Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp623-627)

# FORECASTING WITH A LINEX LOSS: A MONTE CARLO STUDY

Yasemin Ulu
Department of Economics
American University of Beirut
Bliss Street, 1107-2020, Beirut
LEBANON

*Abstract:*
Using Monte-Carlo Simulation, I compare the forecasts of returns from the optimal predictor (conditional mean predictor) for a symmetric quadratic loss function (MSE) with the pseudo-optimal predictor and optimal predictor for an asymmetric loss function under the assumption that agents have asymmetric loss functions. In particular, I use the LINEX asymmetric loss function with different degrees of asymmetry. I generate GARCH(1,1) processes with different persistence levels both with normally distributed errors. The results strongly suggest not to use the conditional mean predictor when agents have any kind of asymmetry. The reduction in mean loss by using the optimal versus the pseudo-optimal predictor however depends on the degree of asymmetry, and the persistence parameters being used

Keywords:
Volatility, forecasting returns, asymmetric loss, LINEX.

## 1. *Introduction*

In the literature, a widely used forecast evaluation criteria is the MSE, which is a symmetric quadratic loss function. MSE penalizes the positive errors and negative errors of the same magnitude equally. However, in finance forecasters do not necessarily have a quadratic cost function[1].

Studies have avoided using general asymmetric loss functions mainly because most of the time the closed form for the optimal predictor does not exist. Granger (1969) showed that the optimal predictor under asymmetric loss is the conditional mean plus a constant bias term. Chiristoffersen and Diebold (1997, 1996) showed that for conditionally Gausian processes if an agent has an asymmetric loss function, adding a constant term is not sufficient and that time varying second order moments become relevant for optimal prediction. They derived the analytical expression for the optimal predictor for two specific asymmetric, LINLIN and LINEX loss functions. The LINLIN loss function is, first used by Granger (1969), and LINEX loss function is introduced by Varian (1974) and is used by Zellner (1986). For more general loss functions they showed how to approximate the optimal predictor numerically.

I introduce the LINEX asymmetric loss function, and the univariate variance model I use in Section 2. Section 3 Describes the Monte Carlo simulations and results. Section 4 is a conclusion.

## 2. Univariate Forecasting with Asymmetric Loss

### 2.1. LINEX Loss Function:

The LINEX, convex loss function is introduced by Varian (1974) and used by Zellner (1986).

$$L(x) = b[\exp(ax) - ax - 1], \quad a \in \Re \setminus \{0\}, \quad b \in \Re_+$$

The LINEX loss function is approximately

[1] See Granger (1969), Granger and Newbold (1986, p.125) and Stockman (1987).

linear on negative x-axis and approximately exponential on positive x-axis when $a > 0$. The parameter, a, determines the shape of the loss function while parameter b scales it.

Chiristoffersen and Diebold (1997) derived the optimal predictor, a pseudo-optimal predictor and the expected losses associated with each, for the LINEX asymmetric loss function under the assumption of conditional normality. Given $y_{t+h} \mid \Omega_t \sim N(\mu_{t+h|t}, \sigma_{t+h|t})$, they showed the optimal predictor is $\hat{y}_{t+h} = \mu_{t+h|t} + (a/2)\sigma^2_{t+h|t}$, and the pseudo-optimal predictor is $\hat{y}_{t+h} = \mu_{t+h|t} + (a/2)\sigma^2_h$, where $\sigma^2_h$ is the unconditional $h$-step ahead homoscedastic prediction error variance The pseudo-optimal predictor coincides with the optimal predictor when $\sigma^2_h = \sigma^2_{t+h|t}$.

Through out the paper I consider $h = 1$, which corresponds to one-step a-head prediction. I consider different degrees of asymmetry to compare the loss associated with using different predictors. Specifically, I fix $b = 1$ and change the values of $a$. I consider cases up to where $a/b = 10$. This asymmetric penalization scheme is plausible in finance.

## 2.2 GARCH Models:

The most commonly used model for time-varying volatility is the G/ARCH model of Engle (1982) and Bollerslev (1986). A GARCH(1,1) model for the return on a financial asset, $r_t$, can be written

$$r_t = \sigma_t \cdot z_t,\ z_t \sim IID(0,1)$$

(1)

$$\sigma^2_t = \gamma + \alpha\, r^2_{t-1} + \beta\, \sigma^2_{t-1}$$

where $\gamma > 0$, $\alpha \geq 0$ and $\beta \geq 0$. I assume $z_t$ has finite first and second moments. For a normal GARCH(1,1) model, denoted n-GARCH(1,1), I assume an independent normal innovation. The return $r_t$ is weakly stationary if its variance is finite. This will be the case if $\alpha + \beta < 1$. Since $E(r^2_t \mid \mathrm{I}_{t-1}) = \sigma^2_t$, where $\mathrm{I}_{t-1}$ is the information set available at time t-1, the conditional variance $\sigma^2_t$ is the minimum mean square error predictor of the realized volatility $r^2_t$.

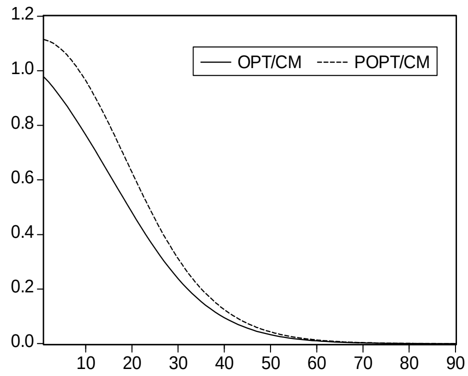
## 3. *Monte Carlo Simulations:*

In this section I study the performance of the optimal, pseudo-optimal and conditional mean predictors when agents have asymmetric LINEX loss function by means of a Monte-Carlo simulation. I look at the average loss ratios associated with each predictor for different degrees of asymmetry.

In the Monte Carlo analysis I consider two scenarios. First I consider normal GARCH(1,1) with different parameter values keeping the "persistence" defined as $\alpha + \beta$ (Bollerslev, 1998) constant. I normalize the unconditional variance to one. The Monte-Carlo simulations are based on 10,000 replications. I then use the simulated n-GARCH(1,1) time-varying conditional standard deviations to compute the optimal, pseudo-optimal and conditional mean predictors for each of the series over the forecast period. I use sample size of 50. This would represent a situation in which one is forecasting weekly data an out-of-sample period of one year. This is typical in empirical work. In the Monte Carlo study, when I compute the optimal, pseudo-optimal and conditional mean predictors, I use the actual value of $h_t$ rather than an estimated value. In empirical work, often the majority of the data is used to estimate the parameters in the conditional variance, and about ten percent of the data is used for the out-of-sample forecasting exercise. The GARCH (1,1) coefficients, alpha and the beta are $\alpha = \{0.3, 0.15, 0.5, 0.10\}$, $\beta = \{.65, .75, .80, .85, .90\}$ respectively. This case presents the performance of the predictors as the alpha and beta coefficients vary for a given persistence level (.95) and different degrees of asymmetry. This high persistence level indicates that market volatility is predictable. I then compute the average loss associated with each predictor and compute the average loss ratios and % reduction in loss by using one type of predictor versus the other.

Secondly, I change the degree of persistence and repeat the same exercise. This case shows the sensitivity of the performance of the predictors to the level of persistence present in the data. For this case I increase the persistence level to .99 by keeping $\beta = \{.65,.75,.80,.85,.90\}$ the same but increasing the ARCH coefficients.

Fig1. $\alpha = .2$, $\beta = .75$, loss ratios from n-GARCH(1,1).




The Figure 1 presents the ratio of the average losses between optimal and conditional mean predictors and the pseudo-optimal and conditional mean predictors when $\alpha = .2$ and $\beta = .75$ for a persistence degree of 0.95. The conditional mean predictor performs the worst. The optimal predictor out performs the pseudo optimal predictor even when a=1. The reduction in loss is around 30% for values of $a$ equal to two and greater. We see that at $a$ equal to five, both lines approach to zero. This is because the conditional mean predictor performs so poorly and the average loss associated with the conditional mean predictor is very large pulling the average loss ratio both with the optimal and pseudo-optimal predictors to zero. Similar results hold for the other simulated series with different GARCH(1,1) parameters, the conditional mean predictor performs the worst, driving the corresponding loss ratios to zero. To avoid this spurious equal performance of optimal and pseudo-optimal predictors only the average loss reduction using the optimal versus the pseudo-optimal predictors will be demonstrated in the proceeding figures, for different GARCH(1,1) models.

Fig 2. Average % Loss reduction for different values of alpha and beta coefficients.

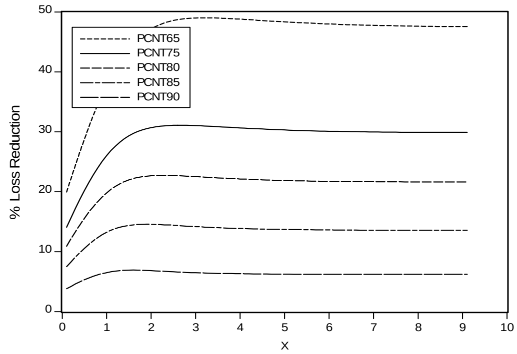



Figure 2 presents the average loss reduction using the optimal versus the pseudo-optimal predictor with different n-GARCH(1,1) parameters. The percent average loss reduction around 8% is minimum for $\beta = .90$ and $\alpha = .05$ for an asymmetry level of two and greater. The percent average loss reduction is maximum for $\alpha = 0.30$ and $\beta = .65$. The percent average loss reduction is maximum for $\alpha = 0.30$ and $\beta = .65$. It is about 20% for a=1 and reaches to 50% for values of a equal to and greater than three. We see that the percent average loss reduction increases with an increasing level of asymmetry and decreasing GARCH(1,1) beta coefficient. The percent average loss reduction is around 30% for $\alpha = .2$ and $\beta = .75$ when the degree of asymmetry is two or greater. Even for low degree of asymmetry the average loss reduction is around 20%. These parameter values are very representative of financial data. So a practitioner might expect around 20% loss reduction using the optimal versus the pseudo-optimal predictor for the given parameter values.

Fig.3, Average % Loss reduction for $\alpha = .99$, $\beta = \{.65,.75,.80,.85,.90\}$.

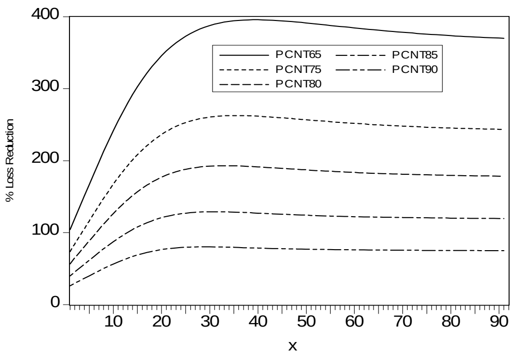



Figure 3 presents the percent loss reduction for the n-GARCH(1,1) for a persistence level of .99 by increasing the values of alpha and keeping $\beta = \{.65,.75,.80,.85,.90\}$ as before. As the persistence increases we see that the percent loss reduction increases as well. The percent average loss reduction that was 8% for $\beta = .90$ and $\alpha = .05$ for a persistence level of .95 before increases to 70 % for an asymmetry level of two and more, for $\beta = .90$ and $\alpha = .09$. This is very interesting because it shows that if you have high persistence using the optimal versus the pseudo-optimal predictor reduces the loss considerably even for low degrees of asymmetry and for any GARCH(1,1) coefficients typical of empirical work. Again the average percent loss reduction is maximum for $\beta = .65$ and $\alpha = .34$ reaching more then 200% even for low degrees of asymmetry.

## 4. Conclusion

I consider the mean losses associated with using the optimal predictor, pseudo-optimal predictor and the conditional mean predictor when agents have asymmetric LINEX loss function.

My results provide strong empirical evidence to the Granger (1969) and CD(1997). The conditional mean predictor performs very poorly compared to the optimal and the pseudo- optimal predictors. For all series, loss associated with using the conditional mean predictor versus using the pseudo or the optimal predictor is considerably higher even for moderate degrees of asymmetry, regardless of the different n-GARCH(1,1) parameters. This result suggests that if agents have any kind of asymmetry, the conditional mean predictor should not be used at all.

The optimal predictor out-performs the pseudo-optimal predictor in all the n-GARCH(1,1) series considered. However, the percentage reduction in loss is very sensitive to the n-GARCH(1,1) parameters being used.

The average percent reduction in loss from using the optimal versus the pseudo-optimal predictor increases with increasing asymmetry, increasing persistence and decreasing beta coefficient. For an empirical representative n-GARCH(1,1) model with parameters alpha=0.2, beta=0.75, the average percent reduction is around 20%. This result suggests that when agents have LINEX type asymmetric loss function, even for low degrees of asymmetry the optimal predictor that incorporates the time varying second moments must be used.

*References:*

[1] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

[2] Bollerslev, T., Engle, R.F., Wooldridge, J.M. (1988), "A capaital Asset Pricing Model with timevarying covariances", *Journal of Political Economy*, 96,116-131.

[3] Chiristoffersen, P. F. and Diebold F. X.(1996), "Further results on Forecasting and Model Selection Under Asymmetric Loss", *Journal of Applied Econometrics*, Vol. 11, No. 5, Special Issue: Econometric Forecasting (Sep.-Oct., 1996), 561-571.

[4] Chiristoffersen, P. F. and Diebold F. X. (1997), "Optimal Prediction Under Asymmetric Loss", Econometric Theory, 13, 808-817.

[5] Diebold, F. X. and Nason J. A (1990), "Nonparametric Exchange Rate Prediction" *Journal of International Economics*, 28, 315-332.
[6] Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, 50, 987-1008.
[7] Engle, R.F., Granger, C.W.J., Kraft, D.F. (1984), " Combining competing forecasts of Inflation using a bivariate ARCH model", *Journal of Economic Dynamics and Control*, 8, 151-165.
[8] Engle, R.F., T. Bollerslev (1986), "Modelling the Persistence of Conditional Variances," *Econometric Reviews* ,5, 1-50.

[9] Granger, C. W. J. (1969), "Prediction with a generalized cost of error function", *Operational Research Quarterly*, 20, 199-207.
[10] Granger, C. W. J. and P. Newbold (1986), *Forecasting Economic Time Series*, 2nd ed. Orlando: Acedemic Press.
[11] Nadaraya, E". A. (1964),"On Estimating Regression", *Theory of Probability and its Applications*, 9,141-142.
[12] Pagan, A. R. and Schewert G. W. (1990) "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics*, 45, 267-290.
[13] Pagan A. and Ullah, A. (1999), "Nonparametric Econometrics" Cambridge.
[14] Silverman, B.W. (1986), "Density Estimation for Statistics and Data Analysis", Chapman and Hall.
[15] Stockman, A.C. (1987), "Economic Theory and Exchange Rate Forecasts", *International Journal of Forecasting*, 3, 3-15.
[16] Watson, G. S. (1964), "Smooth Regression Analysis", *Sankhya* , Series A, 26,359-372.
[17] Varian, H.(1974), " A Bayesian Approach to Real Estate Assessment. In S.E. Feinberg and A. Zellner(eds.), *Studies in Bayesian Econometrics and Statistics in Honor of L.J. Savage*, 195-208.
[18] Zellner, A. (1986), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions", *Journal of American Statistical Association*", Vol. 81, No. 394.