Financial Econometrics Research Centre

WORKING PAPERS SERIES

WP99-20

Forecasting Volatility using LINEX Loss Functions

Soosung Hwang, John Knight and Stephen Satchell

Forecasting Volatility Using LINEX Loss Functions

Soosung Hwang, University of Cambridge
John Knight, University of Western Ontario
Stephen E. Satchell*, University of Cambridge

June 1999

Abstract

This paper applies the LINEX loss functions to volatility forecasting. We derive the optimal one-step-ahead LINEX forecast for various volatility models. Our results suggest that the LINEX loss function may give us better forecasts than conventional ones.

Keywords: LINEX Loss Function, Forecasting, Volatility.

1 Introduction

Empirical evidence suggests that the serial correlation of returns is not strong and returns are not particularly forecastible. However, it is found that volatility, however measured, has strong autocorrelations over time, see Ding, Granger, and Engle (1994). The recent study of Christoffersen and Diebold (1998) confirms that volatility can be forecasted over 10 to 15 days. We refer readers to Day and Lewis (1992), Engle, Hong, Kane, and Noh (1993), Harvey and Whaley (1992), Lamoureux and Lastrapes (1993), Noh, Engle, and Kane (1994), Hwang and Satchell (1998), and Knight and Satchell (1998b) for more details on volatility forecasting.

Forecasting volatility has been a major issue in finance for some time. For example, volatility forecasts are used to price options and to forecast option prices; they can be used to produce confidence intervals for the prices of the underlying

*Corresponding Author: S. E. Satchell, Faculty of Economics and Politics, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, UK., TEL) 44 1223 335281, FAX) 44 1223 335475, E-MAIL) ses11@econ.cam.ac.uk.

assets and the forecasts can be used as a component of multi-period investment strategies. The recent growing concern about risk management and the rapid growth in financial derivative markets has resulted in volatility forecasting attracting a great deal of interest.

The major development in modelling and forecasting volatility has been the introduction of ARCH models by Engle (1982). Since then, numerous conditional volatility models have been suggested and tested, and there have been a number of papers concerning the appropriateness of using mean-squared error calculations to evaluate the efficacy of volatility forecasts, see Christoffersen and Diebold (1996). It is now understood that a comparison between squared returns and the forecasted volatility is not, in general, valid, and that other approaches to volatility forecast evaluation need to be considered.

Following Varian (1975), Zellner (1986), and Christoffersen and Diebold (1996, 1997b), we advocate the use of LINEX optimal forecasts. It turns out that the forecasts can be explicitly computed for a range of currently used volatility models. We extend their results by presenting results for conditional and unconditional one-step-ahead forecasts for GARCH, Exponential GARCH, stochastic volatility, and a moving average conditional heteroskedasticity model.

A general discussion on LINEX forecasting is presented in Section 2, our formulae are derived in Section 3, calculations and conclusions are presented in Sections 4 and 5.

2 Forecasting Returns and Volatility

In this section we consider some alternative procedures for forecasting that take into account the asymmetry of loss. We shall initially consider LINEX loss functions, see Varian (1975), Zellner (1986), and Christoffersen and Diebold (1996, 1997b) for the history and motivation of this method. One of the most significant differences between the most frequently used loss function, i.e., the mean square loss function, and LINEX loss functions is that the mean square loss function is symmetrical, while LINEX loss functions are asymmetic.

The asymmetric LINEX loss function $L(x)$ is given by:

$$L(x) = \exp(-ax) + ax - 1 \tag{1}$$

where x is the loss associated with the predictive error and a is a given parameter. Figure 1 shows the asymmetric properties of the LINEX loss function. With an appropriate LINEX parameter a, we can reflect small (large) losses for underestimation or overestimation. In particular, a negative a will reflect small losses for overestimation and large losses for underestimation. The asymmetrical weights on losses become clear when we compare the conventional mean square loss function, x^2, which is represented by the thick line in figure 1.

A forecast h is computed by carrying out the following optimization

$$\min_h \int L(y-h)pdf(y)dy \tag{2}$$

where y is the variable we wish to forecast. $pdf(y)$ is the unconditional or conditional probability function of y, depending on the context. If we substitute (1) into (2) we see that

$$\int L(y-h)pdf(y)dy = \exp(ha)m_y(-a) + a\mu_y - ah - 1$$

where $m_y(t)$ is the moment generating function of y evaluated at t, $\mu_y = E(y)$. Differentiating the above with respect to h, we find that the optimal h is given by

$$\widehat{h} = -\ell n(m_y(-a))/a \tag{3}$$

This is essentially the result given in equation (3.2) in Zellner (1986).

Consider some fairly general returns process, y_t

$$y_t = \mu_t + \sigma_t e_t \tag{4}$$

where μ_t is a deterministic mean and σ_t^2 is the conditional variance, e_t is $N(0,1)$, the unconditional mgf of $y_t, m_y(-a)$, is given by

$$m_y(-a) = \exp(-a\mu_t)m_{\sigma_t^2}(\frac{a^2}{2})$$

where $m_{\sigma_t^2}(\cdot)$ is the unconditional mgf of the stochastic volatility process.

It follows immediately that the optimal unconditional LINEX forecast h_t is given by

$$\widehat{h}_t = \mu_t - \ell n(m_{\sigma_t^2}(\frac{a^2}{2}))\diagup a. \tag{5}$$

For $a > 0$, the extra term can be positive or negative depending on the distribution of σ_t^2. Furthermore, the expectation may only be defined for some values of a.

To illustrate the above, consider σ_t^2 following a $\chi^2(m)$ distribution, then

$$h_t = \mu_t + \frac{m}{2a}\ell n(1-a^2),\ 0 < a < 1$$

and μ_t deterministic.

In general, from (4)

$$\begin{aligned} \widehat{h} &= -\frac{\ell n(m_y(-a))}{a} \\ &= -\frac{1}{a}\ell n(\exp(-\mu a)E_{\sigma_{t+1}}(m_e(-a\sigma_{t+1})) \\ &= \mu - \frac{1}{a}\ell n E_{\sigma_{t+1}}(m_e(-a\sigma_{t+1})) \end{aligned} \tag{6}$$

where $m_e(-a) = E[\exp(-ae_t)]$. In many cases, particularly in option pricing problems, forecasting the volatility is a topic of direct interest. In what follows, we shall concentrate on LINEX volatility forecasts.

3 Volatility Forecasts

Christoffersen and Diebold (1997) (CD) have examined the properties of LINEX forecasts under the assumption that the statistical process is conditionally normal. We would write this as $y_{t+h}|\Omega_t \sim N(\mu_{t+h|t}, \sigma^2_{t+h|t})$ where Ω_t is the information set up to time t, typically $\Omega_t = \{y_1, ..., y_t\}$, and where $\mu_{t+h|t}$ and $\sigma^2_{t+h|t}$ are the mean and variance of y_{t+h}, conditional on Ω_t, we can write $y_{t+h}|\Omega_t$ as $y_{t+h|t}$.

Upon examination of standard models, however, we find that the above condition rarely holds. For ARCH/GARCH models introduced by Engle (1982) and Bollerslev (1986), for example, $y_{t+1|t}$ is conditionally normal but $y_{t+h|t}$ is not normal for $h > 1$, see Baillie and Bollerslev (1992) and Knight and Satchell (1998b) for detailed discussion. For stochastic volatility (SV) models developed by Taylor (1986) and Harvey and Shephard (1993, 1996), $y_{t+1|t}$ is not normal, even if we expand Ω_t to include the volatilities up to time t. For these reasons we find that the CD analysis, although interesting, applies to very few examples of processes used by economists. The motivation for this paper is to extend CD's results to more general volatility forecasts. In this section we derive, in closed form where possible, conditional and unconditional LINEX forecasts for SV models and for the E-GARCH model of Nelson (1991) and a volatility process due to Knight and Satchell (1998b).

3.1 Conditioning on past information and volatility models.

We shall denote Ω_t as the information set appropriate to the conditioning. Whilst it is obvious that we would include $y_1, ..., y_t$ in Ω_t, it is by no means clear that conditional volatility, $h_1, ..., h_t$, should also be included since these variables are not observed by the econometrician for any of the models that shall be discussed in this section. However, the convenient assumption that the investors know the true parameter values but not the econometrician can be used to give a definition of available information. For this reason we shall adopt the following definition

Definition 1 *We say that conditional volatility of time t, h_t, belongs to the conditioning set Ω_t if h_t can be computed exactly given knowledge of the true parameters, appropriate initial values for the stochastic process governing h_t, and the observed data, $y_1, ..., y_t$.*

We shall apply Definition 1 when considering the different models under consideration. Summarising these future results we note that for a GARCH (1,1), where $h_t = \alpha + \beta h_{t-1} + \gamma y^2_{t-1}$, we could compute $h_1, ..., h_{t+1}$ given h_0, α, β, γ and $\{y_1, ...y_t\}$ so that $h_1, ..., h_{t+1}$ are clearly in Ω_t. Turning now to a stochastic volatility model (SVM), $y_t = z_t e^{(\xi + h_t)/2}$ and $h_t = \lambda + \alpha h_{t-1} + \nu_t$, it is apparent that knowledge of h_0, λ, ξ, α and $\{y_1, ...y_t\}$ is not enough to compute $h_1, ..., h_t$ so that these variables are not in Ω_t. It is interesting to see that Nelson's Exponential GARCH model (Nelson, 1991) has the same properties as GARCH as does the Knight and Satchell (1,1)

model (Knight and Satchell, 1998b). See the following subsections for the definitions of models and further discussions.

3.2 GARCH(p,q) Models

The GARCH(p,q) process is defined by

$$\begin{aligned} y_t &= z_t h_t^{1/2} \\ h_t &= \alpha + \sum_{i=1}^{p} \beta_i h_{t-1} + \sum_{j=1}^{q} \gamma_j y_{t-j}^2 \end{aligned} \tag{7}$$

where $z_t \sim iid\ N(0,1)$. We shall compute conditional forecasts for $\ell n y_t^2$ and y_t. The information set, according to Definition 1, includes $h_1, ..., h_{t+1}$.

Firstly,

$$\begin{aligned} \ell n y_t^2 &= \ell n z_t^2 + \ell n h_t \\ &= \ell n \chi_{(1)}^2 + \ell n h_t \end{aligned} \tag{8}$$

Thus the moment generating function of $\ell n y_t^2$ is

$$E[e^{-a\ell n(y_t^2)}] = E[e^{-a\ell n \chi_{(1)}^2}] E[e^{-a\ell n h_t}] \tag{9}$$

The moment generating function of $\ell n \chi_{(1)}^2$ is

$$\begin{aligned} m_{\ell n \chi_{(1)}^2}(-a) &= E[e^{-a\ell n \chi_{(1)}^2}] \\ &= E[(\chi_{(1)}^2)^{-a}] \\ &= \int_0^\infty x^{-a} \frac{1}{\Gamma(\frac{1}{2}) 2^{1/2}} x^{1/2-1} e^{-x/2} dx \\ &= \frac{1}{\Gamma(\frac{1}{2}) 2^{1/2}} \int_0^\infty x^{-a+1/2-1} e^{-x/2} dx \end{aligned} \tag{10}$$

Transforming from x to $w = x/2$, which implies $dx = 2dw$, we see that

$$\begin{aligned} m_{\ell n \chi_{(1)}^2}(-a) &= \frac{2^{-a}}{\Gamma(\frac{1}{2})} \int_0^\infty w^{-a-1/2} e^{-w} dw \\ &= \frac{2^{-a}}{\Gamma(\frac{1}{2})} \Gamma(-a + \frac{1}{2}) \\ &= 2^{-a} \frac{\Gamma(-a + \frac{1}{2})}{\Gamma(\frac{1}{2})} \end{aligned}$$

where $\Gamma(.)$ is the gamma function and the LINEX parameter a is restricted to be less than $\frac{1}{2}$ since $\frac{1}{2} - a > 0$. We now consider conditional and unconditional forecasts of volatility.

3.2.1 Optimal One-step-ahead Conditional Forecast of $\ell n y_t^2$ and y_t in GARCH Models

The moment generating function conditional on past h_t is

$$\begin{aligned} m_{\ell n(y_t^2)}(-a)\,|_{\Omega_{t-1}} &= E[e^{-a\ell n(y_t^2)}|\ell n h_t] = E[e^{-a\ell n \chi^2_{(1)}}]e^{-a\ell n h_t} \\ &= 2^{-a}\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}h_t^{-a} \end{aligned} \tag{11}$$

Therefore, the LINEX optimal conditional forecast of $\ell n(y_t^2)$ is

$$\begin{aligned} E[\ell n(y_t^2)|\Omega_{t-1}] &= -\frac{\ell n(m_{\ell n(y_t^2)}(-a))}{a}|_{\Omega_{t-1}} \\ &= \ell n(h_t) + \ell n(2) - \frac{1}{a}\ell n[\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}] \end{aligned} \tag{12}$$

where $h_t = \alpha + \sum_{i=1}^{p}\beta_i h_{t-i} + \sum_{j=1}^{q}\gamma_j y_{t-j}^2$.

In addition, for y_t, the conditional mgf is

$$\begin{aligned} m_{y_t}(-a)\,|_{\Omega_{t-1}} &= E[e^{-ay_t}|\Omega_{t-1}] = E[e^{-az_t h_t^{1/2}}|h_t^{1/2}] \\ &= e^{\frac{a^2 h_t}{2}} \end{aligned} \tag{13}$$

Therefore, the one step ahead conditional forecast is

$$\begin{aligned} E[y_t|\Omega_{t-1}] &= -\frac{\ell n(m_{y_t}(-a))}{a}|_{\Omega_{t-1}} \\ &= -\frac{1}{2}\frac{a^2 h_t}{a} \\ &= -\frac{1}{2}a h_t \end{aligned} \tag{14}$$

Equation (14) agrees with the CD result discussed in section 1.

3.2.2 Optimal One-step-ahead Unconditional Forecast of $\ell n y_t^2$ and y_t in GARCH Models

We now carry out unconditional one-step-ahead forecasts. The moment generating function of $\ell n(y_t^2)$ is, from equations (9) and (10), given by

$$\begin{aligned} m_{\ell n(y_t^2)}(-a) &= E[e^{-a\ell n(y_t^2)}] = 2^{-a}\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}E[e^{-a\ell n h_t}] \\ &\equiv 2^{-a}\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}m_{\ell n h}(-a) \end{aligned}$$

where $m_{\ell nh}(-a)$ is the unconditional mgf of ℓnh_t, which is typically unknown. Therefore, we can write the LINEX one-step ahead unconditional forecast of $\ell n(y_t^2)$ as

$$\begin{aligned} E[\ell n(y_t^2)] &= -\frac{\ell n(m_{\ell n(y_t^2)}(-a))}{a} \\ &= -\frac{1}{a}\ell n(m_{\ell nh}(-a)) + \ell n(2) - \frac{1}{a}\ell n[\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}] \end{aligned} \tag{15}$$

where $a < \frac{1}{2}$.

For y_t, the unconditional mgf is

$$\begin{aligned} m_{y_t}(-a) &= E[e^{-ay_t}] = E[E[e^{-az_t h_t^{1/2}}|h_t^{1/2}]] \\ &= E[e^{\frac{a^2 h_t}{2}}] \\ &\equiv m_h(\frac{a^2}{2}) \end{aligned} \tag{16}$$

Hence the unconditional LINEX forecast of y_t is

$$\begin{aligned} E[y_t] &= -\frac{\ell n(m_{y_t}(-a))}{a} \\ &= -\frac{\ell n(m_h(\frac{a^2}{2}))}{a} \end{aligned} \tag{17}$$

3.3 Exponential GARCH

The Exponential GARCH model introduced by Nelson (1991) is given by (18) below. It is interesting to note that following definition (7), $h_1, ..., h_{t+1}$ belongs to the information set. We define y_t by,

$$\begin{aligned} y_t &= \sigma_t z_t \\ \sigma_t &= e^{h_t/2} \\ h_t &= \alpha_t + \sum_{j=1}^{\infty} \beta_j(\theta z_{t-j} + \gamma(\mid z_{t-j} \mid - E \mid z_{t-j} \mid)) \end{aligned} \tag{18}$$

Note that

$$\begin{aligned} \ell n y_t^2 &= \ell n \sigma_t^2 + \ell n z_t^2 \\ &= h_t + \ell n z_t^2 \end{aligned} \tag{19}$$

setting $\alpha_t = 0$ without loss of generality, we have

$$\ell n y_t^2 = h_t + \ell n \chi_{(1)}^2 \tag{20}$$

with

$$h_t = \sum_{j=1}^{\infty} \beta_j(\theta z_{t-j} + \gamma(\mid z_{t-j} \mid -E \mid z_{t-j} \mid))$$

Thus

$$E[\exp(-a\ell n y_t^2)] = E[e^{-ah_t}]E[e^{-a\ell n \chi^2_{(1)}}] \tag{21}$$

since h_t depends only on lagged z_t's.

3.3.1 Optimal One-step-ahead Conditional Forecast of $\ell n y_t^2$ in E-GARCH Models

Using the same method as in the GARCH(p,q) model, the moment generating function of $\ell n y_t^2$ conditioned on h_t is

$$\begin{aligned} m_{\ell n(y_t^2)}(-a)\,\big|_{\Omega_{t-1}} &= E[\exp(-a\ell n y_t^2)] = e^{-ah_t}E[e^{-a\ell n \chi^2_{(1)}}] \\ &= 2^{-a}\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}e^{-ah_t} \end{aligned} \tag{22}$$

Therefore, the LINEX optimal conditional forecast of $\ell n(y_t^2)$ is

$$\begin{aligned} E[\ell n(y_t^2)|\Omega_{t-1}] &= -\frac{\ell n(m_{\ell n y_t^2}(-a))}{a}|_{\Omega_{t-1}} \\ &= h_t + \ell n(2) - \frac{1}{a}\ell n[\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}] \end{aligned} \tag{23}$$

where $h_t = \sum\limits_{j=1}^{\infty} \beta_j(\theta z_{t-j} + \gamma(\mid z_{t-j} \mid -E \mid z_{t-j} \mid))$.

3.3.2 Optimal One-step-ahead Unconditional Forecast of $\ell n y_t^2$ in E-GARCH Models

Now since $z_t \sim iid\ N(0,1)$ we have

$$E[e^{-ah_t}] = \prod_{j=1}^{\infty} E[\exp(-a\theta\beta_j z_{t-j} - a\beta_j\gamma\ |z_{t-j}|)] \cdot \exp(a\beta_j\gamma E\ |z_{t-j}|). \tag{24}$$

Examining $E[\exp(a_1 z_t + b_1|z_t|)]$, with $z_t \sim iid\ N(0,1)$, we have

$$\begin{aligned} E[\exp(a_1 z_t + b_1|z_t|)] &= \int_{-\infty}^{\infty} e^{a_1 z + b_1|z|}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\ &= \int_{-\infty}^{0} e^{a_1 z - b_1 z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\ &\quad + \int_{0}^{\infty} e^{a_1 z + b_1 z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \end{aligned} \tag{25}$$

Consider

$$\int_0^\infty e^{a_1 z + b_1 z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \exp((a_1+b_1)^2/2) \cdot \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(z-(a_1+b_1))^2) dz \tag{26}$$

If we put $q = z - (a_1 + b_1)$, then $dz = dq$, so that we have

$$\begin{aligned} &= \exp((a_1+b_1)^2/2) \cdot \int_{-(a_1+b_1)}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{q^2}{2}} dq \\ &= \exp((a_1+b_1)^2/2) \cdot \Phi(a_1+b_1). \end{aligned}$$

where $\Phi(.)$ is the cumulative density function of the standard normal distribution. Next

$$\begin{aligned} \int_{-\infty}^0 e^{(a_1-b_1)z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz &= \int_0^\infty e^{(b_1-a_1)w} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \\ &= \exp((b_1-a_1)^2/2)\Phi(b_1-a_1) \end{aligned} \tag{27}$$

by putting $w = -z$, then $dz = -dw$. Finally, for $E\,|z|$ when $z \sim N(0,1)$, we require

$$\begin{aligned} E(|z|) &= \int_{-\infty}^\infty |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= -\int_{-\infty}^0 z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \int_0^\infty z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= 2\int_0^\infty z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-w} dw = \sqrt{\frac{2}{\pi}} \end{aligned} \tag{28}$$

Thus

$$\begin{aligned} E[^{-ah_t}] &= \prod_{j=1}^\infty [\exp(a^2\beta_j^2(\theta+\gamma)^2/2)\Phi(-a\beta_j(\theta+\gamma)) \\ &\quad + \exp(a^2\beta_j^2(\theta-\gamma)^2/2)\Phi(-a\beta_j(\gamma-\theta))] \\ &\quad \cdot \exp\left(a\beta_j\gamma\sqrt{\frac{2}{\pi}}\right) \\ &= m_h(-a) \end{aligned} \tag{29}$$

Therefore, using equations (10) and (21), we have

$$E[\exp(-a\ell n y_t^2)] = 2^{-a} \frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})} \cdot m_h(-a) \tag{30}$$

Therefore the optimal LINEX unconditional forecast by $\ell n y_{t+1}^2$ is given by

$$E[\ell n y_t^2] = -\frac{\ell n(m_{\ell n y^2}(-a))}{a} \tag{31}$$

$$
\begin{aligned}
&= -\ell n\{2^{-a}\frac{\Gamma(\frac{1}{2}-a)}{\Gamma(\frac{1}{2})}\cdot m_h(-a)\}/a \\
&= -\frac{1}{a}\{-a\ell n2+\ell n\Gamma(\frac{1}{2}-a)-\ell n\Gamma(\frac{1}{2}) \\
&\quad +\sum_{j=1}^{\infty}[a\beta_j\gamma\sqrt{\frac{2}{\pi}}+\ell n\{\exp(a^2\beta_j^2(\theta+\gamma)^2/2)\Phi(-a\beta_j(\theta+\gamma)) \\
&\quad +\exp(a^2\beta_j^2(\theta-\gamma)^2/2)\Phi(-a\beta_j(\gamma-\theta))\}] \\
&= \ell n2-\frac{1}{a}\ell n[\frac{\Gamma(\frac{1}{2}-a)}{\Gamma(\frac{1}{2})}]-\gamma\sqrt{\frac{2}{\pi}}\sum_{j=1}^{\infty}\beta_j \\
&\quad -\frac{1}{a}\sum_{j=1}^{\infty}\ell n[e^{a^2\beta_j^2(\theta+\gamma)^2/2}\cdot\Phi(-a\beta_j(\theta+\gamma))+e^{a^2\beta_j^2(\theta-\gamma)^2/2}\cdot\Phi(-a\beta_j(\gamma-\theta))]
\end{aligned}
$$

3.4 Stochastic Volatility Model

In this section, we investigate LINEX optimal forecasts of the stochastic volatility model (SVM). This model is discussed in Taylor (1986) and Harvey and Shephard (1993, 1996). The SVM is given by

$$
\begin{aligned}
y_t &= z_t e^{h_t/2} \\
h_t &= \lambda+\alpha h_{t-1}+\nu_t, \qquad \nu_t\sim iid\ N(0,\sigma^2)
\end{aligned} \tag{32}
$$

where $z_t\sim iid\ N(0,1)$ and it is assumed that z_t and ν_t are independent. Note the log-volatility can be represented as $\ell ny_t^2=h_t+\ell nz_t^2$. The moment generating function of ℓny_t^2 is

$$
\begin{aligned}
E[\exp(-a\ell ny_t^2)] &= E[\exp(-ah_t)\exp(-a\ell nz_t^2)] \\
&= E[e^{-ah_t}]e^{-a\xi}2^{-a}\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}.
\end{aligned} \tag{33}
$$

Although not immediately obvious, according to Definition 1, h_1, ..., h_t, h_{t+1} are not in the information set, intuitively because there are two sources of noise.

3.4.1 Optimal One-step-ahead Conditional Forecast of ℓny_t^2 in SVM

The optimal LINEX forecast of ℓny_t^2 conditional on h_t is

$$
\begin{aligned}
E[\ell ny_t^2|\Omega_{t-1}] &= -\frac{\ell n(m_{\ell ny_t^2}(-a))}{a}|_{\Omega_{t-1}} \\
&= E(h_t|\Omega_{t-1})+\ell n2-\frac{1}{2}\ell n[\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}]
\end{aligned} \tag{34}
$$

In general $E(h_t|\Omega_{t-1})$ will depend upon lagged y values, but a simple expression for this term does not appear to be available in the SVM.

3.4.2 Optimal One-step-ahead Unconditional Forecast of $\ell n y_t^2$ in SVM

Note that the unconditional moment generating function of h_t is

$$E[e^{-ah_t}] = \exp(\frac{-a\lambda}{1-\alpha}) \exp(\frac{a^2\sigma^2}{2(1-\alpha^2)}) \tag{35}$$

Therefore, the optimal LINEX prediction of $\ell n y_t^2$ is given by

$$\begin{aligned} E[\ell n y_t^2] &= -\frac{\ell n(m_{\ell n y^2}(-a))}{a} \\ &= -\frac{1}{a}\ell n \left\{ e^{\frac{-a\lambda}{(1-\alpha)}} e^{\frac{\sigma^2 a^2}{2(1-\alpha^2)}} 2^{-a} \frac{\Gamma(\frac{1}{2}-a)}{\Gamma(\frac{1}{2})} \right\} \\ &= -\frac{1}{a}\left\{ \frac{-a\lambda}{1-\alpha} + \frac{\sigma^2 a^2}{2(1-\alpha^2)} - a\ell n 2 + \ell n \left(\frac{\Gamma(\frac{1}{2}-a)}{\Gamma(\frac{1}{2})} \right) \right\} \\ &= \frac{\lambda}{1-\alpha} - \frac{\sigma^2 a}{2(1-\alpha^2)} + \ell n 2 - \frac{1}{a}\ell n \left(\frac{\Gamma(\frac{1}{2}-a)}{\Gamma(\frac{1}{2})} \right) \end{aligned} \tag{36}$$

3.5 Knight-Satchell Modified GARCH(p,q)

This model is presented in Knight and Satchell (1998b). Essentially, it writes h_t as linear in lagged h_t and lagged z_t^2, thereby eliminating the non-linearities in equation (7). The Knight-Satchell (KS) Modified GARCH(p,q) can be represented as

$$\begin{aligned} y_t &= z_t h_t^{1/2} \\ h_t &= \alpha + \sum_{i=1}^{p} \beta_i^2 h_{t-i} + \sum_{j=1}^{q} \gamma_j z_{t-j}^2 \end{aligned} \tag{37}$$

where $z_t \sim iid\ N(0,1)$. See Knight and Satchell (1998) for further discussion on this model. In this model the information set, Ω_{t-1}, contains $h_1, h_2, , , , h_t$.

3.5.1 Optimal One-step-ahead Conditional Forecast of y_t in the KS Modified GARCH(p,q)

The mgf of y_t conditioning on the information set Ω_{t-1} is

$$\begin{aligned} E[e^{-ay_t}|\Omega_{t-1}] &= E[e^{-ah_t^{1/2}z_t} \mid h_t^{1/2}] \\ &= e^{\frac{a^2 h_t}{2}} \end{aligned} \tag{38}$$

where h_t is defined in equation (37). Therefore, the LINEX optimal one-step-ahead forecast is

$$\begin{aligned} E[y_t|\Omega_{t-1}] &= -\frac{\ell n(m_y(-a))}{a}_{|\Omega_{t-1}} \\ &= -\frac{1}{2}ah_t \end{aligned} \tag{39}$$

which is exactly the same as that for the GARCH models in equation (14) except for the different conditional volatility process h_t. Again this is a special case of the CD result.

3.5.2 Optimal One-step-ahead Unconditional Forecast of y_t in the KS Modified GARCH(1,1)

Now let us consider a simple case of $p = 1$ and $q = 1$. The mgf of the conditional volatility of the modified GARCH(1,1) model can be shown to be

$$m_y(-a) = \exp\left(\frac{a^2\alpha}{2(1-\beta)}\right) \cdot \prod_{j=0}^{\infty}(1 - a^2\gamma\beta^j)^{-1/2} \tag{40}$$

The optimal LINEX one-step-ahead unconditional predictor of y_t is given by

$$\begin{aligned} E[y_t] &= -\frac{\ell n(m_y(-a))}{a} \\ &= -\frac{1}{a}\left\{\frac{a^2\alpha}{2(1-\beta)} + \sum_{j=0}^{\infty}\ell n(1 - a^2\gamma\beta^j)^{-1/2}\right\} \\ &= -\frac{a\alpha}{2(1-\beta)} + \frac{1}{2a}\sum_{j=0}^{\infty}\ell n(1 - a^2\gamma\beta^j) \end{aligned} \tag{41}$$

See Knight and Satchell (1998) for proof. The optimal LINEX forecast for the more complicated KS GARCH(p, q) models where p>1 and q>1 will be obtained by an application of the above method.

4 Results

4.1 Forecasting Measures

We first calculate the one-step ahead forecast, $h_{t+j+1} = E_{t+j}(y^2_{t+j+1})$, of the GARCH (1,1) model. Then various measures are computed for the test of forecasting power. We first calculate conventional mean absolute forecast error (MAFE) and mean squared forecast error (MSFE) of the forecast, which are represented as follows.

$$MAFE = \frac{1}{n}\sum_{j=0}^{n-1}\left||y_{t+j+1}| - h_{t+j+1}^{1/2}\right| \tag{42}$$

$$MSFE = \frac{1}{n}\sum_{j=0}^{n-1}[|y_{t+j+1}| - h_{t+j+1}^{1/2}]^2$$

where the number n is the number of forecasts. We let the j run over n periods where we have estimated the model n times. These measures are one of the most frequently used measures to test forecasting power of a model, see Day and Lewis (1992), Engle, Hong, Kane, and Noh (1993), and Hwang and Satchell (1998) for examples. Alternatively, we can use mean absolute log-forecast error (MALFE) and mean squared log-forecast error (MSLFE),

$$MALFE = \frac{1}{n}\sum_{j=0}^{n-1}\left|\ell n(y_{t+j+1}^2) - \ell n(h_{t+j+1})\right| \tag{43}$$

$$MSLFE = \frac{1}{n}\sum_{j=0}^{n-1}[\ell n(y_{t+j+1}^2) - \ell n(h_{t+j+1})]^2$$

See Christodoulakis and Satchell (1998) for these measures.

We now suggest four alternative LINEX forecast measures corresponding to the four conventional measures in equations (42) and (43). For LINEX volatility forecast and LINEX forecast error, the logarithms of correction factor in equation (12), $LCF \equiv \ell n(2) - \frac{1}{a}\ell n[\frac{\Gamma(-a+\frac{1}{2})}{\Gamma(\frac{1}{2})}]$, should be calculated. Note that the gamma function requires $-a+\frac{1}{2} > 0$ and thus we have $a < \frac{1}{2}$. We choose eight values for a: a=0.375, 0.25, 0.125, -0.5, -1, -1.5, -2, and -2.5. To get some idea of the loss function that these numbers imply, readers should inspect figure 1.

For given parameter values of a, we first calculated mean absolute LINEX forecast error ($MAFE_{LINEX}$) and mean squared LINEX forecast error ($MSFE_{LINEX}$)

$$MAFE_{LINEX} = \frac{1}{n}\sum_{j=0}^{n-1}\left||y_{t+1}| - (h_{t+1}\exp(LCF))^{1/2}\right| \tag{44}$$

$$MSFE_{LINEX} = \frac{1}{n}\sum_{j=0}^{n-1}[|y_{t+1}| - (h_{t+1}\exp(LCF))^{1/2}]^2$$

and mean absolute LINEX log-forecast error ($MALFE_{LINEX}$) and mean squared LINEX log-forecast error ($MSLFE_{LINEX}$)

$$MALFE_{LINEX} = \frac{1}{n}\sum_{j=0}^{n-1}\left|\ell n(y_{t+1}^2) - (\ell n(h_{t+1}) + LCF)\right| \tag{45}$$

$$MSLFE_{LINEX} = \frac{1}{n}\sum_{j=0}^{n-1}[\ell n(y_{t+1}^2) - (\ell n(h_{t+1}) + LCF)]^2$$

These four measures can be compared with the four measures in equations (42) and (43). The results will show the effects of LINEX forecast represented in conventional measures.

Finally, the LINEX measure is calculated. Note that the LINEX measure is

$$L(x) = \exp(-ax) + ax - 1$$

as in equation (1). Note that for a LINEX GARCH forecast, $x = \ell n(y_{t+1}^2) - (\ell n(h_{t+1}) + LCF)$, while for a conventional GARCH forecast, $x = \ell n(y_{t+1}^2) - \ell n(h_{t+1})$. We report both of these cases to assess the magnitude of the change in the LINEX forecast measure.

4.2 Data and Procedures

For our data, we use return volatility provided by Datastream. We took a large UK company, Glaxo Wellcome. We shall only calculate results for GARCH(1,1) in what follows. Although we could extend our calculations to all models discussed, we focus our attention on GARCH(1,1) because of its great popularity

The return volatility is calculated from the log-return less the mean log-return. In what follows, we shall use y_t^2 for the return volatility at time t. More formally, y_t^2 is obtained from log-return series, r_t, as follows:

$$y_t^2 = 250[r_t - \overline{\mu}]^2$$

where the number 250 is used to annualise the squared daily return series and $\overline{\mu}$ is the in-sample mean of r_t. A total of 1978 daily log-returns of Glaxo Wellcome from 2 January 1990 to 31 July 1997 is used. We iterate that $\overline{\mu}$ is calculated using only past observations to avoid any look-ahead bias.

We use a rolling sample of the past volatilities. On day t, the conditional volatility of one period ahead, $t+1$, is constructed by using the estimates which are obtained from only the past observations (i.e., 1738 observations in this study). By recursive substitution of the conditional volatility, a one-step ahead forecast is constructed. On the next day $(t+1)$, using 1738 recent observations (i.e., 1738 observations from the second observation to the 1739^{th} observation), we estimate the parameters again and get another one-step ahead forecast. The estimation and forecasting procedures are performed 240 times using rolling windows of 1738 observations.

Estimations are carried out using the Berndt, Hall, Hall, and Hausman (BHHH) algorithm for the maximisation of the log-likelihood of the GARCH (1,1) model. However, daily estimation of a model is time consuming work. Hwang and Satchell (1998) show that there is little difference in forecasting performance between daily

estimation and longer estimation intervals, e.g., weekly, monthly, and quarterly. In this study we estimate the GARCH(1,1) model every 20 days (approximately monthly estimation) to avoid excessive calculation. Then the coefficients obtained from the estimation are used for forecasting the next 20 days.

4.3 Results

The results are reported in table 1. The first four measures, MAFE, MSFE, MALFE, and MSLFE, are conventional as in equations (42) and (43) and are not different across the LINEX parameter a. The second four measures, MAFE_{LINEX}, MSFE_{LINEX}, MALFE_{LINEX}, and MSLFE_{LINEX}, are conventional measures over LINEX forecasts. We may not expect to find any evidence of any superior forecasting power of the LINEX forecast with the conventional measures. However, the table shows that if we choose an appropriate value for the LINEX parameter, a, we can reduce forecast error even in terms of conventional measures. The values of MAFE_{LINEX}, MSFE_{LINEX}, MALFE_{LINEX}, and MSLFE_{LINEX} are all less than those of MAFE, MSFE, MALFE, and MSLFE when $-1 < a < 0.125$. In three out of four cases, we find minimum values of forecast errors at $a = 0.125$ and in one case we have a minimum value at $a = -0.5$.

All the LINEX parameters which give smaller values of forecast errors are larger than -1 and the corresponding values of the logarithmic correction factor are less than zero. This means that the one-step ahead forecast of the GARCH(1,1) model is biased upward, giving support to the that the GARCH model is affected by a small number of large volatilities rather than a large number of small volatilities.

The last two rows in table 1 show the results of the LINEX forecast error measure for the GARCH forecast and the LINEX GARCH forecast. As we expected, all values of the LINEX measures for the LINEX GARCH forecast are less than those for the GARCH forecast except $a = 1$. Note that when $a = 1$, the LINEX forecast error measure is equivalent to conventional mean absolute or mean squared forecast errors.

The above results suggest that the LINEX forecast from GARCH models is preferred to the conventional forecast from the GARCH model in terms of the conventional MSFE and MAFE. Moreover, the LINEX GARCH forecast performs better than the conventional GARCH forecast.

5 Conclusions

This study shows the one-step-ahead optimal LINEX forecasts for various volatility models. In addition, the empirical results in section 4 compares the conventional volatility forecasts with the LINEX forecasts of GARCH(1,1) using the mean squared and absolute forecast measures and the LINEX measure.

Our findings are encouraging.. For the data set considered, the LINEX forecasts outperform the conventional forecasts with an appropriate LINEX parameter.

Further research needs to look at multiperiod LINEX conditional and unconditional forecasts. Other work of interest would be to extend our empirical result to all models. As yet we have no general results as to which models would be especially favoured by LINEX relative to mean squared estimates for an appropriate family of loss functions.

References

Andersen, T. G. and Bollerslev, T., 'Long-Run Volatility Dynamics and High Frequency Returns', *Journal of Finance*, 52 (1997), 975-1005.

Baillie, R. T., and Bollerslev, T., 'Prediction in Dynamic Models with Time-Dependent Conditional Variances', *Journal of Econometrics*, 52 (1992), 91-113.

Bollerslev, T., 'Generalized Autoregressive Conditional Heteroskedasticity', *Journal of Econometrics*, 31 (1986), 307-327.

Day, E. T., and Lewis, C. M., 'Stock Market Volatility and the Information Content of Stock Index Options', *Journal of Econometrics*, 52 (1992), 267-287.

Christodoulakis, G. A., and Satchell, S. E., 'Hashing GARCH: A Reassessment of Volatility Forecasting Performance', in Knight, J. and Satchell, S. E. (ed.), *Forecasting Volatility in the Financial Markets*, Butterworth-Heinemann, 1998.

Christoffersen, P. F. and Diebold, F. X., 'Further Results on Forecasting and Model Selection under Asymmetric Loss', *Journal of Applied Econometrics*, 11 (1996), 561-571.

Christoffersen, P. F. and Diebold, F. X., 'Optimal Prediction under Asymmetric Loss', *Econometric Theory*, 13 (1997), 808-817.

Christoffersen, P. F. and Diebold, F. X., 'How Relevant is Volatility Forecasting for Financial Risk Management?', Working paper 6844, 1998, National Bureau of Economic Research.

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

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

Engle, R. F., Hong, C., Kane, A. and Noh, J., 'Arbitrage Valuation of Variance Forecasts with Simulated Options', *Advances in Futures and Options Research,* 6 (1993), 393-415.

Harvey, A. C. and Shephard, N., 'Estimation and Testing of Stochastic Variance Models', Econometrics discussion paper EM/93/268, 1993, London School of Economics.

Harvey, A. C. and Shephard, N., 'Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns', *Journal of Business and Economic Statistics,* 14 (1996), 429-34.

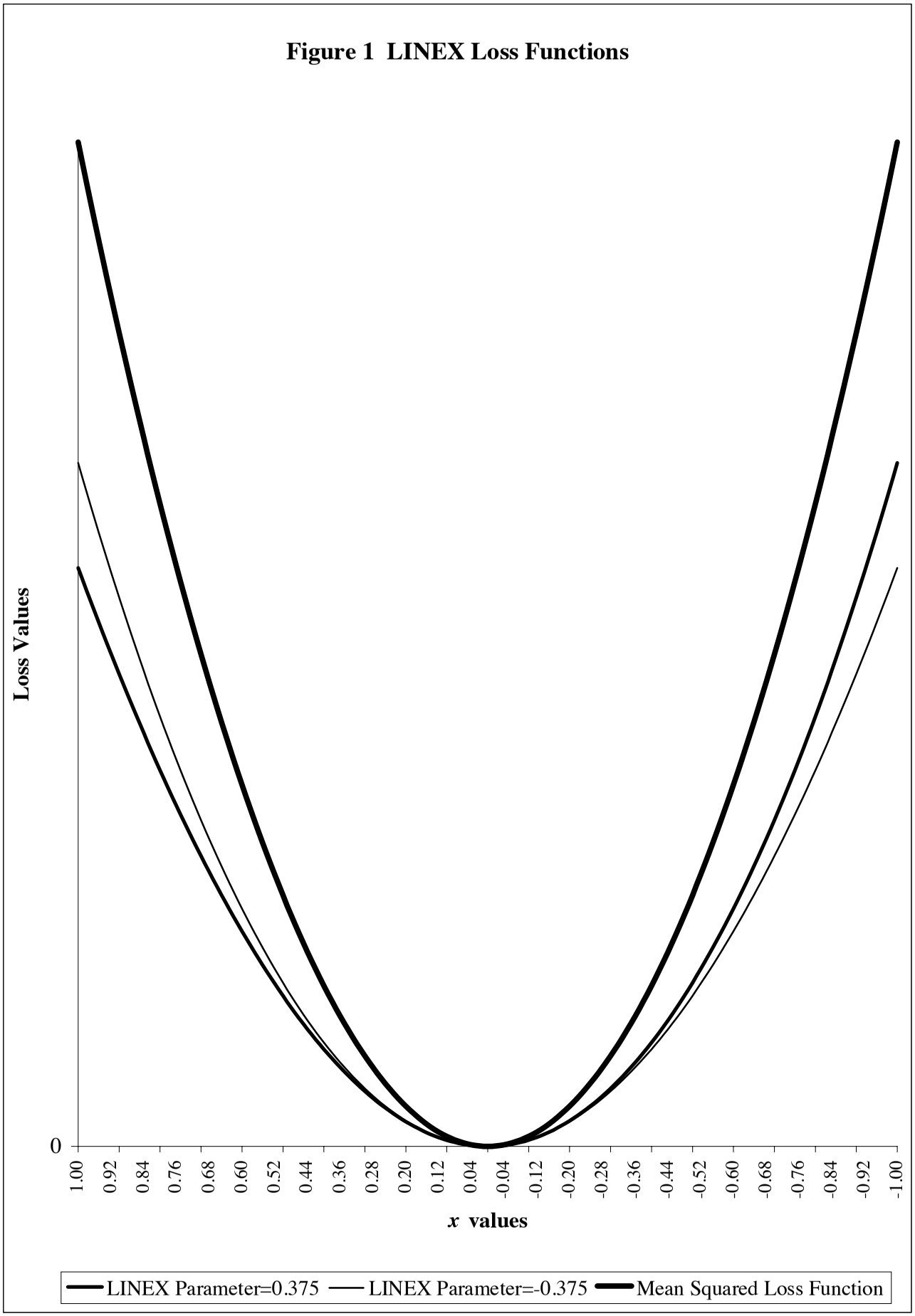
Harvey, C. R., and Whaley, R. E., 'Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market', *Journal of Financial Economics*, 31 (1992), 43-73.
Hwang, S. and Satchell, S. E., 'Implied Volatility Forecasting: A Comparison of Different Procedures Including Fractionally Integrated Models with Applications to UK Equity Options', in Knight, J. and Satchell, S. E. (ed.), *Forecasting Volatility in the Financial Markets*, Butterworth-Heinemann, 1998.
Knight, J. and Satchell, S. E., Forecasting Volatility in the Financial Markets. Butterworth-Heinemann, 1998a.
Knight, J. and Satchell, S. E., 'GARCH Processes - Some Exact Results, Some Difficulties and a Suggested Remedy', in Knight, J. and Satchell, S. E. (ed.), *Forecasting Volatility in the Financial Markets*, Butterworth-Heinemann, 1998b.
Lamoureux, C. G., and Lastrapes, W. D., 'Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities', *Review of Financial Studies*, 6 (1993), 293-326.
Nelson, D. B., 'Conditional Heteroskedasticity in Asset Returns: A New Approach', *Econometrica*, 59 (1991), 347-370.
Noh, J., Engle, R. F., and Kane, A., 'Forecasting Volatility and Option Prices of the S&P 500 Index', *Journal of Derivatives*, (1994), 17-30.
Taylor, S. J., Modeling Financial Time Series, John Wiley: Chichester, 1986.
Varian, Hal R., 'A Bayesian Approach to Real Estate Assessment', in Fienberg, S. E. and Zellner, A. (ed.), *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, Amsterdam: North-Holland, 1975.
Zellner, A., 'Bayesian Estimation and Prediction Using Asymmetric Loss Functions', *Journal of American Statistical Association*, 81 (1986), 446-451.

Table 1 Comparison of Forecasting Results of GARCH(1,1) Model

	a	0.375	0.250	0.125	-0.500	-1.000	-1.500	-2.000	-2.500
Conventional Measures	MAFE of $\lvert y_t \rvert - h_t^{1/2}$	0.1248	0.1248	0.1248	0.1248	0.1248	0.1248	0.1248	0.1248
	MSFE of $\lvert y_t \rvert - h_t^{1/2}$	0.0226	0.0226	0.0226	0.0226	0.0226	0.0226	0.0226	0.0226
	MALFE of $ln(y_t^2)-ln(h_t)$	1.9571	1.9571	1.9571	1.9571	1.9571	1.9571	1.9571	1.9571
	MSLFE of $ln(y_t^2)-ln(h_t)$	7.7717	7.7717	7.7717	7.7717	7.7717	7.7717	7.7717	7.7717
	Logarithms of Correction Factor (LCF)	-3.1657	-2.1695	-1.6325	-0.4516	0.0000	0.3116	0.5493	0.7415
	$MAFE_{LINEX}$ of $\lvert y_t \rvert - (h_t exp(LCF))^{1/2}$	0.1253	0.1097	**0.1013**	0.1058	0.1248	0.1482	0.1709	0.1930
	$MSFE_{LINEX}$ of $\lvert y_t \rvert - (h_t exp(LCF))^{1/2}$	0.0315	0.0255	0.0221	**0.0187**	0.0226	0.0290	0.0370	0.0461
	$MALFE_{LINEX}$ of $ln(y_t^2)-(ln(h_t)+LCF)$	2.4641	1.9390	**1.7441**	1.7745	1.9571	2.1470	2.3085	2.4501
	$MSLFE_{LINEX}$ of $ln(y_t^2)-(ln(h_t)+LCF)$	7.8350	5.6539	**5.3014**	6.5551	7.7717	8.8488	9.8013	10.6538
LINEX Measure	LINEX Measure of $ln(y_t^2)-ln(h_t)$	1.6295	0.4508	**0.0794**	0.5056	1.4577	2.8799	5.4951	11.8090
	LINEX Measure of $ln(y_t^2)-(ln(h_t)+LCF)$	0.5795	0.2212	**0.0479**	0.4619	1.4577	2.7795	4.3608	6.1764

Notes: A total of 1978 daily log-returns of Glaxo Wellcome from 2 January 1990 to 31 July 1997 is used. A rolling sample of the past volatilities is used.
Recent 941 observations are used for estimation of the GARCH(1,1) model. The above restults are based on 240 forecasts.

Figure 1 LINEX Loss Functions



List of other working papers:

1999

1. Yin-Wong Cheung, Menzie Chinn and Ian Marsh, How do UK-Based Foreign Exchange Dealers Think Their Market Operates?, WP99-21
2. Soosung Hwang, John Knight and Stephen Satchell, Forecasting Volatility using LINEX Loss Functions, WP99-20
3. Soosung Hwang and Steve Satchell, Improved Testing for the Efficiency of Asset Pricing Theories in Linear Factor Models, WP99-19
4. Soosung Hwang and Stephen Satchell, The Disappearance of Style in the US Equity Market, WP99-18
5. Soosung Hwang and Stephen Satchell, Modelling Emerging Market Risk Premia Using Higher Moments, WP99-17
6. Soosung Hwang and Stephen Satchell, Market Risk and the Concept of Fundamental Volatility: Measuring Volatility Across Asset and Derivative Markets and Testing for the Impact of Derivatives Markets on Financial Markets, WP99-16
7. Soosung Hwang, The Effects of Systematic Sampling and Temporal Aggregation on Discrete Time Long Memory Processes and their Finite Sample Properties, WP99-15
8. Ronald MacDonald and Ian Marsh, Currency Spillovers and Tri-Polarity: a Simultaneous Model of the US Dollar, German Mark and Japanese Yen, WP99-14
9. Robert Hillman, Forecasting Inflation with a Non-linear Output Gap Model, WP99-13
10. Robert Hillman and Mark Salmon , From Market Micro-structure to Macro Fundamentals: is there Predictability in the Dollar-Deutsche Mark Exchange Rate?, WP99-12
11. Renzo Avesani, Giampiero Gallo and Mark Salmon, On the Evolution of Credibility and Flexible Exchange Rate Target Zones, WP99-11
12. Paul Marriott and Mark Salmon, An Introduction to Differential Geometry in Econometrics, WP99-10
13. Mark Dixon, Anthony Ledford and Paul Marriott, Finite Sample Inference for Extreme Value Distributions, WP99-09
14. Ian Marsh and David Power, A Panel-Based Investigation into the Relationship Between Stock Prices and Dividends, WP99-08
15. Ian Marsh, An Analysis of the Performance of European Foreign Exchange Forecasters, WP99-07
16. Frank Critchley, Paul Marriott and Mark Salmon, An Elementary Account of Amari's Expected Geometry, WP99-06
17. Demos Tambakis and Anne-Sophie Van Royen, Bootstrap Predictability of Daily Exchange Rates in ARMA Models, WP99-05
18. Christopher Neely and Paul Weller, Technical Analysis and Central Bank Intervention, WP99-04
19. Christopher Neely and Paul Weller, Predictability in International Asset Returns: A Re-examination, WP99-03
20. Christopher Neely and Paul Weller, Intraday Technical Trading in the Foreign Exchange Market, WP99-02
21. Anthony Hall, Soosung Hwang and Stephen Satchell, Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models, WP99-01

1998

1. Soosung Hwang and Stephen Satchell, Implied Volatility Forecasting: A Compaison of Different Procedures Including Fractionally Integrated Models with Applications to UK Equity Options, WP98-05
2. Roy Batchelor and David Peel, Rationality Testing under Asymmetric Loss, WP98-04
3. Roy Batchelor, Forecasting T-Bill Yields: Accuracy versus Profitability, WP98-03

4. Adam Kurpiel and Thierry Roncalli , Option Hedging with Stochastic Volatility, WP98-02
5. Adam Kurpiel and Thierry Roncalli, Hopscotch Methods for Two State Financial Models, WP98-01