Journal of Economic and Social Research 11(2) 2009, 1-28

A Comparison of Conditional Volatility Estimators for the ISE National 100 Index Returns

Bülent Köksal *

Abstract. We compare more than 1000 different volatility models in terms of their fit to the historical ISE-100 Index data and their forecasting performance of the conditional variance in an out-of-sample setting. Exponential GARCH model of Nelson (1991) with "constant mean, t-distribution, one lag moving average term" specification achieves the best overall performance for modeling the ISE-100 return volatility. The t-distribution seems to characterize the distribution of the heavy tailed returns better than the Gaussian distribution or the generalized error distribution. In terms of forecasting performance, the best models are the ones that can accommodate a leverage effect. Results from fitting the selected exponential GARCH model to the historical ISE-100 return data indicates that the return volatility reacts to bad news 24% more than they react to good news as a result of a one standard deviation shock to the returns. As the magnitude of shock increases, the asymmetry becomes larger.

JEL Classification Codes: C12, C13, C15, C22, C52, C53, G10, G15, G17.
Keywords: GARCH, Volatility Models, Istanbul Stock Exchange, ISE-100.

1. Introduction

Modeling and forecasting time-varying financial market volatility are important for investors who are interested in the forecast of the variance of a series over the holding period for calculating measures of risk, pricing

---

* Department of Economics, Fatih University, Buyukcekmece, 34500, Istanbul, Turkey. E-mail: bkoksal@fatih.edu.tr . We would like to thank Istanbul Stock Exchange for providing the ISE-100 Index data. Any remaining errors are our own.

derivatives, and hedging. The long-run forecast of the conditional variance would be unimportant for these investors who hold the asset for a certain period only. In a seminal paper, Engle (1982) shows how to model the conditional variance of a time series. Bollerslev (1986) generalizes Engle's work by allowing the conditional variance to be an ARMA process. The literature continued to grow by extending these works to the case of vector processes. Early articles on multivariate extensions are Engle, Granger and Kraft (1986), Diebold and Nerlove (1989), and Bollerslev, Engle and Wooldridge (1988). See Poon and Granger (2003) for an extensive survey. Bollerslev (2008) provides a comprehensive list of different volatility models discussed in the literature.

Our aim in this paper is to determine the best volatility model for modeling the behavior of the Istanbul Stock Exchange-100 (ISE-100) National Index returns. We compare more than 1000 different volatility models in terms of their fit to the historical ISE-100 Index data and their forecasting performance of the conditional variance in an out-of-sample setting. We use the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to evaluate the models' fit. To assess the forecasting performance, we use four different loss functions where the predicted variances are compared to the realized variance.

The exponential GARCH (EGARCH) model of Nelson (1991) has the best overall performance for modeling the ISE-100 return volatility. Specifically, the EGARCH(2,2) model with a "constant mean, t-distribution, one lag moving average term" specification has the best fit and forecasting performance when compared to the other models.

Other major findings are that the widely used GARCH(1,1) model performs well but is outperformed by more sophisticated models that allow for leverage effect. The loss functions select "zero-mean, generalized error distribution, moving average term with one lag" specification, and the information criteria select "constant-mean, t-distribution" specification as the best GARCH(1,1) specification. Overall, the t-distribution seems to characterize the distribution of the heavy tailed returns better than the Gaussian distribution or the generalized error distribution. There are no significant differences between the performances of the three specifications for the expected value of the returns (zero mean, constant mean, and GARCH-in-mean). Models that allow for leverage effects are slightly superior to the models that do not. In terms of forecasting performance, the best models are the ones that can accommodate a leverage effect.

When we fit the selected EGARCH(2,2) model to the historical ISE-100 return data, we find that the return volatility reacts to bad news 24% more than they react to good news as a result of a one standard deviation shock to the returns. As the magnitude of the shock increases, the asymmetry becomes larger. A three standard deviation shock to the returns produces a 91% volatility difference between bad and good news.

The paper is organized as follows. Section 2 describes the volatility models that we consider, while Section 3 describes data and empirical methodology. In Section 4, we discuss our empirical findings and determine the best volatility model. In Section 5, we present the results from fitting the best model to ISE-100 historical returns and then conclude in Section 6.

2. GARCH Models

We compare the performance of different GARCH models in terms of their ability to model the mean and volatility of compounded returns, defined as $r_t = \log(ISE100_t) - \log(ISE100_{t-1})$ where $ISE100_t$ is the closing value of the daily ISE-100 index at time *t*. The basic model that GARCH type models fit is the following:

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= \sigma_t v_t \\ Var(\varepsilon_t \mid I_{t-1}) &= \sigma_t^2 \end{aligned} \tag{1}$$

where $\mu_t = E(r_t \mid I_{t-1})$, I_{t-1} denotes the information set available at time *t*-1 and v_t is a sequence of iid random variables with mean 0 and variance 1. v_t is generally assumed to follow standard normal, standardized Student-t distribution or generalized error distribution. We also define the standardized value of ε_t as $z_t = \varepsilon_t / \sigma_t$.

Several different volatility models are developed in the literature by assuming different structures for σ_t^2.[1] For example, if we specify

[1] See Bollerslev (2008) for a comprehensive list of different GARCH models.

$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$, we get the ARCH model developed by Engle (1982), and if we specify $\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$, we get the familiar and widely used GARCH model developed by Bollerslev (1986) as a generalization of Engle's ARCH model. In this paper we employ different specifications for μ_t, ε_t, σ_t^2 and the distribution of v_t to determine which specification best characterizes the ISE-100 index returns.

We employ three different specifications for the conditional mean: $\mu_t = 0$, $\mu_t = \mu_0$, and $\mu_t = \mu_0 + \mu_1 \sigma_{t-1}^2$. The latter specification is the ARCH-M model of Engle et al. (1987) and allows for the return process to depend on its own conditional variance which is a measure of asset's riskiness. We use three specifications for the distribution of v_t: Gaussian, Student-t distribution and generalized error distribution. As is well-known, financial return data are usually characterized better by heavy-tailed distributions such as t distribution. We also use the generalized error distribution which is a parametric family of symmetric distributions that includes all normal and Laplace distributions as special cases. Generalized error distribution has a shape parameter. When this parameter is equal to two, the distribution becomes the normal distribution and when it is less (greater) than 2, the distribution has fatter (thinner) tails than the normal distribution.[2] Finally, we allow the error to follow a moving average process with one lag, in addition to "no moving average" case.

Different GARCH type volatility models are obtained by using different specifications for σ_t^2. There have been quite a few number of models discussed in the literature. In his "Glossary to ARCH (GARCH)", Bollerslev (2008) lists more than 100 entries. Therefore, it is necessary to be selective in any work comparing these models. In addition, some of these models are just special cases of other more general ones. Accordingly, we select a number of common models that are listed in Table 1. The use of acronyms for different GARCH models is not fully consistent in the existing literature. Therefore, rather than referring to a model by a specific name, we list the specifications that we use for the conditional variance as well as the

[2] See, for example, Nelson (1991) for the density function of the generalized error distribution.

papers that first introduce these models in Table 1. Following Hansen and Lunde (2005), we estimate the volatility models in Table 1 by using four combinations of the lag length parameters $p, q = 1, 2$. Overall, we estimate 1116 different models.

Table 1. Specifications for the Conditional Variance

ARCH: Engle (1982)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$
GARCH: Bollerslev (1986)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
IGARCH: Engle and Bollerslev (1986), Nelson (1990)	$\sigma_t^2 = \omega + \varepsilon_{t-1}^2 + \sum_{i=2}^{q} \alpha_i (\varepsilon_{t-i}^2 - \varepsilon_{t-1}^2) + \sum_{j=1}^{p} \beta_j (\sigma_{t-j}^2 - \varepsilon_{t-1}^2)$
Taylor/Schwert: Taylor (1986), Schwert (1989)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \lvert \varepsilon_{t-i} \rvert + \sum_{j=1}^{p} \beta_j \sigma_{t-j}$
SAGARCH (Simple Asymmetric GARCH): Engle (1990)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} [\alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}] + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
TGARCH (Threshold GARCH): Zakoian (1994)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \left[\alpha_i \lvert \varepsilon_{t-i} \rvert + \gamma_i \lvert \varepsilon_{t-i}^{+} \rvert \right] + \sum_{j=1}^{p} \beta_j \sigma_{t-j}$
GJR GARCH: Glosten et al. (1993)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} [\alpha_i + \gamma_i I_{(\varepsilon_{t-i} > 0)}] \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
GJR power GARCH: Extension of GJR GARCH	$\sigma_t^{\varphi} = \omega + \sum_{i=1}^{q} [\alpha_i + \gamma_i I_{(\varepsilon_{t-i} > 0)}] \varepsilon_{t-i}^{\varphi} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\varphi}$
EGARCH (Exponential GARCH): Nelson (1991)	$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q} [\alpha_i z_{t-i} + \gamma_i (\lvert z_{t-i} \rvert - \sqrt{2/\pi})] + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2)$
PGARCH (Power GARCH): Higgins and Bera(1992)	$\sigma_t^{\varphi} = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^{\varphi} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\varphi}$

NGARCH (Nonlinear GARCH): Special case of NPGARCH below.	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i} - \kappa_i)^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
NGARCHK (Nonlinear GARCH with one shift): Special case of NPGARCHK below.	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i} - \kappa)^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
AGARCH (Asymmetric GARCH): Special Case of A-PGARCH below.	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i (\mid \varepsilon_{t-i} \mid + \gamma_i \varepsilon_{t-i})^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$
A-PGARCH (Asymmetric Power GARCH): Ding et al. (1993)	$\sigma_t^\varphi = \omega + \sum_{i=1}^{q} \alpha_i (\mid \varepsilon_{t-i} \mid + \gamma_i \varepsilon_{t-i})^\varphi + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^\varphi$
NPGARCH (Nonlinear Power GARCH): A more general form of NPGARCHK below.	$\sigma_t^\varphi = \omega + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i} - \kappa_i)^\varphi + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^\varphi$
NPGARCHK (Nonlinear Power GARCH with one shift): Bollerslev et al. (1994)	$\sigma_t^\varphi = \omega + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i} - \kappa)^\varphi + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^\varphi$

The literature for modeling and forecasting time-varying financial market volatility extends the original work by Engel in several ways to better characterize the behavior of the returns. For example, an interesting behavior of asset prices is that new information seems to have an asymmetric effect on volatility, called the *leverage effect*. Glosten et al. (1993), Zakoian (1994) and Nelson (1991) develop models that allow for the asymmetric effect of the news. As a further example, Ding et al. (1993) and Higgins and Bera (1992) model σ_t^φ, rather than σ_t^2, where φ is a parameter to be estimated. Our purpose is to determine the volatility model that best characterizes the volatility of the ISE-100 returns. Therefore, we do not describe all the alternative models that we consider in detail. In Section 5, we discuss the model selected by our methodology and fit this model to ISE-100 return data.

3. Data and Empirical Methodology

We use daily ISE-100 National Index data from 5 January 1998 to 31 December 2008, and calculate the compounded returns as $r_t = \log(ISE100_t) - \log(ISE100_{t-1})$. The sample consists of 2731 daily returns. We then fit the set of GARCH models described above to these data in order to compare the models' performances.

Standard model evaluation criteria, such as Akaike information criterion (AIC) and Bayesian information criterion (BIC), are widely used to compare the performance of different GARCH models. These criteria penalize the decrease in the degrees of freedom when more variables are added and defined as: $AIC = -2 \cdot \log(likelihood) + 2 \cdot k$ and $BIC = -2 \cdot \log(likelihood) + \log(N) \cdot k$, where k is model degrees of freedom and N is the number of observations. These criteria evaluate models based on their fit to the historical data. Since BIC imposes a larger penalty on adding more parameters to the model, it will select a more parsimonious model than AIC does.

If the purpose of the study is to evaluate the behavior of a historical time series, using these methods would be fine. But, if the aim is to select the model that has best forecasting performance, model goodness of fit will be less important. It is common to evaluate the forecasting performance of a model by using the one-step-ahead forecast errors. In this approach, the sample is split into an estimation period and an evaluation period. The model is fit to the data in the estimation period, and then it is used to forecast the values in the evaluation period. The forecast errors are calculated as the difference between actual values and predicted values. Finally, a loss function such as mean squared error or mean absolute error is calculated to compare the forecasting performance of different models. The model with the minimum loss is the best model in terms of forecasting performance.

In this paper, we evaluate different GARCH models by using both approaches above. First we fit the GARCH models to all sample data and use AIC and BIC to compare model performance in terms of their *fit* to historical data. To use the second approach, we first calculate daily realized variance by calculating the variance of intraday 5-minute returns.[3] Then we

[3] See Andersen et al. (2003) for an extensive discussion of volatility measurement.

estimate each GARCH model by using the first 2480 observations and use the parameter estimates to forecast the realized variance for each day in the evaluation period (last 251 observations).[4] Then we compare the predicted variance from each model to the realized variance and calculate the values of the following 4 different loss functions based on squared and absolute errors:[5]

$$MSE_1 = \frac{1}{n}\sum_{i=1}^{n}(\sigma_t - h_t)^2 \qquad MSE_2 = \frac{1}{n}\sum_{i=1}^{n}(\sigma_t^2 - h_t^2)^2$$

$$MAE_1 = \frac{1}{n}\sum_{i=1}^{n}|\sigma_t - h_t| \qquad MAE_2 = \frac{1}{n}\sum_{i=1}^{n}|\sigma_t^2 - h_t^2|$$

Once the values of the loss functions are calculated, it is possible to order the models according to their losses. The model with the minimum loss is the best model. Finally, we use the *superior predictive ability (SPA)* test of Hansen (2005) to test if the models selected by the “minimum loss” criterion outperforms the widely used GARCH(1,1) model. Hansen’s test compares different models to a benchmark model in terms of the expected loss and makes it possible to decide if these models outperform the benchmark model.[6] The null hypothesis is that the benchmark model is as good as any other model in terms of expected loss. A significance test statistic for a specific model indicates that this model outperforms the benchmark model.

When we use all sample data to compare the fit of models, the likelihoods of 2.69% of the 1116 models that we consider did not converge. Similarly, when we estimate the models for the forecast comparison, the likelihoods of 0.9% of the full set did not converge. This is typical since ARCH model likelihoods are often difficult to maximize.

---

[4] Estimation sample covers the period from 5 January 1998 to 31 December 2007 and the out-of-sample evaluation period covers 2 January 2008 to 31 December 2008.

[5] As Hansen and Lunde (2005) state, it is not clear which loss function is best for evaluating different volatility models. We prefer using loss functions based on mean squared and absolute errors.

[6] We implement the SPA test by using Ox Console[tm] version 5.10 (see Doornik (2007)) which is free for academic use. Detailed information for implementation of SPA test as well as a sample code can be found on Hansen’s website.

4. Model Comparison Results

Table 2, Panels A and B report the first 15 models selected by the information criteria.[7] AIC results reported in Panel A show that constant mean model and t-distribution seem to better characterize the return behavior. In addition, including more lags to the volatility models improves the overall fit. The best two models selected by AIC is EGARCH(2,2) model of Nelson (1991) with and without a moving average term, which allows for leverage effects. The third volatility model is NPGARCHK model discussed in Bollerslev et al. (1994). This is a very general model and the volatility equation has a shift parameter κ, which allows the minimum conditional variance to occur at a value of lagged innovations other than zero.

Table 2, Panel B reports the models selected by BIC. As discussed earlier, BIC imposes a larger penalty on adding more parameters to the model and selects a more parsimonious model than AIC. This is evident in Panel B as the best models are the ones with fewer lagged parameters, no moving average terms, and zero mean specification. In this case, the best model is simple asymmetric GARCH (SAGARCH) model of Engle (1990), which again allows for the asymmetric effect of positive and negative innovations. The second and third models are nonlinear GARCH models with shift parameters. NGARCH and NGARCHK are special cases of NPGARCH and NPGARCHK; common to all these models is the characteristic of allowing the minimum conditional variance to occur at a value or values of lagged innovations other than zero.

Hansen and Lunde (2005) find that a GARCH(1,1) model is not outperformed by more complicated models in their analysis of exchange rates, whereas, the GARCH(1,1) is inferior to models that allow for leverage effects in their analysis of IBM returns. To compare our results to theirs, we also look at the performance of a GARCH(1,1) model, by comparing its AIC and BIC values to the rest of the models. Figure 1 displays the distribution of information criteria for all 1086 models for which the likelihood function has a maximum. The x-axis is the positive value of the information criterion, so that larger values imply better models. The dashed line displays the location

[7] 15 models form a very small portion of all models considered in the paper. Below, we report the distribution of AIC and BIC values for all models and determine if some characteristics of the models are better than others. We also discuss the performance of the GARCH(1,1) model, which has been used extensively in the literature.

of the best performing GARCH(1,1) model according to the information criteria, which has “constant mean” structure and the error follows the t-distribution. It is clear from Figure 1 that GARCH(1,1) is outperformed by many more sophisticated models.

Table 2. Models Selected by Information Criteria

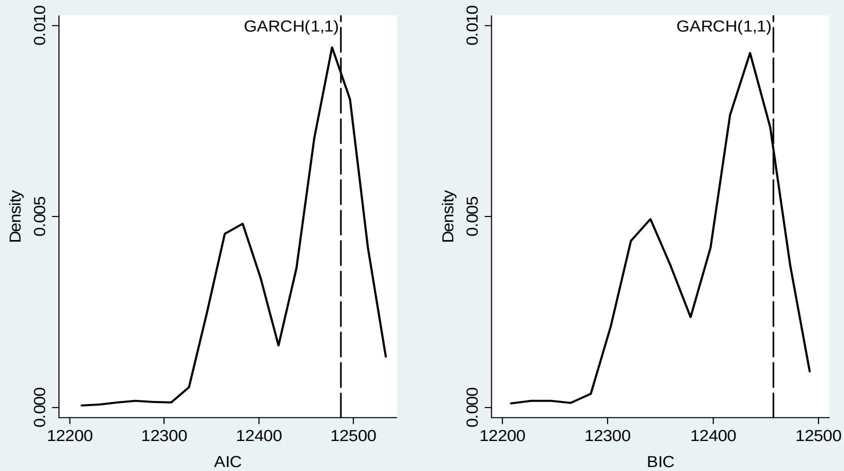
Panel A. Models Selected by AIC

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage
1	EGARCH	2	2	Constant Mean	Yes	t	Yes
2	EGARCH	2	2	Zero Mean	No	t	Yes
3	NPGARCHK	2	2	Zero Mean	No	GED	No
4	GARCH	2	2	GARCH in Mean	No	t	No
5	NGARCH	2	1	Constant Mean	Yes	t	No
6	NGARCH	2	1	Constant Mean	No	t	No
7	SAGARCH	1	1	Constant Mean	Yes	t	Yes
8	NGARCH	1	1	Constant Mean	Yes	t	No
9	NGARCH	1	1	Constant Mean	Yes	t	No
10	SAGARCH	1	1	Constant Mean	No	t	Yes
11	NGARCH	1	1	Constant Mean	No	t	No
12	NGARCH	1	1	Constant Mean	No	t	No
13	NGARCH	2	1	Constant Mean	No	t	No
14	NGARCH	2	1	Constant Mean	Yes	t	No
15	NPGARCHK	1	1	Constant Mean	No	t	No

Panel B. Models Selected by BIC

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage
1	SAGARCH	1	1	Zero Mean	No	t	Yes
2	NGARCH	1	1	Zero Mean	No	t	No
3	NGARCHK	1	1	Zero Mean	No	t	No
4	SAGARCH	1	1	Constant Mean	No	t	Yes
5	NGARCHK	1	1	Constant Mean	No	t	No
6	NGARCH	1	1	Constant Mean	No	t	No
7	GJR GARCH	1	1	Zero Mean	No	t	Yes
8	AGARCH	1	1	Zero Mean	No	t	Yes
9	SAGARCH	1	1	Zero Mean	Yes	t	Yes
10	NGARCH	1	1	Zero Mean	Yes	t	No
11	NGARCHK	1	1	Zero Mean	Yes	t	No
12	NGARCH	2	1	Zero Mean	No	t	No
13	EGARCH	1	1	Zero Mean	No	t	Yes
14	NGARCHK	2	1	Zero Mean	No	t	No
15	NPGARCH	1	1	Zero Mean	No	t	No

Figure 1. Model Performance Based on Information Criteria: All Models

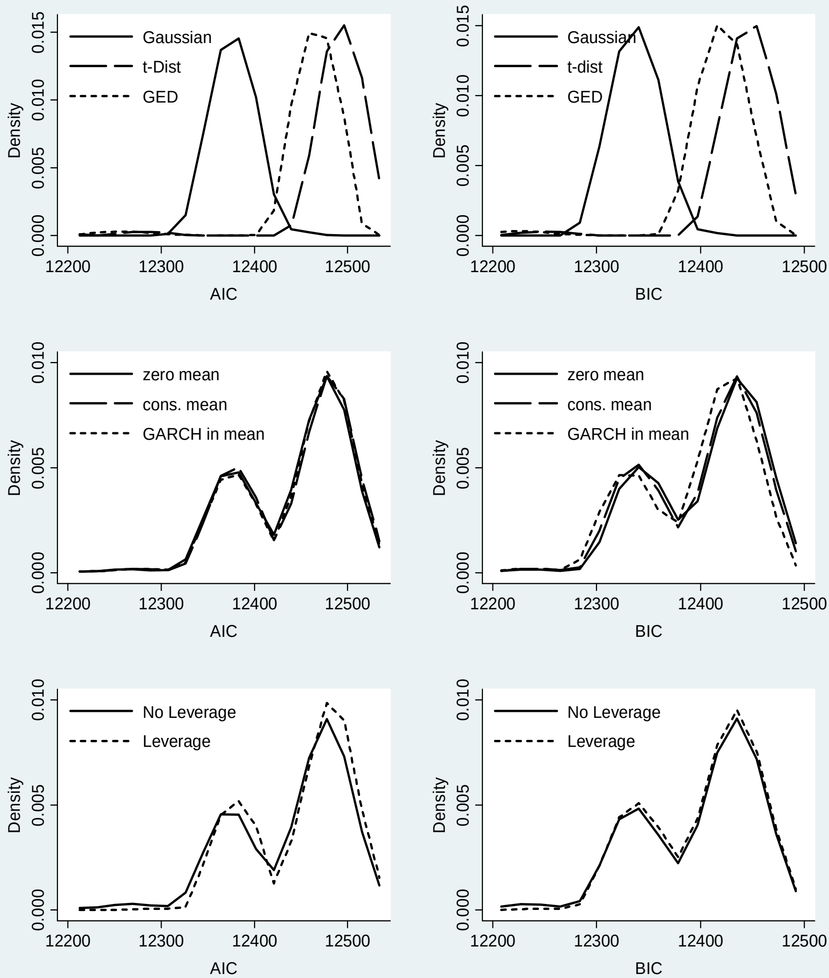


These graphs display density estimates of the information criteria for the models considered. The x-axis is the positive value of the information criterion, so that larger values imply better models. Vertical dashed line displays the location of the best performing GARCH(1,1) model.

We also report the distribution of the information criteria based on the form of the mean equation, the distribution of v_t, and whether the model allows for leverage effects. The first two graphs in Figure 2 show that the t-distribution seems to characterize the behavior of the heavy-tailed returns better than the normal distribution. There are no major differences between the three mean specifications. The distribution of the information criteria for "GARCH-in-mean", "mean only" and "no mean" specifications are almost identical. Finally, according to the AIC, models that allow for leverage effects are slightly superior to those that do not.

As discussed earlier, one's purpose for estimating a volatility model determines the model selection technique that will be used. AIC and BIC are useful for selecting the best model to evaluate the behavior of a historical time series. When the purpose is to *forecast* volatility, forecasting performances of the models should be compared by using possibly different loss functions. We use two forms of the mean squared error and mean absolute deviation criterion, described in the "data and empirical methodology" section above.

Figure 2. Model Performance Based on Information Criteria: Subcategories



These graphs display density estimates of the information criteria for the models considered. The x-axis is the positive values of the information criteria, so that larger values imply better models.

Table 3, Panels A to D report the best 15 models selected by the loss functions that we consider, as well as the SPA test statistics that test whether the selected model outperforms the best GARCH(1,1) model. According to the MSE_1 and MAE_2 results in panels A and D of Table 3, the best model is the EGARCH(2,2) model which has “constant mean, t-distribution, one lag moving average term” specification. Recall that this same model was also selected by AIC as the best model (See Table 2). The values of the SPA test statistics for the EGARCH(2,2) model in Panels A and D are 4.8865 and 5.9349, which are significant at the 1% level, indicating that this model significantly outperforms the GARCH(1,1) model. Inspection of Table 3 shows that different versions of the EGARCH(2,2) model are selected among the best models. Another result from Table 3 is that “constant mean” specification seems to be the best approach for modeling the expected value of the returns. Finally, all loss functions select the models that can accommodate a leverage effect; a result which implies that in order to examine the behavior of ISE-100 returns, a model that allows for a leverage effect should be used.[8]

We now turn to the density estimates of the loss functions for all models to see if certain specifications are better than others overall. We first look at the position of the best performing GARCH(1,1), which has “zero-mean, generalized error distribution, moving average term with one lag” specification. Figure 3 shows that GARCH(1,1) model performs quite well and is outperformed by only a small number of other more general models in terms of forecasting performance.

Figures 4 to 6 report the distribution of the loss functions according to different types of models. There are no major differences between models according to the error distribution and the mean specification; overall, consistent with the results in Figure 2, models that allow for leverage effects seem to be slightly superior to models that do not.

[8] Nelson’s EGARCH model uses the level of standardized value of the error term, leading to a more natural interpretation of the size and persistence of shocks, since the standardized value is a unit-free measure.

Table 3. Models Selected by the Loss Functions

Panel A. Models Selected by MSE_1

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage	SPA Test
1	EGARCH	2	2	Constant Mean	Yes	t	Yes	4.8865 ***
2	EGARCH	2	2	Constant Mean	No	t	Yes	4.8641 ***
3	EGARCH	2	2	Constant Mean	Yes	GED	Yes	4.8778 ***
4	EGARCH	2	2	Constant Mean	No	GED	Yes	4.8554 ***
5	GJR GARCH	2	2	Zero Mean	No	GED	Yes	5.2952 ***
6	GJR GARCH	2	2	GARCH in Mean	Yes	Gaussian	Yes	5.4913 ***
7	GJR GARCH	2	2	Zero Mean	Yes	Gaussian	Yes	5.4285 ***
8	EGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	4.9242 ***
9	AGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	5.4685 ***
10	AGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	5.5441 ***
11	AGARCH	2	2	Constant Mean	No	Gaussian	Yes	5.5463 ***
12	SAGARCH	2	2	Constant Mean	No	Gaussian	Yes	5.2943 ***
13	SAGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	5.3144 ***
14	SAGARCH	2	2	Zero Mean	No	Gaussian	Yes	5.2178 ***
15	GJR P. GARCH	2	2	GARCH in Mean	Yes	GED	Yes	4.5219 ***

***, ** and * denote significance of the SPA test statistic at the 1%, 5%, and 10% levels, respectively.

Panel B. Models Selected by MSE_2

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage	SPA Test
1	GJR GARCH	2	2	GARCH in Mean	Yes	Gaussian	Yes	3.4552 **
2	GJR GARCH	2	2	Zero Mean	Yes	Gaussian	Yes	3.4350 **
3	GJR GARCH	2	2	Zero Mean	No	GED	Yes	3.2915 **
4	EGARCH	2	2	Constant Mean	Yes	t	Yes	2.5102
5	AGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	3.4479 **
6	EGARCH	2	2	Constant Mean	No	t	Yes	2.4969
7	AGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	3.4057 **
8	EGARCH	2	2	Constant Mean	Yes	GED	Yes	2.4956
9	AGARCH	2	2	Constant Mean	No	Gaussian	Yes	3.4130 **
10	EGARCH	2	2	Constant Mean	No	GED	Yes	2.4832
11	EGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	2.6358 *
12	SAGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	2.6079 *
13	SAGARCH	2	2	Constant Mean	No	Gaussian	Yes	2.5867
14	GJR P. GARCH	2	2	GARCH in Mean	Yes	GED	Yes	2.4831
15	EGARCH	1	1	Zero Mean	No	Gaussian	Yes	2.4831

***, ** and * denote significance of the SPA test statistic at the 1%, 5%, and 10% levels, respectively.

Table 3. (cont'd)

Panel C. Models Selected by MAE_1

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage	SPA Test
1	GJR P. GARCH	2	2	GARCH in Mean	Yes	GED	Yes	6.7602 ***
2	EGARCH	1	1	Zero Mean	No	Gaussian	Yes	6.7602 ***
3	GJR GARCH	2	2	Zero Mean	No	GED	Yes	6.9769 ***
4	GJR GARCH	2	2	GARCH in Mean	Yes	Gaussian	Yes	7.2326 ***
5	GJR P. GARCH	2	2	Constant Mean	Yes	GED	Yes	6.6626 ***
6	EGARCH	2	2	Constant Mean	Yes	t	Yes	6.2684 ***
7	GJR P. GARCH	2	2	Zero Mean	Yes	GED	Yes	6.6211 ***
8	GJR GARCH	2	2	Zero Mean	Yes	Gaussian	Yes	7.2300 ***
9	SAGARCH	2	2	Constant Mean	No	Gaussian	Yes	6.8173 ***
10	SAGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	6.8335 ***
11	EGARCH	2	2	Constant Mean	No	t	Yes	6.2138 ***
12	AGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	7.2488 ***
13	AGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	7.2872 ***
14	GJR P. GARCH	2	2	GARCH in Mean	Yes	Gaussian	Yes	6.6775 ***
15	AGARCH	2	2	Constant Mean	No	Gaussian	Yes	7.2265 ***

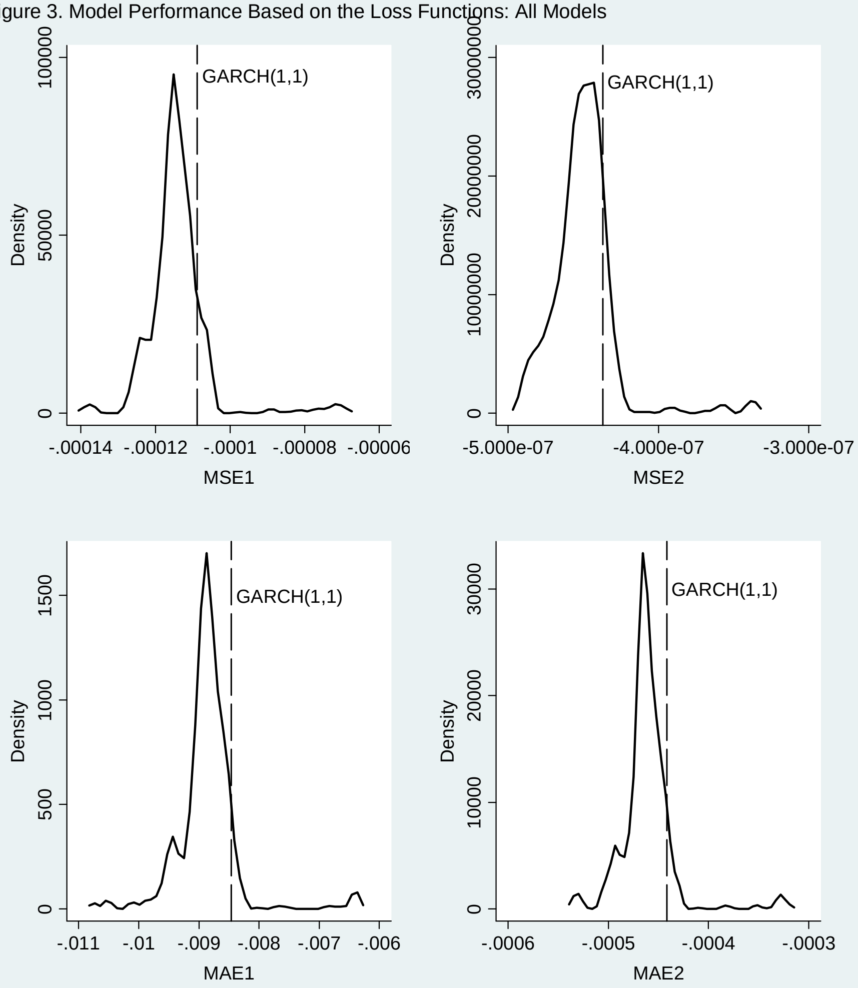
***, ** and * denote significance of the SPA test statistic at the 1%, 5%, and 10% levels, respectively.

Panel D. Models Selected by MAE_2

No	Model	q	p	Mean Model	MA Term	Distribution	Leverage	SPA Test
1	EGARCH	2	2	Constant Mean	Yes	t	Yes	5.9349 ***
2	EGARCH	2	2	Constant Mean	No	t	Yes	5.9048 ***
3	EGARCH	2	2	Constant Mean	Yes	GED	Yes	5.9184 ***
4	EGARCH	2	2	Constant Mean	No	GED	Yes	5.8863 ***
5	GJR GARCH	2	2	Zero Mean	No	GED	Yes	6.2371 ***
6	GJR GARCH	2	2	GARCH in Mean	Yes	Gaussian	Yes	6.4914 ***
7	GJR GARCH	2	2	Zero Mean	Yes	Gaussian	Yes	6.4458 ***
8	SAGARCH	2	2	Constant Mean	No	Gaussian	Yes	6.4868 ***
9	SAGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	6.4970 ***
10	AGARCH	2	2	Zero Mean	Yes	Gaussian	Yes	6.4996 ***
11	AGARCH	2	2	GARCH in Mean	No	Gaussian	Yes	6.5272 ***
12	AGARCH	2	2	Constant Mean	No	Gaussian	Yes	6.5119 ***
13	GJR P. GARCH	2	2	GARCH in Mean	Yes	GED	Yes	5.7722 ***
14	EGARCH	1	1	Zero Mean	No	Gaussian	Yes	5.7722 ***
15	SAGARCH	2	2	Zero Mean	No	Gaussian	Yes	6.3817 ***

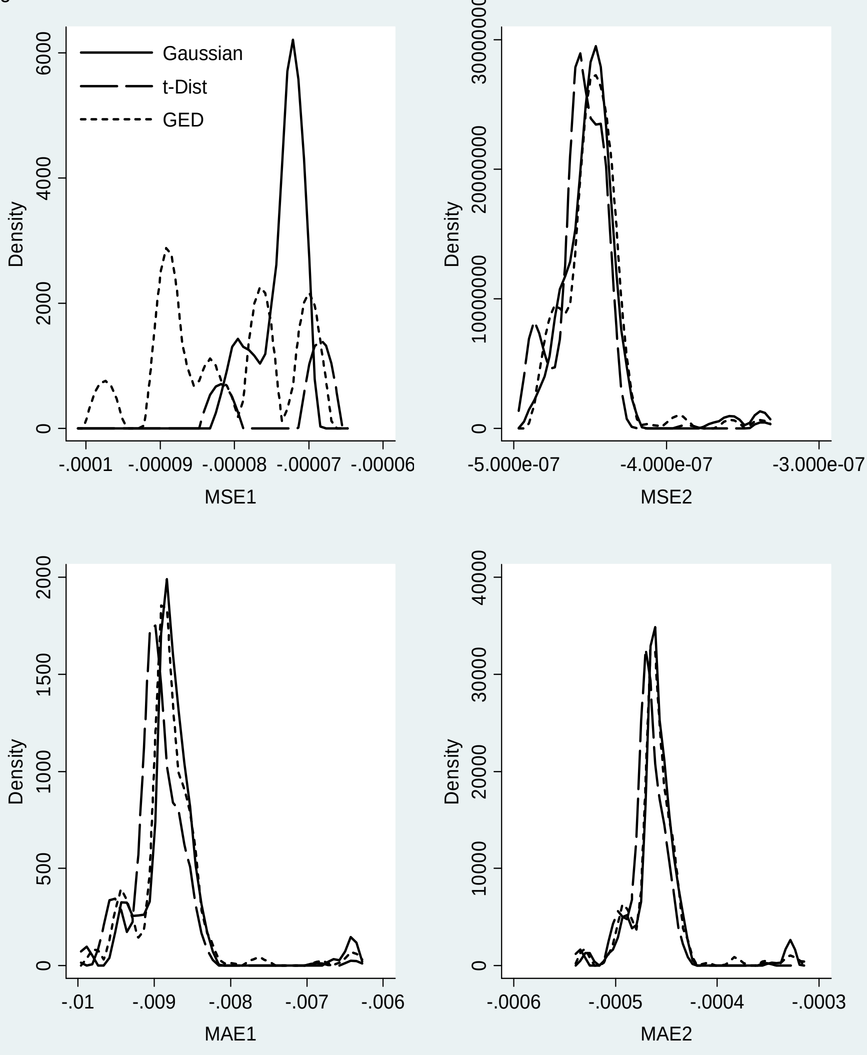
***, ** and * denote significance of the SPA test statistic at the 1%, 5%, and 10% levels, respectively.

Figure 3. Model Performance Based on the Loss Functions: All Models



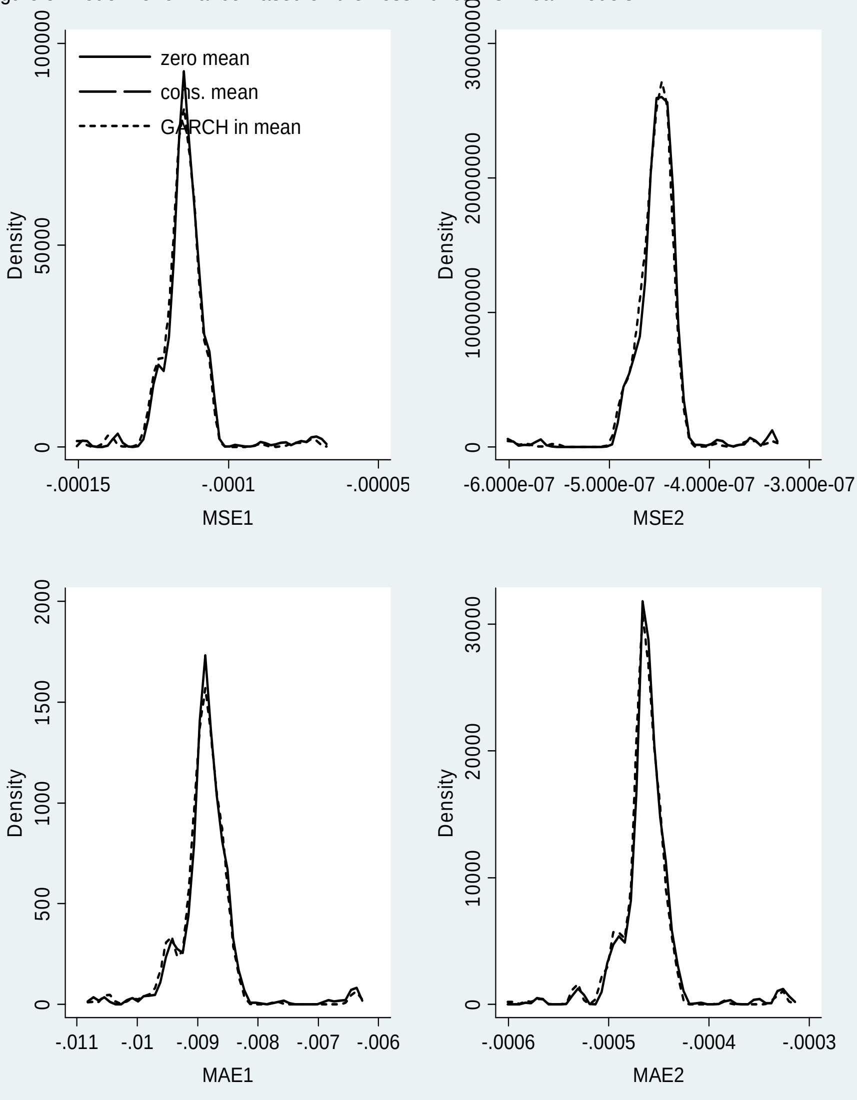
These graphs display density estimates of the losses for the models considered. The x-axis is the negative value of the loss, so that larger values imply better models. Vertical dashed line displays the location of the best performing GARCH(1,1) model.

Figure 4. Model Performance Based on the Loss Functions: Distributions



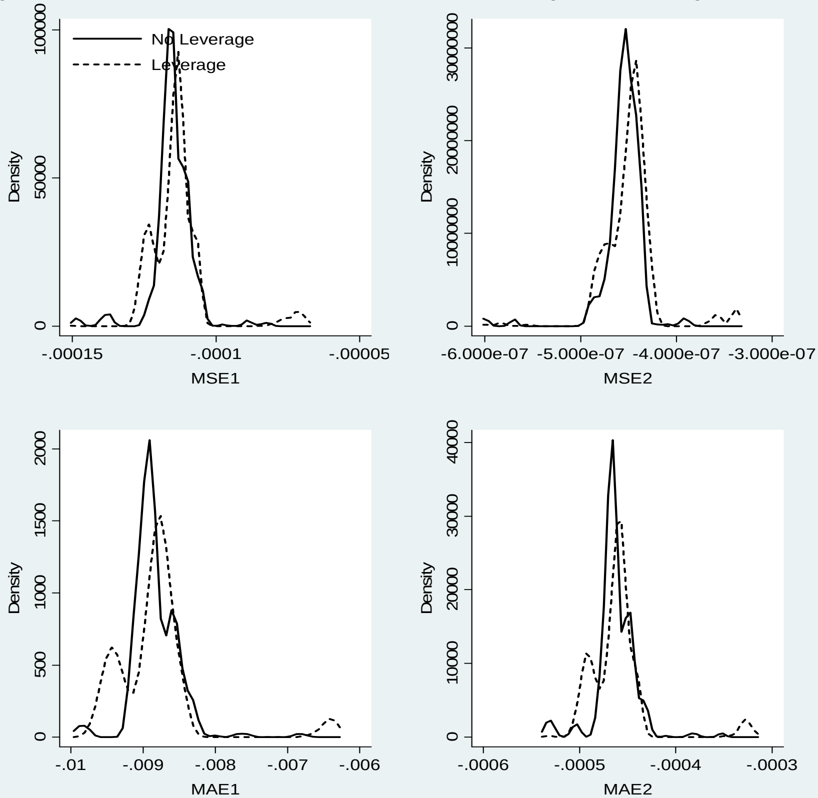
These graphs display density estimates of the losses for the models considered.
The x-axis is the negative values of the losses, so that larger values imply better models.

Figure 5. Model Performance Based on the Loss Functions: Mean Models



These graphs display density estimates of the losses for the models considered.
The x-axis is the negative values of the losses, so that larger values imply better models.

Figure 6. Model Performance Based on the Loss Functions: Leverage vs. No Leverage



These graphs display density estimates of the losses for the models considered.
The x-axis is the negative values of the losses, so that larger values imply better models.

5. An Evaluation of the ISE-100 Returns

In this section, we use the best model selected by the information criteria and the loss functions to examine the behavior of ISE-100 return volatility. We use daily ISE-100 National Index data from 5 January 1998 to 31 December 2008. Figure 7 displays the time series behavior of ISE-100 index over the last 10 years. The effects of the 2000-2001 financial crisis on the Turkish stock market are clearly seen.[9] The index dropped from the levels of 20000s to as low as 8000s. The market performed well after 2004 and reached a

[9] See Ozatay and Sak (2002) and Turhan (2008).

peak of 58231.9 on 15 October 2007. The Turkish economy was adversely affected from the 2008 global financial crisis and the index lost more than half of its value by the end of 2008. Figure 8 displays the historical behavior of the compounded returns. Volatility is high during the 1998-2004 period which includes 2000-2001 financial crisis. The index returns are relatively stable after 2004 and their volatility increase as a result of the 2008 global financial crisis.[10]

Figure 7. ISE-100 Index

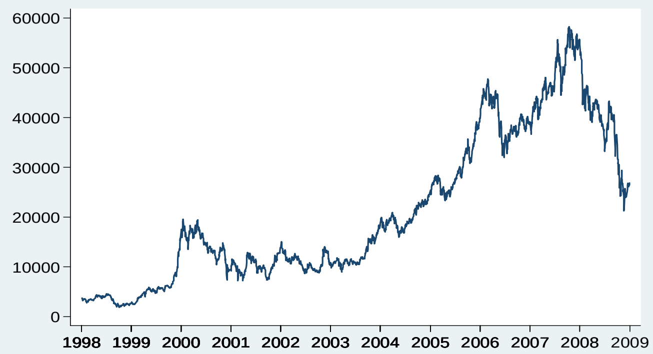
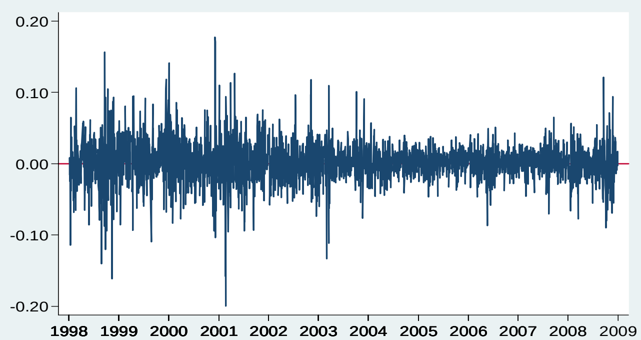


Figure 8: Compounded ISE-100 returns



Results from the previous section indicate that, overall, Nelson's EGARCH(2,2) model is the best model both in terms of fit and forecasting performance. Accordingly, we employ this model to analyze the behavior of

[10] Köksal (2009) finds that the volatility of the ISE-100 index decreases after the 2004 local elections.

ISE-100 index returns and return volatility. Specifically, we estimate the following exponential GARCH model:

$$
\begin{gathered}
r_t = \mu + \varepsilon_t + \rho\varepsilon_{t-1} \\
\varepsilon_t = \sigma_t v_t \\
Var(\varepsilon_t \mid I_{t-1}) = \sigma_t^2 \\
\log(\sigma_t^2) = \omega + \sum_{i=1}^{2}[\alpha_i z_{t-i} + \gamma_i(| z_{t-i} | - \sqrt{2/\pi})] + \sum_{j=1}^{2}\beta_j \log(\sigma_{t-j}^2)
\end{gathered}
\tag{2}
$$

where $\mu = E(r_t \mid I_{t-1})$, I_{t-1} denotes the information set available at time *t*-1, v_t is a sequence of iid random variables that follow the standardized Student-t distribution with υ degrees of freedom, and z_t is the standardized value of the error term defined as $z_t = \varepsilon_t / \sigma_t$. The EGARCH model allows for the asymmetric effects of the shocks on the volatility as follows: If z_{t-i} is positive, i.e., there is a positive shock or good news, the effect on the log of the conditional variance is $\alpha_i + \gamma_i$. If z_{t-i} is negative, i.e., there is a negative shock or bad news, the effect of the shock on the log of the conditional variance is $-\alpha_i + \gamma_i$. If $\gamma_i > 0$ and $\alpha_i < 0$, volatility reacts to bad news more than it does to good news.

Full results from estimating the model above are reported in Table 4. Expected return is estimated to be 0.00086 and significant at the 5% level. Almost all parameters in the conditional variance equation are significant at the 0.1% level. The estimated volatility equation is reproduced below for convenience.

Table 4. Full Results from the EGARCH(2,2) Model

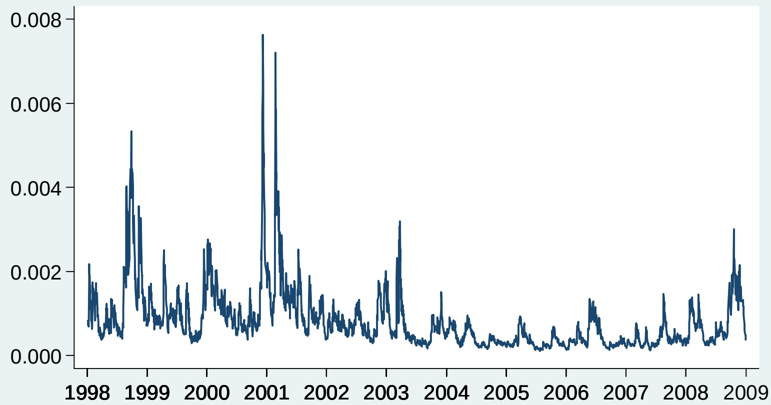
Mean Model	Estimate	Robust Standard Error	t-Statistic	Significance Level	95% Confidence Interval	
μ	0.00086	0.00041	2.08	0.0370	0.00005	0.00166
ρ	0.02533	0.01898	1.33	0.1820	-0.01186	0.06252
Volatility Model						
ω	-0.33559	0.10635	-3.16	0.0020	-0.54403	-0.12715
α_1	-0.04901	0.01584	-3.09	0.0020	-0.08005	-0.01796
α_2	-0.05834	0.01587	-3.68	0.0000	-0.08943	-0.02724
γ_1	0.20957	0.03219	6.51	0.0000	0.14648	0.27266
γ_2	0.22336	0.03137	7.12	0.0000	0.16187	0.28485
β_1	-0.00819	0.01485	-0.55	0.5810	-0.03729	0.02091
β_2	0.96084	0.01482	64.86	0.0000	0.93180	0.98988
υ	7.04897	0.85306			5.62565	9.03104
N	2730					

$$\log\left(\sigma_t^2\right) = -0{,}33559 - 0{,}04901 z_{t-1} + 0{,}20957\left(\left|z_{t-1}\right| - \sqrt{2/\pi}\right) - 0{,}05834 z_{t-2} + 0{,}22336\left(\left|z_{t-2}\right| - \sqrt{2/\pi}\right) - 0{,}00819\log\left(\sigma_{t-1}^2\right) + 0{,}960841\log\left(\sigma_{t-2}^2\right) \quad (3)$$

Equation 3 indicates that current volatility and new shocks are effective for two periods. The predicted values of the conditional variance are displayed in Figure 9, which presents the behavior of the volatility more clearly. The volatility is high until 2004 and reaches a peak during the 2000-2001 financial crisis. It stabilizes after 2004 as a result of the stable economy and increases towards the end of 2008 due to the global financial crisis.

Since $\hat{\gamma}_i > 0$ and $\hat{\alpha}_i < 0$ for i=1,2, volatility of the ISE-100 returns reacts to bad news more than it does to good news. This effect can be illustrated by the news response function or the news impact curve. This curve is a plot of $\hat{\sigma}_t^2$ against $z_{t-1} = \varepsilon_{t-1} / \sigma_{t-1}$ and shows how volatility reacts to good and bad news.

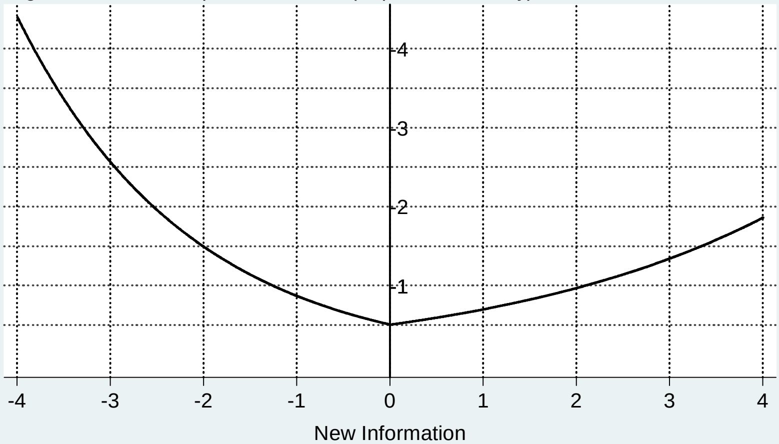
Figure 9: Predicted Conditional Volatility



This graph displays the values of the conditional variance predicted by the EGARCH(2,2) model.

Figure 10 displays the news response function that illustrates the asymmetric effect of bad and good news clearly. For example, when $z_{t-1} = 1$, i.e., there is a one standard deviation positive shock to the returns, the predicted volatility increases by 0.6999, whereas when there is bad news ($z_{t-1} = -1$), the volatility increases by 0.8708. The returns react to bad news 24% more than they react to good news. As magnitude of the shock increases, asymmetry becomes larger. For example, when there is a three standard deviation shock to the returns, volatility increases by 1.3422 for good news, and by 2.5656 for bad news, implying a percentage difference of approximately 91%. Finally, when there is no news (i.e., $z_{t-1} = 0$), expected volatility is 0.5073.

Figure 10. News Response Function (Expected Volatility)



This graph displays the expected volatility for positive and negative shocks to the returns. The x-axis is the standardized error term in the previous period and y-axis is the predicted value of the conditional variance from the EGARCH(2,2) model.

6. Conclusion

We compare more than 1000 GARCH type models in terms of their ability to fit to the historical data and to forecast the conditional variance in an out-of-sample setting. The main findings are that even though widely used GARCH(1,1) model performs well, it is still outperformed by more sophisticated models that allow for leverage effect. The loss functions select the “zero-mean, generalized error distribution, moving average term with one lag” specification, and the information criteria select “constant-mean, t-distribution” specification as the best GARCH(1,1) specification. Overall, the t-distribution seems to characterize the distribution of the returns better than the Gaussian distribution or the generalized error distribution. There are no significant differences between the three specifications for the expected value of returns. Models that allow for leverage effects are slightly superior to models that do not. In terms of forecasting performance, the best models are the ones that can accommodate a leverage effect.

When we look at the best model in terms of the fit and the values of the loss functions, the exponential GARCH (EGARCH) model of Nelson (1991) seems to be the winner for modeling the ISE-100 returns. Specifically, EGARCH(2,2) model that has “constant mean, t-distribution, one lag moving average term" specification has the best fit and forecasting performance when compared to the models we consider. When we fit this model to the historical ISE-100 return data, we find that as a result of a one standard deviation positive shock to the returns, the predicted volatility increases by 0.6999, whereas a negative shock that has the same magnitude increases the volatility by 0.8708. The return volatility reacts to bad news 24% more than it reacts to good news. As the magnitude of the shock increases, the asymmetry becomes larger. When there is a three standard deviation shock to the returns, the volatility increases by 1.3422 for good news, and by 2.5656 for bad news which implies a percentage difference of approximately 91%.

References

Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2003) Parametric and Nonparametric Volatility Measurement. In Ait-Sahalia, Y. & Hansen, L. P. (Eds.) *Handbook of Financial Econometrics*. Amsterdam, Elsevier-North Holland.

Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31: 307-27.

Bollerslev, T. (2008) Glossary to Arch (Garch). School of Economics and Management, University of Aarhus.

Bollerslev, T., Engle, R. F. & Nelson, D. B. (1994) "Arch Models," In Engle, R. F. & McFadden, D. L. (Eds.), *Handbook of Econometrics. Volume 4*: 2959-3038 London and New York, Elsevier North-Holland.

Ding, Z., Granger, C. W. J. & Engle, R. F. (1993) "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance* 1: 83-106.

Doornik, J. A. (2007) *Object-Oriented Matrix Programming Using Ox*. London, Timberlake Consultants Press and Oxford: www.doornik.com.

Engle, R. F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50: 987-1007.

Engle, R. F. (1990) "Stock Volatility and the Crash of '87: Discussion." *Review of Financial Studies* 3: 103-06.

Engle, R. F. & Bollerslev, T. (1986) "Modelling the Persistence of Conditional Variances." *Econometric Reviews* 5: 1-50.

Engle, R. F., Lilien, D. M. & Robins, R. P. (1987) "Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model." *Econometrica* 55: 391-407.

Glosten, L. R., Jagannathan, R. & Runkle, D. E. (1993) "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance* 48: 1779-1801.

Hansen, P. R. (2005) "A Test for Superior Predictive Ability." *Journal of Business and Economic Statistics* 23: 365-80.

Hansen, P. R. & Lunde, A. (2005) "A Forecast Comparison of Volatility Models: Does Anything Beat a Garch(1,1)?" *Journal of Applied Econometrics* 20: 873-889.

Higgins, M. L. & Bera, A. K. (1992) "A Class of Nonlinear Arch Models." *International Economic Review* 33: 137-58.

Köksal, B. (2009) "Political Business Cycles, Partisan Politics and the Effects of Political Events on the Stock Market: Evidence from a Developing Economy." Working Paper, Fatih University.

Nelson, D. B. (1990) "Stationarity and Persistence in the Garch(1,1) Model." *Econometric Theory* 6: 318-34.

Nelson, D. B. (1991) "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59: 347-70.

Ozatay, F. & Sak, G. (2002) "Banking Sector Fragility and Turkey's 2000-01 Financial Crisis," In Collins, S. M. & Rodrik, D. (Eds.), *Brookings Trade Forum 2002*: 121-172 Washington DC, Brookings Institution Press.

Poon, S.-H. & Granger, C. W. J. (2003) "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41: 478-539.

Schwert, G. W. (1989) "Why Does Stock Market Volatility Change over Time?" *Journal of Finance* 44: 1115-53.

Taylor, S. J. (1986) Modelling Financial Time Series. New York, John Wiley & Sons.

Turhan, M. I. (2008) "Why Did It Work This Time?: A Comparative Analysis of Transformation of Turkish Economy after 2002." *Asian-African Journal of Economics and Econometrics* 8: 255-280.

Zakoian, J.-M. (1994) "Threshold Heteroskedastic Models." *Journal of Economic Dynamics and Control* 18: 931-55.