Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp628-631)

PREDICTION UNDER MULTIVARIATE ASYMMETRIC LOSS: A MONTE CARLO STUDY

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

Abstract:
This paper analyzes the loss reduction due to using the optimal predictor in a multivariate framework under the assumption that agents have asymmetric loss functions. A Monte Carlo study shows that the achievable loss reduction is in fact considerably greater then the empirically realized loss reduction. The results are sensitive to the parameter values being used.

Keywords: Volatility, forecasting returns, asymmetric loss, LINLIN, multivariate loss

1.Introduction

Asymmetric loss functions are a better fit to explain the agents' decision problems because they take in to account the fact that over predictions and under predictions of the same magnitude can be penalized differently. The optimal predictor under symmetric loss is simply the conditional mean, the optimal predictor under asymmetric loss however is not. This point is made clear in Granger (1969), Chiristofferson and Diebold (CD henceforth) 1996, and 1997. In fact, CD (1996,1997) consider two types of asymmetric loss functions, linlin and linex, and give the expressions for the optimal, pseudo-optimal and conditional mean predictors when agents have asymmetric loss functions[1]. They conduct a Monte Carlo study to demonstrate the likely loss reduction due to using the optimal predictor versus the pseudo-optimal predictor under the linlin loss function, for a univariate variance model, the GARCH(1,1). They conclude that around 30% loss reduction is achievable, due to using the optimal predictor versus the pseudo optimal predictor when agents have asymmetric loss functions.

The aforementioned study considers a univariate loss function and a univariate variance model by considering a loss associated with one entry at a time. However, we know that in finance, most of the time we are interested in the overall loss. If an agent has the same asymmetric loss function for all the assets in his portfolio, his overall loss will be the sum of his losses due to over predictions and underpredictions. Ulu (2005) considers this issue and extends the CD (1996, 1997) to a multivariate framework. It uses illustrative examples from three

[1] They also point out that the optimal predictor for a general loss function might be very hard or impossible to obtain in general, since the closed expression for the optimal predictor may not exist. For this case they show how to approximate it computationally.

major exchange rates and three major indices. The results for the multivariate case indicate that 8%-10% loss reduction, sometimes reaching to 30% loss reduction is achievable depending on the parameter values of the conditional variance covariance matrix. Although, Ulu (2005) presents illustrative examples, the results are subject to parameter uncertainty. In this paper by means of a Monte Carlo experiment, we will demonstrate the likely realized loss reductions when we know the parameter values rather then estimating them. We consider three panels with different variance covariance parameters that are likely to be found in empirical finance. In the next section we present the LinLin loss function and the multivariate variance covariance matrix. We present the Monte Carlo set up and the results in section 3. Finally, Section 4 concludes.

2. Univariate Forecasting with asymmetric Loss

2.1. *LINLIN Loss Function*:

The LINLIN loss function in general for h-step a head prediction is written as:

$$L(y_{t+h}-\hat{y}_{t+h})=\begin{cases} a\,|\,y_{t+h}-\hat{y}_{t+h}\,| & if\,(y_{t+h}-\hat{y}_{t+h})>0 \\ b\,|\,y_{t+h}-\hat{y}_{t+h}\,| & if\,(y_{t+h}-\hat{y}_{t+h})<=0 \end{cases},$$

where y_{t+h} is the realized value of y, $t+h$ periods a head and $\hat{y}_{t+h}$ is the predicted value of y_{t+h}. The ratio a/b measures the cost of under predicting relative to the cost of over predicting. Through out the paper I consider h=1, which corresponds to one step a head prediction.

CD (1997) gives the expressions the optimal predictor, a pseudo-optimal predictor for the LINLIN asymmetric loss function under the assumption of conditional normality. Given $y_{t+h}\,|\,\Omega_t \sim N(\mu_{t+h|t},\sigma_{t+h|t})$, they find that the optimal predictor for y_{t+h} is, $y_{t+h}=\mu_{t+h|t}+\sigma_{t+h|t}\Phi^{-1}(\frac{a}{a+b})$, and the pseudo-optimal predictor is $\mu_{t+h|t}+\sigma_h\Phi^{-1}(\frac{a}{a+b})$, where σ_h^2 is the h-step a head homoscedastic prediction error variance and $\Phi(z)$ is the N-(0,1) c.d.f[2].

3. Multivariate forecasting with asymmetric loss

3.1 Multivariate Loss function

Ulu(2005) extends the theory to a multivariate framework in which more than one series is to be forecasted[3] and following Zellner (1986), if we assume additively separable loss function in the *n* prediction errors. Then we have the following result.

$$L(Y_{t+h}-\hat{Y}_{t+h})=\sum_{i=1}^{n}L_i(y_{i,t+h}-\hat{y}_{i,t+h})$$

.We will choose $L_i(\cdot)$ to be linlin loss function.

$$L(y_{i,t+h}-\hat{y}_{i,t+h})=\begin{cases} a_i\,|\,y_{i,t+h}-\hat{y}_{i,t+h}\,| \\ \quad \text{if } y_{i,t+h}-\hat{y}_{i,t+h}>0 \\ b_i\,|\,y_{i,t+h}-\hat{y}_{i,t+h}\,| \\ \quad \text{if } y_{i,t+h}-\hat{y}_{i,t+h}\le 0 \end{cases}.$$

Ulu (2005) gives the expressions for the optimal, pseudo-optimal and conditional mean predictors under conditional normality. The optimal

---

[2] However, as CD (1997) point out, the conditionally Gaussian assumption can be relaxed. The optimal predictor is obtained by substituting the appropriate conditional CDF.

[3] We direct the interested reader to Ulu (2005) for detail.

Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp628-631)

predictor vector can be written as $\hat{y}_{i,t+h} = \mu_{i,t+h} + \sigma_{ii,t+h} \cdot \Phi^{-1}(a_i /(a_i + b_i))$. The coefficients a_i and b_i might differ for each coordinate but all b_i's are set to one and a_i's are chosen to be the same for each coordinate for simplicity.

3.2 Multivariate Variance Models

A univariate GARCH model can be generalized to a n-dimensional multivariate GARCH model as $r_t \mid \Psi_{t-1} \sim \mathrm{N}(0, \Sigma_t)$, where r_t is the n-dimensional zero mean random variable, Σ_t is the variance covariance matrix that based on information set available at t-1. Σ_t depends on q lagged values of squares and cross products of r_t and p lagged values of Σ_t. The extension of a univariate GARCH model to a n-dimensional multivariate GARCH model require some restrictions on the conditional variance–covariance matrix Σ. There are different parameterazations of the variance covariance matrix. We will use BEKK model of Engle and Kroner (1995) which imposes restrictions on the conditional variance-covariance matrix (Engle, Kroner, (1995) to ensure positive definiteness.

For the multivariate case I estimate the n-diagonal -BEKK multivariate GARCH(1,1) model . The BEKK variance model can be written as $\Sigma_t = C'C + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'\Sigma_{t-1}B$.

3. *Monte Carlo Simulations:*

In this section I study the performance of three predictors, optimal, pseudo-optimal and conditional mean predictors when agents have asymmetric multivariate linlin 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.

I use sample size 50 which corresponds to forecasting with weekly data using an out-of-sample period of one year. This is typical in empirical work. In the Monte Carlo study, I use the actual values A, B, and C and the generated Σ and the data from the normal-BEKK-GARCH(1,1) model to compute the average losses associated with each predictor. The Monte Carlo experiment is based on 10000 replications.

I consider a set of parameter values for A and B, 3X3 diagonal matrices. These values are similar to what is found in empirical work. The diagonal elements of A, {0.2,0.3,0.4}and the corresponding diagonal elements of B {0.89,0.95,0.86}. The results are presented in Figures one

Fig1.

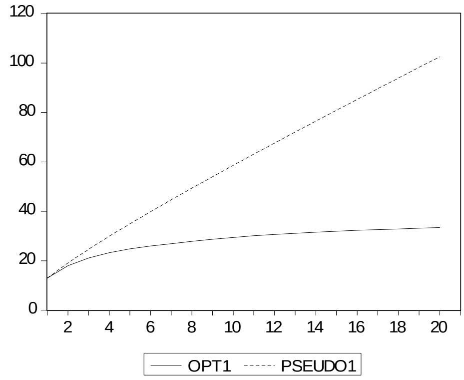

Figure one presents the average losses from using the pseudo-optimal predictor (opt1) versus the optimal predictor (pseudo1) with the parameter matrices of A and B. Clearly considerable loss reduction is achievable using the optimal predictor that incorporates the time varying second moments. This loss reduction reaches to 67% for high degrees of asymmetry.

4.Conclusion

The average loss reduction due to using the optimal predictor versus the pseudo-optimal predictor is analyzed for a set of parameter values similar to those found in empirical work. The simulation result demonstrates that incorporating the time varying second moments in to the predictor reduces the loss up to 68% when agents have asymmetric loss functions.

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] Engle, R. F. (1982), “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation”, *Econometrica*, 50, 987-1008.

[6] 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.

[7] Engle, R.F., T. Bollerslev (1986), “Modelling the Persistence of Conditional Variances,” *Econometric Reviews* ,5, 1-50.

[8] Granger, C. W. J. (1969), “Prediction with a generalized cost of error function”, *Operational Research Quarterly*, 20, 199-207.

[9] 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.

[10] Zellner, A. (1986), “Bayesian Estimation and Prediction Using Asymmetric Loss Functions”, *Journal of American Statistical Association*”, Vol. 81, No. 394.

[11] Ulu, Yasemin (2005), “Optimal Prediction under Asymmetric Loss, Empirical Evidence”, unpublisehed manustcript under review. Extended abstract version (ICSSM 2005).