On Forecasting Exchange Rate Volatility

Christian M. Hafner

In an efficient market, foreign exchange rates have to guarantee absence of triangular arbitrage. This note shows that the no-arbitrage condition can be exploited for forecasting the volatility of a single rate by using the information contained in the other rates. Linearly transforming the volatility forecasts of a bivariate model is shown to be more efficient than using a univariate model for the cross-rate.

Triangular Arbitrage

Suppose we have a system of three foreign currencies, the Euro (EUR), the US Dollar (USD), and the British Pound Sterling (GBP). Denote by USD/EUR the Dollar-Euro exchange rate, i.e., the amount of US Dollar one has to pay for one Euro. Likewise, we denote by EUR/GBP the Euro-Pound exchange rate, and by USD/GBP the Dollar-Pound exchange rate. Then the following condition has to hold in an efficient market:

$$\frac{USD}{EUR}\frac{EUR}{GBP} = \frac{USD}{GBP} \tag{1}$$

In words, the Dollar-Pound exchange rate has to be equal to the product of the Dollar-Euro rate and the Euro-Pound rate. If condition (1) did not hold, then there would be riskless arbitrage opportunities, if transaction costs are sufficiently small. If, for example, the right-hand side of (1) was larger than the left-hand side, then an arbitrage trader could sell one Dollar to buy Pounds at the USD/GBP rate, sell those Pounds for Euros at the GBP/EUR rate, and sell those Euros at the EUR/USD rate to receive more than one Dollar, making a riskless profit. This should be excluded in a sufficiently transparent, liquid and efficient market.

Taking logarithms, we can express (1) equivalently as

$$\log\frac{USD}{EUR} + \log\frac{EUR}{GBP} = \log\frac{USD}{GBP} \tag{2}$$

Since (1) and (2) have to hold at any time t, there is also a no-arbitrage condition for the returns, defined as the first difference of the log rates. Let us denote by $y_{t,1}$ the return on the Dollar-Euro rate, i.e.,

$$y_{t,1} = \log\left(\frac{USD}{EUR}\right)_t - \log\left(\frac{USD}{EUR}\right)_{t-1}$$

and likewise we define $y_{t,2}$ and z_t to be the returns of the EUR/GBP and USD/GBP rates, respectively. Then, the triangular arbitrage condition can be stated in terms of the returns as

$$z_t = y_{t,1} + y_{t,2} \tag{3}$$

for all t.

Suppose now that the returns can be considered as stochastic processes and that we are interested in the volatility, or the variance, of a single rate, e.g. the series z_t. From what we learnt in our first semester of econometrics, we know that

$$\mathrm{Var}(z_t) = \mathrm{Var}(y_{t,1}) + \mathrm{Var}(y_{t,2}) + 2\mathrm{Cov}(y_{t,1}, y_{t,2}) \tag{4}$$

which is yet another form of the same triangular arbitrage condition. Because of the deterministic form of the relation (4), it does not make sense to model the volatilities of the three series jointly in a stochastic model. Rather, only the variances and covariance of two series should be considered as stochastic, while the variance of the third one follows deterministically.

In forecasting the volatility of the third rate, one now has the following two possibilities.

1. Aggregate the bivariate system $(y_{t,1}, y_{t,2})$ according to (3) to obtain z_t and use a univariate volatility model for prediction.
2. Predict the volatility of the bivariate system $(y_{t,1}, y_{t,2})$ and then use (4) to aggregate the forecasts.

The claim is now that the mean squared prediction error of the second method is smaller, or at least not larger, than the mean squared prediction error of the first method. This - perhaps surprising - result is well known in the forecasting of the mean with VARMA models, see e.g. the monograph by Lütkepohl (1987). The intuitive reason is that the second method uses a larger information set. The first method aggregates first, by which information is lost that could be useful for forecasting. In the following, we will investigate this further using a particular volatility model.

An Example using GARCH

To give numerical examples, we use the popular GARCH modelling class, which for the multivariate case has been studied e.g. by Engle and Kroner (1995). Let us denote by H_t the variance matrix of the system $y_t = (y_{t,1}, y_{t,2})$ conditional on the available information set I_{t-1}, i.e. $H_t = \mathrm{Var}(y_t \mid I_{t-1})$. That is, H_t is a (2 x 2) matrix where the two

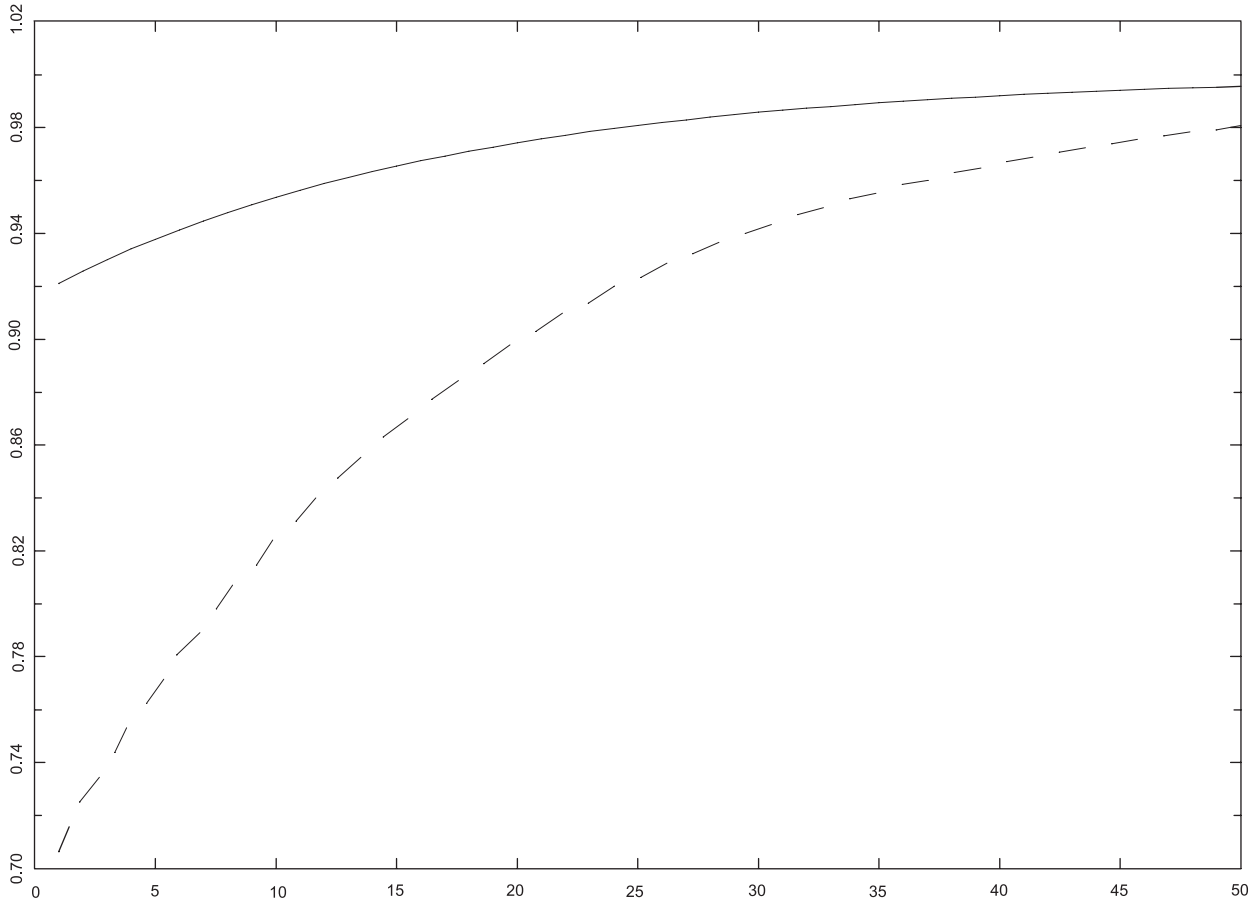


Figure 1: Mean squared errors (MSE) of the prediction of DEM/GBP volatility as a function of the prediction horizon. The solid line is the method using the univariate series DEM/GBP, the dashed one is using the series DEM/USD and USD/GBP and transforms forecasts of a bivariate volatility model. Both MSEs are standardized by the unconditional volatility.

diagonal elements are the conditional variances and the off-diagonal element is the conditional covariance. A bivariate GARCH(1,1) model in the so-called vec form can be written as

$$h_t = \omega + A\eta_{t-1} + Bh_{t-1} \tag{5}$$

where $h_t = \text{vech}(H_t)$, $\eta_t = \text{vech}(y_t y_t')$ and where vech denotes the operator that stacks the lower triangular part of a symmetric matrix into a vector. In (5), ω is a 3 x 1 parameter vector and A and B are 3 x 3 parameter matrices.

We assume that conditions for positivity of the variances hold, as well as that all fourth moments exist. As (5) is a linear system, we can easily obtain the multi-step volatility forecasts $h_t(k) = \text{E}[h_{t+k} \mid I_t]$. These are given recursively by $h_t(1) = h_{t+1}$ and for $k \geq 2$,

$$h_t(k) = \omega + (A + B)h_t(k - 1)$$

Now, it can be shown that the model in (5) has the equivalent representation

$$\eta_t = \sigma + \sum_{i=0}^{\infty} \Phi_i u_{t-i} \tag{6}$$

where $\sigma = (I_3 - A - B)^{-1}\omega$, $u_t = \eta_t - h_t$ is a martingale difference sequence (and therefore white noise), and where the matrices Φ_i are given recursively by $\Phi_0 = I_3$ and $\Phi_i = (A + B)^{i-1}A$, $i \geq 1$. Equation (6) is the vector moving average or VMA(∞) representation of the multivariate GARCH(1,1) process. We can now easily give an expression for the mean squared prediction error as

$$\Sigma_h(k) = \sum_{i=0}^{k-1} \Phi_i \Sigma_u \Phi_i'$$

where $\Sigma_u = \text{E}[u_t u_t']$, for which convenient expressions have been given Hafner (2003), e.g. for the case of Gaussian innovations.

We are now interested in the mean squared prediction error of the variance of the aggregated series z_t. Defining the transformation vector $F = (1, 2, 1)$, one obtains the forecast for the variance of z_t as $Fh_t(k)$. Thus, the mean squared prediction error is

$$\Sigma_h^*(k) = F\Sigma_h(k)F'$$

Alternatively, the first method aggregates the series in a first step and then builds forecasts based on the aggregated

series. For the class of multivariate GARCH models, Nijman and Sentana (1996) have shown that linear combinations are univariate GARCH models, although in general of higher order. Their discussion and Corollary 4.2.2 of Lütkepohl (1987) shows that an upper bound on the order of the model is given by (3, 3). Thus, we can write

$$\psi_t = c + \sum_{i=1}^{3} \alpha_i z_{t-i}^2 + \sum_{j=1}^{3} \beta_j \psi_{t-j}$$

which is a univariate GARCH(3, 3) model. A point to note however is that ψ_t is in general not the conditional variance of z_t but rather the best linear predictor in terms of past values of z_t and z_t^2. This so-called *weak* GARCH model is therefore not providing the best forecast in a mean squared error sense but only the best *linear* predictor. This may be considered as another drawback of the first method. Nevertheless, z_t^2 still has a univariate MA(∞) representation, so that we can easily obtain a predictor for ψ_{t+k} and the mean squared prediction error, which we will denote by $\Sigma_\psi(k)$. The claim is now formally written as $\Sigma_h^*(k) \leq \Sigma_\psi(k)$ for all prediction horizons k.

In the following, we compare the two alternatives by using daily returns of the Deutsche Mark (DEM) and British pound sterling (GBP) versus the US-Dollar (USD) from December 31, 1979 to April 1, 1994, providing 3720 observations, as analyzed by Hafner and Herwartz (2001). They estimate a restricted version of a bivariate GARCH(1,1) model for the returns of DEM/USD and GBP/USD, which can be transformed in a corresponding one for DEM/USD and USD/GBP. For the cross-rate, DEM/GBP, we simply estimate a univariate GARCH(1,1) model, since higher order terms turned out be insignificant. We then calculate the mean squared prediction errors and standardize them with the unconditional variance of the DEM/GBP returns. That is, both mean squared prediction errors will approach 1 as the horizon k goes to infinity. Figure 1 shows the plot of $\Sigma_h^*(k)$ and $\Sigma_\psi(k)$ for k = 1, ..., 50. Clearly, transforming the bivariate volatility forecasts is much better than using the univariate forecasts, especially for small prediction horizons. For this example, the mean squared error can be reduced by about 23% at one-step prediction.

Finally, it should be noted that all the above considerations assume that the stochastic process is known. If parameters need to be estimated the results may change somewhat. The reason is that estimation of the bivariate model involves more parameter uncertainty than estimation of the univariate model.

Conclusion

The objective of this note is to show that in a system of three currencies one can exploit the triangular arbitrage condition for forecasting exchange rate volatility. Only one bivariate model needs to be estimated, and then a linear transformation can be applied to obtain the forecast for the volatility of the cross-rate. This is shown to be more efficient than using the information in the cross-rate only.

About the author

Christian M. Hafner works at the Econometric Institute, Erasmus Universiteit Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

E-mail: chafner@few.eur.nl

References

Engle, R.F., Kroner, K.F. (1995), "Multivariate Simultaneous Generalized ARCH", *Econometric Theory*, 11, 122-150.

Hafner, C.M. (2003) Fourth moment structure of multivariate GARCH models, *Journal of Financial Econometrics*, forthcoming.

Hafner, C.M. & H. Herwartz (2001) Volatility impulse response functions for multivariate GARCH models. CORE discussion paper, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Lütkepohl, H. (1987) *Forecasting aggregated vector ARMA processes*, Lecture Notes in Economics and Mathematical Systems 284, Springer Verlag.

Nijman, T. & Sentana, E. (1996) Marginalization and contemporaneous aggregation in multivariate GARCH processes, *Journal of Econometrics* 71, 71—87.