Economics Department

Economics Working Papers

The University of Auckland *Year* 1999

Forecasting Volatility in the New Zealand Stock Market

Jun Yu
University of Auckland, yujun@smu.edu.sg

Forecasting Volatility in the New Zealand Stock Market*

Jun Yu†

November 10, 1999

Abstract

This paper evaluates the performance of nine alternative models for predicting stock price volatility using daily New Zealand data. The competing models contain both simple models such as the random walk and smoothing models and complex models such as ARCH-type models and a stochastic volatility model. Four different measures are used to evaluate the forecasting accuracy. The main results are the following: 1) the stochastic volatility model provides the best performance among all the candidates. 2) ARCH-type models can perform well or badly depending on the form chosen; the performance of the GARCH(3,2) model, the best model within the ARCH family, is sensitive to the choice of assessment measures. 3) the regression and exponentially weighted moving average models do not perform well according to any assessment measure, in contrast to the results found in various markets.

keywords: Forecasting; Volatility; ARCH; Stochastic Volatility Model

1 Introduction

Volatility in financial markets has attracted growing attention by academics, policy makers and practitioners during the last two decades. Firstly, volatility receives a great deal of concern from policy makers and financial market participants because it can be used as a measurement of risk. Secondly, greater volatility in the stock, bond

*We would like to thank the NZSE for providing the data set. Thanks also go to P.C.B. Phillips, Pierre Siklos and Julian Wright for helpful comments and the seminar participants of the New Zealand Econometric Study Group Workshop at the University of Auckland.

†Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand. Email: j.yu@auckland.ac.nz. Tel: 64-9-373-7599 ext. 5770; fax: 64-9-373-7427.

and foreign exchange markets raises important public policy issues about the stability of financial markets and the impact of volatility on the economy. For example, Garner (1990) finds that the stock market crash in 1987 reduced consumer spending in the US. Maskus (1990) finds that the volatility in foreign exchange markets has an impact on trade. Thirdly, from a theoretical perspective, volatility plays a central role in the pricing of derivative securities. According to the Black-Scholes formula, for instance, the pricing of an European call option is a function of volatility. Therefore, option markets can be regarded as a place where people trade volatility. Finally, for the purpose of forecasting return series, forecast confidence intervals may be time-varying, so that more accurate intervals can be obtained by modelling volatility of returns.

There is a large literature on forecasting volatility. Many econometric models have been used, however, no single model is superior. Using US stock data, for example, Akgiray (1989), Pagan and Schwert (1989) and Brooks (1998) finds the GARCH models outperforms most competitors. Brailsford and Faff (1996) (hereafter BF) find that the GARCH models slightly superior to most simple models for forecasting Australian monthly stock index volatility. Using data sets from Japanese and Singaporean markets respectively, however, Tse (1991) and Tse and Tung (1992) find that the exponentially weighted moving average models provide more accurate forecasts than the GARCH model. Dimson and Marsh (1990) find in the UK equity market more parsimonious models such as the smoothing and simple regression models perform better than less parsimonious models, although the GARCH models are not among the set of competing models considered.[1]

The purpose of this paper is to compare the performance of nine models for predicting volatility in the New Zealand stock market. Our paper contributes to this literature in three aspects. First, we use a data set from a country not previously considered in the literature. Although New Zealand does not have a big and liquid stock market, New Zealand economy is one of the least regulated economies and the New

[1]Knight and Satchell (1998) give more details on volatility forecasting in financial markets.

Zealand stock market is one of the freest sharemarkets in the world. Liberalisation of New Zealand financial markets makes them unparalleled internationally. On the other hand, however, little work has been reported specific to New Zealand's financial markets including the New Zealand stock market. Second, we include a stochastic volatility (SV) model into the competing candidates. Unlike the ARCH-type model which has only one error term, the SV model involves two noise processes and hence is supposed to describe financial time series better than the ARCH-type model. However, to our knowledge, no comparison of its performance of volatility forecasts has yet been made for any financial time series. Third, in additional to the assessment measures used in the literature such as the RMSE and MAE, another two measures, the Theil-U statistic and the LINEX loss function, are employed to evaluate the forecast accuracy.[2] U-statistic is a desirable measure to evaluate a forecasting method since it is invariant to any linear transformation (see Armstrong and Fildes, 1996). The LINEX loss function is asymmetric and hence can evaluate positive errors more (or less) than negative errors (see Christoffersen and Diebold, 1997).

The paper is organized as following. In Section 2, we review the unique features of the New Zealand stock market and describe the data set. Section 3 outlines the nine competing models used in this paper for volatility forecasts. We then present the measures used to assess the performance of the candidate models in Section 4. Section 5 describes the empirical results and Section 6 concludes.

2 The New Zealand Stock Market and NZSE40

The New Zealand stock market is one of the least regulated market. In Asia, the Stock Exchanges are primarily arms of government, controlled by government appointees. In the United States, the government acts as an overall market regulator of competitive exchanges. Australia has developed a closely monitored infrastructure with well-defined

[2] Although the Theil-U statistic is a standard measure used to evaluate a forecasting method in macroeconomics, it is used much less frequently in the literature of volatility forecasting.

linkages between the market and outside regulators. Since 1984 New Zealand has conducted a program to deregulate the economy including its financial markets. The reform has established minimal government intervention, under which the NZSE has developed a self-regulatory model that is unparalleled internationally. For example, New Zealand does not impose statutory controls on the Stock Exchange's listing rules, in contrast to most other countries. Also, in the NZSE regulation and oversight of the market rely on contractual principles and New Zealand's take over code, organized by the Exchange and largely self-regulated. Moreover, different from many other markets, insider trading in the NZSE is a civil, not a criminal offense.

Several indices are available for New Zealand. The data set we use is the NZSE40 capital index, which cover 40 largest and most liquid stocks listed and quoted on the New Zealand Stock Market Exchange (NZSE), weighted by the market capitalisation without dividends reinvested. The sample consists of 4741 daily returns over the period from 1 January 1980 to 31 December 1998. Returns are defined as the natural logarithm of price relatives; that is, $r_t = \log \frac{X_t}{X_{t-1}}$, where X_t is the daily capital index.

The data set is used to forecast monthly stock market volatility using various models. In the literature there are a number of ways to obtain monthly volatility series. The first one is proposed by Merton (1980) and Perry (1982) who calculate the volatility in a month simply as the sum of squared daily returns in that month; that is,

$$\sigma_T^2 = \sum_{t=1}^{N_T} r_t^2, \tag{2.1}$$

where r_t is the daily return on day t and N_t is the number of trading days in month T. Akgiray (1989), however, uses a different formula

$$\sigma_T^2 = \sum_{t=1}^{N_T} (r_t - \bar{r}_t)^2 \left[1 + 0.1 \sum_{j=1}^{N_T - 1 - j} \phi^j \right], \tag{2.2}$$

where $\bar{r}_t$ is the mean and ϕ is the first-lag autocorrelation. Of note is that expressions (2.1) and (2.2) share the same spirit; that is, the squared daily return is used as the proxy of the daily volatility. Ding, Granger and Engle (1993) advocate the third way to measure the volatility series where the absolute values of daily stock returns is used. Another possibility is to use the difference between the highest and lowest daily prices (Parkinson, 1980). Although the last method provides a more efficient volatility estimator in terms of approximating the diffusion term in small sample, it is subject to more biases (for example, due to the closure of the stock exchange over night; see Garman and Klass, 1980). The third method is interesting since it generates a series which may have different long memory properties and consequently have a bearing on forecastability. However, it is used less frequently since the long memory models receive little attention in the literature of volatility forecasting. The second method typically provides very similar results as the first method. Hence, we only present the results based on expression (2.1).

In total we have 228 monthly volatilities. Figure 1 plots the series. From this graph, we can easily identify two particularly volatile periods. The first one corresponds to the 1987 crash while the second one occurred on October 1997, the period for the Asian financial crisis. Table 1 shows the mean, median, maximum, kurtosis and part of the first seven autocorrelations of the entire sample. The sample maximum is 0.052157 which happened on October 1987. The sample kurtosis is 77.94 and suggests that the unconditional distribution of volatility is not a normal distribution. The autocorrelations in the volatilities are not very small and consistently positive for the first six orders. This is the evidence of volatility clustering and suggests that the volatility is predictable. To test for possible unit roots we calculate the augmented Dickey-Fuller (ADF) statistic and the results are also presented in Table 1. The ADF statistic for the entire sample is -5.06, which is smaller than -2.57, the critical value at a 10% significance level. Hence, we have to reject the hypothesis that the monthly volatility in the NZSE40 index over the period from 1980 to 1998 has a unit root. Due to the two obvious outliers in the entire sample we need to be concerned with the role of these two possible breaks. In Table 1 we further presents the results of the unit root test for three sub-samples where the entire sample are split by the two crashes. The ADF statistics are -3.32, -9.63 and -11.00 respectively. Hence, no sub-sample involves a unit root. Furthermore, we can identify an October effect in the series. This is not surprising since both crashes occurred on October. However, there is no significant January seasonal as in the US market or July and August seasonals as in the Australian market (see Brown *et. al.*, 1983).

After obtaining the monthly volatility series, we have to choose the forecasting horizon. In this paper we perform 1-month ahead forecasts. Furthermore, we have to choose a period for estimating parameters and a period for predicting volatility. The first 15 years of data are used to fit the models. Thus the first month for which an out-of-sample forecast is obtained is January 1994. As we roll over the sample, we reestimate the models and make sequential 1-month ahead forecasts. Hence, in total we forecast 48 monthly volatilities. With this setup, we require the candidate models to predict volatility in a period when volatility was very large using the sample with an extremely volatile period.[3]

3 Competing Models

In this section we summarize all nine candidate models used in the paper.

3.1 Random Walk

The random walk model is the simplest possible model and is defined as $\hat{\sigma}^2_{T+1} = \sigma^2_T$, $T = 180, ..., 227$. Hence it assumes that the best forecast of next month's volatility is this month's volatility.

[3]By examining the volatility series of Dow Jones composite over the period from 1978-1988, Brooks (1998) claims that the 1987 crash was exception and has not been repeated at such magnitude since. Consequently, he finds the performance of the competing models is quite different for the sample with the 1987 crash and the sample without. In this paper, however, we do not exclude the 1987 crash since the 1997 crash can be regarded, more or less, as a recurring event of the 1987 crash. By the manner we hope the competing models can predict volatility in a volatile period.

3.2 Historical Average

If we assume the conditional expectation of volatility is constant, the optimal forecast of future volatility would be a historical average; that is, $\hat{\sigma}^2_{T+1} = \frac{1}{T}\sum_{t=1}^{T}\sigma^2_t$, $T = 180, ..., 227$. This is the model used most often in the past to predict volatility. However, more recent evidence suggests that the conditional expectation of volatility is time-varying (Bollerslev, Chou and Kroner, 1992) and hence challenges the validity of the historical average model.

3.3 Moving Average

According to the historical average model, all past observations receive equal weight. In the moving average model, however, more recent observations receive more weight. In the paper, we use two moving average models: a five-year and a ten-year moving average. The five-year model is defined as $\hat{\sigma}^2_{T+1} = \frac{1}{60}\sum_{j=1}^{60}\sigma^2_{T+1-j}$, $T = 180, ..., 227$.

3.4 Simple Regression

This is a one-step ahead forecast based on the simple linear regression of the volatility at period $T + 1$ on the volatility at period T. The expression is given by

$$\hat{\sigma}^2_{T+1} = \beta_1 + \beta_2\sigma^2_T,\ T = 180, ..., 227. \tag{3.3}$$

There are two methods to obtain parameter estimates. In the first method, when the new data arrives, we keep the sample size fixed at 180 and hence discard the least recent data. In the second method, however, we use all the observations available to us and thus the sample size gets larger and larger as new data becomes available. We find the results from these two methods are very close to each other. Consequently, we only report the results for the fixed sample size.

3.5 Exponential Smoothing

Exponential smoothing is a simple method of adaptive forecasting. Unlike forecasts from regression models which use fixed coefficients, forecasts from exponential smoothing methods adjust based upon past forecast errors. Single exponential smoothing forecast is given by $\hat{\sigma}^2_{T+1} = (1-\alpha)\hat{\sigma}^2_T + \alpha\sigma^2_T$, where $0 < \alpha < 1$ is the damping (or smoothing) factor. By repeated substitution, we can rewrite the recursion as $\hat{\sigma}^2_{T+1} = \alpha\sum_{t=1}^{T}(1-\alpha)^t\sigma^2_{T+1-t}$, $T = 180, ..., 227$. This shows why this method is called exponential smoothing – the forecast of σ^2_{T+1} is a weighted average of the past values of σ^2_{T+1-t}, where the weights decline exponentially with time. The value of α is chosen to produce the best fit by minimizing the sum of the squared in-sample forecast errors. Dimson and Marsh (1990) and BF select the optimal α annually. In this paper we choose the optimal α in every month so as to provide better forecasts.

3.6 Exponentially Weighted Moving Average (EWMA)

If we combine the exponential smoothing and moving average models, we have the EWMA model. According to the EWMA model, the forecast is obtained by $\hat{\sigma}^2_{T+1} = (1-\alpha)\hat{\sigma}^2_T + \alpha\frac{1}{L}\sum_{j=1}^{L}\sigma^2_{T+1-j}$, $T = 180, ..., 227$. In this paper, we choose $L = 60, 120$ respectively. The value of α is chosen to produce the best fit by minimizing the sum of the squared in-sample forecast errors. BF select the optimal α annually. In this paper we update the optimal α in every month, again so as to provide better forecasts.

3.7 ARCH

The ARCH(q) model is proposed by Engle (1982) and defined by

$$\begin{cases} r_t = & \mu + \sigma_t\varepsilon_t \\ \sigma_t^2 = & \lambda + \alpha_1(r_{t-1}-\mu)^2 + \cdots + \alpha_q(r_{t-q}-\mu)^2 \end{cases}, \tag{3.4}$$

where $\varepsilon_t \sim iidN(0,1)$. Hence the volatility σ^2_{t+1} can be represented by

$$E((r_{t+1}-\mu)^2|I_t) = \sigma^2_{t+1} = \lambda + \alpha_1(r_t-\mu)^2 + \cdots + \alpha_q(r_{t+1-q}-\mu)^2, \tag{3.5}$$

where I_t is the information set at the end of period t. This is an AR(q) model in terms of $(r_t - \mu)^2$. Therefore, the optimal 1-day ahead forecast of period $t+1$ volatility can be obtained based on the returns on the most recent q days. In general, an h-day ahead step forecast can be formed as follows:

$$\hat{\sigma}^2_{t+h} = \lambda + \alpha_1(\widehat{r}_{t+h-1} - \mu)^2 + \cdots + \alpha_q(\widehat{r}_{t+1-q} - \mu)^2, \tag{3.6}$$

where $\widehat{r}_{t+h-j} = r_{t+h-j}$ if $1 \le h \le j$ and $(\widehat{r}_{t+h-j} - \mu)^2 = \hat{\sigma}^2_{t+h-j}$ if $h > j$. The selection of q is an important empirical question. In this paper we choose q using the LM test proposed by Engle (1982). As in the regression model, we keep the sample size fixed for the ARCH model. For NZSE40 the LM test picks up an ARCH(9) specification. After we obtain the daily volatility forecasts across all trading days in each month, we can calculate monthly volatility forecasts using the expression

$$\hat{\sigma}^2_{T+1} = \sum_{t=1}^{N_{T+1}} \hat{\sigma}^2_t, \; T = 180, ..., 227. \tag{3.7}$$

3.8 GARCH

For the ARCH(q) model, in most empirical studies, q has to be large. This motivates Bollerslev (1986) to use the GARCH(p, q) specification which is defined as

$$\begin{cases} r_t = & \mu + \sigma_t \varepsilon_t \\ \sigma_t^2 = & \lambda + \sum_{j=1}^{q} \alpha_j (r_{t-j} - \mu)^2 + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j}. \end{cases} \tag{3.8}$$

Define $s_t = r_t - \mu, m = \max\{p, q\}$, $\alpha_i = 0$ for $i > q$ and $\beta_i = 0$ for $i > p$. Following Baillie and Bollerslev (1992), the optimal h-day ahead forecast of the volatility can be calculated by iterating on

$$\begin{aligned} \hat{\sigma}^2_{t+h} &= \lambda + \sum_{i=1}^{m} (\alpha_i + \beta_i) \hat{\sigma}^2_{t+h-i} - \beta_h \widehat{w}_t - ... - \beta_m \widehat{w}_{t+h-m}, \text{ for } h = 1, ...p \\ &= \lambda + \sum_{i=1}^{m} (\alpha_i + \beta_i) \hat{\sigma}^2_{t+h-i}, \text{ for } h = p+1, ..., \end{aligned}$$

$$
\begin{aligned}
\hat{\sigma}_\tau^2 &= s_\tau^2, \text{ for } 0 < \tau \le t, \\
\hat{\sigma}_\tau^2 &= s_\tau^2 = T^{-1} \sum_{i=1}^{T} s_i^2, \text{ for } \tau \le 0, \\
\widehat{w}_\tau &= s_\tau^2 - E(s_\tau^2 | I_{\tau-1}), \text{ for } 0 < \tau \le t, \\
\widehat{w}_\tau &= 0, \text{ for } \tau \le 0.
\end{aligned}
$$

With the daily volatility forecasts across all trading days in each month, we can calculate monthly volatility forecasts using expression (3.7).

Again, the selection of p and q is an important empirical question. As in the ARCH model, the LM test is used to choose p and q. The GARCH(1,1) model has been found to be adequate in many applications and hence is used as a candidate model. However, for NZSE40 we found a GARCH(3,2) specification is preferred to the GARCH(1,1) model. Consequently, we also assess the prediction performance of the GARCH(3,2) model.

3.9 SV Model

The SV model used in this paper is defined by

$$
\begin{cases} r_t = & \sigma_t \varepsilon_t = \exp(0.5 h_t) \varepsilon_t \\ h_t = & \lambda + \alpha h_{t-1} + v_t \end{cases}, \tag{3.9}
$$

where $\varepsilon_t \sim iidN(0,1), v_t \sim iidN(0, \sigma_v^2)$, and $corr(\varepsilon_t, v_t) = 0$. Like the ARCH-type models, the SV model also models conditional mean and conditional variance by two different equations. As an alternative setup to the ARCH-type models, however, the SV model is supposed to describe financial time series better than the ARCH-type models, since it essentially involves two noise processes (ε_t and v_t). This added dimension makes the model more flexible (for further discussion, see Ghysels, Harvey and Renault, 1996). Unfortunately, the density function for the SV model has no closed form and hence neither does the likelihood function. This is true even for the simplest version of the SV model such as the one defined by equation (3.9). It is a consequence of this that direct maximum-likelihood estimation is infeasible. Probably due to this reason, despite its intuitive appeal, the SV model has received little attention in the literature on forecasting volatility.

Recently several methods have been proposed to estimate the SV model. Such methods include quasi-maximum likelihood (QML) proposed by Ruiz (1994), simulated maximum likelihood (SML) by Danielsson (1994), GMM by Andersen and Sorensen (1996), Markov Chain Monte Carlo (MCMC) by Jacquier, Polson and Rossi (1994), and the empirical characteristic function (ECF) method by Knight, Satchell and Yu (1998). Some of these methods, such as QML and MCMC, not only obtain the estimates of the model, but also produce forecasts of volatility as by-products. MCMC provides the exact optimal predictors of volatility, however, it is computationally more difficult to implement. The QML method approximates a logarithmic chi-square process by a Gaussian process and hence uses the quasi-likelihood to approximate the full likelihood. Despite its inefficiency, the QML method is consistent and very easy to implement numerically. In this paper, we use QML to estimate parameters in the SV model and obtain h-day ahead volatility forecasts. The algorithm employs a Kalman filter and the formulas are given in the Appendix. As in the ARCH-type model, with the daily volatility forecasts across all trading days in each month, we can calculate monthly volatility forecasts using expression (3.7).

4 Evaluating Measures

We use four measures to evaluate the forecast accuracy, namely, the root mean square error (RMSE), the mean absolute error (MAE), the Theil-U statistic and the LINEX loss function. They are defined by

$$\text{RMSE} = \sqrt{\frac{1}{I}\sum_{i=1}^{I}(\hat{\sigma}_i^2 - \sigma_i^2)^2}, \tag{4.10}$$

$$\text{MAE} = \frac{1}{I}\sum_{i=1}^{I} |\hat{\sigma}_i^2 - \sigma_i^2|, \tag{4.11}$$

$$\text{Theil-U} = \frac{\sum_{i=1}^{I}(\hat{\sigma}_i^2 - \sigma_i^2)^2}{\sum_{i=1}^{I}(\sigma_{i-1}^2 - \sigma_i^2)^2}, \tag{4.12}$$

$$\text{LINEX} = \frac{1}{I}\sum_{i=1}^{I}[\exp(-a(\hat{\sigma}_i^2 - \sigma_i^2)) + a(\hat{\sigma}_i^2 - \sigma_i^2) - 1], \tag{4.13}$$

where a in the LINEX loss function is a given parameter.

The RMSE and MAE are two of the most popular measures to test forecasting power of a model. Despite their mathematical simplicity, however, both of them are not invariant to scale transformations. Also, they are symmetric, a property which is not very realistic and inconceivable under some circumstances (see BF).

In the Theil-U statistic, the error of prediction is standardized by the error from the random walk forecast. For the random walk model, which can be treated as the benchmark model, the Theil-U statistic equals 1. Of course, the random walk is not necessarily a naive competitor, particularly for many economic and financial variables, so that the value of the Theil-U statistic close to 1 is not necessarily an indication of bad forecasting performance. Several authors, such as Armstrong and Fildes (1995), have advocated using U-statistic and close relatives to evaluate the accuracy of various forecasting methods. One advantage of using U-statistic is that it is invariant to scalar transformation. The Theil-U statistic is symmetric, however.

In the LINEX loss function, positive errors are weighed differently from the negative errors. If $a > 0$, for example, the LINEX loss function is approximately linear for $\hat{\sigma}_t^2 - \sigma_t^2 > 0$ ('over-predictions') and exponential for $\hat{\sigma}_t^2 - \sigma_t^2 < 0$ ('under-predictions'). Thus, negative errors receive more weight than positive errors. In the context of volatility forecasts, this implies that an under-prediction of volatility needs to be taken into consideration more seriously. Similarly, negative errors receive less weight than positive errors when $a < 0$. BF argue an under-estimate of the call option price, which corresponds an under-prediction of stock price volatility, is more likely to be of greater concern to a seller than a buyer and the reverse should be true of the over-predictions. Christoffersen and Diebold (1997) provide the analytical expression for the optimal LINEX prediction under assumption that the process is conditional normal. Using a series of annual volatilities in the UK stock market, Hwang, Knight and Satchell (1999) show that the LINEX forecasts outperform the conventional forecasts with an appropriate LINEX parameter, a. In this paper, four values for a are used, namely, 20, 10, -10 and -20. Obviously, $a = -10, -20$ penalise over-predictions more heavily while $a = 10, 20$ penalise under-predictions more heavily. BF also adopt asymmetric loss functions to evaluate forecasting performance. An important reason why the LINEX function is more popular in the literature is it provides the analytical solution for the optimal prediction under conditional normality, while the same argument can not be applied to the loss functions used by BF.

5 Results

The main results of the paper are presented in Tables 2 and 3. In Table 2 we report the value and ranking of all nine competing models under the RMSE, MAE and Theil-U while Table 3 presents the value and ranking under the four LINEX loss functions.

From the examination of Table 2 we note that the RMSE statistic indicates that the SV model provides the most accurate forecasts while the GARCH(3,2) model ranks seconds. Despite its simplicity, the random walk model could sometimes offer very good forecasts within the univariate family. For example, Stock and Watson (1998) find that for the US macroeconomic series the random walk model performs the best among many candidate models. However, the random walk model is not a very good method to forecast volatility of the NZSE40 index according to the RMSE. It ranks eleventh and is 26.9% less accurate than the SV model. This finding is consistent with that findings from some other stock markets (see, for example, Brooks (1998) for the

US market and BF for the Australian market). Another salient feature of the results is that the marginal difference in the RMSE between the first and tenth position is very small (3.3%).

The MAE statistic fovours the exponential smoothing model while the SV model is now second best. The EWMA models does not perform very well although Tse (1991) and Tse and Tung (1992) show that the EWMA model is superior in Japanese and Singaporean markets respectively under the RMSE and MAE. For example, the EWMA(10) ranks the last and is 40.3% and 38.7% less accurate than the exponential smoothing and SV models respectively. Unlike the RMSE, the MAE ranks the GARCH models rather poorly. In particular, the GARCH(3,2) model, which has been ranked second by the RMSE, is now ranked eighth. It is 31.8% and 30.0% less accurate than the exponential smoothing and SV models respectively.

Under the Theil-U statistic, only one model performs worse than the random walk model. This model is the ARCH model and it is evidenced by the Theil-U statistic of 1.1065 which is larger than 1. All the other models have the Theil-U statistic less than 1. The best performer is again the SV model, followed by the GARCH(3,2) model.

The common feature of the above three error statistics is that they assume the underlying loss function is symmetric. In Table 3 the same models are evaluated under asymmetric loss functions, where four LINEX loss functions are used ($a = 20$, 10, -10 and -20).

LINEX with $a = 20$ identifies the GARCH(3,2) model as the best performer while the ARCH model and the random walk model provides the worst forecasts. The SV model ranks a close second. Also note that some models which had reasonably good performance according to symmetric loss functions, perform poorly according to the asymmetric loss functions. For example, the exponential smoothing model which was ranked first by the MAE, is now ranked ninth according to LINEX with $a = 20$. It is 5.5% and 4.6% less accurate than the GARCH(3,2) model and the SV model respectively.

When a smaller positive number, 10, is assigned to a in the LINEX function, LINEX picks up the SV model as the most accurate model while the GARCH(3,2) model now ranks second. The results suggest that the GARCH (3,2) model tends to over-predict the volatility. The reason relates to the fact that when a is a smaller positive number, although the under-predictions are still penalised more heavily than the over-predictions, the penalty attached to the under-predictions is smaller.

As we mentioned in the previous section, when a is a negative number, the over-predictions are penalised more heavily than the under-predictions. LINEX with $a = -10$ ranks the SV model first. Together with the findings from the positive values of a, the SV model can be viewed as the most "unbiased" forecast model. This argument is reinforced by LINEX with $a = -20$. According to this statistic, the SV model ranks first once again while the GARCH(3,2) model now ranks fourth. Furthermore, the marginal differences between the SV model and most competing models increase as a decreases. For example, the SV model is 0.2%, 2.4% and 3.3% more accurate than the closest competitor when $a = 10, -10, -20$ respectively. Moreover, the SV model is 0.2%, 2.4% and 3.5% more accurate than the GARCH(3,2) model when $a = 10, -10, -20$ respectively.

In summary, although the SV model has been estimated inefficiently it is the best model overall. It ranks first by the RMSE, Theil-U statistic and three LINEX functions and second by the MAE and the other LINEX function. The performance of the SV model is robust under both symmetric and asymmetric loss functions. Furthermore, the performance of the ARCH-type models is quite variable. In general, the GARCH(3,2) model provides more accurate forecasts than both the GARCH(1,1) and ARCH models. Being a less parsimonious model, the ARCH model is the least accurate model overall. The performance of the GARCH model, the favorite model in BF, Pagan and Schwert (1990), Akgiray (1989), and Franses and van Dijk (1996), is sensitive to the choice of error statistic. For instance, the GARCH(3,2) model ranks second, first, second and second under the RMSE and LINEX with $a = 20, 10$ and -10 respectively, but ranks eighth and fourth under the MAE and LINEX with $a = -20$. The rankings of the GARCH(3,2) model under the four LINEX functions suggest that the GARCH(3,2) model tends to over predict the actual volatilities. A seller of a call option who shows a great deal of concern with under-prediction, would favour the GARCH(3,2) model. However, the GARCH(3,2) model is dominated by the SV model in all other cases. Moreover, both EWMA models do not perform well under any statistic, although Tse (1991) and Tse and Tung (1992) show that the EWMA models are superior in Japanese and Singaporean markets respectively according to the RMSE and MAE. Finally, no statistic identifies the simple regression model as a good candidate and it ranks tenth, fifth, tenth, tenth, tenth, eighth and eighth under the RMSE, MAE, Theil-U, and four LINEX functions respectively. This finding is in contrast to the Australian results reached by BF and the UK results reached by Dimson and Marsh (1990), where the regression model is found superior under the RMSE.

6 Conclusion

This paper examined nine univariate models to forecast stock market volatility of the NZSE40 index. One of the important models considered here is the SV model. Despite its intuitive appeal, the SV model has received no attention in this literature. After comparing the forecasting performance of all nine models, we find that the SV model is superior according to the RMSE, Theil-U and three asymmetric loss functions.

To use the asymmetric loss function, the selection of an appropriate LINEX parameter a is an important empirical question. Unfortunately, to our knowledge, nothing has been reported about the choice of a sensible range of a. An empirical study in this regard would be interesting.

All the models examined in this paper belong to the univariate time series family.

In more recent literature, some multivariate models have been considered to forecast volatility. For example, Brooks (1998) uses the lagged market trading volume to forecast volatility. However, he finds that the added information cannot improve the out-of-sample forecasting performance. Whether or not there are some other variables that are useful to forecast volatility, such as inflation rates or numbers of listed companies, is another interesting question to answer.

How the size and the liquidity of a market can affect the quality of volatility forecasts, we believe, is also an interesting and yet open question. One would think the smaller the size of the market the harder the forecast. An international comparison would be interesting in this regard.

Appendix

The SV model given by (3.9) can be represented by a linearised version without losing any information,

$$\begin{cases} y_t = & \ln(r_t^2) = h_t + \ln(\varepsilon_t^2) = -1.27 + h_t + \mu_t \\ h_t = & \lambda + \alpha h_{t-1} + v_t \end{cases}, \tag{A.1}$$

with $E(\mu_t) = 0, Var(\mu_t) = \pi^2/2$. If we approximate the distribution of μ_t by a normal distribution with mean 0 and variance $\pi^2/2$, the linearised SV model can be represented by a State-Space model. We follow the standard notations of Hamilton (1994).

$$\begin{cases} y_t = & A'x_t + H'\xi_t + w_t \\ \xi_{t+1} = & F\xi_t + v_{t+1} \end{cases}, \tag{A.2}$$

with $A' = -1.27 + \frac{\lambda}{1-\alpha}$, $x_t = 1$, $H' = 1$, $\xi_t = h_t - \frac{\lambda}{1-\alpha}$, $F = \alpha$, $Q = \sigma_v^2$, $R = \pi^2/2$. Based on the State-Space representation, the Kalman filter can be applied as,

- Initialisation:

$$\begin{cases} \hat{\xi}_{1|0} = & 0 \\ \Sigma_{1|0} = & \sigma_v^2/(1-\alpha^2) \end{cases}, \tag{A.3}$$

- Sequential updating:

$$\begin{cases} \hat{\xi}_{t|t} = & \hat{\xi}_{t|t-1} + \Sigma_{t|t-1}(\Sigma_{t|t-1} + \pi^2/2)^{-1}(y_t + 1.27 - \frac{\lambda}{1-\alpha} - \hat{\xi}_{t|t-1}) \\ \Sigma_{t|t} = & \Sigma_{t|t-1} - \Sigma_{t|t-1}(\Sigma_{t|t-1} + \pi^2/2)^{-1}\Sigma_{t|t-1} \end{cases}, \tag{A.4}$$

- In-sample sequential prediction:

$$\begin{cases} \hat{\xi}_{t+1|t} = & \alpha\hat{\xi}_{t|t-1} + \alpha(1 + \frac{\pi^2}{2\Sigma_{t|t-1}})^{-1}(y_t + 1.27 - \frac{\lambda}{1-\alpha} - \hat{\xi}_{t|t-1}) \\ \Sigma_{t+1|t} = & \alpha^2\Sigma_{t|t} + \sigma_v^2 \end{cases}, \tag{A.5}$$

$$\begin{cases} \hat{y}_{t+1|t} = -1.27 + \frac{\lambda}{1-\alpha} + \hat{\xi}_{t+1|t} \\ E[(y_{t+1} - \hat{y}_{t+1|t})(y_{t+1} - \hat{y}_{t+1|t})'] = \Sigma_{t+1|t} + \pi^2/2 \end{cases}, \tag{A.6}$$

- Out-of-sample forecasting:

$$\begin{cases} \hat{\xi}_{T+h|T} = & \alpha^h\hat{E}(\xi_T|I_T) = \alpha^h\hat{\xi}_{T|T} \\ \hat{y}_{T+h|T} = & -1.27 + \frac{\lambda}{1-\alpha} + \alpha^h\hat{\xi}_{T|T} \end{cases}, \tag{A.7}$$

- Smoothing:

$$\begin{cases} \hat{\xi}_{t|T} = \hat{\xi}_{t|t} + J_t[\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t}] \\ \Sigma_{t+1|T} = \Sigma_{t|t} + J_t(-\Sigma_{t+1|T} + \Sigma_{t+1|t})J_t' \\ J_t = \Sigma_{t|t}\alpha\Sigma_{t+1|t}^{-1} \end{cases}, \tag{A.8}$$

  with $t = T-1, T-2, \cdots, 1$.

The quasi-likelihood is computed by

$$\ln L(\alpha, \lambda, \sigma_v^2) = -\frac{1}{2}\sum \log(\Sigma_{t|t-1} + \pi^2/2) - \frac{1}{2}\sum \frac{(y_t + 1.27 - \frac{\lambda}{1-\alpha} - \hat{\xi}_{t|t-1})^2}{\Sigma_{t|t-1} + \pi^2/2}.$$

The h-day ahead forecast is computed by $(A.7)$ with the QML estimates plugged in.

References

[1] Akgiray, V. Conditional heteroscedasticity in time series of stock returns: Evidence and forecast. *Journal of Business*, 62:55–80, 1989.

[2] Andersen, T. and B. Sorensen. GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business and Economic Statistics*, 14:329–352, 1996.

[3] Armstrong, J.S. and R. Fildes. On the selection of error measures for comparisons among forecasting methods. *Journal of Forecasting*, 14:67–71, 1995.

[4] Baillie, R. and T. Bollerslev. Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics*, 52:91–113, 1992.

[5] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[6] Brailsford, T.J. and R. W. Faff. An evaluation of volatility forecasting technique. *Journal of Banking and Finance*, 20:419–38, 1996.

[7] Brooks, C. Predicting stock index volatility: Can market volumn help? *Journal of Forecasting*, 17:59–80, 1998.

[8] Brown, P., D. Keim, A. Kleidon, and T. Marsh. Stock return seasonalities and the tax-loss selling hypothesis: Analysis of the arguments and Australian evidence. *Journal of Financial Economics*, 12:105–127, 1983.

[9] Christoffersen, P.F. and F. Diebold. Optimal prediction under asymmetric loss. *Econometric Theory*, 13:808–817, 1997.

[10] Danielsson, J. Stochastic volatility in asset prices, estimation with simulated maximum likelihood. *Journal of Econometrics*, 64:375–400, 1994.

[11] Dimson, E. and P. Marsh. Volatility forecasting without data-snooping. *Journal of Banking and Finance*, 14:399–421, 1990.

[12] Ding, Z., C.W.J. Granger, and R.F. Engle. A long memory property of stock returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[13] Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.

[14] Franses, P.H. and D. van Dijk. Forecasting stock market volatility using non-linear GARCH models. *Journal of Forecasting*, 15:229–235, 1996.

[15] Garman, M.B. and M.J. Klass. On the estimation of security price volatilities from historical data. *Journal of Business*, 53:67–78, 1980.

[16] Garner, C.A. Has the stock market crash reduced consumer spending? In *Financial Market Volatility and the Economy*. Federal Reserve Bank of Kansas City, 1990.

[17] Ghysels, E., A.C. Harvey, and E. Renault. Stochastic volatility. In *Handbook of Stististics-14*, 1996.

[18] Hamilton, J.D. *Time Series Analysis*. Princeton, New Jersey, 1994.

[19] Hwang, S, J.L. Knight, and S.S. Satchell. Forecasting volatility using LINEX loss functions. University of Western Ontario, 1999.

[20] Jacquier, E, N.G. Polson, and P.E. Rossi. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12:371–389, 1994.

[21] Knight, J.L. and S.E. Satchell. *Forecasting volatility in the financial markets.* Oxford: Butterworth-Heinemann, 1998.

[22] Knight, J.L., S.S. Satchell, and J. Yu. Efficient estimation of the stochastic volatility model by the empirical characteristic function method. The University of Western Ontario, 1998.

[23] Maskus, K.E. Exchange rate risk and U.S. trade: A sectoral analysis. In *Financial Market Volatility and the Economy*. Federal Reserve Bank of Kansas City, 1990.

[24] Merton, R. On estimating the expected return on the market: An explanatory investigation. *Journal of Financial Economics*, 8:323–361, 1980.

[25] Pagan, A. and G.W. Schwert. Alternative models for conditional stock volatilities. *Journal of Econometrics*, 45:267–290, 1990.

[26] M. Parkinson. The extrem value method for estimating the variance of the rate of return. *Journal of Business*, 53:61–65, 1980.

[27] Perry, P. The time-variance relationship of security returns: Implications for the return-generating stochastic process. *Journal of Finance*, 37:857–870, 1982.

[28] Phillips, P.C.B. Time series regression with a unit root. *Econometrica*, 55:277–310, 1987.

[29] Ruiz, E. Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics*, 63:289–306, 1994.

[30] Stock, J. and M. Watson. A comparison of linear and non-linear univariate models for forecasting macroeconomic time series. NBER Working Paper, No. 6607, 1998.

[31] Tse, Y.K. Stock return volatility in the tokyo stock market. *Japan and the World Economy*, 3:285–298, 1991.

[32] Tse, Y.K. and S.H. Tung. Forecasting volatility in the Singapore stock market. *Asia Pacific Journal of Management*, 9:1–13, 1992.

Table 1: Summary statistics of the series and test for nonstationarity

Summary statistics of the entire sample							
mean	median	maximum	kurtosis	ρ_1	ρ_3	ρ_5	ρ_7
0.002341	0.001417	0.052157	77.94	0.281	0.060	0.146	−0.025

ADF Test for Unit Root			
entire sample	sample before 87	sample for 87-97	sample after 97
−5.06	−3.32	−9.63	−11.00

Note: The entire sample is for the monthly volatility of the NZSE40 index over the period from 1980 through 1998. ρ_i denotes the autocorrelation coefficient of order j. The augmented Dickey-Fuller test statistic is computed as $\hat{\tau} = \hat{\beta}/ase(\hat{\beta})$ in the model $\Delta X_t = \alpha + \beta X_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta X_{t-j} + \varepsilon_t$, where X_t represents the monthly volatility of the NZSE40 index (see, e.g., Hamilton, 1994). The value of p is chosen by AIC. The 10% critical value is −2.57. The 5% critical value is −2.86. The justification for using the Dickey-Fuller table when the residuals are heteroskedastic and possibly serially dependent is provided by Phillips (1987).

Table 2: Forecasting performance of competing models under symmetric loss

	RMSE		MAE		Theil-U	
	value	rank	value	rank	value	rank
Random Walk	0.0059588	11	0.0018413	6	1.000	11
Hist. Average	0.0043990	3	0.0019505	7	0.5450	3
MA(5)	0.0044382	5	0.0016957	3	0.5547	5
MA(10)	0.0044926	8	0.0023627	11	0.5684	8
Regression	0.0045047	10	0.0018014	5	0.5715	10
EWMA(5)	0.0044382	6	0.0016957	4	0.5547	6
EWMA(10)	0.0044926	9	0.0023627	12	0.5684	9
Exp. Smooth	0.0044475	7	0.0014108	1	0.5571	7
ARCH	0.0062680	12	0.0023521	10	1.1065	12
GARCH(1,1)	0.0044088	4	0.0020836	9	0.5474	4
GARCH(3,2)	0.0043870	2	0.0020676	8	0.5420	2
SV	0.0043576	1	0.0014481	2	0.5348	1

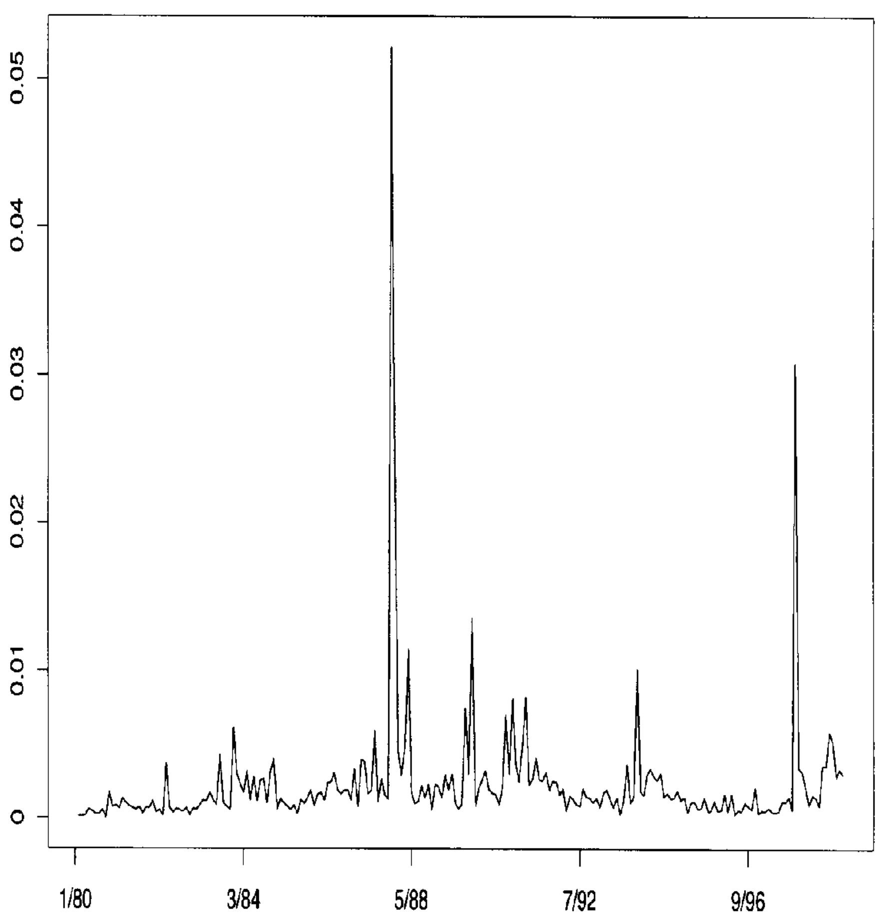
Note: This table lists the value and the ranking of the nine competing models under three measures. The RMSE is defined by (4.10); the MAE is defined by (4.11); the Theil-U statistic is defined by (4.12).

Table 3: Forecasting performance of competing models under asymmetric loss

	LINEX $a = 20$		LINEX $a = 10$		LINEX $a = -10$		LINEX $a = -20$	
	value	rank	value	rank	value	rank	value	rank
Random Walk	7.500179	11	1.812296	11	1.762778	11	7.095735	11
Hist. Average	4.628158	4	1.055301	3	0.891530	3	3.301720	2
MA(5)	4.766143	7	1.080396	5	0.902394	6	3.323553	5
MA(10)	4.743673	5	1.091101	8	0.938119	9	3.505071	9
Regression	4.830648	10	1.103775	10	0.937853	8	3.486059	8
EWMA(5)	4.766157	8	1.080399	6	0.902397	7	3.323568	6
EWMA(10)	4.743676	6	1.091102	9	0.938119	10	3.505074	10
Exp. Smooth	4.829692	9	1.089807	7	0.902277	5	3.309108	3
ARCH	7.915040	12	1.957517	12	1.999728	12	8.260344	12
GARCH(1,1)	4.616085	3	1.056266	4	0.898714	4	3.340270	7
GARCH(3,2)	4.565881	1	1.045357	2	0.890249	2	3.310029	4
SV	4.607941	2	1.043129	1	0.868484	1	3.192824	1

Note: This table lists the value and the ranking of the nine competing models under the four LINEX loss functions where the LINEX loss function is defined by (4.13) $\times$ 1000.

Figure 1



Monthly Volatility of NZSE40 from 1980 to 1998