	F-statistic	4.416576
Durbin-Watson stat	2.165792	Prob(F-statistic)	0.000070
Inverted MA Roots	.98	.19	

Figure 4: Histogram-normality test for GARCH(1,1) on ARMA(2,2) of PX-50 returns

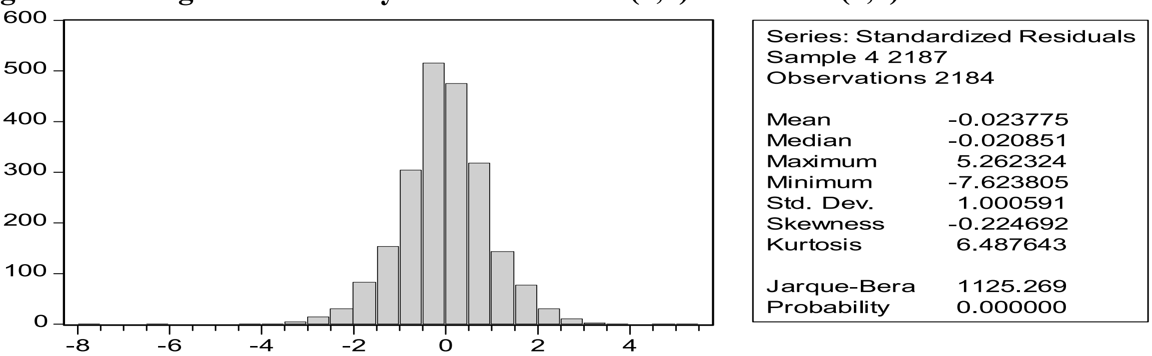
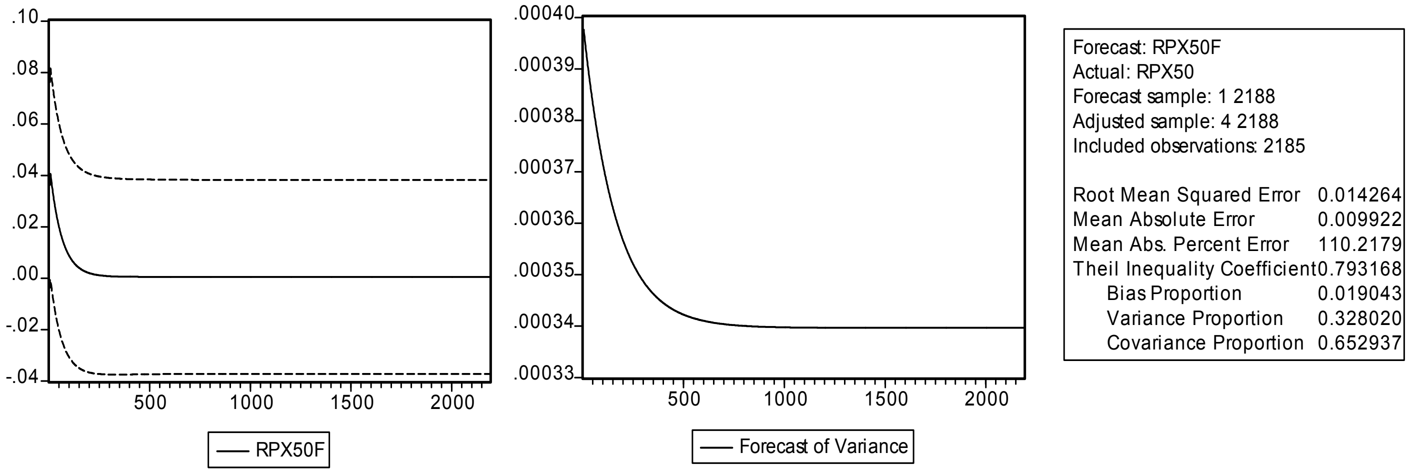
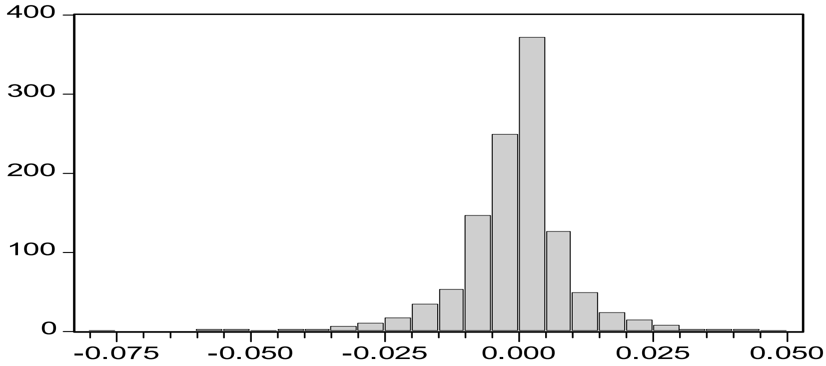


Figure 5: Forecast of PX-50 returns



The 1st sub-sample (before the Russian and Asian Crisis)

Figure 6: Histogram



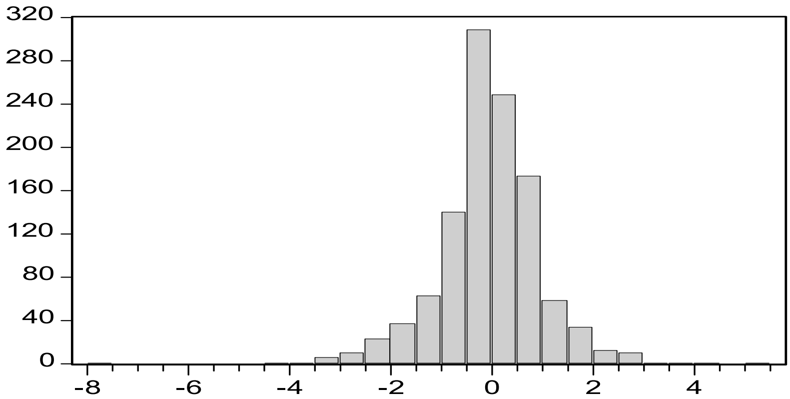
Series: CORRRPX50	
Sample 2 1133	
Observations 1132	
Mean	-0.000836
Median	0.000000
Maximum	0.047037
Minimum	-0.075664
Std. Dev.	0.010725
Skewness	-0.796877
Kurtosis	9.367350
Jarque-Bera	2032.091
Probability	0.000000

Table 5: GARCH(1,1) for the first sub-sample

Dependent Variable: RPX50
Method: ML – ARCH (Marquardt)
Sample(adjusted): 4 1133
Included observations: 1130 after adjusting endpoints
Convergence achieved after 24 iterations
Bollerslev-Wooldrige robust standard errors & covariance
Variance backcast: ON

	Coefficient	Std. Error	z-Statistic	Prob.
C	-9.14E-05	0.000247	-0.370684	0.7109
RPX50(-1)	0.310424	0.035915	8.643352	0.0000
RPX50(-2)	0.099132	0.037815	2.621517	0.0088
RFTSE	0.119925	0.035033	3.423236	0.0006
	Variance Equation			
C	2.66E-06	9.83E-07	2.706252	0.0068
ARCH(1)	0.138215	0.027874	4.958485	0.0000
GARCH(1)	0.844760	0.024404	34.61602	0.0000
R-squared	0.084963	Mean dependent var		-0.000860
Adjusted R-squared	0.080074	S.D. dependent var		0.010704
S.E. of regression	0.010266	Akaike info criterion		-6.646187
Sum squared resid	0.118361	Schwarz criterion		-6.615028
Log likelihood	3762.095	F-statistic		17.37874
Durbin-Watson stat	2.180112	Prob(F-statistic)		0.000000

Table 6: Histogram-normality test



Series: Standardized Residuals	
Sample 4 1133	
Observations 1130	
Mean	-0.036440
Median	-0.030533
Maximum	5.061244
Minimum	-7.643159
Std. Dev.	1.001475
Skewness	-0.433750
Kurtosis	7.771056
Jarque-Bera	1107.190
Probability	0.000000

The second sub-sample (after August, 1998)

Table 7: Correlogram

Sample: 1134 2187
Included observations: 1054

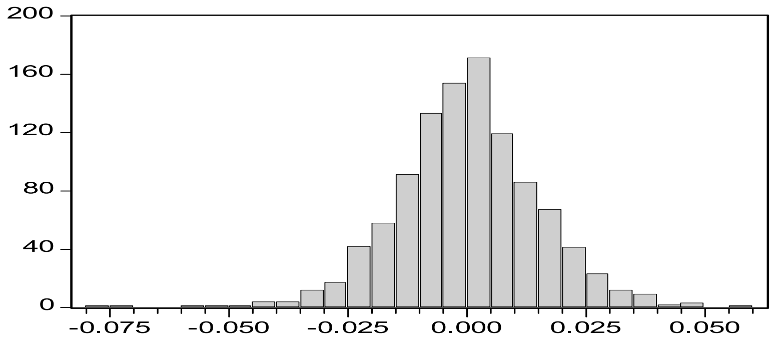
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.| |	.| |	1	0.055	0.055	3.1773	0.075
.| |	.| |	2	0.038	0.035	4.6735	0.097
.| |	.| |	3	0.015	0.011	4.9084	0.179
.| |	.| |	4	0.041	0.039	6.7285	0.151
.| |	*| |	5	-0.053	-0.058	9.6752	0.085
.| |	.| |	6	0.014	0.017	9.8730	0.130
.| |	.| |	7	0.021	0.022	10.328	0.171
.| |	.| |	8	-0.021	-0.025	10.806	0.213
.| |	.| |	9	-0.025	-0.020	11.465	0.245
.| |	.| |	10	0.043	0.043	13.442	0.200
.| |	.| |	11	0.005	0.002	13.470	0.264
.| |	.| |	12	0.026	0.027	14.186	0.289
.| |	.| |	13	0.049	0.044	16.785	0.209
.| |	.| |	14	0.001	-0.012	16.787	0.268
.| |	.| |	15	0.056	0.060	20.174	0.165
.|* |	.|* |	16	0.078	0.070	26.655	0.045
.| |	.| |	17	-0.035	-0.052	27.981	0.045

Table 8: Augmented Dickey-Fuller unit root test

Null Hypothesis: CORRRPX50 has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic based on SIC, MAXLAG=21)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-30.70151	0.0000
Test critical values:	1% level	-3.436336	
	5% level	-2.864071	
	10% level	-2.568169	

Table 9: Histogram



Series: CORRRPX50	
Sample 1134 2187	
Observations 1054	
Mean	4.84E-05
Median	0.000000
Maximum	0.058200
Minimum	-0.075287
Std. Dev.	0.014993
Skewness	-0.178540
Kurtosis	4.497667
Jarque-Bera	104.1051
Probability	0.000000

Table 10: OLS for the second sub-sample

Dependent Variable: RPX50
Method: Least Squares
Sample: 1134 2187
Included observations: 1054
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000151	0.000444	0.341218	0.7330
RDAX(-1)	0.119047	0.042446	2.804649	0.0051
RNASDAQ(-2)	0.057806	0.018826	3.070554	0.0022
RFTSE(-1)	0.099557	0.050495	1.971622	0.0489

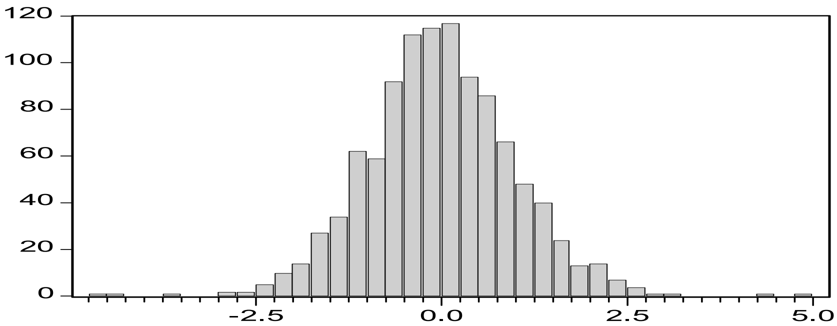
R-squared	0.069562	Mean dependent var	4.84E-05
Adjusted R-squared	0.066904	S.D. dependent var	0.014993
S.E. of regression	0.014482	Akaike info criterion	-5.627991
Sum squared resid	0.220224	Schwarz criterion	-5.609166
Log likelihood	2969.951	F-statistic	26.16689
Durbin-Watson stat	1.917489	Prob(F-statistic)	0.000000

Table 11: GARCH(1,1) for the second sub-sample

Dependent Variable: RPX50
Method: ML - ARCH (Marquardt)
Sample: 1134 2187
Included observations: 1054
Convergence achieved after 11 iterations
Bollerslev-Wooldrige robust standard errors & covariance
Variance backcast: ON

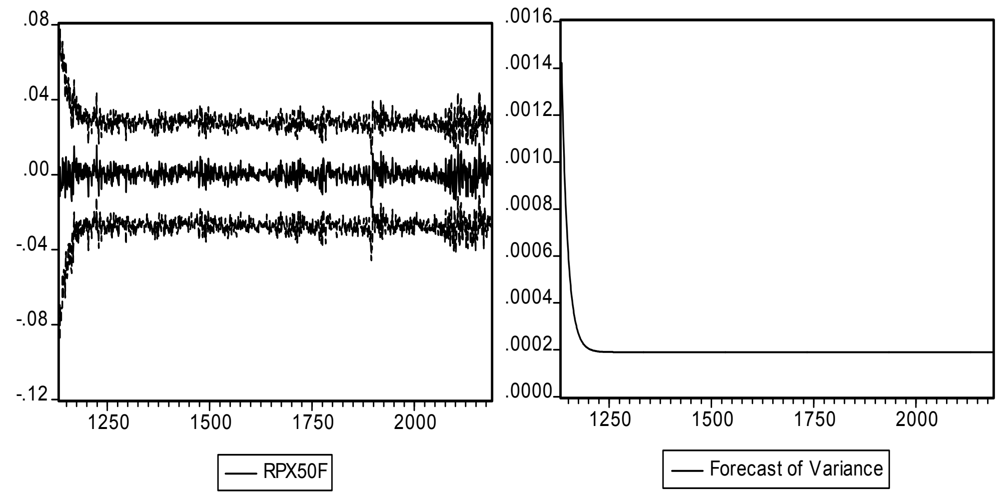
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000164	0.000383	0.427585	0.6690
RNASDAQ(-2)	0.050855	0.017231	2.951411	0.0032
RDAX(-1)	0.127558	0.039901	3.196878	0.0014
RFTSE(-1)	0.111614	0.043310	2.577114	0.0100
	Variance Equation			
C	1.22E-05	4.28E-06	2.841489	0.0045
ARCH(1)	0.069989	0.020922	3.345193	0.0008
GARCH(1)	0.865478	0.033157	26.10248	0.0000
R-squared	0.069077	Mean dependent var		4.84E-05
Adjusted R-squared	0.063742	S.D. dependent var		0.014993
S.E. of regression	0.014507	Akaike info criterion		-5.717824
Sum squared resid	0.220339	Schwarz criterion		-5.684880
Log likelihood	3020.293	F-statistic		12.94834
Durbin-Watson stat	1.916908	Prob(F-statistic)		0.000000

Figure 7: Histogram-normality test



Series: Standardized Residuals	
Sample 1134 2187	
Observations 1054	
Mean	-0.005088
Median	-0.028811
Maximum	4.843051
Minimum	-4.620757
Std. Dev.	1.001177
Skewness	0.059620
Kurtosis	4.350669
Jarque-Bera	80.74185
Probability	0.000000

Table 12: Forecast for the second sub-sample



Forecast: RPX50F	
Actual: RPX50	
Forecast sample: 1134 2187	
Included observations: 1054	
Root Mean Squared Error	0.014459
Mean Absolute Error	0.010900
Mean Abs. Percent Error	125.3223
Theil Inequality Coefficient	0.755189
Bias Proportion	0.000000
Variance Proportion	0.560602
Covariance Proportion	0.439398