Estimation of Value at Risk for the Indian capital market:

Filtered Historical Simulation approach using GARCH model with suitable mean specification

Indrajit Roy
Reserve Bank of India

NIPFP-DEA: 6th Research Meeting, 9-10 March, 2010

Background

Efficient Market Hypothesis (EMH) : Financial markets are informational efficient
In an efficient market, actual price of a security will be a good estimate of its intrinsic value

The weak form of EMH : All past market prices and data are fully reflected in securities prices.
In other words, technical analysis cannot be used to predict and beat a market

The semi strong form of EMH: All publicly available information is fully reflected in securities prices. Therefore, fundamental analysis is of no use. .

Strong form of EMH: Assumes that market reflects even hidden/inside information.
In other words, even insider/hidden information is of no use.

There were numerous study to test the weak form of market efficiency hypothesis.
These studies provided indecisive results.

Background (contd…)

The developed markets were found to be weak form efficient.

However, evidence from emerging markets indicated rejection of the weak form market efficiency hypothesis.

Theoretically, once an anomaly is discovered, investors attempting to profit by exploiting the inefficiency should result its disappearance

The more participants and faster the dissemination of information, the more efficient a market

The contradiction of efficient markets is that if every investor believed a market was efficient, then the market would not be efficient because no one would analyze securities. So efficient markets depend on market participants who believe the market is inefficient and trade securities in an attempt to outperform the market. Therefore, markets are neither perfectly efficient nor completely inefficient. Perhaps, they are efficient to a certain extent.

So the question arises whether the equity returns in Indian markets are predictable.

Background (contd…)

Apart from the form of efficiency, it is the volatility prevailing in the market which influences the return to a large extent

Volatility, which refers to the degree of unpredictable change over time and might be measured by the standard deviation of a sample, often used to quantify the risk of the instrument of portfolio over that time period.

Engle (1982) introduced the concept of Autoregressive Conditional Heteroscedasticity (ARCH) which became a very powerful tool in the modelling of high frequency financial data in general and stock returns in particular. ARCH models allow the conditional variances to change through time as functions of past errors. One significant improvement introduced by Bollerslev (1986) was the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) process.

Background (contd…)

More recently, Janak Raj and Sarat Dhal (2008) investigated the financial integration of India's stock market with that of global and major regional markets. They have used six stock price indices i.e. the 200-scrip index of BSE of India to represent domestic market, stock price indices of Singapore and Hong Kong to represent the regional markets and three stock price indices of U.S., U.K. and Japan to represent the global markets. Based on daily as well as weekly data covering end-March 2003 to end-January 2008 they found that Indian market's dependence on global markets, such as U.S. and U.K., was substantially higher than on regional markets such as Singapore and Hong Kong, while Japanese market had weak influence on Indian market.

Objectives

The paper examines the financial integration of Indian capital market (BSE-SENSEX and NSE-NIFTY) with other global indicators and its own volatility using daily returns covering the period January 2003 to December 2009.

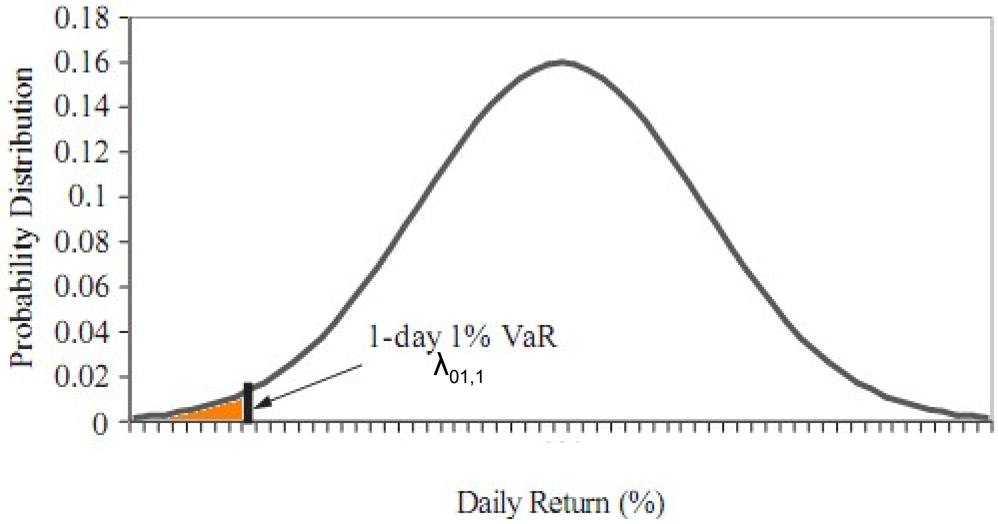
The paper specifies a GARCH framework to model the phenomena of volatility clustering on returns and examines the usefulness of considering lag values of return on (S&P 500, INR-EURO & INR-USD exchange rate, Gold price) as proxies to global financial condition in the specification of the mean equation.

The paper also estimate VaR of return in the Indian capital market based on two composite methods i.e. (a) using univariate GARCH model where in the mean equation lag values of return on (S&P 500, INR-EURO & INR-USD exchange rate, Gold price) have been used and following the filtered historical simulation (FHS) approach to estimate VaR; (b) using ARMA for mean equation, GARCH for volatility and FHS for VaR estimation i.e. ARMA-GARCH-FHS methods; and compare the performance of both the VaR estimates.

Value at Risk

Largest loss expected to be suffered on a portfolio position over a holding period with a given probability

The 100α% one day ahead VaR ($\lambda_{\alpha,t}$) for daily return r_t is defined as $P[r_t <= \lambda_{\alpha,t} \mid r_{t-1}] = \alpha$.



Value at Risk

The Parametric VaR model imposes a strong theoretical assumption on the underlying properties of data; frequently Normal Distribution is assumed because it is well described, can be defined using only the first two moments and it can be understood easily. Other probability distributions may be used, but at a higher computational cost.

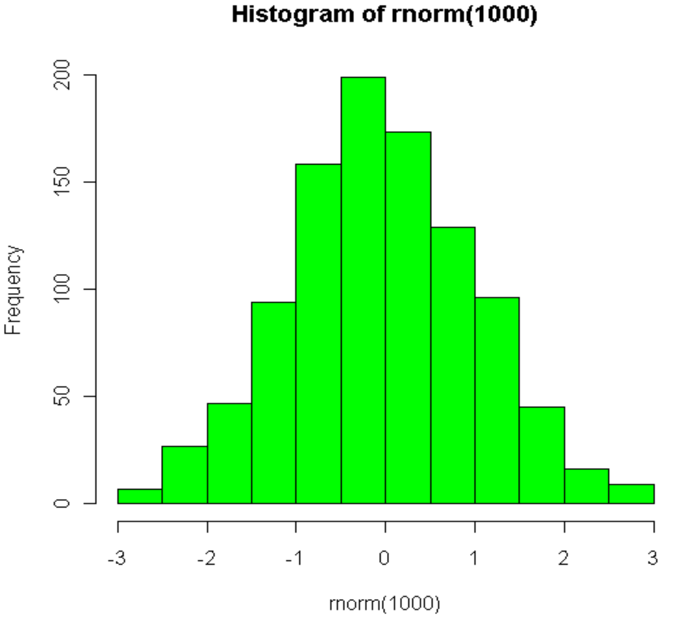
However, empirical evidence indicates that asset price changes, in particular the daily price changes, most of the time does not follow Normal Distribution. In the presence of excess kurtosis, failure rate increases when the VaR is estimated by the Gaussian distribution.

Value at Risk (contd…)

- Historical Simulation (HS)
- Monte Carlo Simulation (MCS)
- Filtered Historical Simulation (FHS)

VaR: Historical Simulation (HS)

Apart from stationarity of the returns, Historical Simulation (HS) does not require any statistical assumption in particular to the volatility. In Historical Simulation method we consider the availability of a past sequence of daily portfolio returns for m days; r_t t=1,2…m. The HS technique simply assumes that the distribution of tomorrow's portfolio returns, r_{t+1}, is well approximated by the empirical distribution of the past *m* observations—that is, $\{r_{t+1-\tau}\}_{\tau=1..m}$. In other words, the distribution of r_{t+1} is captured by the histogram of $\{r_{t+1-\tau}\}_{\tau=1..m}$. Thus, we simply sort the returns in $\{r_{t+1-\tau}\}_{\tau=1..m}$ in ascending order and choose the VaR^p_{t+1} to be the number such that only 100*p*% of the observations are smaller than the VaR^p_{t+1}.



VaR: Monte Carlo Simulation (MCS)

1-day VaR=> Given daily returns $(r_1, r_2 \ldots r_t)$

we need to forecast return for 't+1' day i.e. r_{t+1} and its probability distribution

Step I

Let us consider GARCH(1,1) model

$$r_{t+1}=c+\phi_1 r_t+\phi_2 r_{t-1}+..+\phi_k r_{t+1-k}+\psi_1 x_{1,t}+\psi_2 x_{2,t}+..+\psi_s x_{s,t}+\sigma_{t+1}\eta_{t+1}\ ; \qquad t=1,2\ldots T$$

$$\sigma^2_{t+1}=\omega+\alpha Resid^2_t+\beta\sigma^2_t;$$

where $Resid_t = (r_t - c - \Sigma\phi_i r_{t-l} - \Sigma\psi_j x_{j,t})$; innovation $\{\eta_t\}$ is white noise process, with zero mean and unit variance and $\alpha+\beta<1$. for the sake of simplicity, let us assume η_t follows Normal Distribution N(0,1).

Step II

Based on the above specified GARCH model, at the end of day 't' we can project the variance of day 't+1' i.e. $\sigma^{\wedge 2}_{t+1}$.

VaR: Monte Carlo Simulation (MCS)

Step III

Let $\{\eta_{i,1}^{\wedge};\ i=1,2\ldots L)$ be a set of large number of random numbers drawn from the standard Normal Distribution N(0,1). From these random numbers $\{\eta_{i,1};\ i=1,2\ldots L\}$ we can calculate a set of hypothetical returns for day 't+1' as

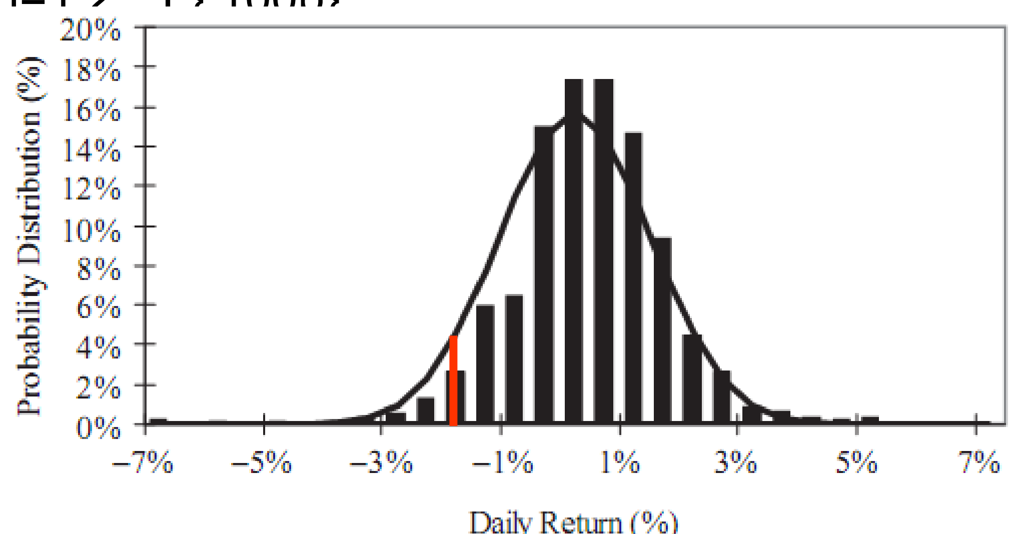
$$r^{\wedge}_{i,t+1} = c + \Sigma\phi_i r_{t+1-i} + \Sigma\psi_j x_{t+1-j} + \sigma^{\wedge}_{t+1}\ \eta_{i,1}^{\wedge}\ ;\ i=1,2\ldots L$$

$$Resid_{i,t+1} = (r^{\wedge}_{i,t+1} - c - \Sigma\phi_i r_{t+1-i} - \Sigma\psi_j x_{t+1-j})$$

Step IV

If we draw a histogram based on these L hypothetical *1*-day returns i.e. $\{r^{\wedge}_{i,t+1:t+k};\ i=1,2\ldots L\}$, then the *1*-day VaR can be calculated as the 100p percentile

$$VaR^{p}_{t+1:t+k} = -\ Percentile\{\{r^{\wedge}_{i,t+1:t+k};\ i=1,2\ldots L\}, 100p\}$$



VaR: Filtered Historical Simulation (FHS)

Historical Simulation (HS) is a non-parametric approach and does not assume any statistical distribution of returns, whereas parametric approach such as the Monte Carlo simulation (MCS) takes the opposite view and assumes parametric models for variance, correlation (if a disaggregate model is estimated), and the distribution of standardized returns.

MCS is good if the assumed distribution is fairly accurate in description of reality. Whereas, HS is sensible as the observed data may capture features of the returns distribution that are not captured by any standard parametric model. HS assumes weak form of market efficiency and does not take into consideration volatility clustering.

The FHS approach attempts to combine the best of the MCS with the best of the HS

VaR: Filtered Historical Simulation (FHS) ………contd…

Let's assume we have estimated a GARCH-type model of our portfolio variance. Although we are comfortable with our variance model (σ), we may not be comfortable making a specific distributional assumption about the (η), such as a Normal or a *t* distribution. Instead of that, we might like the past returns data (r_t) to determine the distribution directly without making further assumptions

Given a sequence of past returns and estimated GARCH volatility $\{r_{t+1-\tau}, \hat{\sigma}_{t+1-\tau}; \tau = 1,2\ldots m\}$

Estimated past standardized returns are given by

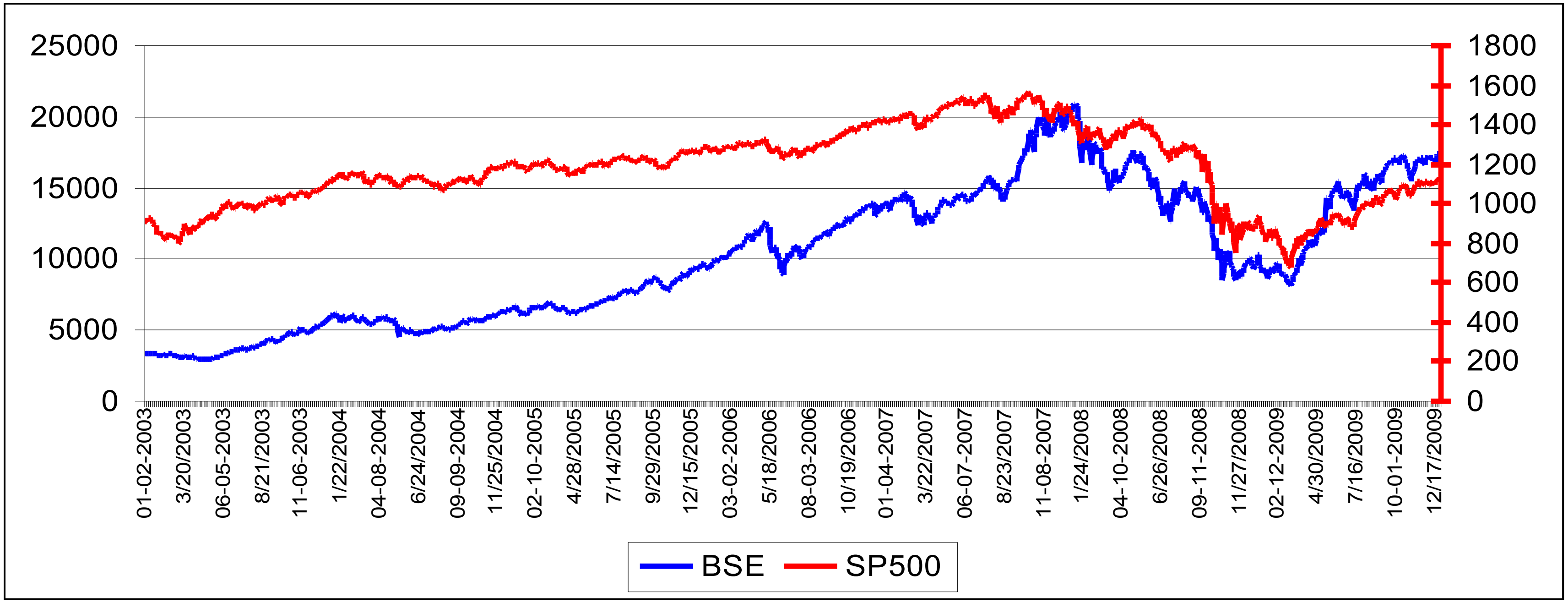
$\hat{\eta}_{t+1-\tau} = (r_{t+1-\tau} - E(r_{t+1-\tau}))/\hat{\sigma}_{t+1-\tau}; \tau = 1,2\ldots m$

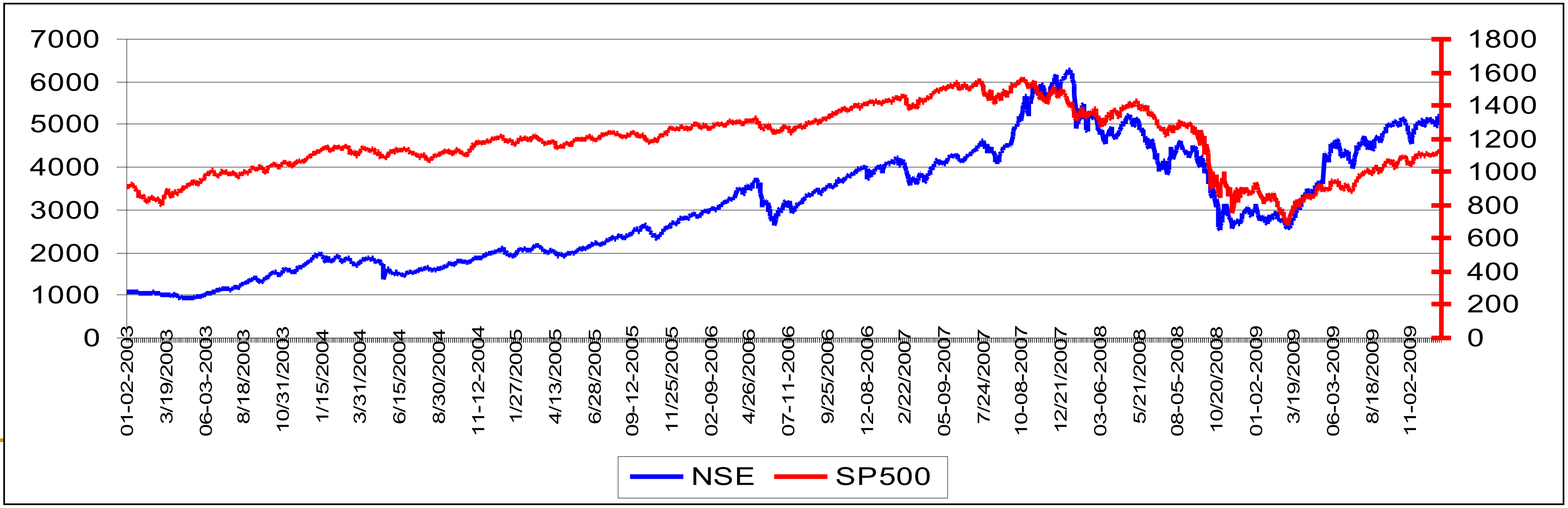
Instead of drawing random $\hat{\eta}$s from a specific probability distribution as it is done in MCS, in FHS method samples are drawn with replacement from $\{\hat{\eta}_{t+1-\tau}; \tau = 1,2\ldots m\}$.

Thereafter, similarly as in the case of MCS, we can get the hypothetical return of 't+k' day

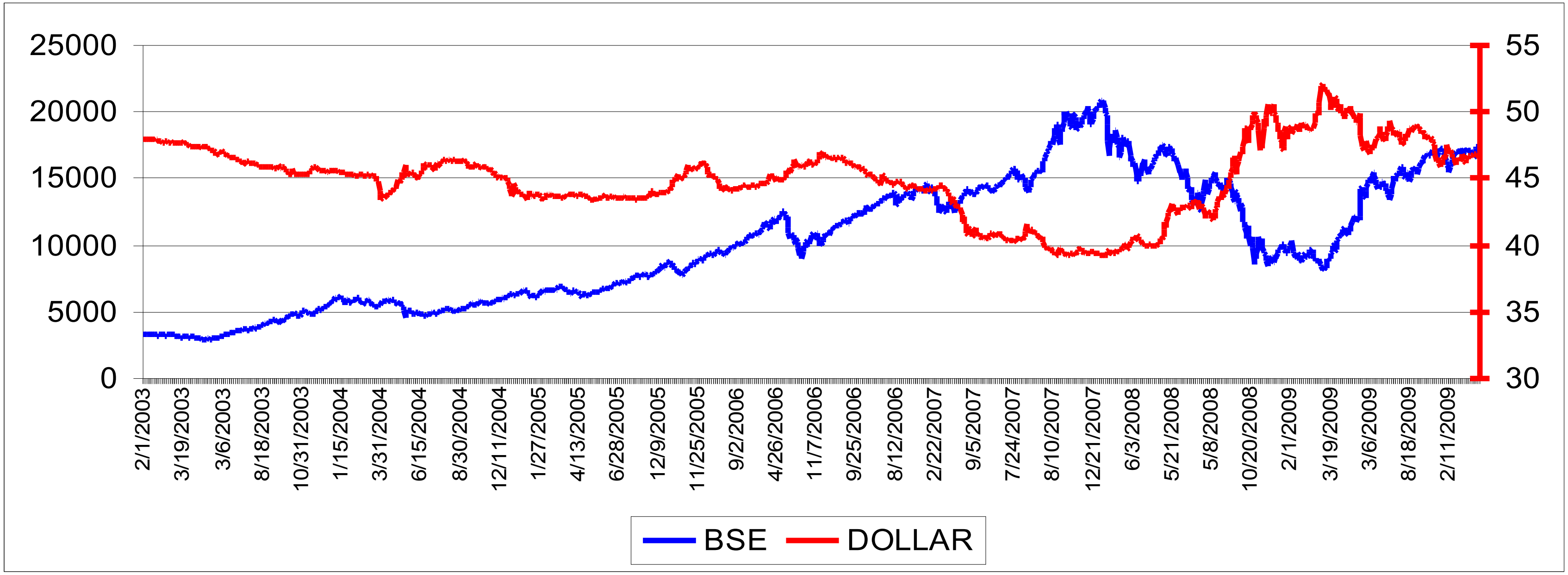
$\hat{r}_{i,t+k} = c + \Sigma\phi_i * r_{i,t+k-1} + \Sigma\psi_j * x_{j,t+k-1} + \hat{\sigma}_{t+k-1} * \hat{\eta}_{i,k}; i = 1,2\ldots L.$

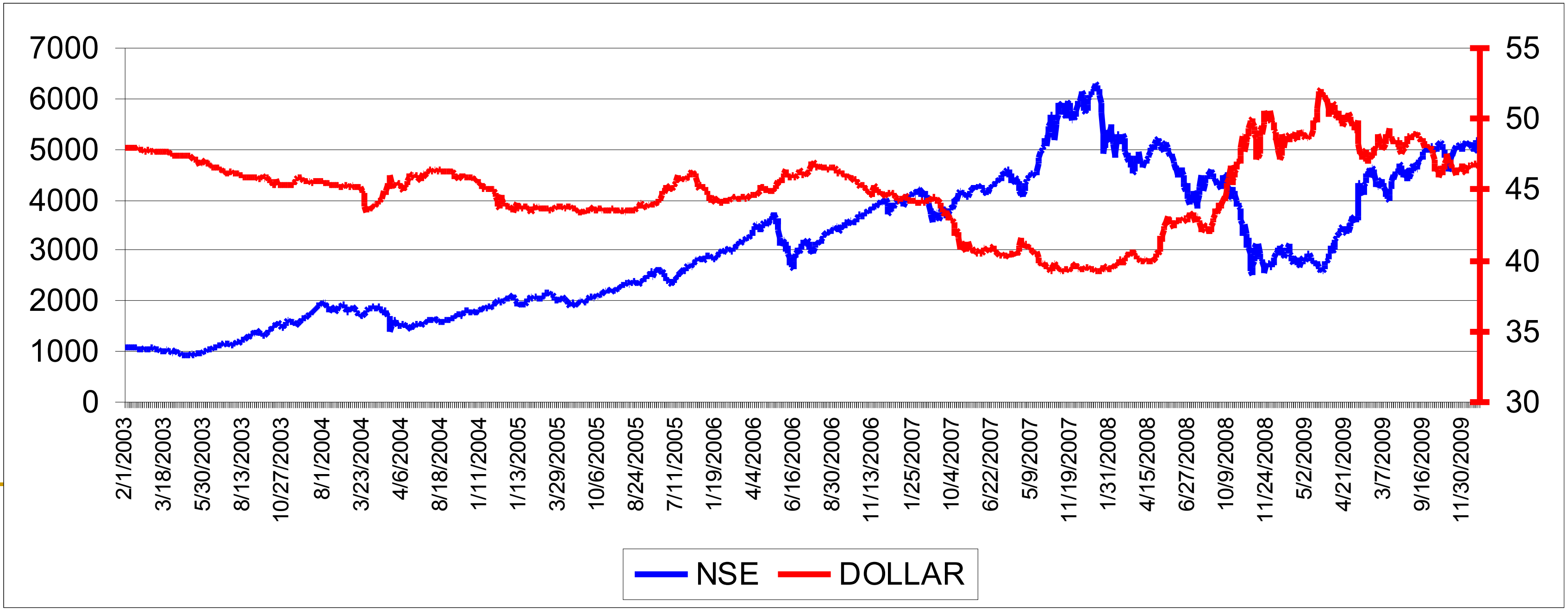
Movement of Indian capital market and S&P 500 (level)



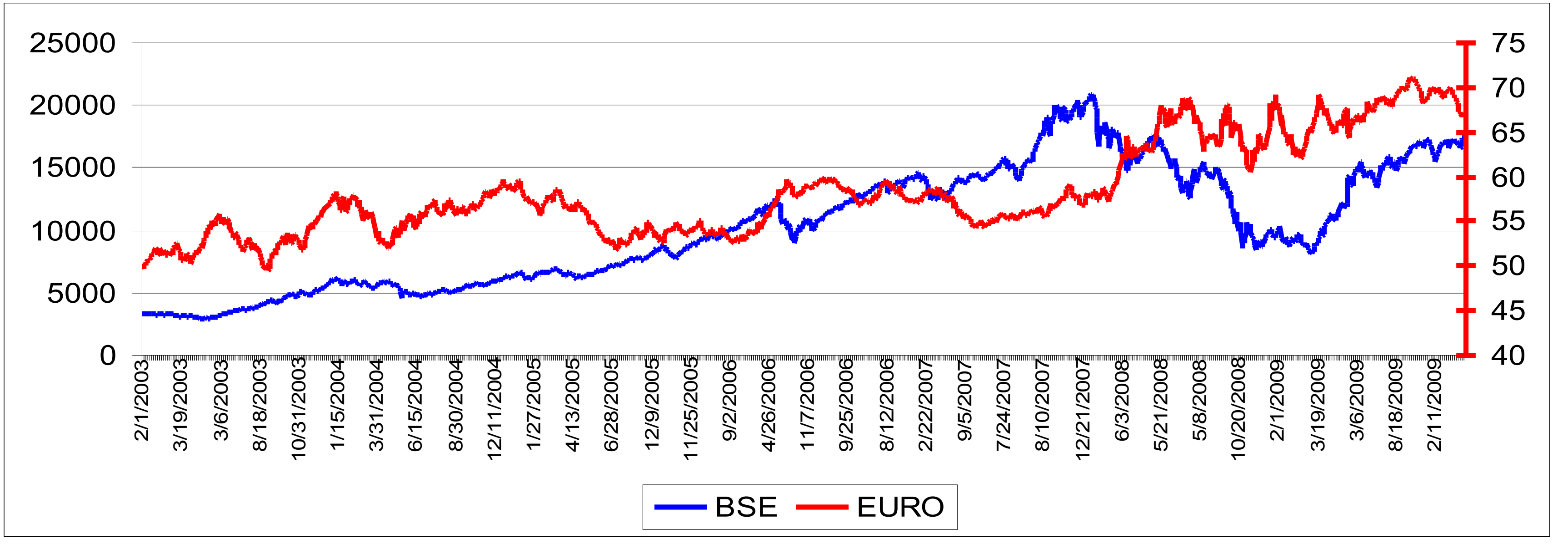


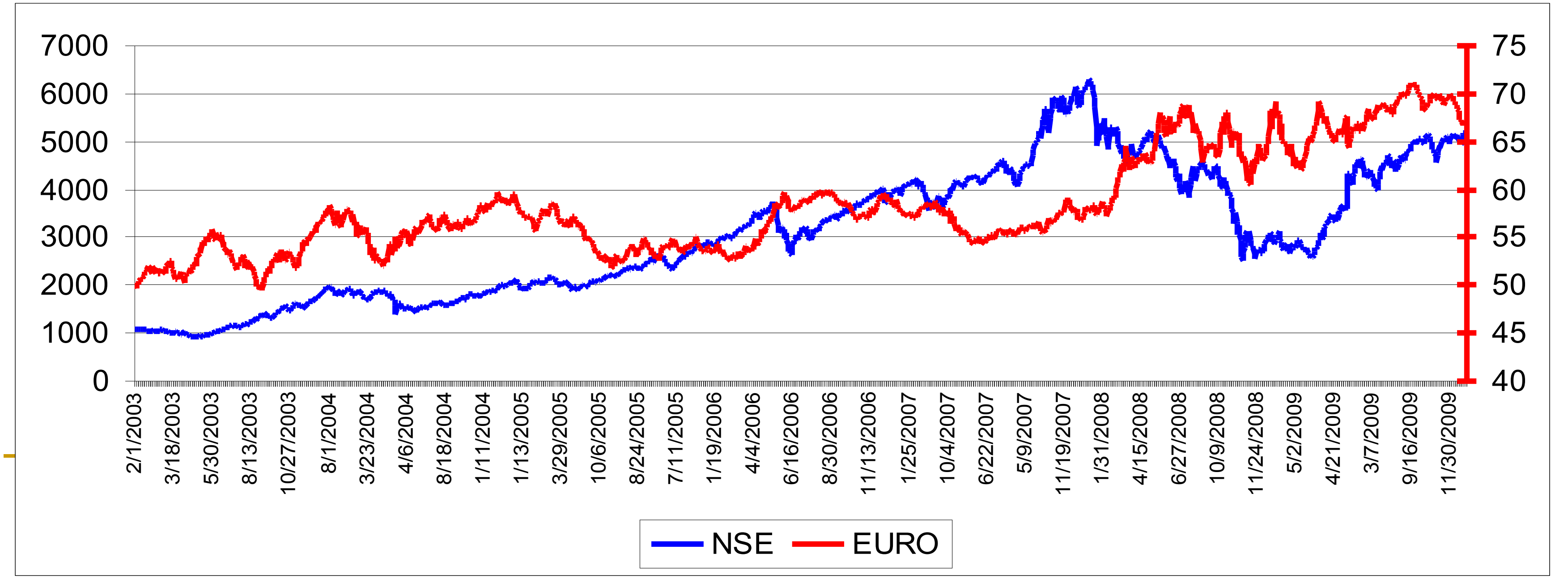
Movement of Indian capital market and INR-USD exchange rate (level)



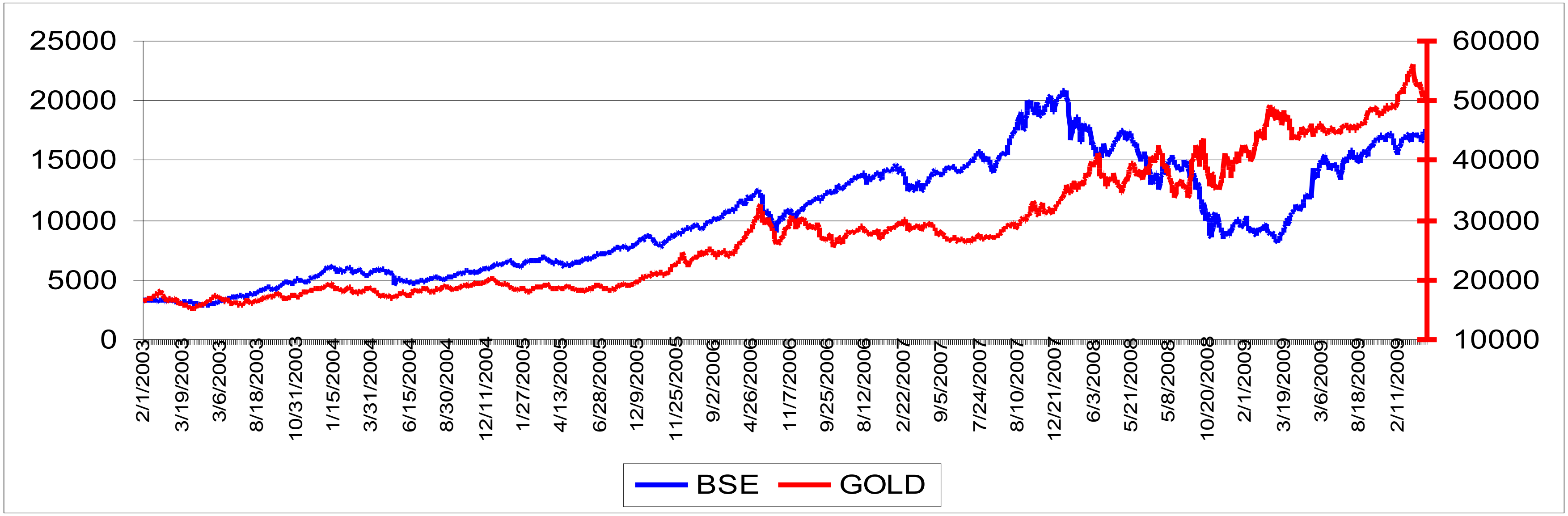


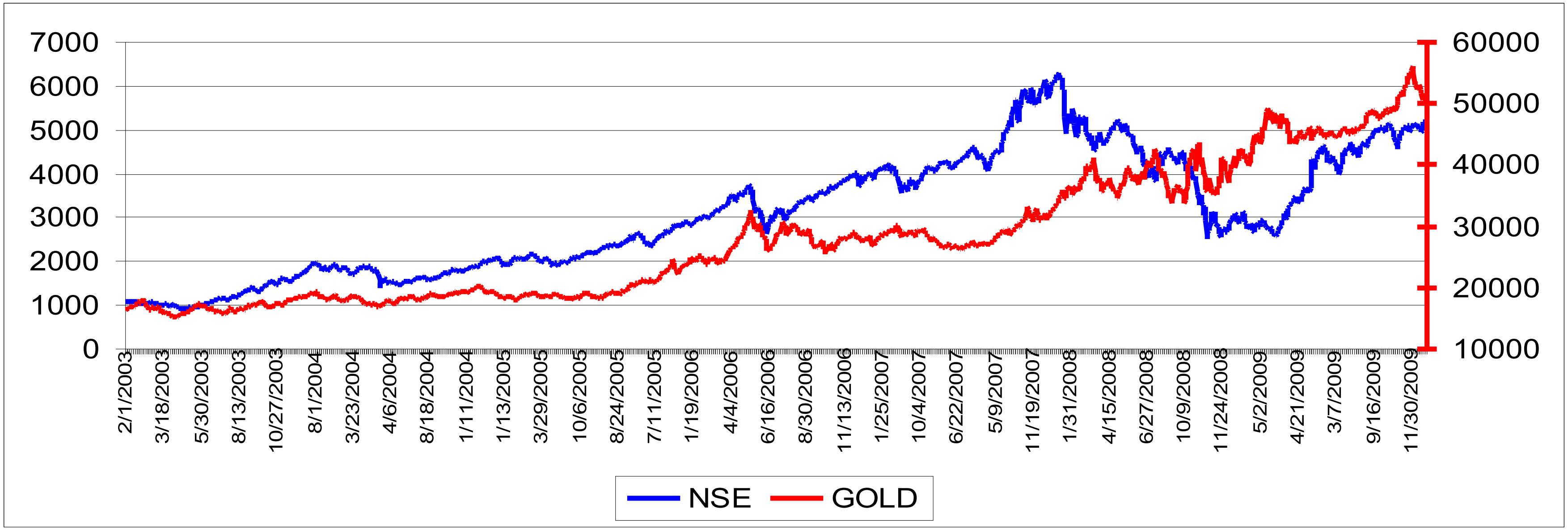
Movement of Indian capital market and INR-EURO exchange rate (level)



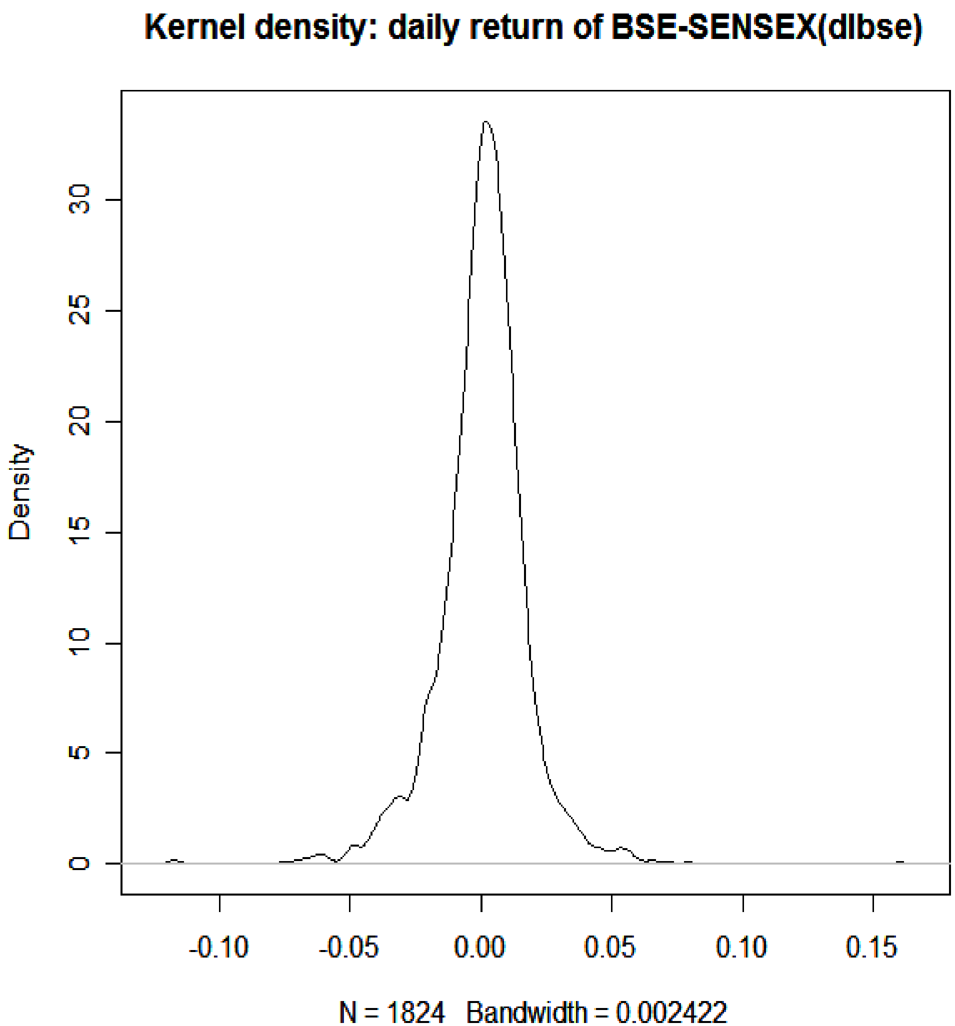


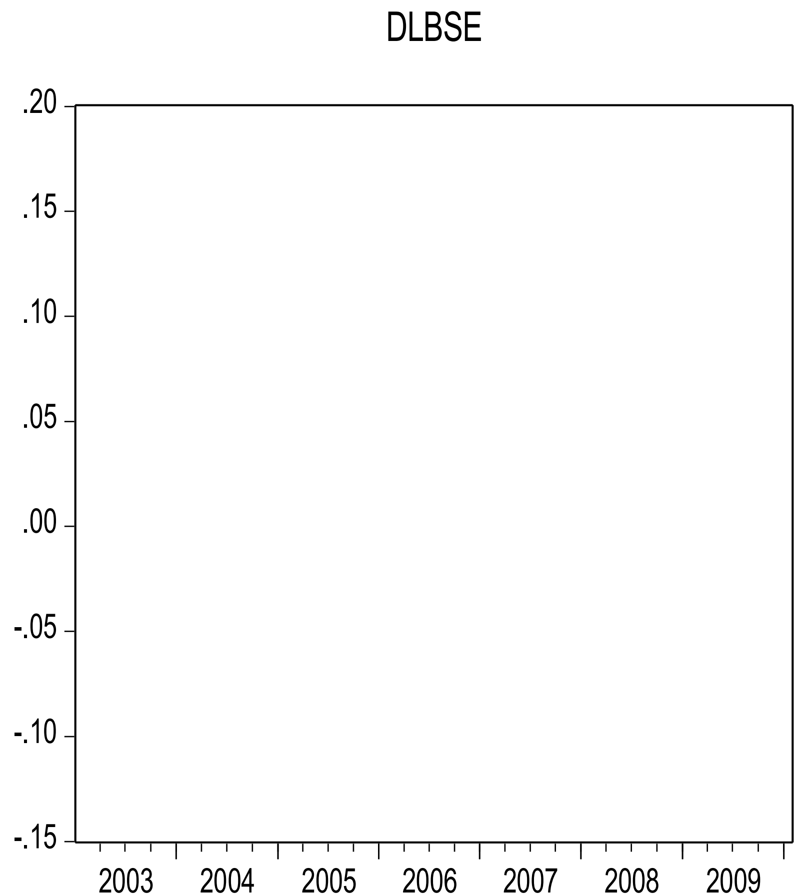
Movement of Indian capital market and Gold price (level)



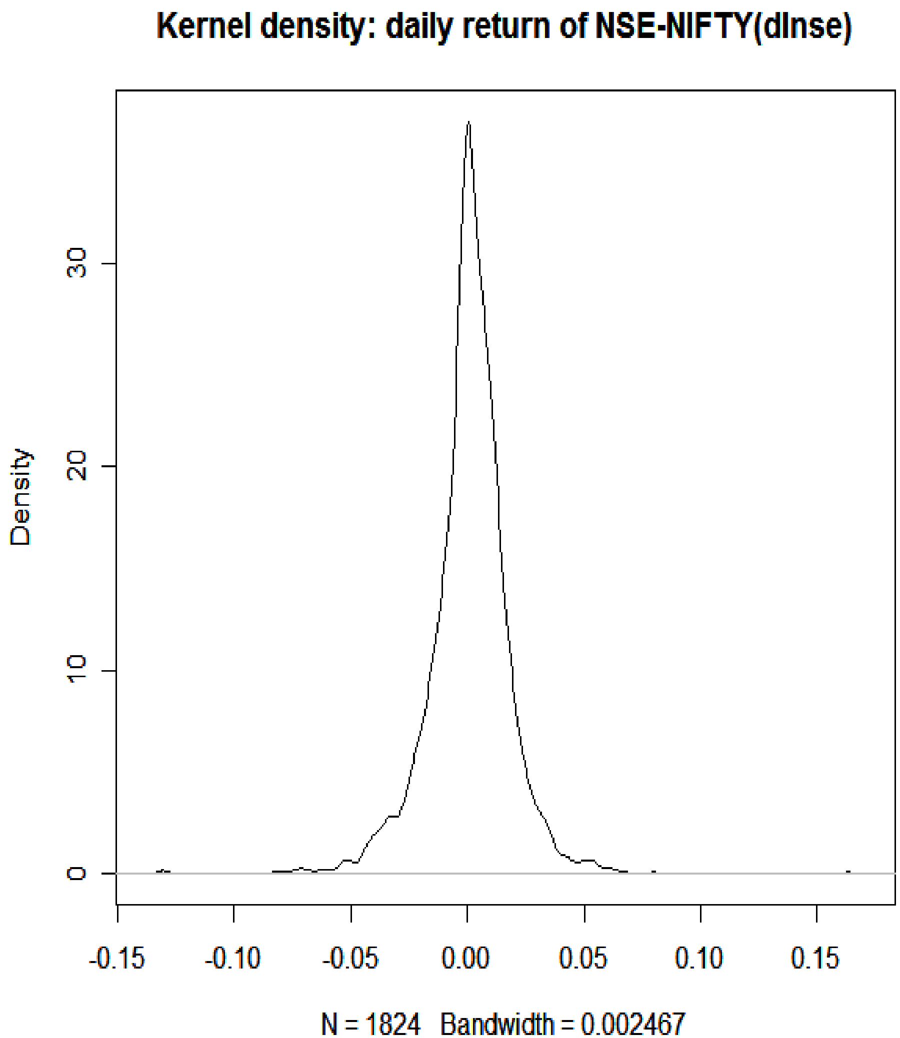


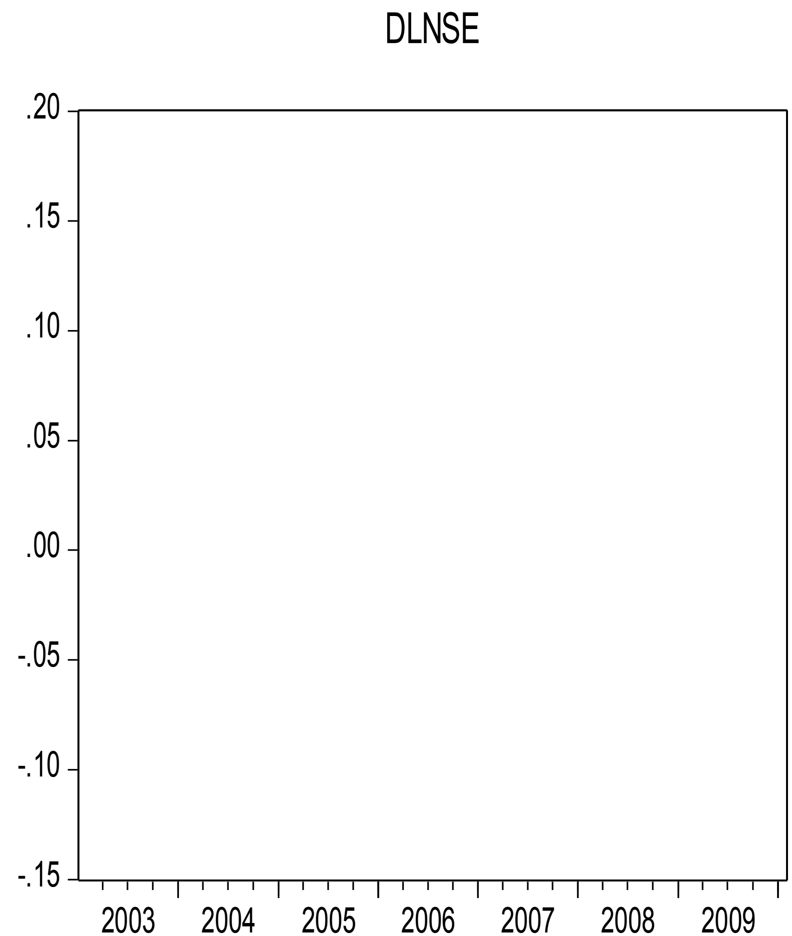
Daily return of BSE SENSEX: does not follow normal distribution





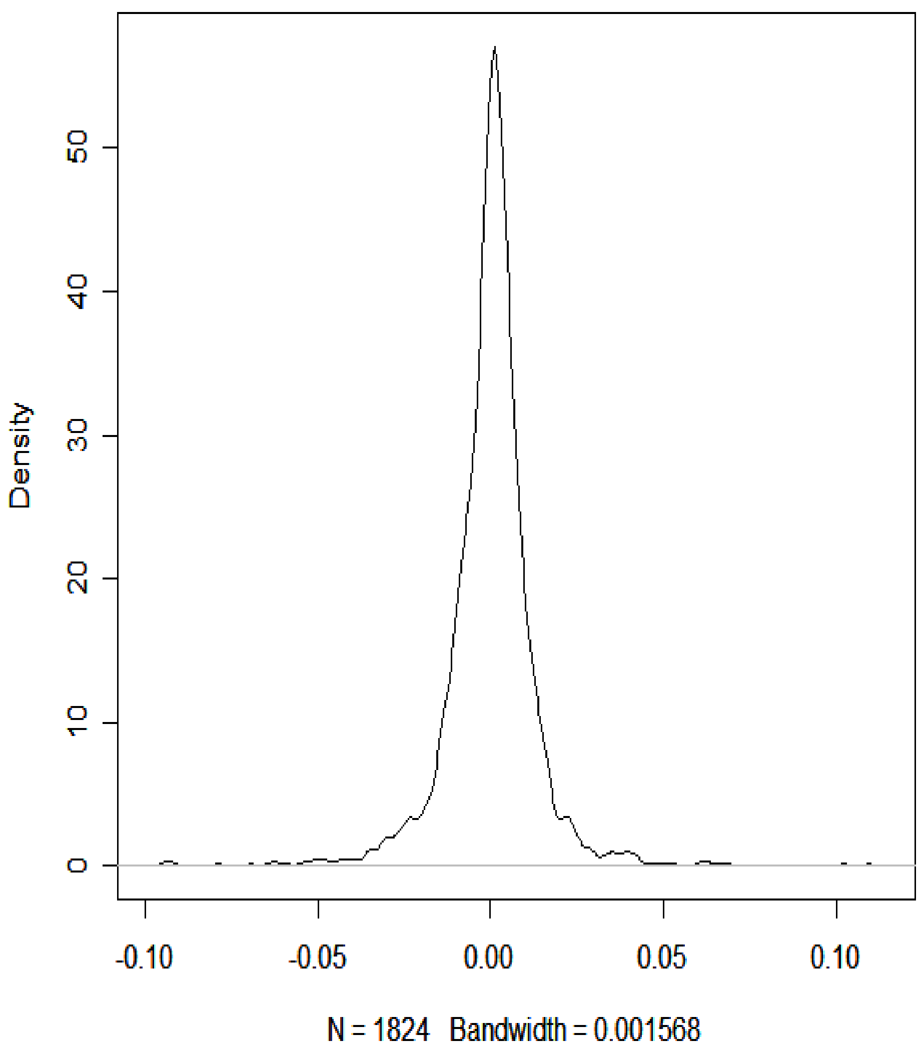
Daily return of NSE NIFTY: does not follow normal distribution

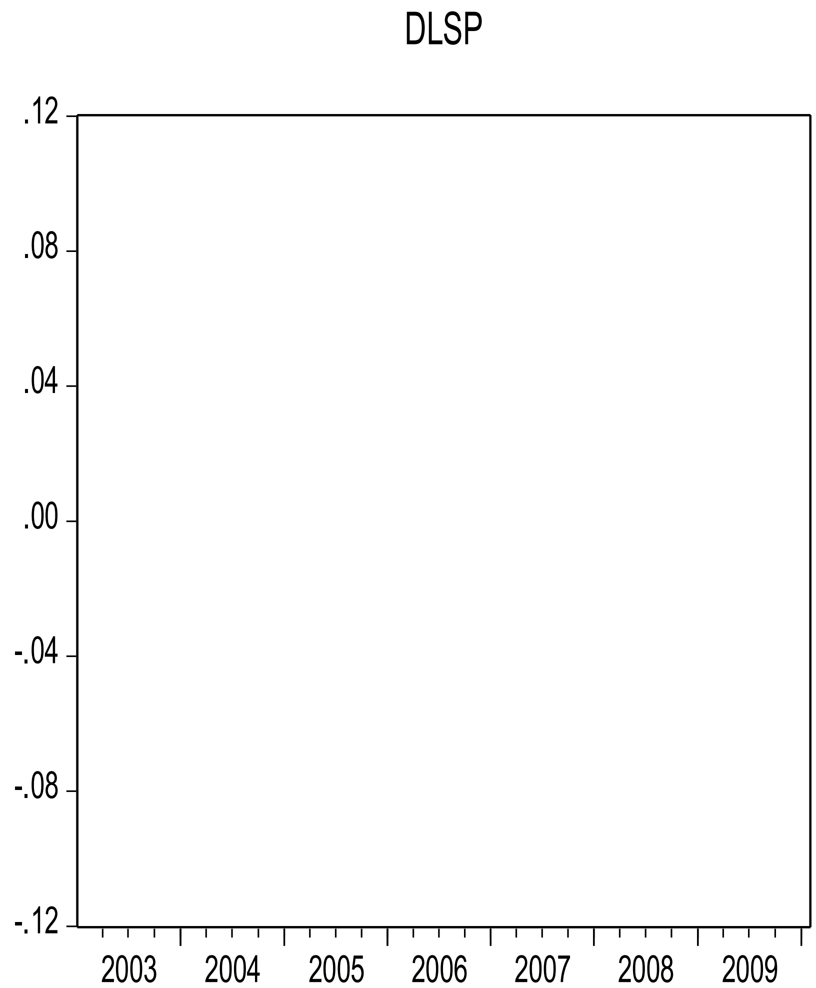




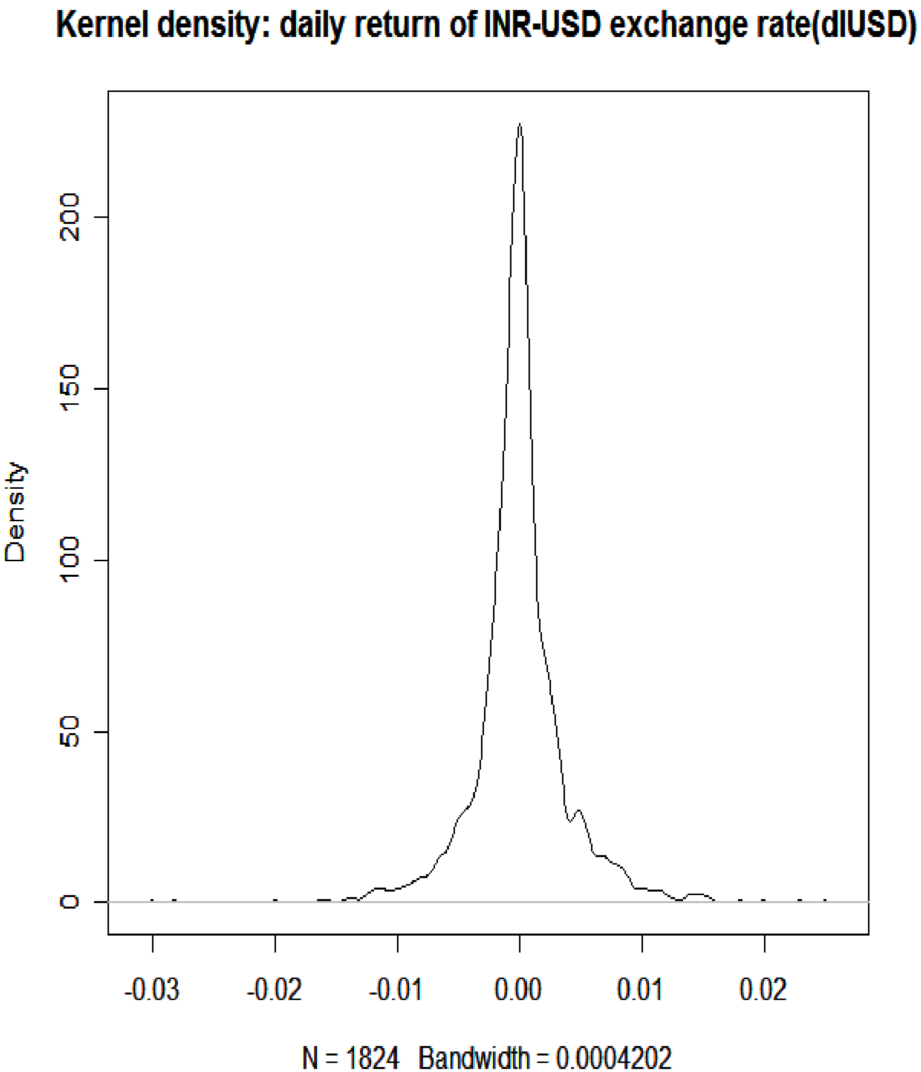
Daily return of S&P 500: does not follow normal distribution

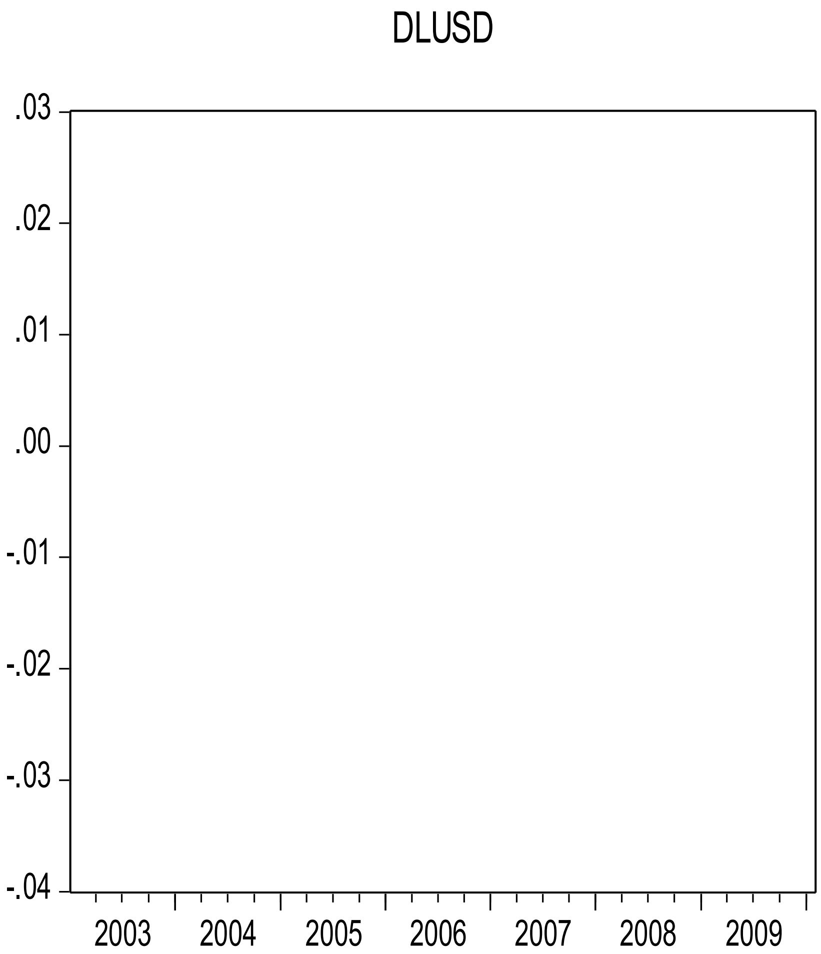
Kernel density: daily return of S&P 500 stockprices(dlsp)



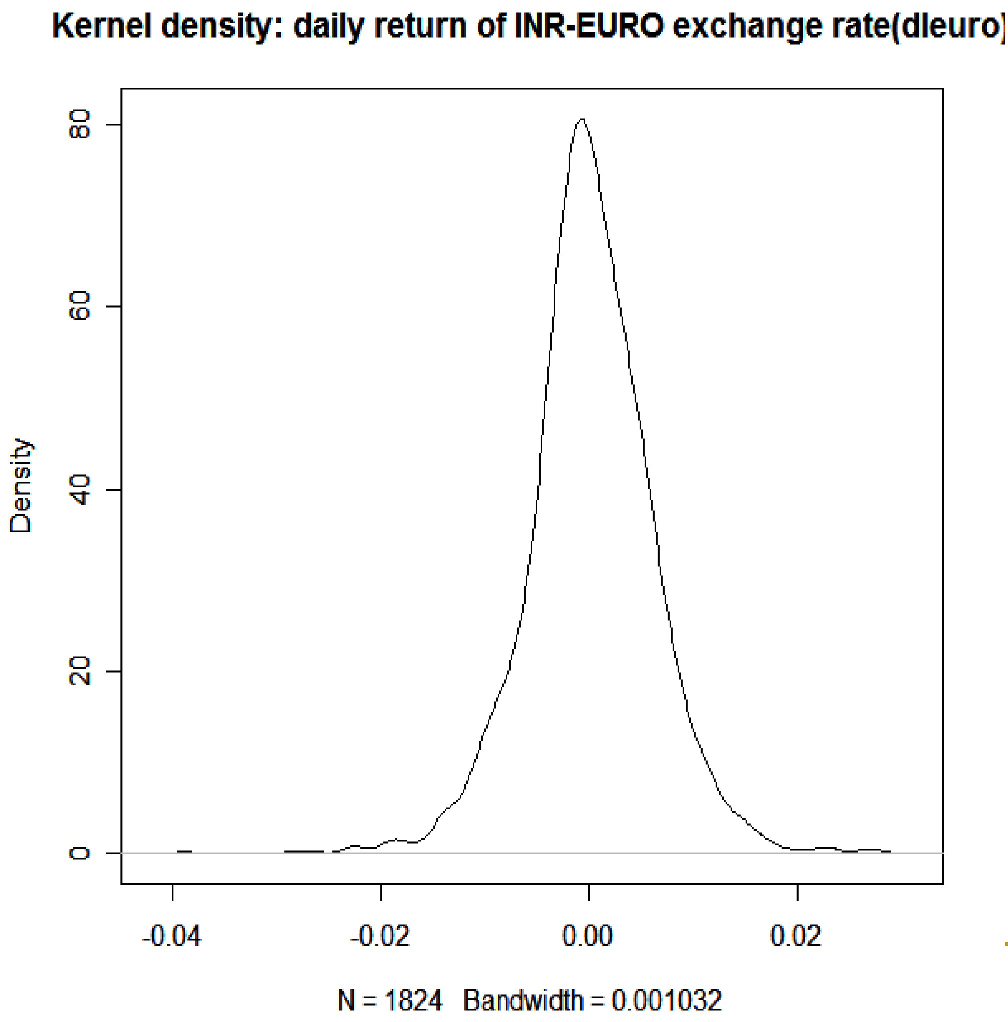


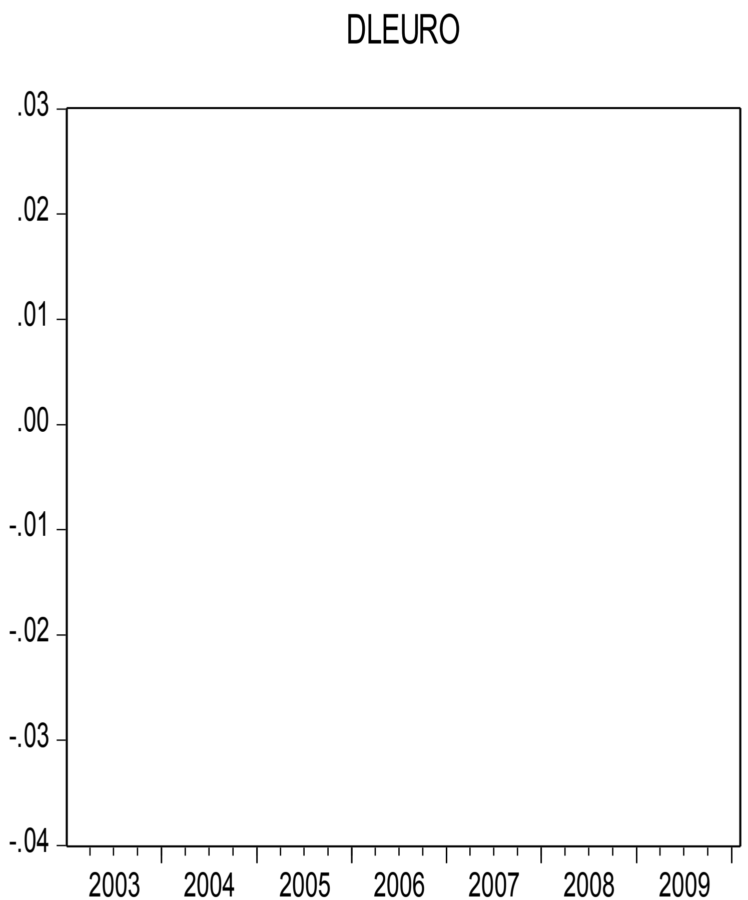
Daily return of INR-USD exchange rate: does not follow normal distribution



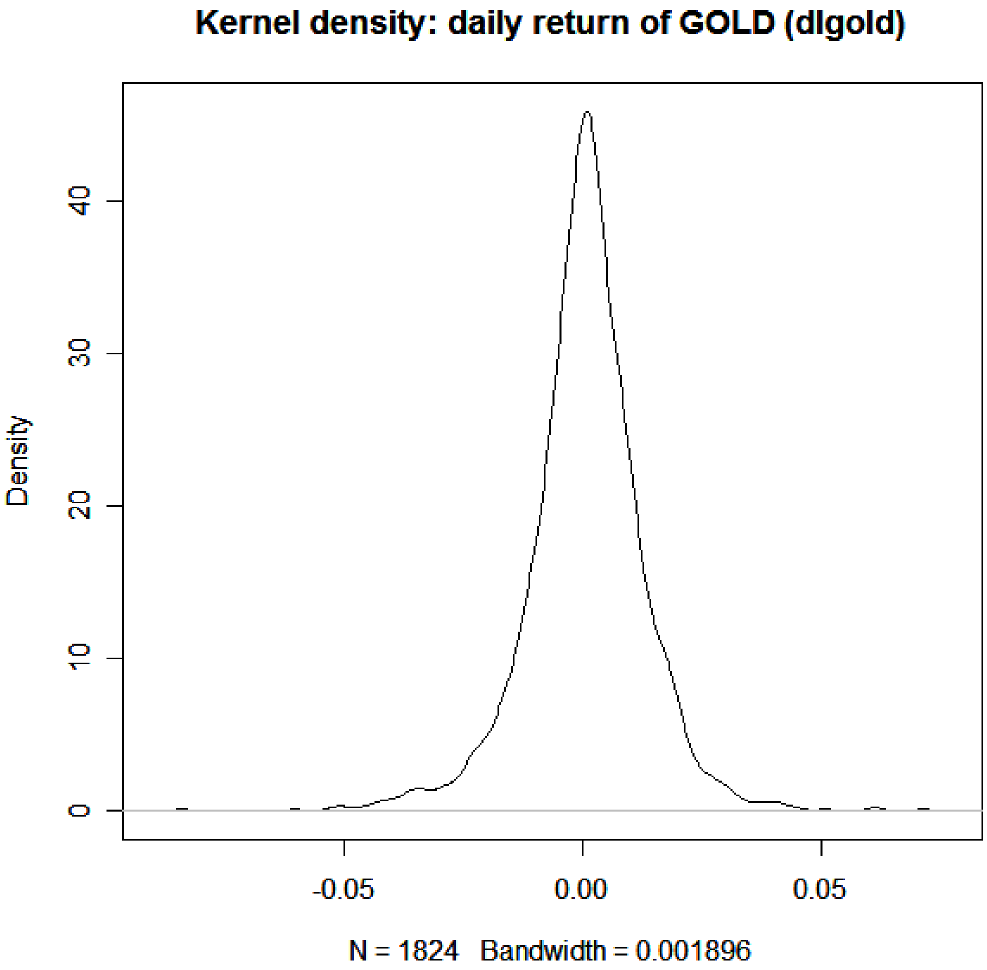


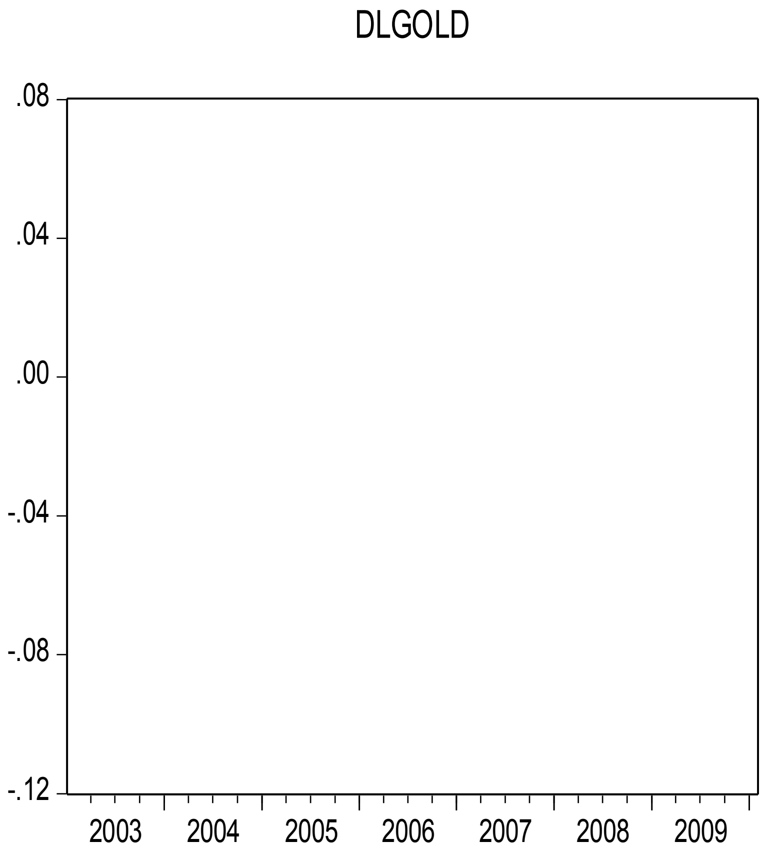
Daily return of INR-EURO exchange rate: does not follow normal distribution





Daily return of GOLD price: does not follow normal distribution





Descriptive statistics

	DLNSE	DLBSE	DLSP	DLUSD	DLEURO	DLGOLD
Mean	0.0008	0.0008	0.0001	0.0000	0.0001	0.0006
Median	0.0011	0.0016	0.0008	0.0000	0.0000	0.0009
Maximum	0.1633	0.1599	0.1096	0.0249	0.0279	0.0713
Minimum	-0.1305	-0.1181	-0.0947	-0.0301	-0.0389	-0.0840
Std. Dev.	0.0175	0.0172	0.0133	0.0039	0.0061	0.0125
Skewness	-0.3193	-0.1124	-0.2320	-0.0225	-0.1407	-0.3072
Kurtosis	12.0731	11.0435	15.1397	10.8258	5.5835	6.9446
Jarque-Bera	6370.1470	4985.6290	11364.1900	4715.8520	520.0318	1211.2270
Sum	1.4859	1.5693	0.1937	-0.0374	0.2476	1.1261
Sum Sq. Dev.	0.5655	0.5443	0.3263	0.0275	0.0684	0.2842
Observations	1848	1848	1848	1848	1848	1824

Results….

Dependent Variable: D(LOG(BSE))
Method: ML - ARCH
Date: 02/11/10 Time: 02:47
Sample (adjusted): 1/09/2003 12/23/2009
Included observations: 1815 after adjustments
Convergence achieved after 21 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(12) + C(13)*RESID(-1)^2 + C(14)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001516	0.000270	5.605564	0.0000
D(LOG(SP500(-1)))	0.325582	0.026625	12.22838	0.0000
D(LOG(SP500(-2)))	0.167160	0.026299	6.356212	0.0000
D(LOG(SP500(-3)))	0.133932	0.029617	4.522167	0.0000
D(LOG(SP500(-4)))	0.100051	0.027628	3.621330	0.0003
D(LOG(BSE(-2)))	-0.060436	0.025798	-2.342648	0.0191
D(LOG(BSE(-3)))	-0.048912	0.020935	-2.336401	0.0195
D(LOG(GOLD(-2)))	0.057216	0.023705	2.413654	0.0158
D(LOG(EURO(-3)))	0.158242	0.053619	2.951257	0.0032
D(LOG(DOLLAR(-3)))	-0.246202	0.088501	-2.781915	0.0054
D(LOG(DOLLAR(-4)))	0.157666	0.084305	1.870196	0.0615
Variance Equation				
C	5.05E-06	1.00E-06	5.032612	0.0000
RESID(-1)^2	0.152794	0.012646	12.08255	0.0000
GARCH(-1)	0.838035	0.012368	67.75576	0.0000

R-squared	0.079684	Mean dependent var	0.000901
Adjusted R-squared	0.073041	S.D. dependent var	0.017291
S.E. of regression	0.016648	Akaike info criterion	-5.720290
Sum squared resid	0.499150	Schwarz criterion	-5.677836
Log likelihood	5205.163	Hannan-Quinn criter.	-5.704625
F-statistic	11.99514	Durbin-Watson stat	2.045551
Prob(F-statistic)	0.000000		

Results….

Dependent Variable: D(LOG(NSE))
Method: ML - ARCH
Date: 02/11/10 Time: 05:10
Sample (adjusted): 1/09/2003 12/23/2009
Included observations: 1815 after adjustments
Convergence achieved after 22 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(10) + C(11)*RESID(-1)^2 + C(12)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001341	0.000292	4.592404	0.0000
D(LOG(SP500(-1)))	0.324685	0.027540	11.78939	0.0000
D(LOG(SP500(-2)))	0.168209	0.027508	6.114933	0.0000
D(LOG(SP500(-3)))	0.102682	0.030620	3.353393	0.0008
D(LOG(NSE(-2)))	-0.038768	0.026039	-1.488849	0.1365
D(LOG(GOLD(-2)))	0.061592	0.025277	2.436714	0.0148
D(LOG(EURO(-3)))	0.177160	0.051925	3.411806	0.0006
D(LOG(DOLLAR(-3)))	-0.349376	0.088119	-3.964831	0.0001
D(LOG(DOLLAR(-4)))	0.190597	0.087568	2.176547	0.0295
Variance Equation				
C	5.53E-06	9.61E-07	5.750372	0.0000
RESID(-1)^2	0.130725	0.010611	12.31977	0.0000
GARCH(-1)	0.857389	0.011140	76.96483	0.0000

R-squared	0.068311	Mean dependent var	0.000855
Adjusted R-squared	0.062627	S.D. dependent var	0.017624
S.E. of regression	0.017064	Akaike info criterion	-5.633237
Sum squared resid	0.524976	Schwarz criterion	-5.596848
Log likelihood	5124.162	Hannan-Quinn criter.	-5.619810
F-statistic	12.01773	Durbin-Watson stat	2.059930
Prob(F-statistic)	0.000000		

Results….

Dependent Variable: D(LOG(BSE))
Method: ML - ARCH
Date: 02/07/10 Time: 10:55
Sample (adjusted): 1/07/2003 10/30/2009
Included observations: 1779 after adjustments
Convergence achieved after 23 iterations
MA Backcast: 1/02/2003 1/06/2003
Presample variance: backcast (parameter = 0.7)
GARCH = C(7) + C(8)*RESID(-1)^2 + C(9)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001605	0.000314	5.115677	0.0000
AR(1)	0.525339	0.062328	8.428613	0.0000
AR(2)	-0.870264	0.061490	-14.15291	0.0000
MA(2)	0.798233	0.066763	11.95614	0.0000
MA(3)	0.125830	0.026230	4.797224	0.0000
MA(1)	-0.422631	0.066514	-6.354028	0.0000
		Variance Equation		
C	6.29E-06	1.07E-06	5.880678	0.0000
RESID(-1)^2	0.155941	0.012945	12.04653	0.0000
GARCH(-1)	0.830730	0.013657	60.82617	0.0000

R-squared	0.016262	Mean dependent var	0.000878
Adjusted R-squared	0.013487	S.D. dependent var	0.017366
S.E. of regression	0.017249	Akaike info criterion	-5.655195
Sum squared resid	0.527499	Schwarz criterion	-5.627452
Log likelihood	5039.296	Hannan-Quinn criter.	-5.644948
F-statistic	3.663544	Durbin-Watson stat	2.036460
Prob(F-statistic)	0.000302		

Inverted AR Roots	.26+.90i	.26-.90i	
Inverted MA Roots	.28-.89i	.28+.89i	-.14

Results….

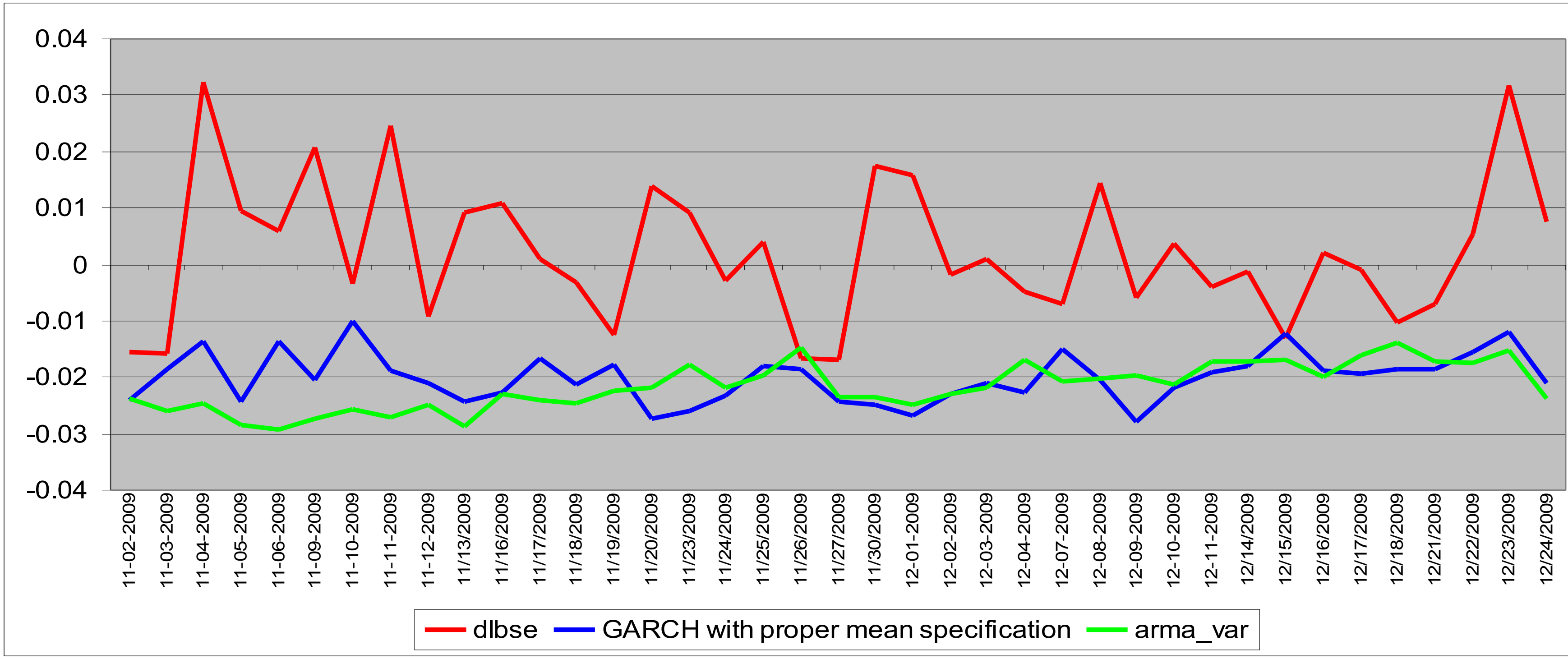
Dependent Variable: D(LOG(NSE))
Method: ML - ARCH
Date: 02/11/10 Time: 23:34
Sample (adjusted): 1/09/2003 12/23/2009
Included observations: 1815 after adjustments
Convergence achieved after 30 iterations
MA Backcast: 1/03/2003 1/08/2003
Presample variance: backcast (parameter = 0.7)
GARCH = C(8) + C(9)*RESID(-1)^2 + C(10)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001600	0.000317	5.047347	0.0000
AR(2)	-0.455719	0.084090	-5.419443	0.0000
AR(4)	-0.613579	0.096555	-6.354705	0.0000
AR(1)	0.498441	0.134215	3.713759	0.0002
MA(2)	0.422532	0.079148	5.338499	0.0000
MA(4)	0.672877	0.092978	7.236974	0.0000
MA(1)	-0.433857	0.127459	-3.403883	0.0007
Variance Equation				
C	7.61E-06	1.12E-06	6.785284	0.0000
RESID(-1)^2	0.137741	0.011880	11.59414	0.0000
GARCH(-1)	0.843305	0.012652	66.65310	0.0000

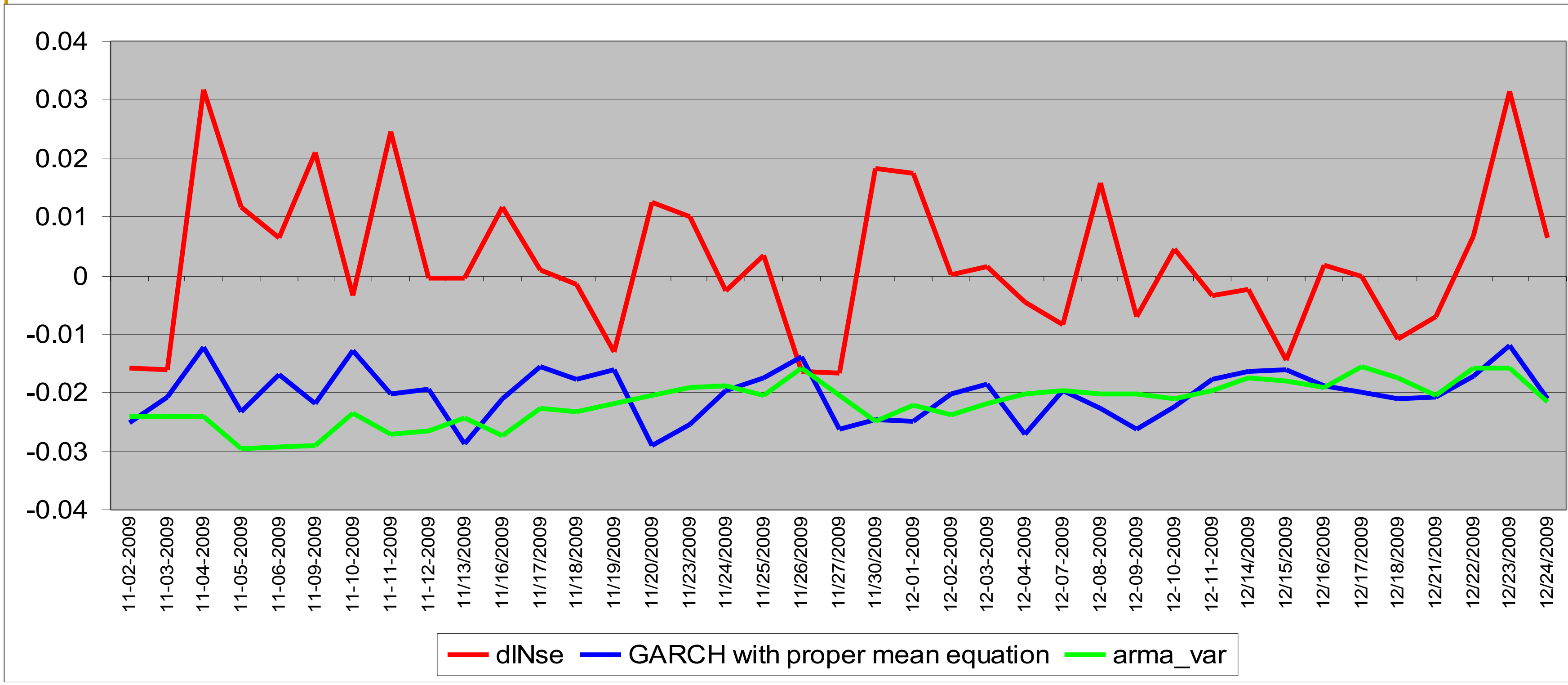
R-squared	0.007860	Mean dependent var	0.000855
Adjusted R-squared	0.004567	S.D. dependent var	0.017624
S.E. of regression	0.017584	Akaike info criterion	-5.571782
Sum squared resid	0.559039	Schwarz criterion	-5.541457
Log likelihood	5066.392	Hannan-Quinn criter.	-5.560592
F-statistic	1.591396	Durbin-Watson stat	1.991630
Prob(F-statistic)	0.112254		

Inverted AR Roots	.68-.73i	.68+.73i	-.43-.66i	-.43+.66i
Inverted MA Roots	.68-.73i	.68+.73i	-.46-.68i	-.46+.68i

Daily return on BSE and corresponding VaR based on Model A and Model B



Daily return on NSE and corresponding VaR based on Model A and Model B



Let the dispersion of VaR at 5% significant level based on model A (${}^{A}VaR_t^{.05}$)

from the actual return (r_t) be $D^A = \Sigma(r_t - {}^{A}VaR_t^{.05})^2$ and $D^B = \Sigma(r_t - {}^{B}VaR_t^{.05})^2$ for mode B.

It is observed that

($D^A_{BSE} = 0.02625$, $D^B_{BSE} = 0.03022$), ($D^A_{NSE} = 0.02673$, $D^B_{NSE} = 0.03050$)

Since $D^A_{BSE} < D^B_{BSE}$; $D^A_{NSE} < D^B_{NSE}$, we conclude that for both BSE-SENSEX and NSE-NIFTY price indices, model A performs better in estimating the VaR.

Conclusion

(I) It is found that global financial situation (lag values of return on S&P 500, INR-EURO & INR-USD exchange rate, Gold price used as proxies to global financial condition) has significant impact on Indian capital market.

(II) VaR of return in the Indian capital market estimated based on GARCH with suitable mean specification outperforms the ARMA-GARCH method

Thank you