Volatility Forecasting Models for CEE Stock Exchange Indices

MSc student Spătaru Marina
Coordinator Professor Moisă Altăr

Academy of Economic Studies Bucharest, Doctoral School of Finance and Banking

2014

Foreword

”Many investors think volatility is the same thing as risk, but it’s not. Being risk-averse doesn’t mean avoiding volatility. Don’t fear the market’s gyrations. Volatility is the best friend of the unemotional, patient, debt-free investor. A wildly fluctuating market means that solid businesses will occasionally be available for you to buy at irrationally low prices.”

Warren Buffett

Contents

Goals of the paper

- **To model and to forecast volatility using stochastic volatility models and GARCH models**

- **To analyze comparatively the predictive ability of these models**

- **To perform empirical exercises for five blue-chips indices from CEE capital markets, namely <u>BET</u> (Romania), <u>BUX</u> (Hungary), <u>BELEX15</u> (Serbia), <u>PX</u> (Czech Republic) and <u>WIG20</u> (Poland)**

- **To derive conclusions on the specific volatility of each market and their particular features**

Volatility in economics

Definition: *"Volatility is commonly allied to risk, in that it provides a measure of the possible variation or movement in a particular economic variable or some function of that variable."*

Frank Knigh, "Risk, Uncertainty, and Profit", 1921

General volatility measures

Standard deviation

- measure of absolute volatility
- shows how much an investment's return varies from its average return over time

Beta

- measure of relative volatility
- indicates the price variance of an investment compared to the market as a whole

General approach:

Higher expected returns can only occur with correspondingly higher risk (Portfolio Selection Theory developed by Markowitz and Capital Asset Pricing Model, developed by Sharpe)

Importance of volatility forecast

- **Volatility risk is considered as one of the prime and hidden risk factors on capital markets**

- **Forecasting accurately future volatility is essential to**
  - **derivatives pricing**
  - **optimal asset allocation**
  - **portfolio risk management**
  - **dynamic hedging**
  - **input for Value-at-Risk models**

- **The importance of volatility forecasting was highlighted when in 2003 Professor R.F. Engle was awarded the Noble prize for his contribution in modeling volatility dynamics**

Competing models

Main class of models that accommodate time-variation in variance:

GARCH models:

- Proposed by Bollerslev in 1986, based on the framework defined by Engle in 1982
- The conditional variance at time t is actually a function of three terms: i) the log-term average, ii) the news on past volatility measured by the ARCH term and iii) the previous conditional variance denoted by the GARCH term

Stochastic Volatility models:

- Introduced by Taylor in 1987
- The variance at time t is given by an unobserved variable that evolves accordingly to an autoregressive process

+ Very popular

+ Simple

+ Fast

+ Linear

- Large moves reduce the forecasting performance (R. Reider, 2009)

- Parameters resulted using different time scales give inconsistent results (M. Caporin, 2011)

+ Better out-of sample forecasting performance (Kim, Shephard, Chib, 1998)

- Less popular

- More complicated

- Non-linear

GARCH models

General framework GARCH (p,q):

$$\begin{cases} y_t = \mu + \varepsilon_t, & \text{for } t = 1, \dots T \quad (1) \\ \varepsilon_t = z_t \sigma_t, \quad z_t \sim N(0,1), & \text{for } t = 1, \dots T \quad (2) \\ \sigma^2_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon^2_{t-i} + \sum_{j=1}^{q} \beta_j \sigma^2_{t-j}, & \text{for } t = 2, \dots T \quad (3) \end{cases}$$

Conditions:

1. $\alpha_i > 0$ and $\beta_j > 0$ – ensures the positivity of the volatility process
2. $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ – avoids explosive variance and makes (3) stationary

Unconditional variance: $var(\varepsilon_t) = \frac{\omega}{1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j}$

Conditional variance: σ^2_t

Alternative GARCH (p,q) models

1. z_t is Gaussian, $\mathbf{f(z_t)} = \frac{1}{\sigma^2\sqrt{\pi}} e^{-\frac{(z_t-\mu)^2}{2\sigma^2}}$, where μ is the mean and σ is the standard deviation of z_t

2. z_t follows a Student-t distribution, $\mathbf{f(z_t)} = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} (1+ \frac{t^2}{\nu})^{-\frac{\nu+1}{2}}$, where ν is the number of degrees of freedom

3. z_t follows a GED, $\mathbf{f(z_t)} = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-(|z_t-\mu|/\alpha)^{\beta}}$, where μ is the location, α is the scale and β is the shape

Standard Stochastic Volatility model (SV)

General framework:

$$\begin{cases} y_t = \mu + e^{\frac{1}{2}ht}\,\varepsilon_t\,, & \varepsilon_t \sim N(0,1) \quad \text{for } t = 1, \ldots T \quad (4) \\ h_t = \mu_h + \phi_h\,(h_{t-1} - \mu_h) + \zeta_t, & \zeta_t \sim N(0,\sigma^2_h) \quad \text{for } t = 2, \ldots T \quad (5) \end{cases}$$

ε_t, ζ_t - i.i.d.

Linearized framework:

$$\begin{cases} y_t^* = h_t + \varepsilon_t^*,\ \varepsilon_t^* \mid s_t \sim N(\mu_{s_t} - 1.2704,\ \sigma^2_{s_t}) & \text{for } t = 1, \ldots T \quad (6) \\ h_t = \mu_h + \phi_h\,(h_{t-1} - \mu_h) + \zeta_t, \quad \zeta_t \sim N(0,\sigma^2_h) & \text{for } t = 2, \ldots T \quad (7) \end{cases}$$

$y_t^* = \log((y_t - \mu)^2 + c)$, $\varepsilon_t^* = \log \varepsilon_t^2$, $s_t \in \{1, 2, 3, 4, 5, 6, 7\}$

$\varepsilon_t^* \sim \log\text{-}\chi^2(1)$, but It can be estimated using a seven-component Gaussian mixture (Kim, Shepherd, Chib,1998)

Condition:

$|\ \phi_h\ | < 1$ – ensures that the AR(1) process in (5) is stationary

Conditional variance: $Var\ (y_t / h_t) = e^{h_t}$

Moving Average Stochastic Volatility Model (MASV)

General framework:

$$\begin{cases} y_t = \mu + u_t, & & \text{for } t = 1, \ldots T & (8) \\ u_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \ldots + \psi_q \varepsilon_{t-q}, & \varepsilon_t \sim N(0, e^{h_t}) & \text{for } t = 1, \ldots T & (9) \\ h_t = \mu_h + \phi_h (h_{t-1} - \mu_h) + \zeta_t, & \zeta_t \sim N(0, \sigma^2_h) & \text{for } t = 2, \ldots T & (10) \end{cases}$$

εt, ζt - i.i.d.

SV and MASV are nested: for $\psi_1 = \psi_2 = \ldots = \psi_q = 0$, MASV turns into SV

Conditions:

1. $| \phi_h | < 1$ – ensures that the AR(1) process in (10) is stationary
2. $\psi_p < 1$, $p = \overline{1, q}$ – makes the MA(q) process in (9) invertible

Conditional variance: $\text{Var}(y_t | \mu, \psi, \mathbf{h}) = e^{h_t} + \psi_1^2 e^{h_{t-1}} + \ldots + \psi_q^2 e^{h_{t-q}}$

Conditional covariance: $\text{Cov}(y_t, y_{t-j} | \mu, \psi, \mathbf{h}) = \begin{cases} \sum_{i=0}^{q-j} e^{h_{t-i}} \psi_{i+j} \psi_i, & j = 1, \ldots, q \\ 0 & , j > q \end{cases}$,

$\boldsymbol{\psi} = (\psi_{1,\ldots,} \psi_q)'$, $\mathbf{h} = (h_1, \ldots, h_T)'$

Estimation of the Stochastic Volatility Models

Independent priors: $p(\mu, \psi, \mu_h, \phi_h, \sigma^2_h) = p(\mu)\, p(\psi)\, p(\mu_h)\, p(\phi_h)\, p(\sigma^2_h)$

$\psi \sim N(\psi_0, V_\psi)\, 1(|\psi|<1)$

$\mu \sim N(\mu_0, V_\mu)$

$\mu_h \sim N(\mu_{h0}, V_{\mu h})$

$\phi_h \sim N(\phi_{h0}, V_{\phi_h})\, 1(|\phi_h|<1)$

$\sigma^2_h \sim IG(\nu_h, S_h)$

Initialization of the state: $h_1 \sim N(\mu_h, \sigma^2_h/(1-\phi^2_h))$

Bayesian analysis:

$$\underbrace{p(\mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h|\mathbf{y})}_{\text{Joint posterior distribution}} = \underbrace{p(\mathbf{y}|\mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h)}_{\text{Likelihood function}}\, p(\mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h)$$

$p(\mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h|\mathbf{y}) =$

Gibbs sampler

1. $p(\mu \mid \mathbf{y}, \mathbf{h}, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h) = p(\mu \mid \mathbf{y}, \mathbf{h}, \boldsymbol{\psi})$;
2. $p(\mathbf{h} \mid \mathbf{y}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h, \sigma^2_h)$;
3. $p(\boldsymbol{\psi} \mid \mathbf{y}, \mathbf{h}, \mu, \mu_h, \phi_h, \sigma^2_h) = p(\boldsymbol{\psi} \mid \mathbf{y}, \mathbf{h}, \mu)$;
4. $p(\mu_h \mid \mathbf{y}, \mathbf{h}, \mu, \boldsymbol{\psi}, \phi_h, \sigma^2_h) = p(\mu_h \mid \mathbf{h}, \phi_h, \sigma^2_h)$;
5. $p(\phi_h \mid \mathbf{y}, \mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \sigma^2_h) = p(\phi_h \mid \mathbf{h}, \mu_h, \sigma^2_h)$;
6. $p(\sigma^2_h \mid \mathbf{y}, \mathbf{h}, \mu, \boldsymbol{\psi}, \mu_h, \phi_h) = p(\sigma^2_h \mid \mathbf{h}, \mu_h, \phi_h)$

Evaluating the likelihood function

$u_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \ldots + \psi_q \varepsilon_{t-q},\ \varepsilon_t \sim N(0, e^{h_t}) \quad \Rightarrow \quad \mathbf{u} = \mathbf{H}_\psi\, \boldsymbol{\varepsilon}$ (11)

$$\mathbf{H}_\psi = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \psi_1 & 1 & \vdots & 0 \\ \psi_2 & \psi_1 & \vdots & 0 \\ \vdots & \psi_2 & \vdots & 0 \\ \psi_q & \vdots & \vdots & 0 \\ 0 & \psi_q & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad |\mathbf{H}_\psi| = 1$$

$\mathbf{u} \sim N(\mathbf{0}, \Sigma_\mathbf{y})$

$\Sigma_\mathbf{y} = \mathbf{H}_\psi \mathbf{S}_\mathbf{y} \mathbf{H}_\psi'$, $\quad \mathbf{S}_\mathbf{y} = \mathrm{diag}(e^{h_1}, \ldots, e^{h_T})$, $\quad |\Sigma_\mathbf{y}| = |\mathbf{S}_\mathbf{y}| = \exp\{\sum_{t=1}^{T} h_t\}$

$(\mathbf{y} \mid \mu, \psi, \mathbf{h}) \sim N(\mu 1, \Sigma_\mathbf{y})$

$$\log p(\mathbf{y} \mid \mu, \psi, \mathbf{h}) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} h_t - \frac{1}{2}(\mathbf{y} - \mu 1)' \Sigma_\mathbf{y}^{-1} (\mathbf{y} - \mu 1)$$

Sampling the state

$$p(\mathbf{h}|\mathbf{y}, \mu, \psi, \mu_h, \phi_h, \sigma^2_h) \propto p(\mathbf{y}|\mathbf{h}, \mu, \psi)\, p(\mathbf{h}|\mu_h, \phi_h, \sigma^2_h) = \underbrace{p(\mathbf{y}^*|\mathbf{h}, s, \psi)}_{1.}\, \underbrace{p(\mathbf{h}|\mu_h, \phi_h, \sigma^2_h)}_{2.}$$

1. $\ln p(\mathbf{y}^*|\mathbf{h}, s, \psi) = -\frac{1}{2}(\mathbf{y}^* - \mathbf{h} - \mathbf{d})' \Sigma_{\mathbf{y}^*}^{-1} (\mathbf{y}^* - \mathbf{h} - \mathbf{d}) + c1,$

$\mathbf{d} = (\mu_{s1} - 1.2704, \dots, \mu_{sT} - 1.2704)'$, $\Sigma_{\mathbf{y}^*} = \mathrm{diag}(\sigma^2_{s1}, .., \sigma^2_{sT})$

2. $\log p(\mathbf{h}|\mu_h, \phi_h, \sigma^2_h) = -\frac{1}{2}(\mathbf{h} - \mathbf{H}_{\phi h}^{-1} \tilde{\alpha})' \mathbf{H}_{\phi h}' \Sigma_{\mathbf{h}}^{-1} \mathbf{H}_{\phi h} (\mathbf{h} - \mathbf{H}_{\phi h}^{-1} \tilde{\alpha}) + c2$

$$\Sigma_{\mathbf{h}} = \mathrm{diag}(\sigma^2_h/(1-\varphi^2_h), \sigma^2_h, \dots, \sigma^2_h),\ \tilde{\alpha} = \begin{pmatrix} \mu_h \\ (1-\phi h)\mu_h \\ \vdots \\ (1-\phi h)\mu_h \end{pmatrix},\ \mathbf{H}_{\varphi h} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi h & 1 & 0 & \cdots & 0 \\ 0 & -\phi h & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\phi h & 1 \end{pmatrix}.$$

$\log p(\mathbf{h}|\mathbf{y}, \mu, \psi, \mu_h, \phi_h, \sigma^2_h) = -\frac{1}{2}(\mathbf{h}'\mathbf{K_h}\,\mathbf{h} - 2\mathbf{h}\,\mathbf{K_h}\,\hat{\mathbf{h}}) + c3,$

$\mathbf{K_h} = \mathbf{H}_{\phi h}' \Sigma_{\mathbf{h}}^{-1} \mathbf{H}_{\phi h} + \Sigma_{\mathbf{y}^*}^{-1}$ and $\hat{\mathbf{h}} = \mathbf{K_h^{-1}} (\mathbf{H}_{\phi h}' \Sigma_{\mathbf{h}}^{-1} \tilde{\alpha} + \Sigma_{\mathbf{y}^*}^{-1} (\mathbf{y}^* - \mathbf{d}))$

Data description

y_t = Daily returns of blue-chips indices from CEE stock exchanges
$= Index_t / Index_{t-1} - 1$

Index	Stock exchange	No. of stocks aggregated	Period	No. of observations	Source
BET	Bucharest Stock Exchange	10	Jan 3, 2005 – May 9, 2014	2,341	BSE
PX	Prague Stock Exchange	14	Jan 3, 2005 – May 9, 2014	2,436	Erste Equity Research Fact Set
BELEX15	Belgrad Stock Exchange	15	Oct 5, 2005 – May 9, 2014	2,216	Erste Equity Research Fact Set
BUX	Budapest Stock Exchange	up to 25	Jan 3, 2005 – May 9, 2014	2,433	Erste Equity Research Fact Set
WIG20	Warsaw Stock Exchange	20	Jan 3, 2005 – May 9, 2014	2,341	Erste Equity Research Fact Set

Descriptive statistics for the indices' returns time series

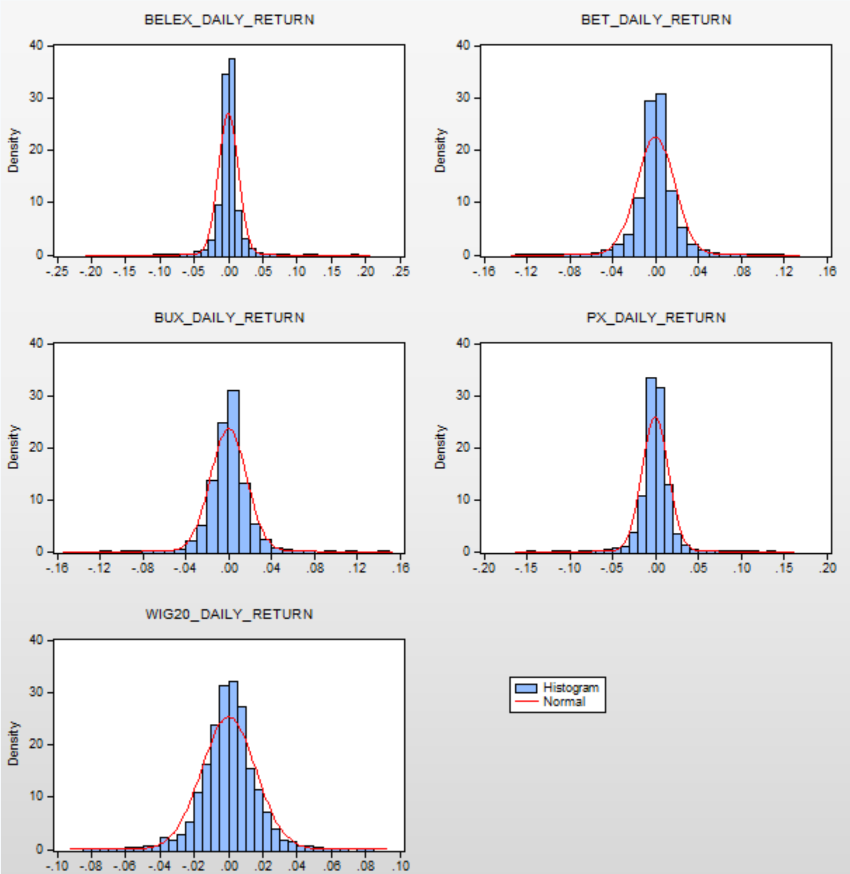
Indicator	Blue-chips index				
	BET	PX	BELEX15	BUX	WIG20
Mean	0.000331	0.000108	0.999864	0.000228	0.000209
Maximum	0.111427	0.131609	0.18912	0.140854	0.084966
Minimum	-0.122929	-0.149435	-0.102923	-0.118817	-0.080962
Std. Dev.	0.017684	0.015417	0.014748	0.016905	0.015779
Skew ness	-0.372859	-0.165682	1.334625	0.142016	-0.184303
Kurtosis	9.813614	16.0382	27.07185	9.685639	5.819956
Jarque-Bera	4582.647	19351.51	54160.79	4542.126	788.919

Source: own computes

Data features – Normal density test

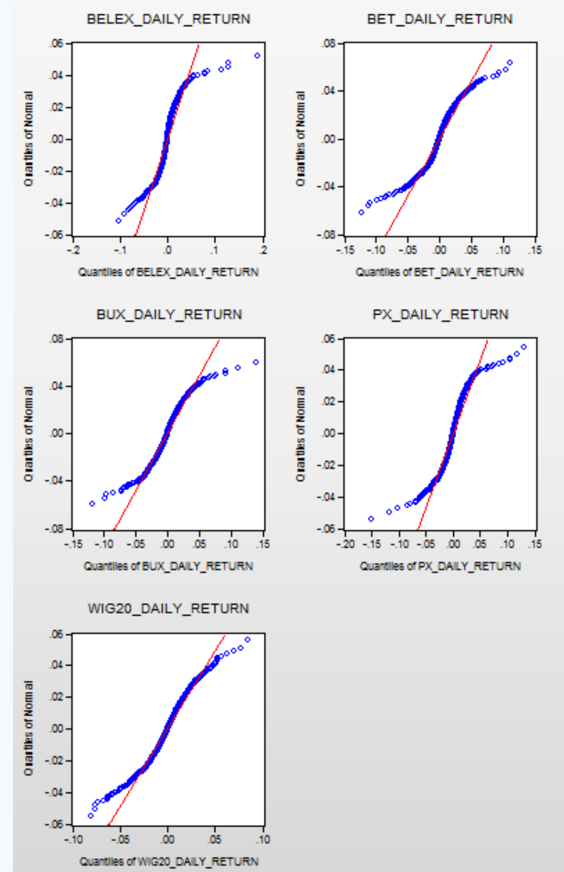
Positive excess kurtosis confirms the usual leptokurtic distributions of stock prices' returns

Empirical density of daily returns vs. the normal density



Source: Own computes (E-Views output)

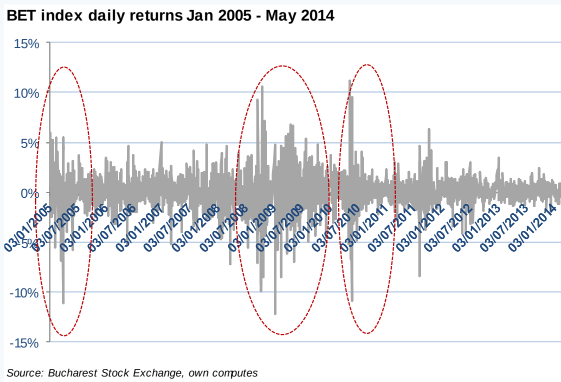
Q-Q plots for indices' daily returns

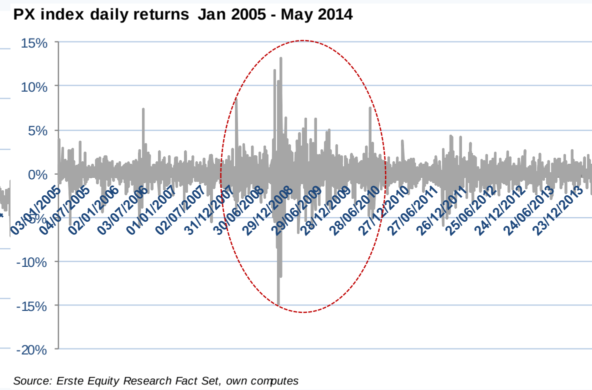


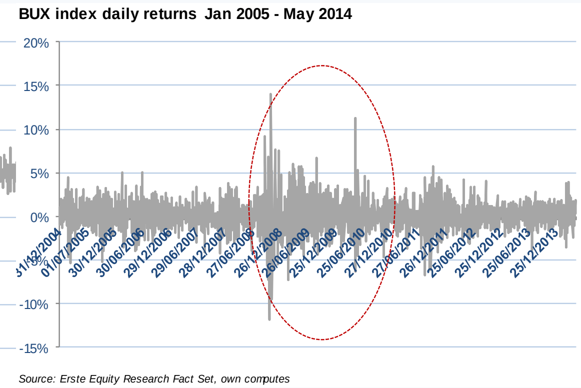
Source: Own computes (E-Views output)

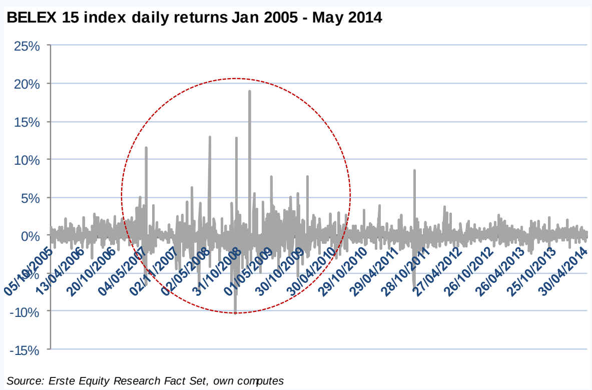
Data features – Heteroskedasticity

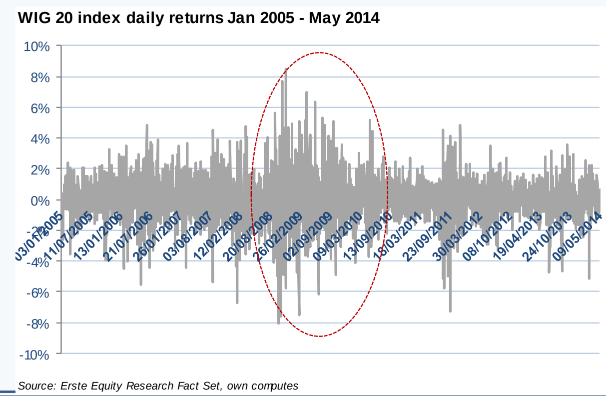
Daily returns development for CEE blue-chips indices











Time series do not have constant variance over time => the need of modeling the data in a time-varying framework

Large volatility periods are roughly the same for all the five CEE markets considered (volatility clusters mostly overlap)

Data features – Serial independence test

Autocorrelation coefficients of daily returns

Correlogram of BET_DAILY_RETURN

Date: 05/29/14 Time: 14:49
Sample: 12/31/2004 5/09/2014
Included observations: 2341

Lag	AC	PAC	Q-Stat	Prob
1	0.092	0.092	19.710	0.000
2	-0.010	-0.019	19.962	0.000
3	-0.033	-0.030	22.480	0.000
4	-0.020	-0.014	23.399	0.000
5	0.037	0.040	26.626	0.000
6	-0.014	-0.023	27.099	0.000
7	0.002	0.005	27.106	0.000
8	0.052	0.053	33.395	0.000
9	0.027	0.018	35.111	0.000
10	-0.004	-0.010	35.157	0.000
11	0.037	0.044	38.369	0.000
12	0.006	0.001	38.450	0.000
13	0.028	0.025	40.343	0.000
14	0.046	0.044	45.257	0.000
15	0.053	0.050	51.966	0.000

Correlogram of PX_DAILY_RETURN

Date: 05/29/14 Time: 14:52
Sample: 12/31/2004 5/09/2014
Included observations: 2436

Lag	AC	PAC	Q-Stat	Prob
1	0.082	0.082	16.292	0.000
2	-0.075	-0.082	30.072	0.000
3	-0.057	-0.044	37.908	0.000
4	0.007	0.010	38.044	0.000
5	0.043	0.035	42.629	0.000
6	-0.036	-0.045	45.793	0.000
7	-0.006	0.008	45.878	0.000
8	0.017	0.015	46.550	0.000
9	0.012	0.005	46.915	0.000
10	-0.030	-0.031	49.091	0.000
11	-0.024	-0.013	50.507	0.000
12	0.060	0.059	59.223	0.000
13	0.004	-0.013	59.269	0.000
14	0.022	0.031	60.454	0.000
15	-0.032	-0.028	62.953	0.000

Correlogram of BELEX_DAILY_RETURN

Date: 05/29/14 Time: 14:53
Sample: 12/31/2004 5/09/2014
Included observations: 2216

Lag	AC	PAC	Q-Stat	Prob
1	0.274	0.274	167.06	0.000
2	0.111	0.039	194.41	0.000
3	0.019	-0.022	195.25	0.000
4	0.032	0.031	197.57	0.000
5	-0.007	-0.023	197.68	0.000
6	0.013	0.018	198.06	0.000
7	0.028	0.025	199.78	0.000
8	0.067	0.055	209.85	0.000
9	0.054	0.022	216.38	0.000
10	0.014	-0.016	216.84	0.000
11	0.011	0.007	217.12	0.000
12	0.033	0.029	219.52	0.000
13	0.063	0.050	228.27	0.000
14	0.099	0.073	250.33	0.000
15	0.100	0.051	272.76	0.000

Correlogram of BUX_DAILY_RETURN

Date: 05/29/14 Time: 14:54
Sample: 12/31/2004 5/09/2014
Included observations: 2433

Lag	AC	PAC	Q-Stat	Prob
1	0.059	0.059	8.5513	0.003
2	-0.081	-0.085	24.704	0.000
3	-0.023	-0.013	25.971	0.000
4	0.088	0.084	44.932	0.000
5	0.028	0.014	46.819	0.000
6	-0.065	-0.055	57.027	0.000
7	-0.006	0.009	57.109	0.000
8	0.014	-0.001	57.586	0.000
9	-0.027	-0.034	59.309	0.000
10	-0.044	-0.030	64.072	0.000
11	0.020	0.024	65.074	0.000
12	0.029	0.015	67.125	0.000
13	-0.001	0.002	67.131	0.000
14	-0.039	-0.028	70.947	0.000
15	-0.014	-0.015	71.456	0.000

Correlogram of WIG20_DAILY_RETURN

Date: 05/29/14 Time: 14:56
Sample: 12/31/2004 5/09/2014
Included observations: 2341

Lag	AC	PAC	Q-Stat	Prob
1	0.040	0.040	3.8009	0.051
2	-0.056	-0.057	11.063	0.004
3	-0.003	0.002	11.086	0.011
4	-0.002	-0.005	11.093	0.026
5	-0.009	-0.009	11.271	0.046
6	-0.006	-0.005	11.343	0.078
7	0.001	0.000	11.345	0.124
8	-0.032	-0.033	13.797	0.087
9	0.019	0.022	14.619	0.102
10	0.015	0.009	15.147	0.127
11	-0.015	-0.014	15.702	0.153
12	-0.020	-0.017	16.606	0.165
13	0.007	0.007	16.733	0.212
14	-0.012	-0.014	17.052	0.253
15	0.041	0.044	21.106	0.133

Source: Own computes (E-Views output)

=>Up to lag 15 (at least), daily returns seem to be auto-correlated (confirmation of the volatility clustering phenomenon)
=>Exception: WIG20 for which auto-correlation is confirmed only up to lag 5

Data features – Stationarity tests

ADF and KPSS tests for stationarity

Null Hypothesis: BET_DAILY_RETURN has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-44.20562	0.0001
Test critical values:	1% level	-3.432949	
	5% level	-2.862574	
	10% level	-2.567366	

Null Hypothesis: BET_DAILY_RETURN is stationary
Exogenous: Constant
Bandwidth: 11 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.143856
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

Null Hypothesis: PX_DAILY_RETURN has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-36.33074	0.0000
Test critical values:	1% level	-3.432842	
	5% level	-2.862527	
	10% level	-2.567341	

Null Hypothesis: PX_DAILY_RETURN is stationary
Exogenous: Constant
Bandwidth: 6 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.122681
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

Null Hypothesis: BELEX_DAILY_RETURN has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-35.50128	0.0000
Test critical values:	1% level	-3.433106	
	5% level	-2.862643	
	10% level	-2.567403	

Null Hypothesis: BELEX_DAILY_RETURN is stationary
Exogenous: Constant
Bandwidth: 14 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.221117
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

Null Hypothesis: BUX_DAILY_RETURN has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-23.43465	0.0000
Test critical values:	1% level	-3.432848	
	5% level	-2.862529	
	10% level	-2.567342	

Null Hypothesis: BUX_DAILY_RETURN is stationary
Exogenous: Constant
Bandwidth: 0 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.123634
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

Null Hypothesis: WIG20_DAILY_RETURN has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-46.44174	0.0001
Test critical values:	1% level	-3.432949	
	5% level	-2.862574	
	10% level	-2.567366	

Null Hypothesis: WIG20_DAILY_RETURN is stationary
Exogenous: Constant
Bandwidth: 14 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.122270
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

Source: Own computes (E-Views output)

Augmented Dickey Fuller (ADF) and Kwiatokski-Phillips-Schmidts-Shin Unit Root (KPSS) tests have opposite null hypothesis, which strengthens the result that all the five data series are stationary

GARCH models comparison

Hansen P., Lunde A. (2004) => **there is no significant evidence that higher-order GARCH outperforms GARCH(1,1) model**

Error distribution of GARCH (1,1)		Blue-chips index				
		BET	PX	BELEX15	BUX	WIG20
Akaike info criterion	Gauss	-5.703709	-6.009348	-6.188887	-5.617342	-5.687982
	t-Student	-5.751796	-6.052126	-6.374738	-5.641363	-5.707365
	GED	-5.746878	-6.047423	-6.367566	-5.646466	-5.706903
Schwarz criterion	Gauss	-5.691406	-5.997439	-6.176014	-5.605421	-5.675675
	t-Student	-5.737032	-6.037835	-6.359289	-5.627057	-5.692596
	GED	-5.732114	-6.033132	-6.352118	-5.63216	-5.692134
Log likelihood	Gauss	6678.339	7318.377	6859.193	6832.88	6657.096
	t-Student	6735.602	7371.438	7066.022	6863.076	6680.764
	GED	6729.847	7365.714	7058.079	6869.2979	6680.224
		⇓ t-Student error distribution	⇓ t-Student error distribution	⇓ t-Student error distribution	⇓ GED error distribution	⇓ t-Student error distribution

Source: Own computes

AIC = $\log(\sigma^2) + \frac{2k}{T}$

k = no. of parameters

T = no. of observations

SBIC = $\log(\sigma^2) + \frac{k}{T}\log(T)$,

k = no. of parameters

T = no. of observations

Lnorm= $\frac{1}{2}\sum_{t=1}^{T}[\log(2\pi) + \log(\sigma^2) + z^2]$, $z^2 = \frac{\varepsilon_t}{\sigma_t}$

Lstud = T {log $\Gamma(\frac{\nu+1}{2})$-log $\Gamma(\frac{\nu}{2})$- $\frac{1}{2}\log[(\nu-2)]\} - \frac{1}{2}\sum_{t=1}^{T}[\log(\sigma^2) + (1 + \quad)\log(1 + \frac{z}{\nu-2})]$, ν = no. of degrees of freedom

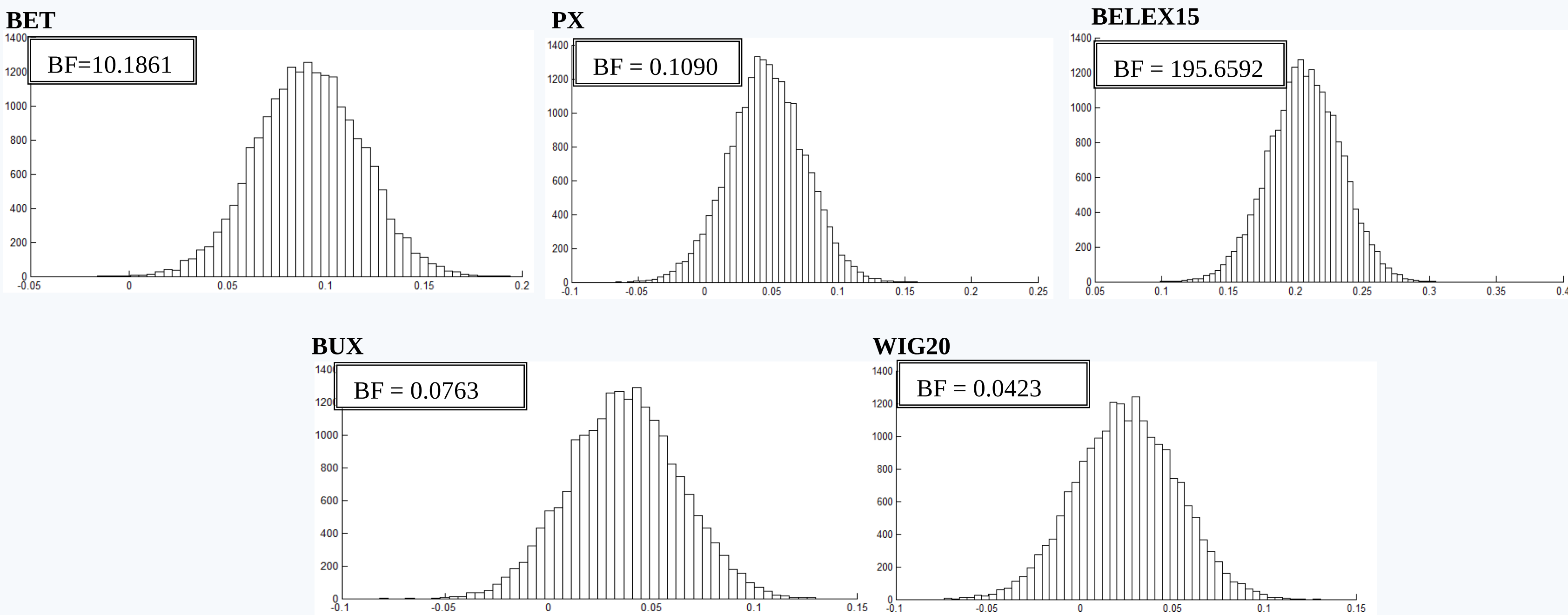
LGED = $\sum_{t=1}^{T}[\log\left(\frac{\nu}{\lambda}\right) - 0.5\left|\frac{z}{\lambda}\right| - (1+\lambda^{-1})\log(2) - \log\Gamma(\frac{1}{\nu}) - 0.5\log(\sigma^2)]$, $\lambda = \sqrt{\frac{\Gamma(\frac{1}{\nu})2^{\frac{-2}{\nu}}}{\Gamma(\frac{3}{\nu})}}$

Stochastic Volatility models comparison

SV and MASV (q=1) are **nested models** => if $\psi = 0$, MASV becomes SV

$$\mathbf{BF} = \frac{p(\psi=0)}{p(\psi=0|y)} = \frac{\text{a priori probability to have } \psi=0}{\text{a posteriori probability to have } \psi=0}$$

Estimate of $p(\psi|y)$ within the MASV model



Source: Own computes (Matlab output)

BET, BELEX 15 => MASV models
PX, BUX, WIG20 => SV model

Parameters of the models

BET

MASV

parameter	posterior mean	posterior stdev.	5%-tile	95%-tile
μ	0.0007	0.0004	0.0001	0.0013
μ_h	-7.7713	0.1142	-8.3034	-7.1386
ϕ_h	0.9878	0.0052	0.9794	0.9971
σ^2_h	0.0122	0.0024	0.0088	0.0168
ψ	0.0916	0.0266	0.0482	0.1348

GARCH(1,1), Student-t Errors

parameter	mean	stdev	5%-tile	95%-tile
μ	0.0006	0.0002	0.0001	0.0011
AR(1)	0.0806	0.0213	0.0387	0.1224
ω	53.5×10^{-7}	12.5×10^{-7}	28.9×10^{-7}	78×10^{-7}
α	0.1841	0.0215	0.1419	0.2263
β	0.8125	0.0184	0.7765	0.8486

PX

SV

parameter	posterior mean	posterior stdev.	5%-tile	95%-tile
μ	0.0005	0.0003	0.0001	0.0010
μ_h	-8.2450	0.1763	-8.4793	-8.0079
ϕ_h	0.9825	0.0051	0.9738	0.9902
σ^2_h	0.0123	0.0021	0.0093	0.0160

GARCH(1,1), Student-t Errors

parameter	mean	stdev	5%-tile	95%-tile
μ	0.0007	0.0002	0.0003	0.0011
AR(2)	-0.0455	0.0211	-0.0868	-0.0041
ω	45.9×10^{-7}	11.4×10^{-7}	23.7×10^{-7}	8.2×10^{-7}
α	0.1299	0.0173	0.0960	0.1637
β	0.8496	0.0182	0.8139	0.8853

BELEX15

MASV

parameter	posterior mean	posterior stdev.	5%-tile	95%-tile
μ	0.0001	0.0004	0.0000	0.0007
μ_h	-8.2886	0.8414	-8.6346	-8.0869
ϕ_h	0.9820	0.0062	0.9716	0.9915
σ^2_h	0.0139	0.0027	0.0100	0.0189
ψ	0.2067	0.0269	0.1623	0.2500

GARCH(1,1), Student-t Errors

parameter	mean	stdev	5%-tile	95%-tile
μ	0.0000	0.0002	0.9996	1.0004
AR(1)	0.2135	0.0214	0.1715	0.2555
ω	63×10^{-7}	13.2×10^{-7}	37×10^{-7}	88.5×10^{-7}
α	0.2436	0.0328	0.1833	0.3120
β	0.7541	0.0218	0.7116	0.7968

BUX

SV

parameter	posterior mean	posterior stdev.	5%-tile	95%-tile
μ	0.0004	0.0003	0.0001	0.0010
μ_h	-0.8023	0.2206	-8.2398	-7.8078
ϕ_h	0.9827	0.0052	0.9738	0.9905
σ^2_h	0.0101	0.0018	0.0075	0.0134

GARCH(1,1), GED

parameter	mean	stdev	5%-tile	95%-tile
μ	0.0004	0.0002	0.9996	1.0005
AR(2)	-0.0413	0.0205	-0.0815	-0.0012
ω	47.4×10^{-7}	13.1×10^{-7}	21.7×10^{-7}	73×10^{-7}
α	0.0936	0.0132	0.0678	0.1195
β	0.8897	0.0151	0.8602	0.9193

WIG20

SV

parameter	posterior mean	posterior stdev.	5%-tile	95%-tile
μ	0.0004	0.0004	0.0001	0.0010
μ_h	-8.0144	0.6120	-8.2928	-7.8311
ϕ_h	0.9840	0.0054	0.9751	0.9923
σ^2_h	0.0084	0.0015	0.0062	0.0110

GARCH(1,1), Student-t Errors

parameter	mean	stdev	5%-tile	95%-tile
μ	0.0005	0.0002	0.0000	0.0010
AR(2)	-0.0417	0.0218	-0.0844	-0.0009
ω	21.2×10^{-7}	7.8×10^{-7}	58.9×10^{-7}	15.5×10^{-7}
α	0.0592	0.0093	0.0411	0.0774
β	0.9325	0.0104	0.9122	0.9530

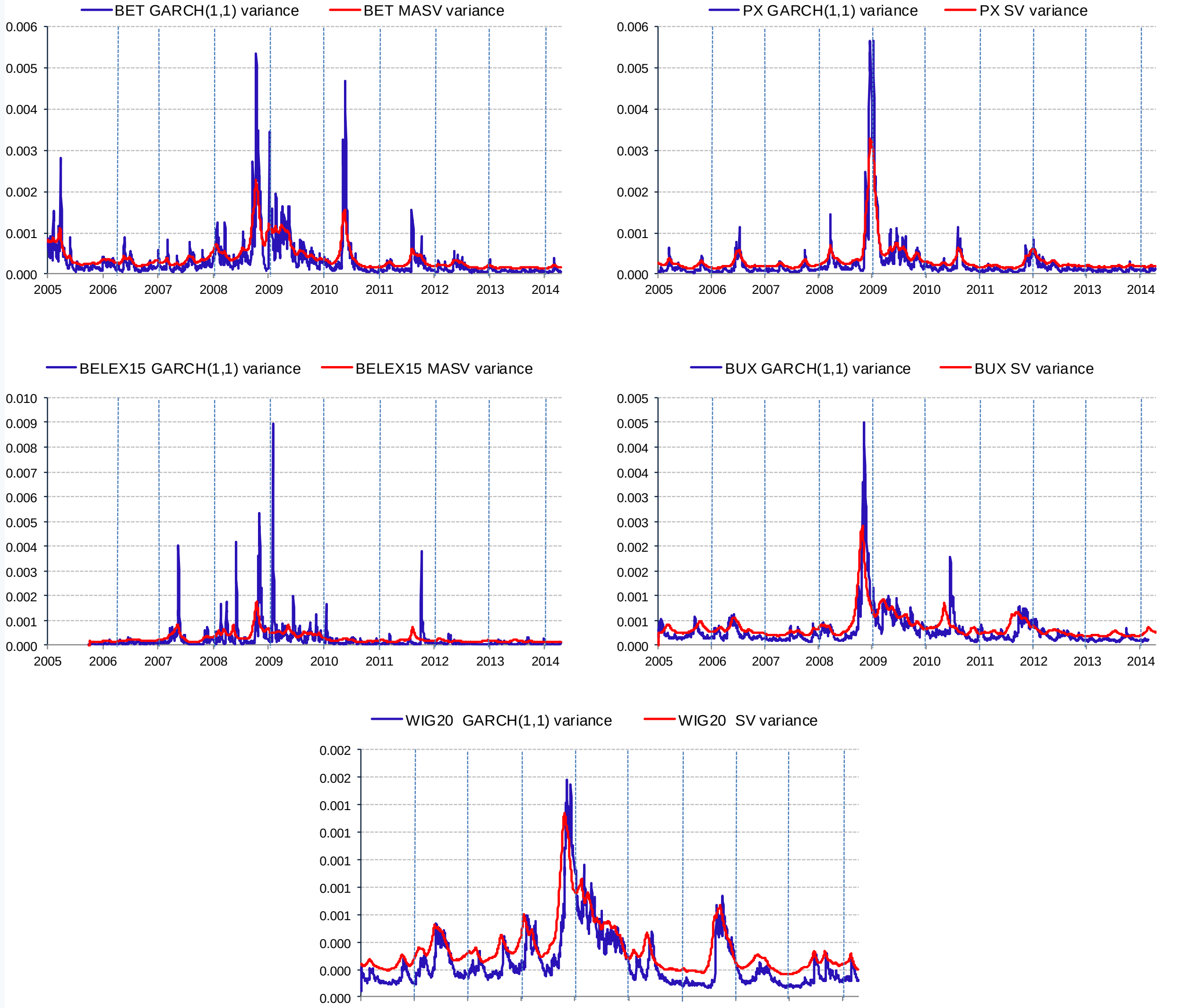
=> For all the five GARCH models, **α + β is below 1**, which means that there are no explosive phenomena in the framework

=> **Parameters are quite precisely estimated**, with low standard deviation and narrow confidence interval

=> the **Romanian BET index brings one of the highest returns** (μ) among the five indices considered, as modeled by both GARCH and SV/MASV

Conditional variance output

Variances modeled with stochastic volatility models and GARCH(1,1) models



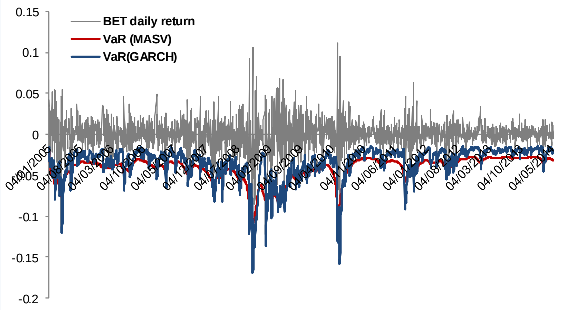
Source: Own computes

The use of past errors in modeling the volatility by GARCH models, **induces higher conditional variance** compared to the output from Stochastic Volatility models

GARCH vs. SV – backtesting volatility forecasts

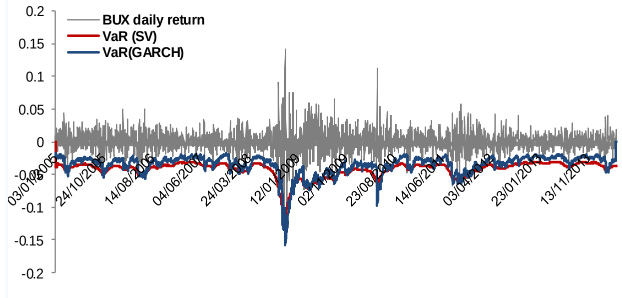
VaR metric calculated by GARCH and Stochastic Volatility models

BET development and limits-violation backtesting for VaR of 1%



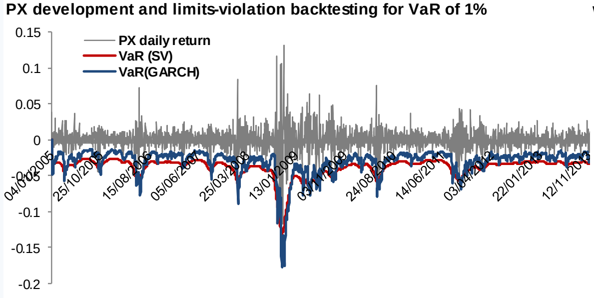
Source: BSE, own computes

BUX development and limits-violation backtesting for VaR of 1%



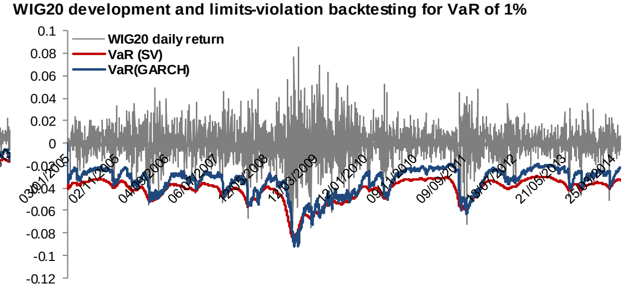
Source: Erste Fact Set, own computes

PX development and limits-violation backtesting for VaR of 1%



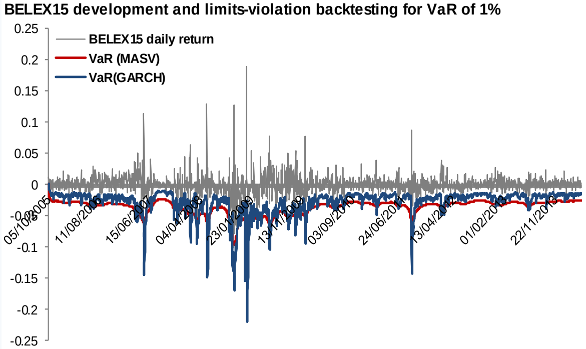
Source: Erste Fact Set, own computes

WIG20 development and limits-violation backtesting for VaR of 1%



Source: Erste Fact Set, own computes

BELEX15 development and limits-violation backtesting for VaR of 1%



Source: Erste Fact Set, own computes

Index	Model	Limits-violations (% of total sample)
BET	MASV	0.52%
	GARCH(1,1) -t Student	1.98%
PX	SV	0.39%
	GARCH(1,1) -t Student	1.89%
BELEX15	MASV	0.45%
	GARCH(1,1) -t Student	1.31%
BUX	SV	0.60%
	GARCH(1,1) -GED	1.55%
WIG20	SV	0.52%
	GARCH(1,1) -t Student	1.67%

Comparing method: testing the violation of the VaR limits given by the number of excesses outside the confidence interval

VaR <= normal distribution, 1% significance level, one-day-ahead volatility forecast

VaRs computed with <u>one-day-ahead volatility forecast</u> from SV/MASV had fewer violations than the ones calculated with the GARCH

GARCH vs. SV – out-of-sample volatility forecast (I)

Out -of- sample horizon: February 2, 2014 – May 9, 2014 (66-68 observations)

Descriptive statistics for the indices' daily returns
in-sample period

Indicator	Blue-chips index				
	BET	PX	BELEX15	BUX	WIG20
Mean	0.000328	0.000104	-0.000147	0.00025	0.00021
Std. Dev.	0.017887	0.015557	0.014501	0.016984	0.015887
Skew ness	-0.366821	-0.165442	1.32185	0.138133	-0.176609
Kurtosis	9.64473	16.64321	26.41562	9.728508	5.768601
Jarque-Bera	4232.578	18391.86	497977.01	4472.56	738.4192

Source: own computes

Descriptive statistics for the indices' daily returns
out-of-sample period

Indicator	Blue-chips index				
	BET	PX	BELEX15	BUX	WIG20
Mean	0.000427	0.000255	-0.000233	0.000233	0.000182
Std. Dev.	0.008529	0.009102	0.004369	0.013859	0.011546
Skew ness	-1.231324	-0.142936	0.9216328	0.3027	-0.851902
Kurtosis	7.910103	14.344603	13.861707	3.778663	7.5488256
Jarque-Bera	85.49229	94.992164	92.518022	62.675271	64.885601

Source: own computes

GARCH vs. SV – out-of-sample volatility forecast (II)

Comparing method: minimum Mean Absolute Error (MAE), Mean Square Error (MSE) and Heteroskedasticity-adjusted Mean Square Error (HMSE)

MAE=$\frac{1}{m}\sum_{t=T-m}^{T}|RVt - FVt|$

MSE=$\frac{1}{m}\sum_{t=T-m}^{T}(RVt - FVt)^2$

HMSE=$\frac{1}{m}\sum_{t=T-m}^{T}(1 - \frac{FVt}{RVt})^2$

RV = realized volatility

= sum of 1-hour squared returns

FV = forecasted volatility (k days ahead)

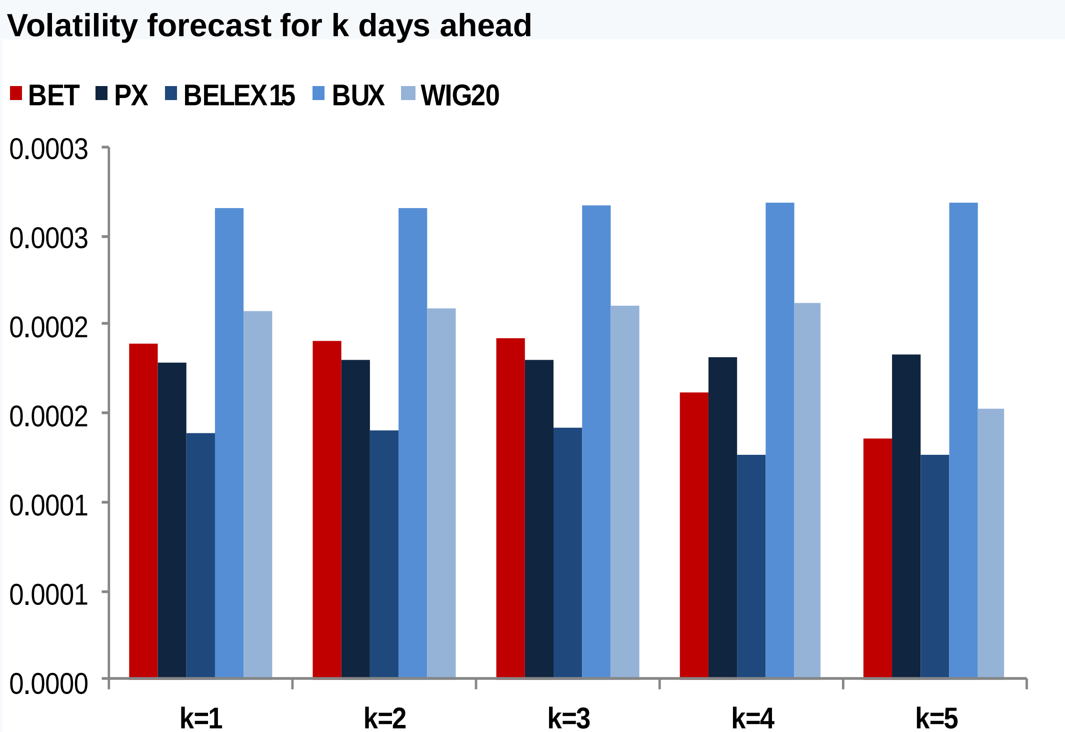
Index	IndicatorSV/ Indicator GARCH	k days ahead				
		k=1	k=2	k=3	k=4	k=5
BET	MAE	0.525	0.505	0.503	1.044	1.062
	MSE	0.624	0.640	0.645	1.061	1.070
	HMSE	0.160	0.131	0.121	1.072	1.142
PX	MAE	0.868	0.898	0.897	0.901	0.889
	MSE	0.831	0.855	0.860	0.882	0.834
	HMSE	0.826	0.988	0.875	0.953	0.950
BELEX15	MAE	0.784	0.817	0.791	1.049	1.068
	MSE	0.864	0.878	0.858	1.149	1.224
	HMSE	0.895	1.036	0.912	1.088	1.006
BUX	MAE	0.768	0.762	0.760	0.753	0.757
	MSE	0.608	0.599	0.601	0.601	0.598
	HMSE	0.887	0.856	0.901	0.887	0.822
WIG20	MAE	0.647	0.046	0.066	0.288	1.853
	MSE	0.900	0.001	0.002	0.004	1.139
	HMSE	0.920	0.000	0.000	0.003	1.106

For all the five indices, on the short run, Stochastic Volatility models outperform GARCH models in forecasting volatility

For BET, BELEX15 and WIG20, the GARCH seems to perform better in forecasting volatility for 4 to 5 days ahead

Volatility forecast results for CEE blue chips indices

Illustrative example: Using the parameters estimated for the state equation (in the case of stochastic volatility models) or for the volatility equation (for GARCH models), it can be ran volatility forecasts on a five-day horizon.



Source: Own computes

Giving the current context, the most volatile stock exchange in the coming five days seems to be Hungary's market. In the opposite corner are the Serbian market, the Czech one and the Romanian one.

The importance of such volatility forecasting results comes from their integration in larger volatility analysis.

Concluding remarks

- Stochastic Volatility models deliver better volatility forecasts than GARCH models on the short run (1-3 trading days ahead) for all the five indices considered
- GARCH outperforms the Stochastic Volatility model in the case of BET and BELEX15 for forecasting 4 and 5 days ahead volatility (the two indices are frontier market indices according to MSCI classification, which makes them more unpredictable on longer time frames), and in the case of WIG20 for forecasting 5 days ahead volatility (the mutation seen on the Polish capital market in September 2013 when the Polish state decided to partially nationalize pension funds, which ensured in 2013 almost 40% of the market liquidity, gave to the Polish capital market a higher-risk profile, translated into larger, unpredictable fluctuations).
- Further research: - Testing the models on larger out-of-sample horizons
  - Applying the model for a stock (instead of an index) for which there are options in order to compute the implied volatility
  - Testing SV models for which the errors are not assumed to be Gaussian
  - Testing SV models with conditional mean of returns

Selected references

[1] Berger, J., Pericchi, L., "The Intrinsic Bayes Factor for Model Selection and Prediction", Journal of the American Statistical Association, Vol. 91, p 109-122, 1996
[2] Bollerslev T., "Generalized Autoregressive Conditional Heteroskedasticity", "Journal of Econometrics", Vol. 31, p 307-328, 1986
[3] Chan C., Jeliazkov I., "Efficient Simulation and Integrated Likelihood Estimation in State Space Models", International Journal of Mathematical Modeling and Numerical Optimization, Vol. 1, p 101-120
[4] Chan C., "Moving Average Stochastic Volatility Models with Application to Inflation Forecast", Journal of Econometrics, Vol. 176, p 162-172, 2013
[5] Fleming J., Kirby C., "A Closer Look at the Relation between GARCH and Stochastic Autoregressive Volatility", Journal of Financial Econometrics, Vol. 1, p 365-419, 2003
[6] Galdi F.C. and Pereira L.M., "Value at Risk (VaR) Using Volatility Forecasting Models: EWMA, GARCH and Stochastic Volatility", Brazilian Business Review, Vol. 4, p 74-94
[7] Hansen p., Lunde A., "A forecast comparison of volatility models: does anything beat a GARCH(1,1)?", Journal of Applied Econometrics, Vol. 20, p. 873-889, 2001
[8] Kim S., Shepherd N., Chib S., "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models", review of Economic Studies, Vol. 65, p 361-393, 1998
[9] Verdinelli I., Wasserman L., "Computing Bayes Factors Using a Generalization of the Savage-Dickey Density Ratio", Journal of the American Statistical Association, Vol. 90, p 514-618, 1995