Intl. Trans. in Op. Res. xx (2022) 1–31

Prioritizing of volatility models: a computational analysis using Data Envelopment Analysis

Konstantinos Petridis [a,*], Nikolaos E. Petridis [b], Ali Emrouznejad [c] and Fouad Ben-Abdelaziz [d]

[a]*Department of Business Administration, University of Macedonia, Thessaloniki, Greece*
[b]*Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece*
[c]*Aston Business School, Aston University, Birmingham, UK*
[d]*Neoma Business School, Rouen, France*
E-mail: k.petridis@uom.edu.gr [K. Petridis]; nikolaos.e.petridis@uom.edu.gr [N. Petridis]; a.emrouznejad@aston.ac.uk [E. Emrouznejad] Fouad.BEN.ABDELAZIZ@neoma-bs.fr [F. B. Abdelaziz]

Abstract

Economic crisis and uncertainty in global status-quo affect stock markets around the world. This fact imposes improvement in the development of volatility models. However, the comparison among volatility models cannot be made based on a single error measure as a model can perform better in one error measure and worst in another. In this paper we propose a two stage approach for prioritizing volatility models which in the first stage we develop a novel Slacks-Based Data Envelopment Analysis to rank volatility models. The robustness of the proposed approach has also been investigated using cluster analysis. In a second stage analysis, it is investigated whether the efficiency scores depend on models characteristics. These attributes concern the time needed in order the model to be estimated, the value of Akaike Information Criterion, the number of models' significant parameters, groups and bias terms and the error sum of squares. Further, dummy variables have been introduced to the regression model in order to find whether the employed model includes an in-mean affect, whether the assumed distribution is skewed and whether the employed model belongs to the fGARCH family. The main findings of this research show that the number of models' statistically significant coefficients, error sum of squares and in-mean effects tend to increase the efficiency scores, while time elapsed, number of statistically significant bias terms and skewed error distributions tend to decrease the efficiency score.

Keywords: Statistical Distributions; Forecasting; Mathematics for computing;Data Envelopment Analysis ; Ranking; β- regression;Finance

*Author to whom all correspondence should be addressed (e-mail: k.petridis@uom.edu.gr).

1. Introduction

Every second, millions of financial transactions are performed around the world. The tools that are used in order to handle information about stocks, bonds or any financial product over time, are volatility models. These models are applied to time-series data and provide valuable insight by forecasting the future values. The identification of the best volatility technique is based on error measures. There is a plethora of error measures that provides decisions as to whether a volatility model is suitable or not. However, error measures may provide contradictory results.

Applying two volatility models on a specific dataset, let us assume Model 1 and Model 2, based on the fact that the best or the most suited volatility model cannot be determined, creates the need of a methodology that provides a single measure based on which volatility models are ranked. Using the data of IBM prices (yahoo.com) as explained in Section 3, error measures for 198 volatility models are derived. Nevertheless, the problem that arises, lies on the choice of the most representative forecasting error measure, based on which a volatility model is proved to be accurate.

Even though some error measures are frequently used more than others, it cannot be concluded that these error measures are representative across all volatility models and data sets. A single aggregated measure can be determined if weights can be assigned to each error measure in order to rank volatility methods, nevertheless this ranking would be a subjective, as by changing the weights on the error measures, different scores would derive which may lead to different rankings. In this paper a novel application of Slacks-Based Data Envelopment Analysis (hereafter SBM - DEA) model is presented. With this SMB-DEA model, the volatility models are ranked based on the efficiency score obtained from the linear programming. A comparison of the proposed method is conducted with econometric approaches and the rankings are statistically analyzed. The efficiency scores of each volatility model subject a second stage analysis in order to examine which exogenous factors affect the score. Reviewing relative literature, the problem of forecasting models' ranking has been also examined with the use of Multi-Criteria Decision Making (MCDA). However, the problem with the use of MCDA models is that these models require from the user(s) to provide specific rules based on which, a criterion is better than another one by providing pairwise comparisons. In this context, a group of experts could provide different pairwise comparisons from another panel of experts. This leads to different rankings due to the fact of different subjective judgments. On the contrary, the model proposed in this paper provides an objective score based on which the ranking of each technique/model is conducted.

The rest of the paper is organized as follows: in Section 2 the literature review is presented providing all the relevant works in the field and background information. In Section 3, the methodology is presented providing an introduction to classical DEA models in order to make a transition to the proposed model while in Section 4 the case study demonstrates the applicability of the proposed method. This is followed by discussion of the results and comparison with other statistical approaches in Section 5. In Section 6, the second stage analysis is presented, including clustering analysis. Conclusions and direction for future directions are given in Section 7.

2. Literature review

Prioritizing or ranking forecasting models based on one or more criteria has been examined in the literature (1). Duong (1988) (2), has applied Analytical Hierarchy Process (AHP) in order to prioritize forecasting models based on AIC (Akaike Information Criterion), MSE_S (Mean Forecast Squared Error one - step ahead), $MAXE$ (Maximum Forecast Error one - step ahead), MSE_L Mean Forecast Squared Error 10 or 11 - step ahead) and $\overline{MSE}$ (Average Mean Squared Forecast error, from 1 to 11 steps). Decision based models have also been applied for the evaluation of forecasting models based on decision makers and agents (3). However applying a MDCM methodology like AHP would result in different rankings if the pairwise comparison matrix is constructed from another set of decision maker(s). Accuracy measures are the ones to be used as a means of comparing forecasting models and indicated their importance (4). A comparison of the forecasting models has also been conducted based on a variety of criteria; experts provided their insight for forecasting models based on accuracy measures and ease of use. Results demonstrate that researchers, from the group of experts, rated higher the accuracy component in comparison to ease of use. On the contrary, decision makers consider higher the ease of use of each forecasting technique rather than the accuracy (5).

The problem of comparing forecasting models, and in particular volatility models, has been examined using a selection of different methodologies. Out of sample volatility models are compared based on the values of R^2 from a Mincer – Zarnowitz (MZ) regression or logarithmic version which has been noted that is less sensitive to outliers ((6), (7)). However, taking R^2 from MZ regression is not a representative measure as it was proved not to penalize any biased terms. Some studies the supremacy of a volatility technique (or the corresponding family) towards other models has been demonstrates through data application and empirical results. Using naive p – values Hansen and Lunde (2005) (8) compared $GARCH(1, 1)$ and $ARCH(1)$ volatility models. Similar approached have been proposed for forecasting volatility (9). The results show $GARCH(1, 1)$ yield the best results in all criteria except for MSE_2. However the aforementioned technique is based on the comparison of the all corresponding error measures of each technique. In the case where there is better performance in some measures of one volatility technique but not in others then no agreement is extracted. A comparison between ARMA, ARFIMA and GARCH models has also been proposed which applied to exchange rates of many currencies (e.g. Pound, Mark and Yen) to the US Dollar. The comparison of these models has been assessed based on single error measure (MSE) regressed against R^2 (10). In the case where more than one error measures is used, a technique could perform better in one and worse in another measure hence there would be no technique that would outperform in comparison to the others.

Comparisons of hybrid forecasting – Artificial Neural Network (ANN) models have been applied on wind speed data (11), on environmental science (river flow forecasting) (12) and in short term load ((13), (14)), but none of these studies use all measurement errors, hence the need for a single score that would aggregate all error measures is necessary.

Besides the study of volatility models comparison, the study of stochastic volatility (SV) models comparison has also been examined through a WinBUGS implementation using Bayesian Statistics (15). In this study SV models have been applied to extract the error measures of exchange rate of New Zealand (NZ) dollar and US Dollar. The empirical results demonstrate that the best specifications are those that allow for time-varying correlation coefficients. Stochastic Volatility models have also been examined in the prism of an Euler discretization technique to compare different methodologies (16).

Searching through literature, the main approaches used for a comparison amongst two or more volatility or forecasting models based on a single or multiple error measures are statistical, MCDM and judgmental decision support systems. However, when comparing volatility models based on error measures, there is very little probability that a technique provides better results in almost all error measures against other volatility models (17). An alternative way of comparing volatility models is to measure the performance based on economic significance which could be proxied by the gain or revenue that the decision maker will incur after applying different volatility models. Nevertheless, to the best of our knowledge, no technique has been proposed that would aggregate all error measures to provide a ranking based on the efficiency score derived from this analysis. The proposed approach provides a unified score for all volatility models which is very easy to be reproducible and can be applied to other models based on their error measures. The proposed methodology has exhausted almost all volatility models and compares each other providing an efficiency score, based on which, each technique is ranked.

3. Methodology

3.1. Introduction to DEA

Data Envelopment Analysis (DEA) is the most frequently used model for measuring efficiency and productivity of decision making units (DMU), which been firstly been introduced by Charnes, Cooper, & Rhodes (1978) (18) and Banker, Charnes, & Cooper (1984) (19). Based on this technique, a Linear Programming (LP) model is solved for each unit under investigation, hereafter called Decision Making Unit (DMU). Assuming that there are j DMUs ($j = 1, 2, ..., n$) each consumes i ($i = 1, 2, ..., m$) inputs ($x_{i,j}$) and produces r ($r = 1, 2, ..., s$) outputs ($y_{r,j}$) in order to assess efficiency, LP model (1) is solved for each DMU j. There are two main DEA models, depending on the nature of the data, a decision maker aims at minimizing the inputs to produce more outputs (input oriented DEA model) or to maximize the outputs while retaining the inputs to the same levels (output oriented models). LP model (1) is solved n times (one for each DMU). The levels of decisions that are derived after each iteration, are efficiency scores for each DMU under investigation ($\theta^*_{j_o}$) and the λ^*_j for each DMU. Efficiency score $\theta^*_{j_o}$ yields a number in the range of $[0, 1]$; a DMU with $\theta^*_{j_o} = 1$ is efficient whereas a DMU with $\theta^*_{j_o} < 1$ is inefficient. Model (1) is a Variable Returns to Scale (VRS) model due to constraint $\left(\sum_{j=1}^{n} \lambda_j = 1\right)$. This constraint implies that an increase in the inputs does not imply same proportional increase in the outputs. Variables s_i^- and s_r^+ are slack variables which are associated with inputs and outputs, correspondingly. They denote the further reduction in inputs and the increase in outputs in order to be fully efficient. The variables are penalized in the objective function by an infinitesimal positive number, ϵ (whereas $\epsilon \simeq 10^{-5}$). Once the LP model (1) is solved, the projected values for inputs and outputs are given as follows: $\hat{x}_{i,j} = x_{i,j} \cdot \theta^* - s_i^{-,*}$ and $\hat{y}_{r,j} = y_{r,j} + s_r^{+,*}$ (projected inputs, outputs are denoted with $\hat{x}$ and $\hat{y}$ correspondingly while the optimal solutions for slack variables are denoted with $s_i^{-,*}, s_r^{+,*}$). A DMU is fully efficient if $\theta^*_{j_o} = 1$ and $s_i^{-,*} = 0, s_r^{+,*} = 0$. In this instance, the DMU cannot improve any more the inputs and the outputs and it belongs to the benchmark.

$$
\begin{aligned}
& min\ \theta_{j_o} - \epsilon \cdot \Big(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \Big) \\
& \quad s.t. \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot x_{i,j} + s_i^- = \theta_{j_o} \cdot x_{i,j_o}, i = 1, 2, ..., m \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot y_{r,j} - s_r^+ = y_{r,j_o}, r = 1, 2, ..., s \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \\
& \quad \lambda_j \geq 0, j = 1, .., n \\
& \quad \theta\ free
\end{aligned} \tag{1}
$$

The dual model of LP model (1) is presented below; variables v_i and u_r are dual variables assigned to inputs and outputs.

$$
\begin{aligned}
& max\ \sum_{r=1}^{s} u_r \cdot y_{r,j_o} + \omega \\
& \quad s.t. \\
& \quad \sum_{i=1}^{m} v_i \cdot x_{i,j} = 1, j = 1, 2, ..., n \\
& \quad \sum_{r=1}^{s} u_r \cdot y_{r,j} - \sum_{i=1}^{m} v_i \cdot x_{i,j} + \omega \leq 0, j = 1, 2, ..., n \\
& \quad u_r \geq \epsilon, r = 1, ..., s \\
& \quad v_i \geq \epsilon, i = 1, ...m \\
& \quad \omega\ free
\end{aligned} \tag{2}
$$

In this context, each DMU represents a volatility model where accuracy measures are the inputs. DEA assesses volatility models based on a single objective score, unlike MCDM models that assess volatility models (2) with multiple criteria or attributes based on the judgmental scores from a panel of experts. The errors based on which volatility models are assessed are the following (20):

1. Mean Absolute Percentage Error (MAPE) $= \frac{1}{n} \sum_{t=1}^{T} |100 \cdot \frac{e_t}{Y_t}|$
2. Root Mean Square Error (RMSE) $= \sqrt{\frac{1}{T} \sum_{t=1}^{T} e_t^2}$
3. Mean Square Error (MSE) $= \frac{1}{T} \sum_{t=1}^{T} e_t^2$

4. Symmetric Mean Absolute Percentage Error (sMAPE) $= \frac{1}{n}\sum_{i=t}^{T}\frac{100\cdot|e_t|}{Y_t+F}$
5. Geometric Mean Absolute Error (GMAE) $= \sqrt[T]{\prod_{t=1}^{T}|e_t|}$
6. Median Absolute Percentage Error (MdAPE) = $median(\sum_{t=1}^{T}|100\cdot\frac{e_t}{Y_t}|)$
7. Root Mean Squared Percentage Error (RMSPE) = $\sqrt{\frac{1}{n}\sum_{t=1}^{T}|100\cdot\frac{e_t}{Y_t}|}$
8. Root Median Absolute Percentage Error (RMdAPE) = $\sqrt{median(\sum_{t=1}^{T}|100\cdot\frac{e_t}{Y_t}|)}$

The error measures used in this analysis represent ratios; therefore application of standard DEA models may fail to provide correct results. It has been shown in the literature that when using ratio data the Production Possibility Set (PPS) is multiplicative and not additive ((21), (22)). If the PPS is denoted as $P = \{(x,y) : x \to y\}$ whereas x are the inputs which produce outputs (y), according to convexity axiom any two points $P_A(x_A, y_A), P_B(x_B, y_B)$ are represented with the following combination $P' = \lambda \cdot P_A + (1-\lambda)\cdot P_B$. Assigning values to λ, the new point that is created may not belong to the standard PPS (Figure 1) ((23)).

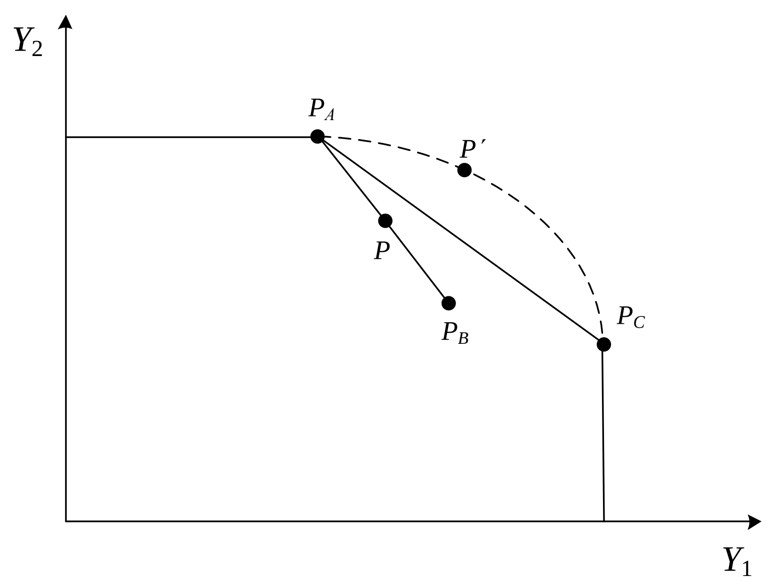

Fig. 1. An output productivity possibility set.

The mathematical formulation for an input oriented without explicit output is shown in the following context:

$$
\begin{aligned}
& min\ \theta_{j_o} \\
& \quad s.t. \\
& \quad \prod_{j=1}^{n} \Big((x_{i,j})^{\lambda_j} \Big) \leq \theta_{j_o} \cdot x_{i,j_o}, i = 1, 2, ..., m \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \\
& \quad \lambda_j \geq 0, j = 1, .., n \\
& \quad \theta\ \textit{free}
\end{aligned}
\tag{3}
$$

Model (3) is non-linear and uses products instead of summations as constraints in the model. In order to transform this model to a linear one using log-linear transformation (4).

$$
\begin{aligned}
& log_{10}\Big(\prod_{j=1}^{n} \Big((x_{i,j})^{\lambda_j} \Big) \Big) \leq log_{10}\Big(\theta_{j_o} \cdot x_{i,j_o} \Big) \iff \\
& \sum_{j=1}^{n} (log_{10}(x_{i,j}) \cdot \lambda_j \leq log_{10}(\theta_{j_o}) + log_{10}(x_{i,j_o})
\end{aligned}
\tag{4}
$$

Model (5) is a log-linear representation of a Geometrical DEA model. The log_{10} efficiency is denoted with $\tilde{\theta}_{j_o}$ while the log_{10} of inputs and outputs are denoted with $\tilde{x}$ and $\tilde{y}$ correspondingly. Using this transformation, LP model (5) is a log-linear transformation of model (3). The efficiency of each DMU is calculated as $10^{\tilde{\theta}^*}$ (whereas $\tilde{\theta}^*$ is the optimal solution of LP model (5)).

$$
\begin{aligned}
& min\ \tilde{\theta}_{j_o} \\
& \quad s.t. \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot \tilde{x}_{i,j} \leq \tilde{x}_{i,j_o} + \tilde{\theta}_{i,j_o}, i = 1, 2, ..., m \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \\
& \quad \lambda_j \geq 0, j = 1, .., n \\
& \quad \tilde{\theta}\ \textit{free}
\end{aligned}
\tag{5}
$$

3.2. Geometrical Slacks-Based Model DEA without outputs

In this instance, we extend the standard DEA log linear model to the log linear Geometrical Slacks-Based Model. The notion behind the use of such a model is because it is non-radial which takes into account non-zero slack values. The mathematical formulation for input-orinted SBM is presented below (6). A DMU is efficient if $\rho = 1$ and inefficient if $\rho < 1$ (24), (25).

$$
\begin{aligned}
& min\ \tilde{\rho}_{j_o} = 1 - \frac{1}{m} \cdot \sum_{i=1}^{m} s_i^- \\
& \quad s.t. \\
& \quad \sum_{j=1}^{n} \lambda_j \cdot \tilde{x}_{i,j} + s_i^- = \tilde{x}_{i,j_o}, i = 1, 2, ..., m \\
& \quad \sum_{j=1}^{n} \lambda_j = 1 \\
& \quad \lambda_j \geq 0, j = 1, .., n \\
& \quad \tilde{\rho}\ free
\end{aligned} \tag{6}
$$

In this case study, each volatility model is considered as a DMU while the inputs are the error measures after applying each one of the volatility models to financial data set described in Figure 2.

The volatility models used are shown in Table 1 while the conditional distributions assumed are shown in Table 2. Combining each one of the volatility model with each conditional distribution produces a unique entity. For example, the $ARCH(1)$ assuming normal conditional distribution produces a different model (in our context, a DMU) to $ARCH(1)$ combined with skewed normal conditional distribution. Thus, there are 11 volatility models without an *in-mean-effect*, and 11 volatility models assuming an *in-mean-effect* both categories assuming 9 conditional distributions. The total combinations lead to $11 \cdot 2 \cdot 9 = 198$ volatility models (DMUs).

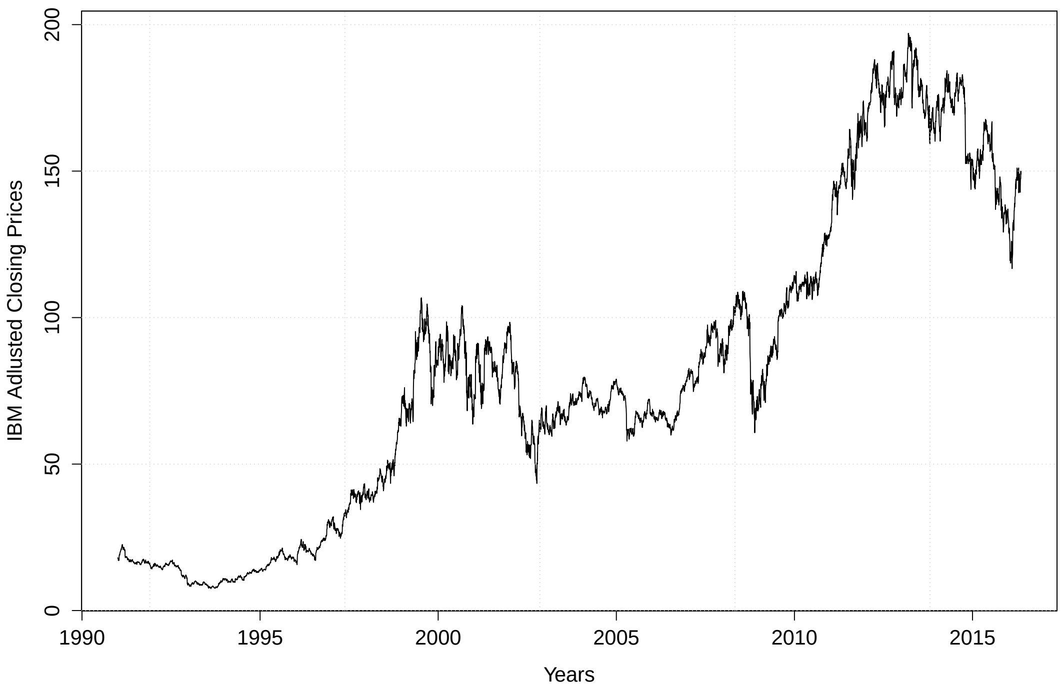

Fig. 2. *IBM* Adjusted Closing Prices

Table 1
Volatility models

Author	Model name	Equations
Engle,R.F.,(1982) (26)	ARCH(p)	$y_t = x_t^T \beta + e_t$ $e_t^2 = a_0 + \sum_{i=1}^{p} a_i \cdot e_{t-i}^2 + v_t$ $v_t \sim N(0, \sigma_v^2)$
Bollerslev,T.,(1986) (27)	GARCH(p,q)	$y_t = x_t^T \beta + e_t$ $\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i \cdot e_{t-i}^2 +$ $\sum_{j=1}^{q} \sigma_{t-j}^2 + u_t$ $u_t \sim N(0, \sigma_u^2)$
Engle, R.F.et al.,(1987) (28)	ARCH-M(p)	$y_t = \mu_t + e_t$ $\mu_t = \beta + \delta \cdot \sigma_t^2$ $\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i e_{t-i}^2 + r_t$ $r_t \sim N(0, \sigma_r^2)$
Bollerslev,T.,(1986) (27)	GARCH-M(p,q)	$y_t = \lambda_0 + \delta_0 \sigma_t(\theta_0)$ $\epsilon_t = \sigma_t(\theta_0)\eta_t$ $\eta_t \sim IID(0,1)$ $\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_{t-i}\epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(\theta_0)$
Ding.et al.,(1987) (29)	APARCH(m,s)	$y_t = \mu_t + \alpha_t$ $\alpha_t = \sigma_t \epsilon_t$ $\epsilon_t \sim D(0,1)$ $\sigma_t^\delta = \omega + \sum_{i=1}^{m} \alpha_i \left(\mid \alpha_{t-i} \mid + \gamma_i \alpha_{t-i} \right)^\delta$ $+ \sum_{j=1}^{s} \beta_j \sigma_{t-j}^\delta$
Taylor,S.J.,(2007) (30), Schwert,W.G.(1990) (31)	AV-GARCH(p,q)	APARCH(m,s) when $\delta = 1$ and $\gamma_i = 0$
Glosten,R.L et al., (1990) (32)	GJR - GARCH(p,q)	APARCH(m,s) when $\delta = 2$
Zakoian,J.M,(1993) (33)	T-GARCH(p,q)	APARCH(m,s) when $\delta = 1$
Bera,A.K, and Higgins,M.L (1993) (34)	N-GARCH(p,q)	APARCH(m,s) when $\gamma_i = 0$ and $\beta_j = 0$

Table 2
Conditional distributions used

Distribution	Parameters	Density function
Normal Distribution	mean $=\mu$, variance = σ^2	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{(x_i-\mu)^2}{\sigma^2}$
Student's - t	location $=\alpha$, scale = β, shape=ν	$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\beta\nu\pi}\Gamma(\frac{\nu}{2})}\left(1+\frac{(x_i-\alpha)^2}{\beta\nu}\right)^{-\frac{\nu+1}{2}}$
Generalized Error Distribution	location $=\alpha$, scale = β, shape=κ	$f(x) = \frac{\kappa e^{-0.5\lvert\frac{x_i-\alpha}{\beta}\rvert^{\kappa}}}{2^{1+\kappa^{-1}}\beta\Gamma(\kappa^{-1})}$
Skewed Normal Distribution	skewness $=\xi$	$f(z \mid \xi) = \frac{2}{\xi+\xi^{-1}}\left[f(z\xi)H(-z)+f(\xi^{-1}z)H(z)\right]$
Skewed Generalized Error Distribution	mean $=\mu$, variance=σ^2 , skewness=λ, scaling = k	$f(x) = \frac{C}{\sigma} e^{-\frac{\lvert x_i-\mu+\delta\sigma\rvert^k}{\left[1-sgn\left(x_i-\mu+\delta\sigma\right)\lambda\theta\sigma\right]^k}}$
Skewed Student's - t	shape = β	$f(x) = \frac{2\beta}{1+\beta^2}\left[t_v(\beta x_i)I(x_i<0)+t_v(\frac{x_i}{\beta})I(x0)\right]$
Normal Inverse Gaussian	location=μ, scale =δ, shape =α,β	$f(x) = \frac{\delta\alpha e^{\left(\delta\sqrt{\alpha^2-\beta^2}\right)} K_1\left(\alpha\sqrt{\delta^2+\left(x_i-\mu\right)^2}\right) e^{beta(x_i-\mu}}{\pi\sqrt{\delta^2+\left(x_i-\mu\right)^2}}$
Generalized Hyperblic Distribution	location=μ, scale =δ, shape =α,λ,asymmetry=β	$f(x) = \frac{\frac{\sqrt{\alpha^2-\beta^2}}{\delta^{\lambda}}}{\sqrt{2\pi}K_{\lambda}(\delta\sqrt{\alpha^2-\beta^2})}$
Johnson's SU Distribution	skewness = ν,kurtosis =τ, mean = μ, variance = σ^2	$f(x) = \frac{\tau}{\sigma}\frac{e^{\frac{1}{2\left[\nu+\tau log\left[r+\left(r^2+1\right)^{\frac{1}{2}}\right]\right]}}}{\left(r^2+1\right)^{\frac{1}{2}}\sqrt{2\pi}}$

4. Case study

The applicability of the proposed methodology is demonstrated via a case study. In this example, 198 volatility models have been considered as seen in Table 1. The accuracy measures that have been examined are MAPE, MSE, RMSE, sMAPE, GMAE, MdAPE, RMSPE, and RMdAPE (the mathematical formula of which has been given above). One of the problems of forecasting/volatility models is that their performance is based on accuracy measures in which not all models may perform well. For example assuming that there are two volatility models $VOL1$, and $VOL2$ and two accuracy measures ($AM1$ and $AM2$). If $VOL1$ is performing better than $VOL2$ in $AM1$ and worst than $VOL2$ in $AM2$ then unless there is a weighting in the accuracy measures, there is no way to conclude which of the volatility models should be used. The problem of selecting the best volatility technique can be addressed with the proposed methodology as by solving the model for each DMU a single score is derived.

LP model (6) is analytically written below as follows using the accuracy presented above. As all of the accuracy measures should be minimized, the accuracy measures presented above are considered as inputs.

$$
\begin{aligned}
& min\ \rho_{j_o} = 1 - \frac{1}{8} \cdot \left(s^-_{MAPE} + s^-_{MSE} + s^-_{RMSE} + s^-_{sMAPE} + s^-_{GMAE} + s^-_{MdAPE} + s^-_{RMSPE} + s^-_{RMdAPE} \right) \\
& s.t. \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{MAPE}_j + s^-_{MAPE} = \widetilde{MAPE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{MSE}_j + s^-_{MSE} = \widetilde{MSE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{RMSE}_j + s^-_{RMSE} = \widetilde{RMSE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{sMAPE}_j + s^-_{sMAPE} = \widetilde{sMAPE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{GMAE}_j + s^-_{GMAE} = \widetilde{GMAE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{MdAPE}_j + s^-_{MdAPE} = \widetilde{MdAPE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{RMSPE}_j + s^-_{RMSPE} = \widetilde{RMSPE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j \cdot \widetilde{RMdAPE}_j + s^-_{RMdAPE} = \widetilde{RMdAPE}_{j_o} \\
& \sum_{j=1}^{198} \lambda_j = 1 \\
& \lambda_j \geq 0, j = 1, .., 198 \\
& \tilde{\rho}\ free
\end{aligned} \tag{7}
$$

A profiling of the data used is shown in Figure 3 and in Figure 4. In Figure 3 MAPE, MSE, RMSE and sMAPE error measures are shown while in Figure 4, GMAE, MdAPE, RMSPE and RMdAPE. In both figures the common outcome is that there is not a single criterion based on which the volatility models can be assessed and ranked. For example, in Figure 3 in MAPE error measure, the first 20 volatility models show large fluctuation in comparison with the corresponding values to MSE, RMSE and sMAPE error measures. Also, regarding sMAPE it can be seen that almost all volatility models provide low scores except for two outliers. In Figure 4 the scores of volatility models demonstrate a similar pattern but still there is not an error measure where all volatility models (shown in the x axis) provide better results than another one.

In Table 3, a summary of descriptive statistics is provided per each error measure. It can be seen that the produced volatility models' mean absolute percentage error (MAPE) is quite small, with minimum value which reaches 0.83, mean value equal to 0.97 and maximum value equal to 1.19. All forecasting error measures for the volatility models demonstrate no extreme values and small deviation except for symmetric mean absolute percentage error (sMAPE), with a minimum value of 165.5 and a maximum value of 2061.

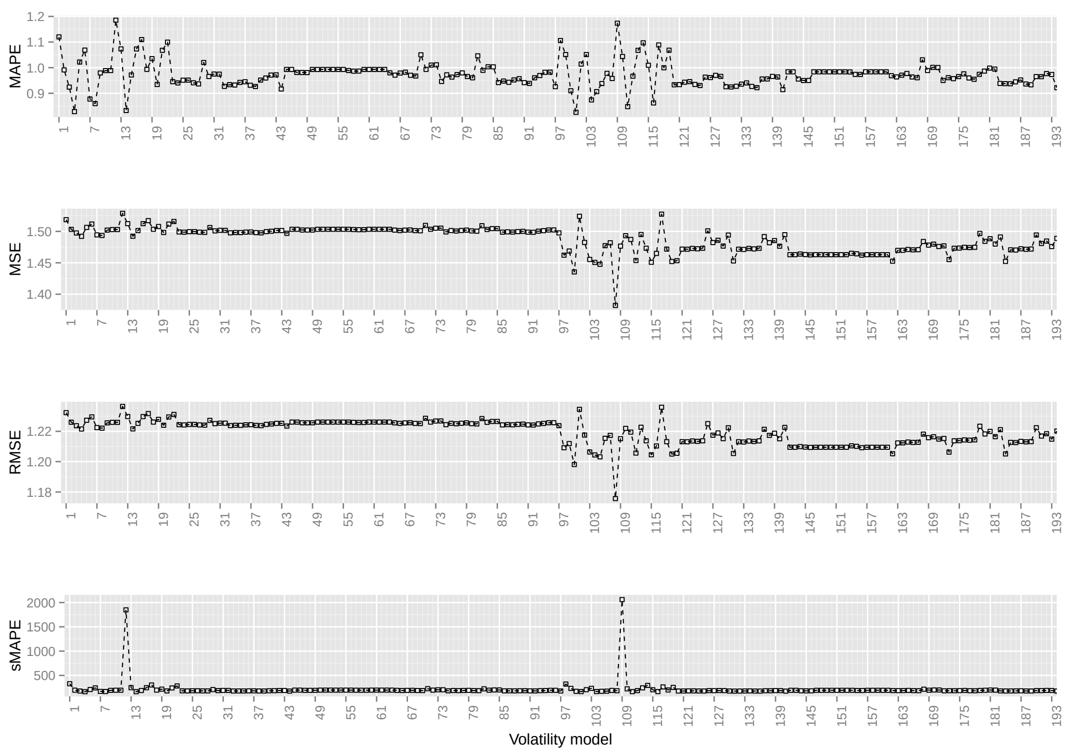

Fig. 3. Line plots of MAPE, MSE and RMSE and sMAPE error measures

5. Results

5.1. Prioritizing volatility models

In this section the results of the application of SBM. In Figure 5 a profile of the efficiency is shown; the DMUs that are efficient are denoted with a red circle. The DMUs that are efficient are:

- ARCH(1)-skew student t
- ARMA(1,0)-csGARCH(1,1)-skew student t

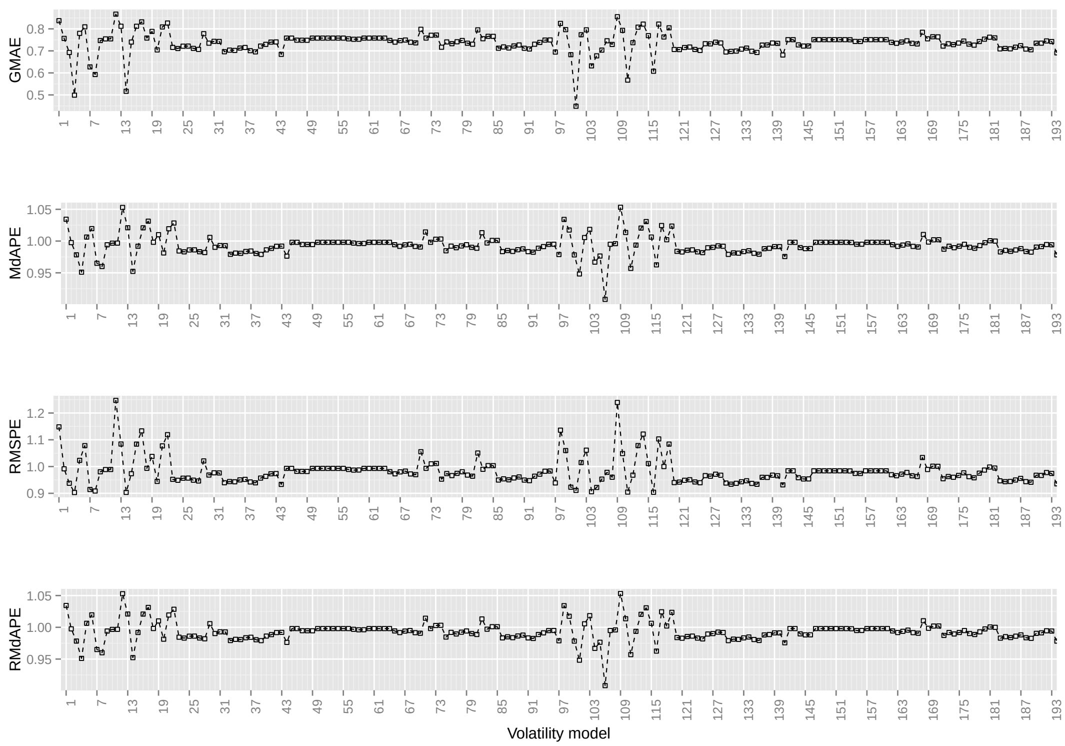

Fig. 4. Line plots of GMAE, MdAPE, RMSPE and RMdAPE error measures

- ARMA(1,0)-csGARCH(1,1)-normal inverse
- ARMA(1,0)-csGARCH(1,1)-Johnson's SU
- apARCH(1,1)-skew normal
- ARMA(1,0)-apARCH(1,1)-normal

It can be seen that most of the volatility models with higher efficiency scores are the ones that belong to the $ARMA(1,0) - csGARCH(1,1)$ family. This is attributed to the fact that, by construction, component GARCH ($csGARCH$) is a more flexible specification compared with ARCH and GARCH models (35). Specifically, the volatility component in the long - run is considered to be stochastic, while the short run, which is also called transitory volatility part, is defined as the difference between the conditional variance and the trend. Therefore, it is developed to account for long - run dependencies, assuming unconditional variance, as opposed to other GARCH models which impose a constant variance constraint. Except for the component GARCH, Assymetric Power ARCH family volatility models ($APARCH$) demonstrate improved efficiency scores. Assymetric Power ARCH as described in Table 1, contains an extra parameter δ, which is called *heteroscedasticity coefficient*. This parameter in several studies(e.g (36)) is found to be improving the discrepancies between GARCH high autocorrelations and observed ones.

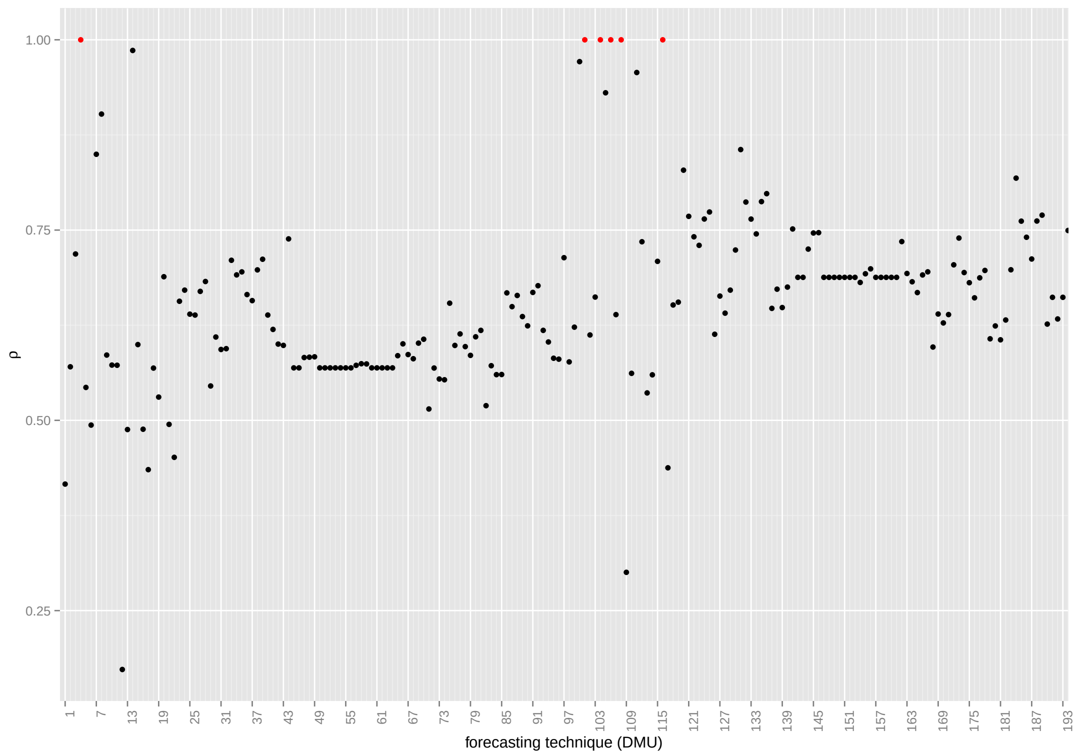

Fig. 5. Efficiency analysis

Accordingly, the ranking for volatility models is presented in Appendix A.

5.2. Cluster analysis

In order to investigation the robustness of the efficiency scores derived from SBM DEA, cluster analysis is conducted. Firstly a distance matrix for the error measures is constructed producing five clusters. The same procedure is carried out for the results of efficiency scores producing the same number of clusters as can be seen in Figure 6. Between the clusters derived by the aforementioned procedures, measures of similarity were applied to investigate the degree of overlapping. Assuming a set of elements ($S = \{\alpha_1, ..., \alpha_{194}, \beta_1, ..., \beta_{194}\}$) two partitions of set S are constructed (A, B); $A = \{\alpha_1, ..., \alpha_{194}\}$ is the set of observations that are clustered with the error measures and $B = \{\beta_1, ..., \beta_{194}\}$ is the set of observations that are clustered with the efficiency scores. Three measures for calculating the degree of overlapping elements were applied which are the Rand Index (37), Adjusted Rand Index (A.R.I) (38) and Normalized Mutual Information (N.M.I) (39). Based on Rand Index measure of similarity, the two partitions from the previously described procedure concentrate a similarity index of 82.8%, based on the A.R.I a score of 44.5% and based on the N.M.I 43.6%. From this finding it is obvious that groups produced by hierarchical cluster analysis which is applied on the error measures are very similar to the ones produced by the cluster analysis on the efficiency scores derived from the proposed SBM DEA approach. This result enhances the reliability of the SBM DEA efficiency scores as it produces same groups as the hierarchical cluster analysis which is a well defined statistical technique.

6. Second stage analysis

6.1. β Regression model

In this section the statistical model which is used to evaluate the factors which affect the efficiency scores is analytically described. Let y_j be the efficiency scores (ρ_j^*) which are produced by applying SBM DEA method. Due to the fact that $y_i \in [0,1]$ and assuming that its mean is associated linearly with a set of predictors, *beta regression* (40) is the most appropriate model to analyze how external factors have an impact on the efficiency scores. The name is given due to the hypothesis that response variable which receives values in the standard unit interval is assumed to be *beta* distributed. (41). One of the commonly used link functions, is the *logit* transformation (i.e $\tilde{y} = log(\frac{y}{1-y})$) however there are several disadvantages, concerning the way that estimated parameters are interpreted and that the response variable suffers from heteroscedasticity. Assuming that the efficiency scores for each combination of volatility model with conditional distributions (y_1,y_2,...,y_n) are *beta* distributed and denoting ϕ the precision parameter which is negatively linked to y's variance, (i.e $y_j \sim B(\mu_j, \phi)$) the *beta* model is the following:

$$g(\mu) = x_j^T \beta = \eta_j \tag{8}$$

where $\beta = (\beta_1, ..., \beta_p)$ denote the vector of unknown parameters to be estimated and $x_j = (x_{j1} \bigoplus, ..., \bigoplus x_{jp})$ the matrix of *p* independent variables and $\bigoplus$ denotes concatenation of variables, while η_i describes the linear predictor. Thanks to model's 8 flexibility several link functions can be implemented. The available options which can best describe the potential association between the response and the independent variables are the: a)the logit link $g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$, b)the probit $g(\mu) = \Phi\left(\mu\right)^{-1}$, c) the log - log function $g(\mu) = \Big(-log\big(-\log(\mu)\big)\Big)$ and other commonly used functions. The selection of the appropriate link function is based on the values of information criteria (i.e *AIC*, *BIC*). Coefficients of 8 are estimated using Maximum Likelihood method (*ML*). The likelihood function which is optimised is given below:

$$l_j(\beta, \phi) = \sum_{j=1}^{n} \Bigg[\log\Gamma - \log\Gamma(\mu_i\phi) - \log\Gamma\Big((1-\mu_j)\phi\Big) + (\mu_j\phi - 1)\log(y_j) + \Big[(1-\mu_i)\phi\Big]\log(1-y_j)\Big]\Bigg] \tag{9}$$

where $\Gamma(.)$ is the Γ function, ϕ is the precision parameter and β unknown coefficients to be estimated.

In order to assess the goodness of fit, McFadden's pseudo R^2 index is used (42). Let LL_{full} be the value of log - likelihood function evaluated on all independent variables and LL_0 the value of log - Likelihood function when the model contains no covariates but the intercept; McFadden's pseudo R^2 is defined as follows:

$$R^2 = 1 - \frac{LL_{full}}{LL_0} \tag{10}$$

McFadden's pseudo R^2 values are considerably lower than those obtained by estimating a linear model using Ordinary Least Squares $\big($*OLS*$\big)$. Typical values which demonstrate excellent fit for a binary logistic model, span from 0.2 to 0.4. (43).

The choice of the McFadden's pseudo R^2 index as a goodness of fit over others(e.g Cox - Snell) (44) is ought to the fact that it satisfies almost all of Kvålseth's eight criteria (45). A proper R^2 statistic should $\big(a\big)$ be easily interpreted and serve as a measure of good fit, $\big(b\big)$ not depend on units of measurements of the variables, $\big(c\big)$ be bounded within a range of values, the maximum of which indicates a perfect fit and the lower, lack of fit, $\big(d\big)$ be generic, namely be able to be applied to any type of type of model, $\big(e\big)$ not be depended on the method by which the model is estimated $\big($e.g maximum likelihood, ordinary least squared$\big)$, $\big(f\big)$ be easily comparable when models are estimated using different datasets, big(g) be comparable with values of other goodness of fit criteria and $\big(h\big)$ take into account equally the sign of residuals.

6.2. *Description of Independent variables*

A total of 198 combinations of volatility models assuming different conditional distribution have been estimated using the *rugarch* package in R (CRAN). The response variable is the efficiency scores which were obtained by applying SBM DEA, while independent variables concern models' specific attributes. In this section the factors which are assumed to be associated with increasing or decreasing models' efficiency scores are analyzed.

Elapsed time in seconds $\big($Time$\big)$: The complexity of the model is negatively associated with the estimation time. However, model's complexity does not always ensure accurate forecasted values, rather than good fit of the model, due to the fact that it explains a considerable amount of unknown variation of dependent variable. Given the fact that out-of-sample forecasting is based on the coefficients that have been calculated using data points in-sample, then model's good fit has an effect on the forecasting accuracy and therefore its efficiency.

Log - Likelihood $\big($Log.Lik$\big)$: The value of Log - Likelihood function evaluated on the vector of estimated coefficients provides information about model's good fit to the available time series data. A value of Log-Likelihood function which is extremely negative signals a good fit to the data. Hence, it is expected that models which exhibit good fit to the data tend to be more efficient in terms of forecasting.

Error Sum of Squares $\big($ESS$\big)$: Error $\big($or residual$\big)$ sum of squares $\big($*ESS* hereafter$\big)$ is measured as the difference between the observed value of dependent variable and the expected one.Let r_t be the actual returns at time *t* and $\hat{r_t}$ the expected returns at time *t* produced by an estimated volatility model using a set of *T* data points. The *ESS* is defined as:

$$ESS = \sum_{t=1}^{T} \left(r_t - \hat{r_t} \right)^2 \tag{11}$$

Error Sum of Squares index measures the amount of dependent's unknown variation which is not explained by the independent variables. Large values of $\big($*ESS*$\big)$ indicate that a model does not fit data appropriately, as the amount of unknown variance increases and potentially the inclusion of exogenous variables is needed. Based on the previous and assuming that *in-sample* estimation affects the out-of-sample forecasting efficiency, models with increased $\big($*ESS*$\big)$ will perform poor forecasting accuracy and will be less efficient.

Number of significant coefficients $\big($No.sig.coeffs$\big)$: Another measure which provides information about the appropriateness of the model to the data is the Number of significant coefficients. Estimated coefficients, may not always have the expected signs, based on the literature and lead to confusing conclusions. In those cases, coefficients are not statistically significant. The number of statistically significant of a model reflects the extent to which attributes of time series are captured by the applied volatility model.

Number of significant bias terms $\big($No.sig.bias$\big)$: Time series of stock markets tend to be characterized by violent breaks, due to spread of positive or negative information. This type of fluctuation is attributed partially to current news however, can be caused either by domestic factors and concern monetary policies announced by countries central banks (46), or fiscal strategies followed by governments. In other cases, they may come from domestic incidents, such as legislation regulating the framework by which firms operate inside a nation's market. These asymmetries $\big($or shocks$\big)$ are incorporated into volatility models and are tested using Enge's test (47), by estimating lagged positive or negative shocks against squared residuals. If z_t be the standardized residuals at time *t*, $I_{u_t - 1 < 0}$ indicate negative shock in the market, while $I_{u_t - 1 \geq 0}$ the positive ones, then the following regression is estimated:

$$z_t^2 = b_0 + b_1 \cdot I_{u_t - 1 < 0} + b_2 \cdot I_{u_t - 1 < 0} \cdot u_{t-1} + b_3 \cdot I_{u_t - 1 \geq 0} \cdot u_{t-1} + \varepsilon_t \tag{12}$$

After estimating 12, coefficients' statistical significance is tested for each parameter separately $\big(H_0 : b_i = 0, i = 1, 2, 3\big)$ and jointly $\big(H_0 : b_1 = b_2 = b_3 = 0\big)$. The number of statistically significant bias terms are recorded.

Number of significant groups $\big($No.sig.groups$\big)$: Except for the number of statistically significant terms which incorporate market's anomalies, the number of statistically significant groups of standardized residuals which are not normally distributed are taken into account as well (48). The evaluation of standardized residuals' distribution is based on the value of χ - square goodness of fit test.

In-mean-effect $\big($in.mean$\big)$: This is a dummy binary variable; if the value of this variable is 1, the forecasts have been based on an *in-mean* model and 0 otherwise. This family of models provide an explicit link between conditional volatility and the optimum forecast of time series. As explained in Table 1 the returns are assumed to be an *ARMA(p,q)* process and the residuals follow several specific *ARCH* or *GARCH* processes.

Skewed distributions $\big($skew$\big)$: In Table 2 several conditional distributions are used in order to perform accurate forecasts. If a distribution is skewed then this variable takes a value of $\big(1\big)$ and $\big(0\big)$ if its central

.

Family GARCH models (fGARCH): The last potential factor which is examined in this paper is the fact that the estimated model which produces forecasts, come from the GARCH family (1) or not (0). This family of models has been firstly introduced by Hentschel (49) and incorporates shifts and rotation concerning the news impact curve. Threshold GARCH (T-GARCH), GJR-GARCH, Non - linear GARCH (N-GARCH), Non - linear Asymmetric GARCH (NA-GARCH) and Absolute Value GARCH model (AV-GARCH) are included in this family of models.

In Table 4 the Pearson pairwise correlation index between the covariates is demonstrated, the effect of which is investigated in this study. Correlation values are low enough and therefore the simultaneous inclusion in the β regression model will not cause multi-collineality (50). Specifically, a relative high negative correlation has been observed between the value of log-likelihood function and the number of significant bias terms (r = -0.63), as well as between the number of significant bias term with the error sum of squares *ESS* (r = -0.58). The values for product moment correlation indices between all other pairs are extremely low (from -0.10 to 0.11). However, correlations may be misleading regarding the intensity of the linear association between two independent variables. For that reason, variance inflation index (*VIF*) is computed.

$$VIF = \frac{1}{\left(1 - R_j^2\right)} \tag{13}$$

where R_j^2 denotes the partial coefficient of determination. It expresses the proportion of explained variance of x_j which is explained by all other independent variables. Values of (*VIF*) above 10 signals existence of multi-collineality which causes inability of model's estimation (50) In table 4, *VIF* values for all independent variables are depicted. The maximum value of *VIF* has been computed for error sum of squares *ESS*, still it is considerably lower than 10 which is the threshold value, above which multi–collineality in the estimated model exists and results are unreliable. For all other covariates which are used in β regression model values of *VIF* do not exceed 2. This suggests that the examined factors can be simultaneously introduced into the model.

6.3. β Regression model estimation results

The estimation of β regression using maximum likelihood method are presented in this section. Four different models have been employed; the first model (*Model 1*) includes only covariates (b) which are associated with the complexity and the appropriateness of the estimated volatility model, namely model's estimation time, log - likelihood's function and residual sum of squares (*ESS*). The second model (*Model 2*) includes independent variables which provide information about the statistical significance of parameters and are the number of statistical significant volatility models' coefficients, the number of statistical significant volatility models' bias terms and the number of statistical significant volatility models' groups of residuals. The third estimated model (*Model 3*) contains factors which potentially have an effect on the efficiency scores of volatility models and are associated with specific characteristics of the models, such as if the estimated volatility model is in - mean, if the conditional distribution assumed is skewed or if the model belongs to the general GARCH family. Lastly, the fourth model (*Model 4*) assesses the effect of all the previous factors simultaneously on the efficiency of volatility model's forecasting capability.

The selection of the link function which best describes the data for each model has been based on information criteria (i.e *AIC*,*BIC*) and the value of log - Likelihood function. Logit and log - log link functions outperformed other possible specifications which associated the mean value with the covariates.

The time (Table 5) needed for software to estimate the available models affects negatively the efficiency of forecasting and this effect is statistically significant $(b = -0.61, p < 0.1)$. The higher the time it takes software to estimates indicates high complexity of the model and inefficiency of the implementation. As stated above, values of log - Likelihood function indicate good fit of the model to the data. Hence, it would be expected that increasing functional adjustment to lead to more accurate forecasts. However, values of log - Likelihood function do not affect efficiency of volatility models' forecasting ability$(b = 0.07, p > 0.1)$ and a good adjustment of a model to the data does not guarantee forecasted values being close to the observed ones. On the contrary, *ESS* tends to increase the volatility models' efficiency scores$(b = 0.01, p < 0.01)$. This finding is interesting in two ways: first it cancels out an implied, though never formally tested, notion that models good fit leads to improved forecasting ability and second that in this case, as unexplained variance increases, models' forecasting efficiency increases as well. Model's 1 goodness of fit is very satisfactory, due to the fact that Mc Fadden's pseudo R^2 exceeds 20%.

Number of model's statistically significant parameters (Model 2) improve the efficiency of forecasting ability $(b = 0.07, p < 0.1)$, however this effect is marginally statistically significant. On the other hand, the number of volatility model's statistically significant bias terms (e.g positive or negative shocks) affect efficiency score of volatility equation forecasting ability negatively $(b = -1.65, p < 0.01)$. The results for p-values imply, that in the case where a model explains abrupt rises or falls of return values, its forecasting ability worsens. The number of statistically significant groups of residuals, which are not independent and identically distributed, does not have any effect on the efficiency scores of volatility models$(b = -0.004, p > 0.1)$. Model's 2 goodness of fit measured by the Mc Fadden's pseudo R^2 is satisfactory as the value of R^2 reaches 21%.

Model 3 investigates whether efficiency scores can be explained by the volatility equations' specific characteristics. Mean values of efficiency scores for ARCH or GARCH - M models (i.e *in mean* effects) are not statistically from those volatility equations which do not assume existence of $ARMA(p, q)$ for the mean of returns $(b = 0.15, p > 0.1)$. In volatility models, where skew conditional distribution is assumed, the efficiency scores are not significantly different from those in which such an assumption is not done $(b = 0.01, p > 0.1)$. However, volatility models which belong to the family of GARCH equations tend to perform lower efficiency scores than those which do not belong to it $(b = -0.30, p < 0.01)$ although this difference is marginally statistically significant. The dummy variables do not seem to explain a great amount of efficiency scores variance due to the fact that R^2 is only 1%, considerably lower than the threshold value which indicates satisfying fit.

In Model 4, the joint effect of all covariates on efficiency scores is examined. The employed model is robust as the coefficient values and their signs do not change. Specifically the time needed to estimate forecasting retains the same value and sign as in Model 1 and the effect on the efficiency score is negative, though not statistically significant $(b = -0.58, p > 0.1)$. Forecasting volatility equations which do not explain a great amount of returns' variability, leading to increased *ESS*, tend to achieve higher efficiency than others. On the contrary, the number of statistically significant equation's terms lead to greater efficiency $\left(b = 0.07, p < 0.05\right)$, however the amount of volatility equation's bias terms which are statistically significant is associated with lower ranking in terms of forecasting efficiency $\left(b = -1.34, p < 0.01\right)$.

A point of great interest is the joint effects of the independent variables on the forecasting efficiency of volatility models when controlling the dummy variables. For that reason, three more models are estimated, which incorporate not only main effects but also interaction terms of the independent variables by the control variables. Model 5 is estimated using the existence of an *in - mean* effect as a control variable. Models 6 and 7 are estimated using as control variables the existence of skewed conditional distribution and whether the model belongs to the family GARCH.

In $\left(\text{Table 6}\right)$, the results of the estimated models are presented. Log - likelihood for volatility equations in which an *in - mean* effect is assumed, tends to increase their efficiency in terms of forecasting errors $\left(b = 15.34, p < 0.01\right)$. Furthermore, volatility equations, which incorporate an in - mean effect and the number of statistically significant coefficients is large, achieve a lower efficiency score for forecasting error measures $\left(b = -0.18, p < 0.05\right)$. Model's 6 fit is satisfying, as Mc Fadden's pseudo R^2 exceeds 20%. When controlling for the existence of skewed conditional distribution $\left(\text{Model 6 }\right)$, it seems that increased volatility equation's log - Likelihood tends to decrease the efficiency of forecasting measures $\left(b = -12.34, p < 0.01\right)$. Given a skewed conditional distribution, volatility equations which have many statistically significant coefficients, bias term or groups of residuals, tend to achieve a lower efficiency in terms of forecasting accuracy and large efficiency when *ESS* increases.

Table 3
Descriptive statistics of the error measures

	MAPE	MAE	MSE	RMSE	sMAPE	GMAE	MdAPE	RMSPE	RMdAPE
Minimum	0.8257	1.055	1.382	1.176	165.6	0.4486	0.9082	0.9031	0.9082
Q_1	0.9452	1.085	1.472	1.213	182.7	0.7159	0.9851	0.9521	0.9851
Median	0.9717	1.093	1.496	1.223	190.1	0.7404	0.9928	0.9735	0.9928
Mean	0.9738	1.092	1.487	1.219	212.8	0.7359	0.9931	0.9811	0.9931
Q_3	0.9930	1.097	1.502	1.226	197.3	0.7580	0.9980	0.9931	0.9980
Maximum	1.1850	1.115	1.529	1.236	2061.0	0.8667	1.0530	1.2470	1.0530

Table 4
Correlation matrix and descriptive statistics for covariates

	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[8]**	**[8]**	**[9]**	**VIF**
Time [1]	1	-0.05	-0.05	-0.15	0.15	-0.11	0.04	-0.06	0.09	1.05
Log.lik [2]	-0.05	1	0.02	0.04	-0.63	0.48	-0.07	0.09	-0.05	1.82
No.sig.coefs [3]	-0.05	0.02	1	0.19	0.09	-0.05	-0.03	0.02	0.20	1.17
No.sig.groups [4]	-0.15	0.04	0.19	1	-0.12	0.26	-0.003	0.005	-0.32	1.24
No.sig.bias [5]	0.14	-0.63	0.09	-0.12	1	-0.58	0.05	-0.03	0.14	2.15
ESS [6]	-0.11	0.48	-0.05	0.26	-0.58	1	-0.02	0.11	-0.58	2.67
in.mean [7]	0.04	-0.07	0.03	-0.003	0.05	-0.02	1	-0.09	-0.02	1.02
skew [8]	-0.06	0.09	0.02	0.005	-0.03	0.11	-0.09	1	-0.10	1.04
fGARCH [9]	0.09	-0.05	0.20	-0.32	0.14	-0.58	-0.02	-0.10	1	1.95
Mean	12	-2829.7	5.45	7.97	1.96	29993.92	-	-	-	
Variance	12.7	$1.34 \cdot 10^5$	1.49	3.09	0.22	19.45	-	-	-	

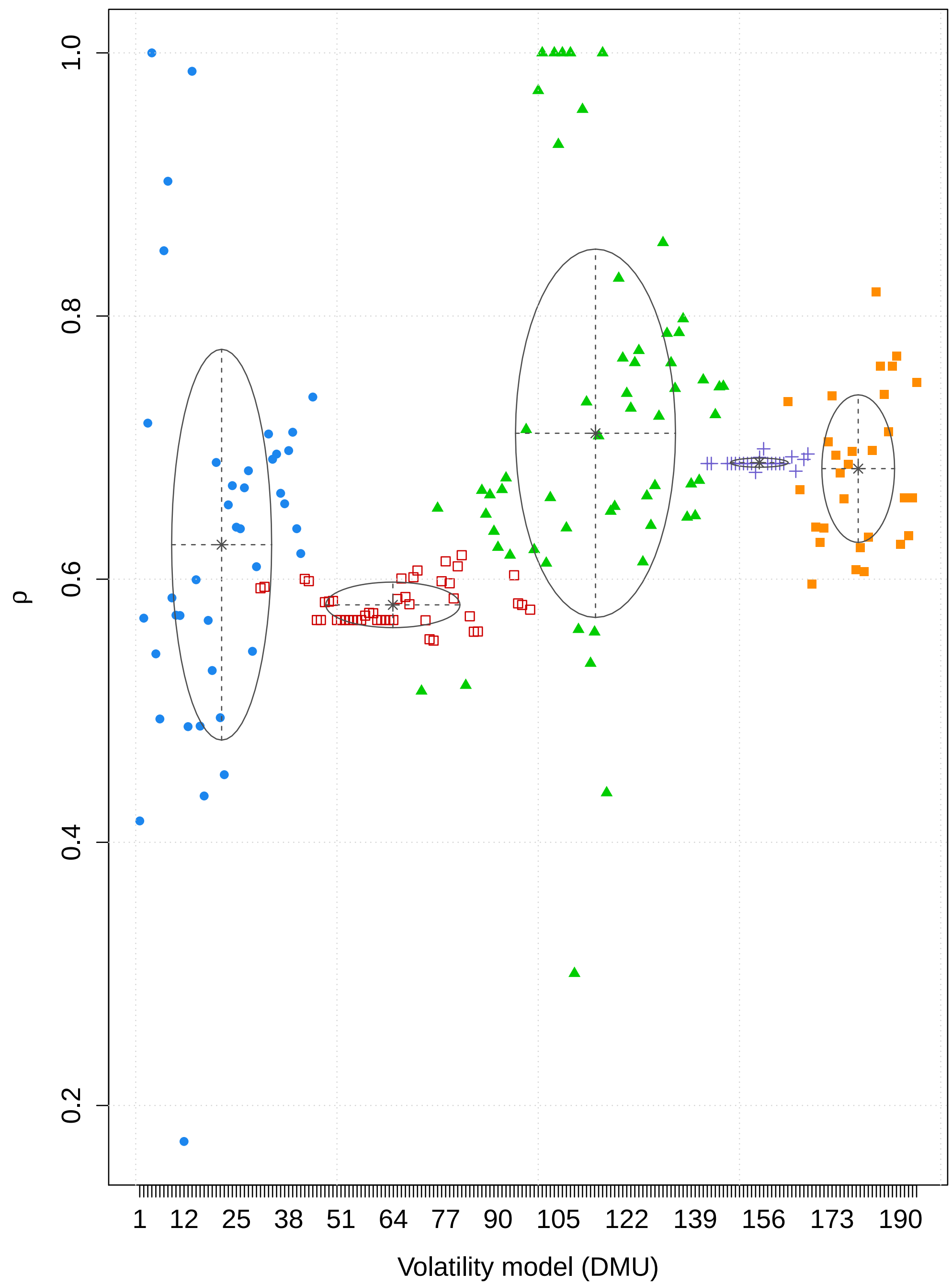

Fig. 6. Cluster analysis plot for ρ

Table 5
Estimation results for β regression models

	Model 1	**Model 2**	**Model 3**	**Model 4**
Time	-0.61^{*}	-	-	-0.58
Log-likelihood	0.07	-	-	-0.02
ESS	0.01^{***}	-	-	0.009^{**}
No.significant coefs	-	0.07^{*}	-	0.07^{**}
No.significant bias terms	-	-1.65^{***}	-	-1.34^{***}
No.significant groups	-	-0.004	-	-0.02
In mean effect(1:yes)	-	-	0.15	0.19^{**}
Conditional distribution(1:skewed)	-	-	0.03	-0.26^{***}
Family GARCH(1:yes)	-	-	-0.30^{***}	0.19
N	198	198	198	198
Mc Fadden's R^2	0.27	0.21	0.01	0.27

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 6
Estimation results for *beta* regression models with interactions

	Model 5	**Model 6**	**Model 7**
Time	-0.66^{*}	-1.47^{***}	-1.02^{*}
Log-likelihood	-0.06	12.12^{***}	-0.022
ESS	0.012^{***}	0.008^{**}	0.016^{***}
No.significant coefs	0.04	0.18^{***}	0.22^{***}
No.significant bias terms	-1.66^{***}	1.64^{***}	-0.84^{**}
No.significant groups	-0.051^{*}	0.014	-4.67×10^{-4}
Log-likelihood × Control	15.34^{***}	-12.34^{***}	2.69
Time × Control	-0.84	0.76	0.39
ESS × Control	-4.73×10^{-5}	$3.28 \times 10^{-4***}$	-0.025^{**}
No.significant coefs × Control	-0.18^{**}	-0.17^{**}	-0.21^{**}
No.significant bias × Control	1.16	-5.20^{***}	0.037^{**}
No.significant groups × Control	0.05	-0.08^{**}	-0.027
N	198	198	198
Mc Fadden's R^2	0.27	0.45	0.30

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

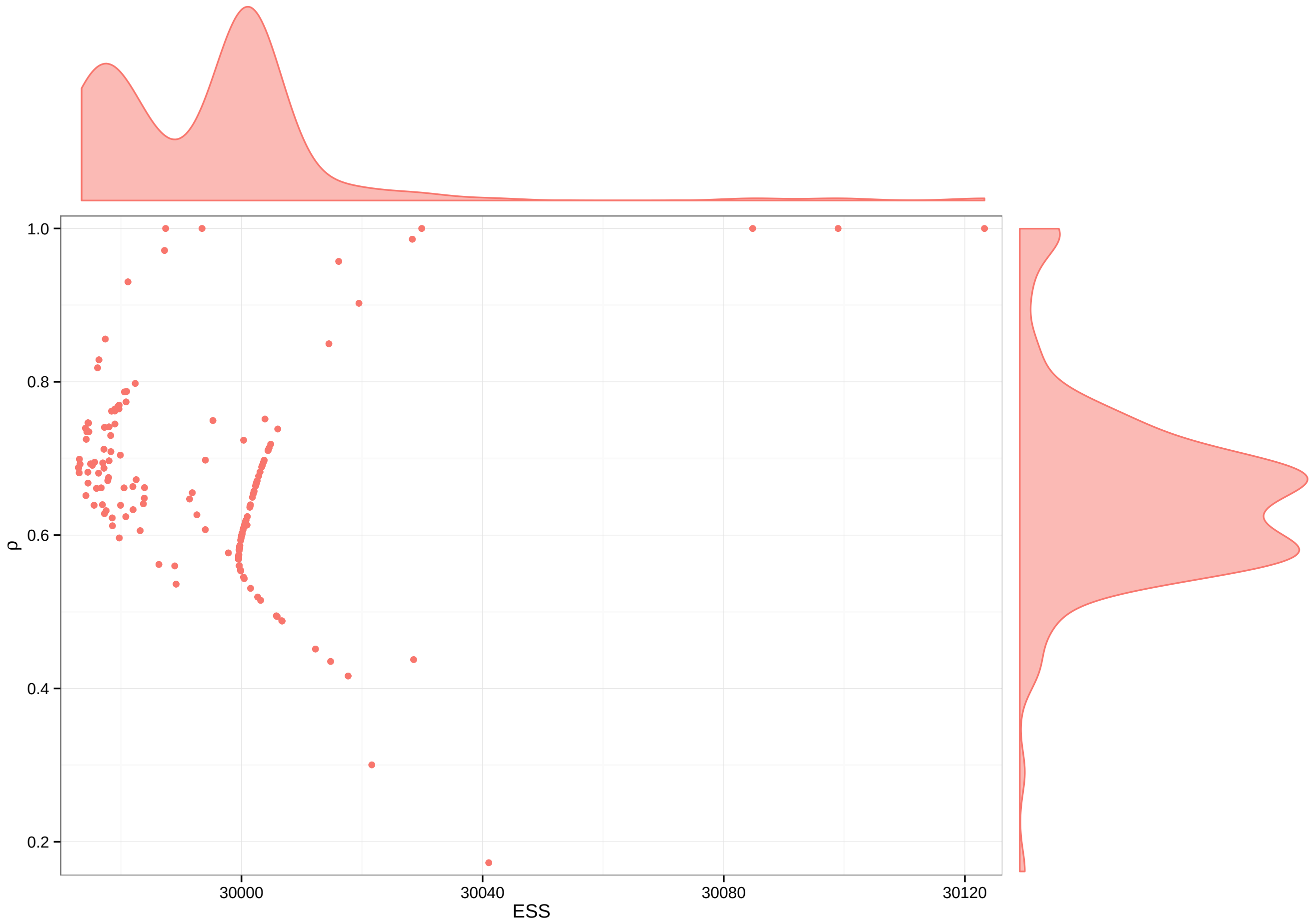

Fig. 7. Joint density distributions of ESS and Efficiency scores (ρ)

Based on the results of second stage analysis derived from 6, *ESS* is statistically significant. The joint density distributions of *ESS* and efficiency scores (ρ) is shown in Figure 7. It can be seen from density plot on top of the figure, that the largest concentration of values of *ESS* variable is reported in the range of $[29972.98, 30000]$. The largest value (mode) equals to 30000, whereas the largest value derived from the density plot on the right-hand side of Figure 7, is 0.57 and the second largest value that is reported is 0.5. It can be also seen from the scatter plot in the center of the figure that the majority of the points lie in the range of $[29972.98, 30000]$ for *ESS* and between $[0.5, 0.8]$ regarding efficiency scores ρ.

7. Conclusions and direction for future research

Global socioeconomic situations around the world affect in a direct or indirect way markets creating uncertainty. The higher the uncertainty, the higher the volatility of the stock prices traded in markets around the world. Based on this fact, a framework to assess and provide the best volatility model based on certain criteria is needed. Several models and methodologies have been proposed to outrank or prioritize forecasting models. One of the methodologies proposed lie in the area of Multi Criteria Decision Making (MCDM) analysis. However, the deficiency of applying this kind of analysis lies in the subjective participation of the decision maker (or the groups of experts), who provides weights on the importance of each criterion. Thus, if different decision makers apply the same technique to the same criteria then another prioritization of volatility models will come up. In the absence of a framework that will serve as a decision support system in order to prioritize the volatility models objectively, an SBM DEA model is proposed. Several combinations of volatility models and conditional distributions have been examined. Values for error measures (namely MAPE, MSE, RMSE, sMAPE, GMAE, MdAPE, RMSPE, RMdAPE) have been derived after having applied all the combinations to *IBM* stock prices data set. Each error measure has been treated as an input while the 198 combinations of volatility models with conditional distributions as DMUs. As the values of the error measures express ratios, the data have been log-normalized using log_10 transformation. Solving the SBM DEA model, minimizing the slack variables for each input, an efficiency score ($\rho \in [0,1]$) is extracted for each DMU; a score of $\rho = 1$ and $s_i^- = 0$ indicates that this DMU is fully efficient whereas if $\rho < 1$ the corresponding DMU is inefficient. From the application to volatility models, the majority of $ARMA(1,0)$ family (models $csGARCH(1,1)$ and $apGARCH$) were prioritized high (1), as the efficiency score was $\rho = 1$. The second stage analysis, shed light to the factors that affect the forecasting efficiency of volatility models. These factors consist of volatility models and attributes, which regard the: (a) appropriateness of fit, (b) dummy variables which indicate the family in which models belong to, (c) skewness of the conditional variance assumed and (d) the existence of an *in - mean* effect. The time needed for a model to be estimated, is consistently negatively associated with the forecasting efficiency of the volatility equations. On the other hand, increasing percent of unexplained total variance leads to higher efficiency. This finding is interesting as it actually means that models, which do not fit data appropriately, tend to exhibit higher efficiency. The value of log - Likelihood does not appear to affect significantly the forecasting efficiency of the volatility models. The number of estimated coefficients, which are statistically significant, seem to have a positive effect on the efficiency, while models whose bias terms are significant, are characterized by decreasing efficiency scores. The number of statistically significant groups of normally distributed standardized residuals is not affecting the efficiency scores. Models that assume an *(*in - mean) effect tend to exhibit increasing efficiency scores, while models in which a skewed unconditional distribution is hypothesized demonstrate lower efficiency scores, compared with those models in which a non - symmetric conditional distribution is assumed. Furthermore, the estimation of β regression models using interaction effects using the dummy variables as control, resulted in some intriguing findings. Values of log - Likelihood for volatility models in which an *(*in - mean) effect is assumed, lead to a greater probability that they perform increasing efficiency scores, while in models in which skewed conditional distribution is used, a potential increase in log - Likelihood results in a higher probability of decreasing efficiency scores. Duration of volatility models' estimation does not seem to affect efficiency scores significantly, in either of the three control variables. On the contrary, there is a consistent pattern of the number of statistically significant parameters. More specifically, the number of statistically significant coefficients of volatility models tend to affect negatively the efficiency scores, independent of the existence of an *(*in - mean) effect, skewed conditional distribution or the fact that the volatility model belongs to $GARCH$ family.

References

[1] J. G. De Gooijer, R. J. Hyndman, 25 years of time series forecasting, International journal of forecasting 22 (3) (2006) 443–473 (2006).

[2] Q. P. Duong, Model selection and ranking: An ahp approach to forecasts combination, Mathematical and Computer Modeling 11 (1988) 282–285 (1988).

[3] M. H. Pesaran, S. Skouras, Decision-based methods for forecast evaluation, A companion to Economic Forecasting (2002) 241–267 (2002).

[4] F. Collopy, J. S. Armstrong, Expert opinions about extrapolation and the mystery of the overlooked discontinuities, International Journal of Forecasting 8 (4) (1992) 575–582 (1992).

[5] J. T. Yokuma, J. S. Armstrong, Beyond accuracy: Comparison of criteria used to select forecasting methods, International Journal of Forecasting 11 (4) (1995) 591–597 (1995).

[6] A. R. Pagan, G. W. Schwert, Alternative models for conditional stock volatility, Journal of Econometrics 45 (1) (1990) 267–290 (1990).

[7] R. F. Engle, A. J. Patton, et al., What good is a volatility model, Quantitative Finance 1 (2) (2001) 237–245 (2001).

[8] P. R. Hansen, A. Lunde, A forecast comparison of volatility models: does anything beat a garch (1, 1)?, Journal of Applied Econometrics 20 (7) (2005) 873–889 (2005).

[9] R. Pezzo, M. Uberti, Approaches to forecasting volatility: Models and their performances for emerging equity markets, Chaos, Solitons & Fractals 29 (3) (2006) 556–565 (2006).

[10] S. Pong, M. B. Shackleton, S. J. Taylor, X. Xu, Forecasting currency volatility: A comparison of implied volatilities and ar (fi) ma models, Journal of Banking & Finance 28 (10) (2004) 2541–2563 (2004).

[11] A. Sfetsos, A comparison of various forecasting techniques applied to mean hourly wind speed time series, Renewable Energy 21 (1) (2000) 23–35 (2000).

[12] A. F. Atiya, S. M. El-Shoura, S. I. Shaheen, M. S. El-Sherif, A comparison between neural-network forecasting techniques-case study: river flow forecasting, Neural Networks, IEEE Transactions on 10 (2) (1999) 402–409 (1999).

[13] I. Moghram, S. Rahman, Analysis and evaluation of five short-term load forecasting techniques, Power Systems, IEEE Transactions on 4 (4) (1989) 1484–1491 (1989).

[14] K. Liu, S. Subbarayan, R. Shoults, M. Manry, C. Kwan, F. Lewis, J. Naccarino, Comparison of very short-term load forecasting techniques, Power Systems, IEEE Transactions on 11 (2) (1996) 877–882 (1996).

[15] J. Yu, R. Meyer, Multivariate stochastic volatility models: Bayesian estimation and model comparison, Econometric Reviews 25 (2-3) (2006) 361–384 (2006).

[16] R. Lord, R. Koekkoek, D. V. Dijk, A comparison of biased simulation schemes for stochastic volatility models, Quantitative Finance 10 (2) (2010) 177–194 (2010).

[17] S.-H. Poon, C. W. Granger, Forecasting volatility in financial markets: A review, Journal of Economic Literature 41 (2) (2003) 478–539 (2003).

[18] A. Charnes, W. W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978) 429–444 (1978).

[19] R. D. Banker, A. Charnes, W. W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, Management Science 30 (9) (1984) 1078–1092 (1984).

[20] W. Liu, D. Zhang, W. Meng, X. Li, F. Xu, A study of dea models without explicit inputs, Omega 39 (5) (2011) 472–480 (2011).

[21] A. Emrouznejad, E. Cabanda, R. Gholami, An alternative measure of the ict-opportunity index, Information & Management 47 (4) (2010) 246–254 (2010).

[22] A. Emrouznejad, B. Rostami-Tabar, K. Petridis, A novel ranking procedure for forecasting approaches using data envelopment analysis, Technological Forecasting and Social Change 111 (2016) 235–243 (2016).

[23] A. Emrouznejad, G. R. Amin, Dea models for ratio data: Convexity consideration, Applied Mathematical Modelling 33 (1) (2009) 486–498 (2009).

[24] K. Tone, A slacks-based measure of efficiency in data envelopment analysis, European journal of operational research 130 (3) (2001) 498–509 (2001).

[25] H. Morita, K. Hirokawa, J. Zhu, A slack-based measure of efficiency in context-dependent data envelopment analysis, Omega 33 (4) (2005) 357–362 (2005).

[26] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, Econometrica: Journal of the Econometric Society (1982) 987–1007 (1982).
[27] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of econometrics 31 (3) (1986) 307–327 (1986).
[28] R. F. Engle, D. M. Lilien, R. P. Robins, Estimating time varying risk premia in the term structure: The arch-m model, Econometrica: Journal of the Econometric Society (1987) 391–407 (1987).
[29] Z. Ding, C. W. Granger, R. F. Engle, A long memory property of stock market returns and a new model, Journal of Empirical Finance 1 (1) (1993) 83–106 (1993).
[30] S. J. Taylor, Modelling financial time series (2007).
[31] G. W. Schwert, Stock volatility and the crash of'87, Review of financial Studies 3 (1) (1990) 77–102 (1990).
[32] L. R. Glosten, R. Jagannathan, D. E. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks, The Journal of Finance 48 (5) (1993) 1779–1801 (1993).
[33] J.-M. Zakoian, Threshold heteroskedastic models, Journal of Economic Dynamics and Control 18 (5) (1994) 931–955 (1994).
[34] A. K. Bera, M. L. Higgins, Arch models: properties, estimation and testing, Journal of Economic Surveys 7 (4) (1993) 305–366 (1993).
[35] G. G. Lee, R. F. Engle, A permanent and transitory component model of stock return volatility, Available at SSRN 5848 (1993).
[36] C. He, T. Teräsvirta, Statistical properties of the asymmetric power arch process (1997).
[37] W. M. Rand, Objective criteria for the evaluation of clustering methods, Journal of the American Statistical association 66 (336) (1971) 846–850 (1971).
[38] L. Hubert, P. Arabie, Comparing partitions, Journal of Classification 2 (1) (1985) 193–218 (1985).
[39] L. Danon, A. Diaz-Guilera, J. Duch, A. Arenas, Comparing community structure identification, Journal of Statistical Mechanics: Theory and Experiment 2005 (09) (2005) P09008 (2005).
[40] F. Cribari-Neto, A. Zeileis, Beta regression in r (2009).
[41] S. Ferrari, F. Cribari-Neto, Beta regression for modelling rates and proportions, Journal of Applied Statistics 31 (7) (2004) 799–815 (2004).
[42] D. McFadden, The revealed preferences of a government bureaucracy: Empirical evidence, The Bell Journal of Economics (1976) 55 –72 (1976).
[43] D. McFadden, Quantitative methods tor analyzing travel behaviour of individuals: Some recent developments (1977).
[44] D. R. Cox, E. J. Snell, Analysis of binary data, Vol. 32, CRC Press, 1989 (1989).
[45] T. O. Kvålseth, Cautionary note about r2, The American Statistician 39 (4) (1985) 279–285 (1985).
[46] X. Liu, H. Fu, Volatility forecasting for interbank offered rate using grey extreme learning machine: The case of china, Chaos, Solitons & Fractals 89 (2016) 249–254 (2016).
[47] R. F. Engle, V. K. Ng, Measuring and testing the impact of news on volatility, The journal of finance 48 (5) (1993) 1749–1778 (1993).
[48] F. C. Palm, Garch models of volatility, Handbook of statistics 14 (1996) 209–240 (1996).
[49] L. Hentschel, All in the family nesting symmetric and asymmetric garch models, Journal of Financial Economics 39 (1) (1995) 71–104 (1995).
[50] J. Wooldridge, Introductory econometrics: A modern approach, Cengage Learning, 2012 (2012).

Appendix

The proposed approach has been applied on a single data set (Historical adjusted close prices of IBM for time period 1995 – 2016), as seen in Figure 2, therefore the ranking extracted for the proposed volatility models concerns only this dataset. The proposed work can be applied to a selection of datasets in order to provide a more robust ranking. Furthermore, as a future direction, composite indicators which would include error measures and would take into account other characteristics of each volatility model could be constructed.

In Table A1, the ranking of the volatility models is presented based on the efficiency derived from SBM model.

Table A1
Ranking of volatility models based on ρ^{*} values.

Volatility Model	SBM Efficiency score	Rank	Volatility Model	SBM Efficiency score	Rank
ARCH(1)-normal	0.416252664	192	ARMA(1,0)-csGARCH(1,1)-normal	0.576798451	149
ARCH(1)-skew normal	0.570361227	156	ARMA(1,0)-csGARCH(1,1)-skew normal	0.622429811	116
ARCH(1)-student t	0.718598849	40	ARMA(1,0)-csGARCH(1,1)-student t	0.971325072	8
ARCH(1)-skew student t	1	1	ARMA(1,0)-csGARCH(1,1)-skew student t	1	1
ARCH(1)-generalized error	0.543240223	180	ARMA(1,0)-csGARCH(1,1)-generalized error	0.612170538	122
ARCH(1)-skew generalized error	0.493762064	186	ARMA(1,0)-csGARCH(1,1)-skew generalized error	0.661960673	90
ARCH(1)-normal inverse	0.849614224	13	ARMA(1,0)-csGARCH(1,1)-normal inverse	1	1
ARCH(1)-generalized hyperbolic	0.902439168	11	ARMA(1,0)-csGARCH(1,1)-generalized hyperbolic	0.930402071	10
ARCH(1)-Johnson's SU	0.585775322	140	ARMA(1,0)-csGARCH(1,1)-Johnson's SU	1	1
ARMA(1,0)-ARCH(1)-normal	0.572601739	152	apARCH(1,1)-normal	0.638924998	106
ARMA(1,0)-ARCH(1)-skew normal	0.572380134	153	apARCH(1,1)-skew normal	1	1
ARMA(1,0)-ARCH(1)-student t	0.172627153	194	apGARCH(1,1)-student t	0.300312457	193
ARMA(1,0)-ARCH(1)-skew student t	0.487924678	188	apARCH(1,1)-skew student t	0.56176432	173
ARMA(1,0)-ARCH(1)-generalized error	0.986009275	7	apARCH(1,1)-generalized error	0.957032382	9
ARMA(1,0)-ARCH(1)-skew generalized error	0.599560585	132	apARCH(1,1)-skew generalized error	0.734673009	36
ARMA(1,0)-ARCH(1)-normal inverse	0.488318341	187	apARCH(1,1)-normal inverse	0.536042039	181
ARMA(1,0)-GARCH(1,1)-generalized hyperbolic	0.435214693	191	apARCH(1,1)-generalized hyperbolic	0.55981384	176
ARMA(1,0)-ARCH(1)-Johnson's SU	0.568662677	172	apARCH(1,1)-Johnson's SU	0.708897722	45
GARCH(1,1)-normal	0.530567459	182	ARMA(1,0)-apARCH(1,1)-normal	1	1
GARCH(1,1)-skew normal	0.688696357	58	ARMA(1,0)-apARCH(1,1)-skew normal	0.437615441	190
GARCH(1,1)-student t	0.494711511	185	ARMA(1,0)-apARCH(1,1)-student t	0.651583281	98
GARCH(1,1)-skew student t	0.451396782	189	ARMA(1,0)-apARCH(1,1)-skew student t	0.65528532	96
GARCH(1,1)-generalized error	0.65646576	95	ARMA(1,0)-apARCH(1,1)-generalized error	0.8287124	14
GARCH(1,1)-skew generalized error	0.671088604	81	ARMA(1,0)-apARCH(1,1)-skew generalized error	0.768036157	21
GARCH(1,1)-normal inverse	0.639510342	104	ARMA(1,0)-apARCH(1,1)-normal inverse	0.741183473	31
GARCH(1,1)-generalized hyperbolic	0.638327836	108	ARMA(1,0)-apARCH(1,1)-generalized hyperbolic	0.729932154	37
GARCH(1,1)-Johnson's SU	0.669444138	83	ARMA(1,0)-apARCH(1,1)-Johnson's SU	0.764541091	22
ARMA(1,0)-GARCH(1,1)-normal	0.682413418	74	TARCH(1,1)-normal	0.773768045	19
ARMA(1,0)-GARCH(1,1)-skew normal	0.545160492	179	TARCH(1,1)-skew normal	0.613112819	121
ARMA(1,0)-GARCH(1,1)-student t	0.609469586	124	TGARCH(1,1)-student t	0.663292074	89
ARMA(1,0)-GARCH(1,1)-skew student t	0.593107874	138	TARCH(1,1)-skew student t	0.640827231	102
ARMA(1,0)-GARCH(1,1)-generalized error	0.594176983	137	TARCH(1,1)-generalized error	0.671085415	82
ARMA(1,0)-GARCH(1,1)-skew generalized error	0.710313125	44	TARCH(1,1)-skew generalized error	0.723809297	39
ARMA(1,0)-GARCH(1,1)-normal inverse	0.691128245	56	TARCH(1,1)-normal inverse	0.855774416	12
ARMA(1,0)-GARCH(1,1)-generalized hyperbolic	0.695068743	52	TARCH(1,1)-generalized hyperbolic	0.78680167	18
ARMA(1,0)-GARCH(1,1)-Johnson's SU	0.665271511	87	TARCH(1,1)-Johnson's SU	0.764446775	23
jgrGARCH(1,1)-normal	0.657314063	94	ARMA(1,0)-TARCH(1,1)-normal	0.744943557	30
jgrGARCH(1,1)-skew normal	0.697704653	49	ARMA(1,0)-TARCH(1,1)-skew normal	0.787391789	17
jgrGARCH(1,1)-student t	0.711677052	43	ARMA(1,0)-TARCH(1,1)-student t	0.797859793	16
jgrGARCH(1,1)-skew student t	0.638361058	107	ARMA(1,0)-TARCH(1,1)-skew student t	0.647111379	101
jgrGARCH(1,1)-generalized error	0.619465518	117	ARMA(1,0)-TARCH(1,1)-generalized error	0.672348733	80
jgrGARCH(1,1)-skew generalized error	0.600191423	131	ARMA(1,0)-TARCH(1,1)-skew generalized error	0.64819414	100
jgrGARCH(1,1)-normal inverse	0.598485014	133	ARMA(1,0)-TARCH(1,1)-normal inverse	0.675095698	79
jgrGARCH(1,1)-generalized hyperbolic	0.738427871	34	ARMA(1,0)-TARCH(1,1)-generalized hyperbolic	0.751448043	26
ARMA(1,0)-jgrGARCH(1,1)-normal	0.568940809	157	AVARCH(1,1)-normal	0.687871527	59
ARMA(1,0)-jgrGARCH(1,1)-skew normal	0.568940809	157	AVARCH(1,1)-skew normal	0.687871527	59
ARMA(1,0)-jgrGARCH(1,1)-student t	0.582529683	145	AVGARCH(1,1)-student t	0.725004287	38
ARMA(1,0)-jgrGARCH(1,1)-skew student t	0.58297213	144	AVARCH(1,1)-skew student t	0.746136118	29
ARMA(1,0)-jgrGARCH(1,1)-generalized error	0.583474777	143	AVARCH(1,1)-generalized error	0.74659859	28
ARMA(1,0)-jgrGARCH(1,1)-skew generalized error	0.568940809	157	AVARCH(1,1)-skew generalized error	0.687871527	59
ARMA(1,0)-jgrGARCH(1,1)-normal inverse	0.568940809	157	AVARCH(1,1)-normal inverse	0.687871527	59
ARMA(1,0)-jgrGARCH(1,1)-generalized hyperbolic	0.568940809	157	AVARCH(1,1)-generalized hyperbolic	0.687871527	59
ARMA(1,0)-jgrGARCH(1,1)-Johnson's SU	0.568940809	157	AVARCH(1,1)-Johnson's SU	0.687871527	59
eGARCH(1,1)-normal	0.568940809	157	ARMA(1,0)-AVARCH(1,1)-normal	0.687871527	59
eGARCH(1,1)-student t	0.568940809	157	ARMA(1,0)-AVARCH(1,1)-student t	0.687871527	59
eGARCH(1,1)-skew student t	0.568940809	157	ARMA(1,0)-AVARCH(1,1)-skew student t	0.687871527	59
eGARCH(1,1)-generalized error	0.572276972	154	ARMA(1,0)-AVARCH(1,1)-generalized error	0.681185715	76
eGARCH(1,1)-skew generalized error	0.574394892	150	ARMA(1,0)-AVARCH(1,1)-skew generalized error	0.692586132	55
eGARCH(1,1)-normal inverse	0.574102625	151	ARMA(1,0)-AVARCH(1,1)-normal inverse	0.699054119	47
eGARCH(1,1)-generalized hyperbolic	0.568940809	157	ARMA(1,0)-AVARCH(1,1)-generalized hyperbolic	0.687871527	59
eGARCH(1,1)-Johnson's SU	0.568940809	157	ARMA(1,0)-AVARCH(1,1)-Johnson's SU	0.687871527	59
ARMA(1,0)-eGARCH(1,1)-normal	0.568940809	157	NAGARCH(1,1)-normal	0.687871527	59
ARMA(1,0)-eGARCH(1,1)-skew normal	0.568940809	157	NAGARCH(1,1)-skew normal	0.687871527	59
ARMA(1,0)-eGARCH(1,1)-student t	0.568940809	157	NAGARCH(1,1)-student t	0.687871527	59
ARMA(1,0)-eGARCH(1,1)-skew student t	0.584951873	142	NAGARCH(1,1)-skew student t	0.734874096	35
ARMA(1,0)-eGARCH(1,1)-generalized error	0.600504299	130	NAGARCH(1,1)-generalized error	0.692975653	54
ARMA(1,0)-eGARCH(1,1)-skew generalized error	0.586440851	139	NAGARCH(1,1)-skew generalized error	0.682076848	75
ARMA(1,0)-eGARCH(1,1)-normal inverse	0.58094634	147	NAGARCH(1,1)-normal inverse	0.667886338	85
ARMA(1,0)-eGARCH(1,1)-generalized hyperbolic	0.601441434	129	NAGARCH(1,1)-generalized hyperbolic	0.69103787	57
ARMA(1,0)-eGARCH(1,1)-Johnson's SU	0.606575042	126	NAGARCH(1,1)-Johnson's SU	0.695121854	51
iGARCH(1,1)-normal	0.514926173	184	ARMA(1,0)-NAGARCH(1,1)-normal	0.596308646	136
iGARCH(1,1)-skew normal	0.568731546	171	ARMA(1,0)-AVARCH(1,1)-skew normal	0.639694868	103
iGARCH(1,1)-student t	0.554333954	177	ARMA(1,0)-NAGARCH(1,1)-student t	0.627984651	112
iGARCH(1,1)-skew student t	0.553302818	178	ARMA(1,0)-AVARCH(1,1)-skew student t	0.638972123	105
iGARCH(1,1)-generalized error	0.653933267	97	ARMA(1,0)-NAGARCH(1,1)-generalized error	0.704347615	46
iGARCH(1,1)-skew generalized error	0.598424383	134	ARMA(1,0)-AVARCH(1,1)-skew generalized error	0.739494253	33
iGARCH(1,1)-normal inverse	0.613553374	120	ARMA(1,0)-NAGARCH(1,1)-normal inverse	0.694196927	53
iGARCH(1,1)-generalized hyperbolic	0.596865796	135	ARMA(1,0)-AVARCH(1,1)-generalized hyperbolic	0.680840876	77
iGARCH(1,1)-Johnson's SU	0.585370395	141	ARMA(1,0)-NAGARCH(1,1)-Johnson's SU	0.660971537	93
ARMA(1,0)-iGARCH(1,1)-normal	0.609760867	123	NGARCH(1,1)-normal	0.687262498	73
ARMA(1,0)-iGARCH(1,1)-skew normal	0.61820696	119	NGARCH(1,1)-skew normal	0.696959399	50
ARMA(1,0)-iGARCH(1,1)-student t	0.519271511	183	NGARCH(1,1)-student t	0.607146225	125
ARMA(1,0)-iGARCH(1,1)-skew student t	0.571788975	155	NGARCH(1,1)-skew student t	0.624043171	115
ARMA(1,0)-iGARCH(1,1)-generalized error	0.560048246	175	NGARCH(1,1)-generalized error	0.605824241	127
ARMA(1,0)-iGARCH(1,1)-skew generalized error	0.560231269	174	NGARCH(1,1)-skew generalized error	0.631843199	111
ARMA(1,0)-iGARCH(1,1)-normal inverse	0.667471663	86	NGARCH(1,1)-normal inverse	0.697853087	48
ARMA(1,0)-iGARCH(1,1)-generalized hyperbolic	0.64937997	99	NGARCH(1,1)-generalized hyperbolic	0.818249977	15
ARMA(1,0)-iGARCH(1,1)-Johnson's SU	0.664078336	88	NGARCH(1,1)-Johnson's SU	0.761725314	25
csGARCH(1,1)-normal	0.636347201	109	ARMA(1,0)-NGARCH(1,1)-normal	0.740555319	32
csGARCH(1,1)-skew normal	0.624178256	114	ARMA(1,0)-AVARCH(1,1)-skew normal	0.711996325	42
csGARCH(1,1)-student t	0.668087969	84	ARMA(1,0)-NGARCH(1,1)-student t	0.761890074	24
csGARCH(1,1)-skew student t	0.676917848	78	ARMA(1,0)-AVARCH(1,1)-skew student t	0.769597439	20
csGARCH(1,1)-generalized error	0.618219442	118	ARMA(1,0)-NGARCH(1,1)-generalized error	0.626477528	113
csGARCH(1,1)-skew generalized error	0.602970292	128	ARMA(1,0)-AVARCH(1,1)-skew generalized error	0.661629878	92
csGARCH(1,1)-normal inverse	0.581598268	146	ARMA(1,0)-NGARCH(1,1)-normal inverse	0.633212836	110
csGARCH(1,1)-generalized hyperbolic	0.580391879	148	ARMA(1,0)-AVARCH(1,1)-generalized hyperbolic	0.66165135	91
csGARCH(1,1)-Johnson's SU	0.71370953	41	ARMA(1,0)-NGARCH(1,1)-Johnson's SU	0.74947053	27