Journal of Financial Econometrics, 2004, Vol. 2, No. 3, 370–389

Asset Allocation by Variance Sensitivity Analysis

SIMONE MANGANELLI
European Central Bank

ABSTRACT
This article provides a solution to the curse of dimensionality associated to multivariate generalized autoregressive conditionally heteroskedastic (GARCH) estimation. We work with univariate portfolio GARCH models and show how the multivariate dimension of the portfolio allocation problem may be recovered from the univariate approach. The main tool we use is "variance sensitivity analysis," the change in the portfolio variance induced by an infinitesimal change in the portfolio allocation. We suggest a computationally feasible method to find minimum variance portfolios and estimate full variance–covariance matrices. An application to real data portfolios implements our methodology and compares its performance against that of selected popular alternatives.

KEYWORDS: dynamic correlations, multivariate GARCH, risk management

Estimates of volatilities and correlations are used for pricing, asset allocation, hedging purposes, and risk management in general. In today's fast changing financial world, it is essential that these measures are easy to understand and implement. Since their introduction by Engle (1982), autoregressive conditionally heteroskedastic (ARCH) models have been used extensively both in academia and by practitioners to estimate the volatility of financial variables. Many articles have been written on the subject, extending the original ARCH model in many directions. The multivariate extension, however, has been met

The idea of this article came about through several discussions I had with Vladimiro Ceci and Walter Vecchiato. In fact, this article builds on an earlier draft (European Central Bank Working Paper [ECB] no. 194) coauthored with Ceci and Vecchiato, who in the meantime have gone on to other pursuits. In addition to Vladimiro Ceci and Walter Vecchiato, I would like to thank Eric Renault (the editor), an anonymous associate editor, two anonymous referees, Lutz Kilian, Luca Lotti, Neil Shephard, Kevin Sheppard, and the editorial board of the ECB Working Paper Series for useful comments and suggestions. The codes to replicate the results in the article can be downloaded from http://www.simonemanganelli.org. The views expressed in this article are those of the author and do not necessarily reflect those of the ECB.
Address correspondence to Simone Manganelli, DG-Research, European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany, or e-mail: simone.manganelli@ecb.int.

doi:10.1093/jjfinec/nbh015
Journal of Financial Econometrics, Vol. 2, No. 3, © Oxford University Press 2004; all rights reserved.

with many difficulties, mainly due to the fact that the number of parameters that need to be estimated increases exponentially as one leaves the univariate domain.

This article suggests looking at the multivariate problem from a different perspective. The key idea is to work with univariate portfolio models and to develop tools to recover the multivariate dimension that is lost in the univariate estimation. This is accomplished by recognizing that the estimated univariate portfolio variance is a function of the weights of the assets that form the portfolio. By taking the derivatives of the variance with respect to these weights, it is possible to obtain information about the local behavior (around the portfolio weights) of the estimated variance.

Estimation of large multivariate generalized ARCH (GARCH) models is notoriously challenging, requiring strong assumptions to make such estimation feasible. For instance, the most general multivariate GARCH model, the GARCH(1,1) vec representation introduced by Engle and Kroner (1995), requires the estimation of 21 parameters to obtain the variance-covariance matrix of just two assets. With 5 assets, there are 465 parameters to estimate, and with 10 assets the number of parameters rises to 6105! Moreover, restrictions need to be imposed on the variance-covariance matrix to ensure its positive definiteness. It is easy to argue that the high level of parameterization and the assumptions on the structure of the variance-covariance matrix are likely to increase the dangers of misspecification and poor performance of the model.

On the other hand, the advantage of fitting variance models directly to the time series of portfolio returns is that they indirectly incorporate any time-varying correlation among the assets. This makes it possible to estimate parsimonious models that summarize the relevant characteristics of the assets entering the portfolio. This is done, for example, by McNeil and Frey (2000) to calculate the value at risk (VaR) of the portfolio. Leaving aside for the moment theoretical considerations, the main empirical drawback of this approach is that the multivariate dimension of the portfolio allocation problem is lost. Given the estimated variance of a portfolio, a risk manager would be unable to determine how this variance changes as the portfolio composition evolves, or to isolate the main sources of risk. It is not clear how to address these issues in an univariate framework. In the following pages we suggest the use of sensitivity measures to overcome this problem.

Recently measures of sensitivity to the weights of the portfolio allocation have been proposed for VaR models. Garman (1996) suggested computing the derivative of the VaR with respect to the individual components of the portfolio, to assess the potential impact of a trade on a firm's VaR. Gourieroux, Laurent, and Scaillet (2000) study the theoretical implication of this exercise on different VaR models. The same type of question can be asked with respect to the variance of a portfolio. When a full variance-covariance matrix is available, this is a straightforward exercise. But when univariate portfolio variances are estimated it is not obvious how to proceed.

The main contribution of this article is to show how to perform variance sensitivity analysis in the context of univariate GARCH models. We derive the sensitivity of the univariate portfolio GARCH variance to the portfolio weights by analytically computing the derivatives of the estimated GARCH variance with respect to these weights. It is important to recognize that not only the portfolio returns, but also the estimated parameters of the GARCH model are a function of the weights. We show how a simple application of the implicit function theorem to the first-order conditions of the log-likelihood maximization problem can be used to overcome this obstacle.

Our sensitivity measure has many interesting practical applications. To start with, risk managers might use the GARCH sensitivity analysis to test whether their actual portfolio has minimum variance. Indeed, the minimum-variance portfolio is characterized by having all first derivatives with respect to the portfolio weights equal to zero. The GARCH sensitivity analysis could also be used to evaluate the impact that each individual (or group of) asset has on the portfolio variance. This would help risk managers to find out the major sources of risk or allow them to evaluate the impact on the portfolio variance of a certain transaction. A third application, proposed in this article, is a new and computationally feasible method to find minimum-variance portfolios. We show how an allocation problem with $(n+1)$ assets can be solved by minimizing a well-behaved function of n variables, whose first and second derivatives are related to the variance sensitivity of the portfolio. Moreover, by exploiting the analytical relationship among variances, covariances, and the variance derivatives with respect to the portfolio weights, we suggest a simple method to estimate full variance-covariance matrices of large portfolios, which are automatically guaranteed to be positive definite. The intuition is that once the minimum-variance portfolio has been found, estimation of the variance-covariance matrix is equivalent to finding the coefficients of a paraboloid with vertex at the minimum variance portfolio and curvature equal to the second derivatives of the portfolio variances.

The plan of the article is the following. Section 1 illustrates our methodology. Section 2 shows how to employ variance sensitivity analysis to find minimum-variance portfolios and estimate full variance-covariance matrices. Section 3 contains an empirical application. Section 4 concludes.

1 VARIANCE SENSITIVITY ANALYSIS

In this section we show how to compute the derivative of the univariate GARCH portfolio variance with respect to portfolio weights. From a pure theoretical perspective, Nijman and Sentana (1996) have shown that GARCH processes are not closed under contemporaneous (or cross-sectional) aggregation. More precisely, they show that a linear combination of variables generated by a multivariate GARCH process will only be a weak GARCH process. Therefore fitting GARCH processes directly to portfolio returns will generally result in misspecified models. In this article we take a more naïve approach and consider any GARCH model as only a rough approximation of the "true" relationship among the observed data. Our results can nevertheless still be interpreted in the "quasi-maximum-likelihood" sense of White (1994). That is—for the chosen parameterization—they provide the closest approximation (in terms of the Kullback-Leibler information criterion) to the true DGP. We believe this to be a reasonable working assumption (at least as reasonable as assuming that the univariate GARCH processes are correctly specified and then deriving the theoretical true GARCH relationship of the aggregated portfolios).

Changing the portfolio weights changes the time series of portfolio returns and thus changes the information set used in the estimation of the univariate GARCH model. As a consequence, the estimated variance is a function of the portfolio weights, both through the vector of portfolio returns and through the estimated parameters (which obviously depend on the time series of portfolio returns used in estimation). Differentiation of portfolio returns with respect to portfolio weights is straightforward. To differentiate the estimated parameters we appeal to the implicit function theorem. The idea is that since the estimated parameters must satisfy the first-order conditions of the log-likelihood maximization problem, if certain continuity conditions are satisfied, the first-order conditions define an implicit function between the estimated parameters and the portfolio weights.

Let y_t be the return of the portfolio composed by $n+1$ assets and let $y_{t,i}$ be the ith asset return, for $t=1, \ldots, T$ and $i=1, \ldots, n+1$. Indicating the weight of asset i by a_i, the portfolio return at time t is $y_t = \sum_{i=1}^{n+1} a_i y_{t,i}$. Note that since the weights a_i have to sum to one, we can write one weight as a function of the others, $a_{n+1} = 1 - \sum_{i=1}^{n} a_i$.

Assume that y_t is modeled as a zero-mean[1] process with a GARCH(p,q) conditional variance h_t:

$$y_t = \sqrt{h_t}\varepsilon_t \qquad \varepsilon_t|\Omega_t \sim (0,1) \tag{1}$$

$$h_t = z_t'\theta, \tag{2}$$

where

$$\underset{mx1}{z_t} = (1, y_{t-1}^2, \ldots, y_{t-q}^2, h_{t-1}, \ldots, h_{t-p})', \quad \underset{mx1}{\theta} = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)', \quad \text{and}$$
$$m = p + q + 1.$$

The information set of this model is $\Omega_t = \{a, [y_{r,1}]_{r=1}^{t-1}, \ldots, [y_{r,n+1}]_{r=1}^{t-1}\}$, where a denotes the n-vector of portfolio weights.[2] Note that the information set

[1] The zero-mean assumption is made only for the sake of simplicity and implies no loss of generality.

[2] The $(n+1)$ weight is given by one minus the sum of the other weights. The corresponding $(n+1)$ asset is the pivotal asset against which the sensitivity analysis is performed. By changing the pivotal asset, one obtains different sensitivity measures. Computing these sensitivity measures for each single asset of the portfolio, it is possible to compute a matrix of sensitivities analogous to the variance-covariance matrix.

includes the time series of the individual assets returns and that a change in the vector of portfolio weights implies a change in the information set. Therefore, to assess the potential impact of a trade on the estimated variance, one would have to reestimate the whole model, given that a, and hence the information set, has changed. The problem is that such a procedure would quickly become cumbersome and impractical as the number of assets increases.

The potential effect of any change in the portfolio weights on the estimated variance could be evaluated by simply computing the first derivative of the variance with respect to the weights. A positive derivative would indicate that the change will increase the variance of the portfolio and vice versa for a negative derivative. Let $\hat{h}_t = \hat{z}_t'\hat{\theta}$ be the estimated variance, where a hat (ˆ) above a variable denotes that the variable is evaluated at the estimated parameter. In computing the derivative of $\hat{h}_t$, one must recognize that not only the vector $\hat{z}_t$, but also the vector of estimated coefficients $\hat{\theta}$, depends on a. By the product rule, the derivative of $\hat{h}_t$ with respect to a is given by

$$\underset{n\times 1}{\frac{\partial \hat{h}_t}{\partial a}} = \underset{n\times m}{\frac{\partial \hat{z}_t'}{\partial a}} \underset{m\times 1}{\hat{\theta}} + \underset{n\times m}{\frac{\partial \hat{\theta}'}{\partial a}} \underset{m\times 1}{\hat{z}_t}. \tag{3}$$

To achieve a clearer picture of the local behavior of the estimated variance with respect to the portfolio allocation, one could determine its concavity by computing the second derivative,

$$\begin{aligned}\underset{n\times n}{\frac{\partial^2 \hat{h}_t}{\partial a \partial a'}} = {} & \underset{n\times m}{\frac{\partial \hat{z}_t'}{\partial a}} \underset{m\times n}{\frac{\partial \hat{\theta}}{\partial a'}} + \underset{n\times m}{\frac{\partial \hat{\theta}'}{\partial a}} \underset{m\times n}{\frac{\partial \hat{z}_t}{\partial a'}} + \underset{n\times mn}{\left(\underset{1\times m}{\theta'} \otimes \underset{n\times n}{I_n}\right)} \underset{mn\times n}{\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial \hat{z}_t'}{\partial a}\right)} \\ & + \underset{n\times mn}{\left(\underset{1\times m}{\hat{z}_t'} \otimes \underset{n\times n}{I_n}\right)} \underset{mn\times n}{\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial \hat{\theta}'}{\partial a}\right)}\end{aligned} \tag{4}$$

where $\otimes$ denotes the Kronecker product and I_n is an $(n \times n)$ identity matrix.

To evaluate Equations (3) and (4), we need to compute $\frac{\partial \hat{\theta}'}{\partial a}$ and $\frac{\partial}{\partial a'}\mathrm{vec}(\frac{\partial \hat{\theta}'}{\partial a})$, the other terms being easily obtained. We compute these derivatives by applying the implicit function theorem to the first-order conditions of the log-likelihood maximization problem. The first-order conditions for Equations (1) and (2) are:

$$T^{-1}\sum_{t=1}^{T} \frac{\partial l_t(\theta)}{\partial \theta}\bigg|_{\theta=\hat{\theta}} = 0, \tag{5}$$

where $l_t(\cdot)$ is the time t component of the log-likelihood function. The following theorem derives the analytical expressions for $\frac{\partial \hat{\theta}'}{\partial a}$ and $\frac{\partial}{\partial a'}\mathrm{vec}(\frac{\partial \hat{\theta}'}{\partial a})$.

Theorem 1 *Assume that $l_t(\cdot)$ is continuously differentiable in a and θ, and define $\hat{I}_{\theta\theta} \equiv T^{-1}\sum_{t=1}^{T} \frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}\Big|_{\theta=\hat{\theta}}$ and $\hat{I}_{\theta a} \equiv T^{-1}\sum_{t=1}^{T} \frac{\partial^2 l_t(\theta)}{\partial\theta\partial a'}\Big|_{\theta=\hat{\theta}}$. If $\hat{I}_{\theta\theta}$ is nonsingular, then*

$$\underset{n\times m}{\frac{\partial\hat{\theta}'}{\partial a}} = -(\hat{I}_{\theta a})'(\hat{I}_{\theta\theta})^{-1} \tag{6}$$

$$\underset{mn\times n}{\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial\hat{\theta}'}{\partial a}\right)} = -\left[(\hat{I}_{\theta\theta})^{-1}\otimes\frac{\partial\hat{\theta}'}{\partial a}\right]\frac{\partial}{\partial a'}\mathrm{vec}(\hat{I}_{\theta\theta}) - \left[(\hat{I}_{\theta\theta})^{-1}\otimes_n\right]\frac{\partial}{\partial a'}\mathrm{vec}(\hat{I}'_{\theta a}) \tag{7}$$

Proof Since the score is continuous and differentiable both in a and θ, if $\hat{I}_{\theta\theta}$ is nonsingular it is possible to apply the implicit function theorem to the first-order conditions. The result follows. ■

In Appendix B we provide the analytical derivation of Equations (6) and (7), assuming that the standardized residuals are normally distributed—that is, $l_t(\theta) = -\frac{1}{2}\ln(h_t) - \frac{1}{2}y_t^2h_t^{-1}$. Both $\hat{I}_{\theta\theta}$ and $\hat{I}_{\theta a}$ can be easily derived analytically, although the algebra might be messy. Note that $\hat{I}_{\theta\theta}$ is just the estimated Hessian of the GARCH model. One may wonder how it is possible to compute the sensitivity of GARCH variances from the simple series of portfolio returns. In fact, Equations (3) and (4) and theorem 1 make use not only of portfolio returns, but also of the returns of the individual assets entering the information set. We illustrate this point with a simple example. Let $y_t = ay_{t,1} + (1-a)y_{t,2}$, where a and $(1-a)$ are the weights associated with assets 1 and 2, respectively. Suppose that an ARCH(1) model is estimated, so that the parametric form of the estimated variance is $\hat{h}_t = \hat{\theta}y_{t-1}^2$.[3] Then one can show that

$$\hat{I}_{\theta a} \equiv \hat{h}_t^{-2}[y_{t-1}^2 y_t(y_{t,1} - y_{t,2}) + 2y_t^2 y_{t-1}(y_{t-1,1} - y_{t-1,2})]. \tag{8}$$

Hence both Equations (3) and (4) and theorem 1 exploit not only the information contained in $\{y_t\}_{t=1}^T$, but also that contained in the individual series $\{y_{t,1}\}_{t=1}^T$ and $\{y_{t,2}\}_{t=1}^T$.

2 SENSITIVITY ANALYSIS AND ASSET ALLOCATION

The sensitivity analysis idea developed in the previous section can be used for estimating large variance-covariance matrices and for optimal conditional portfolio allocation in a mean-variance context. In this section we discuss how this can be accomplished.

The typical portfolio allocation problem under Markovitz's mean-variance framework can be expressed as follows:

$$\max_a E_t[u(y_t)] = E_t[y_t] - k(\mathrm{var}_t(y_t) + E_t[y_t]^2), \tag{9}$$

[3] We left out the constant for the sake of simplicity.

where $y_t = \sum_{i=1}^{n+1} a_i y_{t,i}$ denotes the portfolio return at time t, a is the n-vector of weights, and E_t and var_t denote, respectively, the conditional expectation and conditional variance at time t, given the information set Ω_t.[4] This problem involves maximizing a function of the conditional mean and the conditional variance with respect to portfolio weights. The first-order conditions of this problem will obviously depend on the derivatives of the conditional mean and the conditional variance with respect to a. Any modeling choice of the conditional mean (including a possible GARCH in mean component) can be easily incorporated into the GARCH framework of Equations (1) and (2). The first and second derivatives of the mean and the variance of the portfolio returns with respect to the weights can therefore be derived from the corresponding likelihood using exactly the same procedure described in Section 1. This implies that the maximization problem of Equation (9) can be solved by simply maximizing a function of n variables (the portfolio weights), of which we know the first and second derivatives. This is a relatively straightforward task—provided that the function is sufficiently well behaved—and does not involve the estimation of any variance-covariance matrix. If the portfolio GARCH models were correctly specified, the function would be quadratic in the weights, and the corresponding optimization process would be trivial. In reality, as discussed at the beginning of Section 1, univariate portfolio GARCH models are bound to be misspecified. The degree of misspecification, and its impact on the shape of our function, is impossible to tell on pure theoretical grounds. In the empirical application of Section 3, we find the degree of misspecification to be relatively unimportant. Optimization for a portfolio of 30 assets is extremely robust to the choice of the initial conditions, suggesting a sufficiently well-behaved objective function. Obviously further research along these lines is needed to tell how generally our procedure can be applied.

Sensitivity analysis can be used to estimate large variance-covariance matrices as well. The procedure involves three simple steps:

> Step 1. Minimize the portfolio variance with respect to the weights. This is just a special case of Equation (9), when we set the conditional mean equal to zero. For daily observations, this can be considered a reasonable approximation.
>
> Step 2. Compute the second derivatives of portfolio variance with respect to the weights. In theory, second derivatives should be constant and independent of the values of the weights at which they are computed. In practice, this will not be the case because of the misspecification of the GARCH(1,1) model applied to portfolios. One possibility is to take averages of second derivatives computed for different portfolio weights, or simply to take the second derivative corresponding to the minimum-variance portfolio computed in Step 1.

[4] As usual, we assume that the weights sum to one: $a_{n+1} = 1 - \sum_{i=1}^{n} a_i$.

Step 3. Define $w=[a', 1-\iota' a]'$ as the $(n+1)$-vector of weights corresponding to each asset entering the portfolio, where ι is an n-vector of ones. Compute the variance-covariance matrix (Σ_t) as the solution to the following system:

(i) $w^{*\prime}\Sigma_t w^* = h_t(a^*)$

(ii) $\left.\dfrac{\partial w'\Sigma_t w}{\partial a}\right|_{a=a^*} = 0$

(iii) $\left.\left(\dfrac{\partial^2 w'\Sigma_t w}{\partial a \partial a'}\right)\right|_{a=a^*} = K_t,$

where K_t is an $(n \times n)$ matrix containing the estimated second derivatives and a^*, w^* are the optimal weights associated to the minimum-variance portfolio found in Step 1.

Note that Σ_t has a total of $(n+2)(n+1)/2$ parameters to be estimated. Condition (i) gives one parameter, condition (ii) n parameters, and conditions (iii) $n(n+1)/2$ parameters, for a total of exactly $(n+2)(n+1)/2$ parameters. Note also that the solution to conditions (i)–(iii) gives the coefficients of a paraboloid with vertex corresponding to the minimum-variance portfolio and curvature K_t. This implies, since the minimum variance is strictly positive, that the estimated variance-covariance matrix is automatically positive definite. To see this, consider the generic quadratic form $w'\hat{\Sigma}_t w$, where $\hat{\Sigma}_t$ has been estimated with the above procedure and $w'\begin{bmatrix} \iota \\ 1 \end{bmatrix} = 1$. By conditions (ii) and (iii), this quadratic form achieves a minimum at $w=w^*$, and condition (i) ensures that this minimum is strictly positive. It follows that $w'\hat{\Sigma}_t w > 0$ for any w.

It is straightforward to find an analytical solution to the above system. Partition Σ_t as follows: $\Sigma_t = \begin{bmatrix} A_t & b_t \\ b_t' & c_t \end{bmatrix}$, where A_t is $(n \times n)$, b_t is $(n \times 1)$, and c_t is a scalar. Then

$$\begin{aligned} w'\Sigma_t w &= [a', 1-\iota' a]\begin{bmatrix} A_t & b_t \\ b_t' & c_t \end{bmatrix}\begin{bmatrix} a \\ 1-\iota' a \end{bmatrix} \\ &= a'(A_t - \iota b_t' - b_t \iota' + c_t \iota\iota')a + 2(b_t' - c_t \iota')a + c_t. \end{aligned} \tag{10}$$

Therefore, we have

$$\frac{\partial w'\Sigma_t w}{\partial a} = 2(A_t - \iota b_t' - b_t \iota' + c_t \iota\iota')a + 2(b_t - c_t \iota) \tag{11}$$

$$\frac{\partial^2 w'\Sigma_t w}{\partial a \partial a'} = 2(A_t - \iota b_t' - b_t \iota' + c_t \iota\iota'). \tag{12}$$

Combining these relationships with conditions (i)–(iii), simple calculations give

$$A_t = 0.5K_t + h_t(a^*)\iota\iota' + 0.5a^{*\prime}K_t a^* \iota\iota' - 0.5\iota a^{*\prime}K_t - 0.5K_t a^{*\prime}\iota \tag{13}$$

$$b_t = h_t(a^*)\iota + 0.5a^{*\prime}K_t a^* \iota - 0.5K_t a^* \tag{14}$$

$$c_t = h_t(a^*) + 0.5a^{*\prime}K_t a^*, \tag{15}$$

which completely determine the variance-covariance matrix at time t.

3 EMPIRICAL APPLICATION

In this section we implement our methodology on a selected sample of stocks. We first estimate the sensitivity of GARCH variances on a two-stock portfolio, as described in Section 1. Then we find the minimum-variance portfolios for five different portfolios using the methodology outlined in Section 2, and compare its performance to that of a few popular alternatives (namely, dynamic conditional correlation, orthogonal GARCH, and exponentially weighted moving average).

3.1 Sensitivity of GARCH Variance

We estimate the first and second derivatives of GARCH variances as described in Section 1, using a two-asset portfolio composed of General Motors (GM) and IBM. Daily data are taken from Bloomberg and run from January 2, 1992, through March 11, 2002.

We estimate univariate GARCH(1,1) models for 31 portfolios constructed from these two assets, with the GM weight (a) ranging from -1 to 2, with increments of 0.1. For each estimated GARCH model we compute the first and second derivatives of the estimated variance with respect to the weight a.

In Figure 1 we plot the estimated variances on March 11, 2002, for the 31 portfolios as a function of the weight, together with their first and second derivatives. Note that the variance corresponding to $a=0$ is the variance of IBM, while the variance corresponding to $a=1$ is the variance of GM. The portfolios with a weight greater than one or less than zero are short on IBM or GM, respectively. The estimated variance plotted in Figure 1 is a parabolic and convex function of the portfolio weights a, suggesting that diversification produces

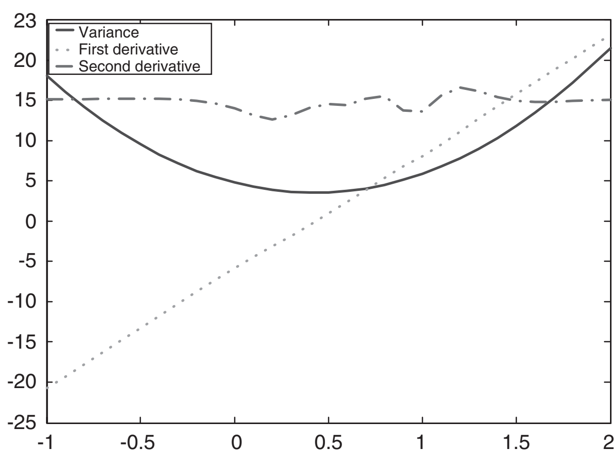

Figure 1 Plot of estimated variance, first and second derivative on March 11, 2002, for 31 portfolios constructed from GM and IBM. On the horizontal axis is the portfolio weight for GM, which ranges from -1 to 2, with increments of 0.1. The variance is computed by reestimating a GARCH(1,1) model for each of the 31 portfolios. The first and second derivatives are computed analytically, as described in Section 1.

significant gains in terms of risk reduction. If the true variance-covariance matrix were available and one computes the portfolio variances as a weighted sum of the individual asset variances and covariances, this function would be exactly a parabola. The fact that fitting univariate GARCH models to the time series of portfolios produces results very close to those one would expect in theory indicates that these univariate GARCH models provide a reasonable approximation of the true (but unknown) model. This intuition is confirmed by the shape of the first and second derivatives. If the function were truly a parabola, then the first derivative would be a straight, positively sloped line and the second derivative a flat line. The plots in Figure 1 show that both the first and second derivatives are very close to their theoretical shape.

In Figure 2 we show the time series of the first derivatives of the estimated variance, $\frac{\partial \hat{h}_t(a)}{\partial a}$, for the two degenerate portfolios, that is, for IBM ($a = 0$) and GM ($a = 1$). The picture indicates by how much the variance would decrease or increase over time if one diversifies away from the portfolios composed of only GM or IBM. Similar pictures can be drawn for any portfolio weight, thus giving the risk manager a precise indication about the consequences—in terms of risk—of changing the composition of the current portfolio.

A second interesting feature of Figure 2 is that the first derivative is always positive for GM and almost always negative for IBM. This implies that the minimum-variance portfolio during the period considered in this analysis was formed by a convex combination of these two assets. The fact that for a few days

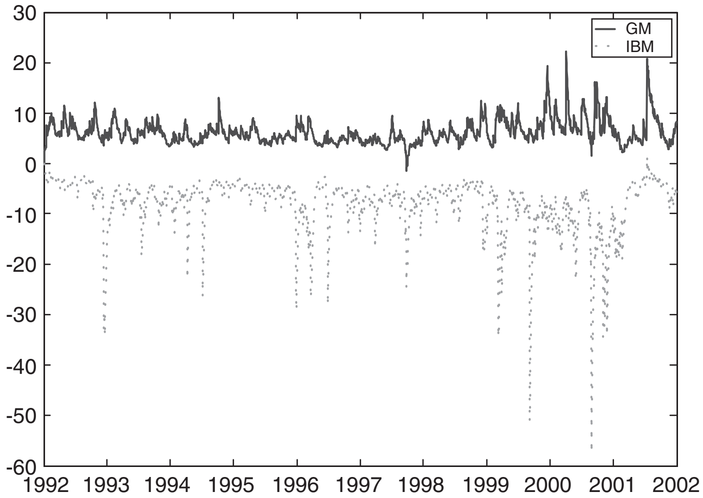

Figure 2 Plot of the first derivative of the estimated variance for the two degenerate portfolios of GM and IBM. The first derivative indicates by how much the variance would increase or decrease over time if one increases the weight of GM (a) in the portfolio. In the case where the portfolio is composed of only GM ($a = 1$), buying an extra share of GM and going short on IBM will increase the overall portfolio variance (upper line). Vice versa, diversifying away from a portfolio composed of only IBM ($a = 0$) would decrease the variance (bottom line).

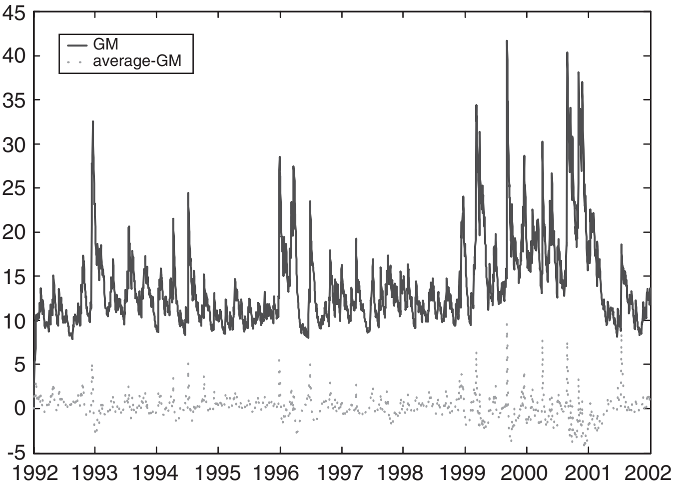

Figure 3 Plot of the estimated second derivatives, computed from the degenerate GM portfolio (upper graph) and its difference w.r.t. the average second derivatives of 31 different portfolios (lower graph). Under correct model specification, the second derivative should not depend on the portfolio weight. Hence the lower graph should be exactly equal to zero. The fact that the dotted line hovers mostly around zero confirms the results obtained in Figure 1, that is, the GARCH(1,1) model provides a reasonable approximation of the portfolio variance process.

toward the end of the sample both first derivatives were positive signals that during those days the risk manager would have had to short GM to construct the minimum-variance portfolio.

Figure 2 also provides insight into the major sources of risk in a portfolio. Indeed, the greater (in absolute value) the first derivative, the greater the risk reduction following a portfolio reallocation. Figure 2 shows that the first derivative of the portfolio containing only the IBM asset is much higher on average (in absolute value) than the first derivative corresponding to the GM portfolio.[5] This implies that during the 1990s an investor could achieve greater variance reduction by diversifying away from the portfolio with only IBM (the "new economy" stock) than from the GM portfolio (the "old economy" stock). In the case of a portfolio with more than two assets, one could compute the variance sensitivity corresponding to each asset and gain in this way an insight about the major sources of risk in the portfolio. In order to reduce the risk, the risk manager should sell the assets with the highest first derivative and buy those with the lowest one.

In Figure 3 we report the time series of the second derivatives for GM, together with its difference from the average second derivatives computed over all the 31 portfolios considered in this exercise. In theory, for a correctly specified model, the second derivative should not depend on the portfolio composition: when plotted against the portfolio weights at any given point in time, it should be

[5] The average first derivative for IBM is −7.86 and for GM is 6.26. That is, the variance sensitivity of the portfolio containing only IBM was about 25% higher than that of the GM portfolio.

a flat line. Therefore the difference between the average second derivatives and that of GM should be zero. Consistent with the findings of Figure 1, this difference is hovering around zero throughout the sample, thus providing further evidence that portfolio univariate GARCH models provide a good approximation of the true variance. The difference is not exactly zero because, as we noted in Section 1, univariate GARCH models are surely misspecified.

The second derivative, being the slope of the first derivative, tells the risk manager by how much the variance sensitivity will change after a change in the portfolio allocation. The greater the magnitude of the second derivative, the greater the change in the variance sensitivity, implying that a smaller portfolio reallocation will be necessary to achieve a given size of variance reduction. Figure 3 shows that in the last couple of years portfolio reallocations had a much greater impact on variance than during the 1990s. The average of the second derivative was 12.36 between 1992 and 1999, rising to 17.17 from 1999 to 2002. In other words, these results show that the concavity of the portfolio variance (as a function of weight a) has increased dramatically over the past few years for GM and IBM. This has obviously important consequences for managing the risk of a portfolio composed of these two assets.

3.2 Minimum-Variance Portfolio Allocation

In this subsection we implement the methodology described in Section 2 to find the allocation minimizing the portfolio variance. We tested our methodology on different subsamples of the Dow Jones index, with the same time span as before, that is, from January 2, 1992, through March 11, 2002.

We compare the performance of our methodology with three alternative multivariate models: dynamic conditional correlation (DCC), Orthogonal GARCH (OGARCH), and exponentially weighted moving average (EWMA). These are some of the most popular methods used in the industry to estimate large variance-covariance matrices.[6] The DCC model was recently proposed by Engle and Sheppard (2001) and Engle (2002). This model can be seen as a generalization of the constant conditional correlation model, originally proposed by Bollerslev (1990). In the DCC model, conditional correlations are directly parameterized rather than assumed constant. Engle (2002) shows that the estimation of the multivariate model can be drastically simplified by using a two-step procedure. First, the univariate GARCH models are estimated for each of the assets. Then the conditional correlation specification is fitted to the standardized residuals obtained in the first step. In the simplest mean-reverting model proposed by Engle (2002)—which is the one we estimate in our empirical application—the same pair of parameters is estimated for all the correlations considered (implying that all correlations have the same degree of persistence). More parameterized models can be used, at the cost, however, of increasing the computational burden.

[6] Ledoit, Santa-Clara, and Wolf (2003) and Brandt and Diebold (2004) provide two interesting alternative strategies to variance-covariance matrix estimation that will not be discussed here.

Alexander (2000) proposed the OGARCH model, based on a principal component GARCH methodology. First, she constructs unconditionally uncorrelated factors, which are linear combinations of the original returns. Then she fits univariate GARCH models to the principal components. Since the conditional variance-covariance matrix of the principal component series is diagonal (i.e., conditional correlations are zero), it is possible to recover the original assets' variance-covariance matrix through a fixed mapping matrix.

The last method we consider is the EWMA, popularized by RiskMetrics. With this method, the variance-covariance matrix at time t is simply computed as a convex combination of the variance-covariance matrix in the previous period, $t-1$, and the matrix of squared and cross-product lagged returns. The weight is usually set equal to 0.94 or 0.97. In the following application we set the decay coefficient for the EWMA equal to 0.94.

For these three methods we first estimate the variance-covariance matrix as of March 11, 2002. Then we analytically compute the weights that give the minimum-variance portfolio. If the variance-covariance matrix is partitioned as in Equation (10), then the optimal weights can be found by setting Equation (11) equal to zero, giving

$$a^* = (\hat{A}_t - \iota\hat{b}_t' - \hat{b}_t\iota' + \hat{c}_t\iota\iota')^{-1}(\hat{c}_t\iota - \hat{b}_t). \tag{16}$$

Finally, we estimate the univariate GARCH variance associated to this portfolio and report the annualized estimated volatility in Table 1.[7] We repeat this procedure for each model under consideration and for five different portfolios with 2, 5, 10, 20, and 30 assets, respectively.[8]

The variance sensitivity analysis (VSA) model is estimated by minimizing directly the univariate GARCH variance with respect to portfolio weights. We use the function *fminunc* in Matlab, providing as input the first and second analytical derivatives computed in Section 1. Convergence is very rapid and very robust to the choice of the initial conditions.[9] This suggests that the objective function is well behaved even for high-dimensional problems. To produce the results in Table 1, we chose as initial conditions of the VSA model the optimal weights of EWMA.

Since the VSA method is designed to find the minimum-variance portfolio as measured by univariate GARCH, it is not surprising that it outperforms all the other models. What is more interesting is that the outperformance (as measured by the percentage difference in annualized volatility) increases with the number of assets. While with a two-asset portfolio the differences in the minimum variances are negligible, these differences rise monotonically with the number of assets for the DCC and OGARCH methods. With five assets, DCC and OGARCH overestimate the minimum-variance portfolio by about 7%. With 10 assets, the difference rises to 10% and 18%, respectively, while with 20 and 30 assets it ranges from

---

[7] Annualized volatility is given by the square root of (252* daily variance).

[8] Assets are progressively aggregated in the order reported in Appendix A.

[9] Convergence for a 30-asset portfolio occurs in less than 15 iterations for randomly chosen initial conditions.

Table 1 Comparison between the VSA methodology and the most popular multivariate GARCH models, as of March 11, 2002.

	Portfolio with 2 assets			Portfolio with 5 assets			Portfolio with 10 assets			Portfolio with 20 assets			Portfolio with 30 assets		
	Min. vol.	% VSA	Seconds	Min. vol.	% VSA	Seconds	Min. vol.	% VSA	Seconds	Min. vol.	% VSA	Seconds	Min. vol.	% VSA	Seconds
DCC	30.64%	0.01	9	28.26%	7.20	16	18.77%	10.19	121	16.35%	23.06	185	12.92%	25.99	351
OGARCH	30.68%	0.15	5	28.06%	6.45	10	20.15%	18.34	18	17.55%	32.06	35	14.23%	38.84	53
EWMA	30.66%	0.09	2	27.53%	4.42	2	17.69%	3.86	2	15.17%	14.20	2	11.00%	7.34	2
VSA	30.64%	—	71	26.36%	—	214	17.03%	—	105	13.29%	—	169	10.25%	—	465

DCC, dynamic conditional correlation; OGARCH, orthogonal GARCH; EWMA, exponentially weighted moving average.
For each portfolio we report the univariate GARCH annualized volatility associated with the minimum-variance weights implied by the estimated variance-covariance matrix, the percentage difference w.r.t. VSA, and the computation time to estimate the model.

23% to almost 40%. Clearly this is an indication that as the number of assets increases, the restrictions imposed by the standard multivariate GARCH models become more and more binding.

A different picture emerges from the results of the EWMA. Its performance does not seem to deteriorate as much and as fast as the other two competing models. Although surprising at first sight, this result is due to the way EWMA estimates are constructed. Using the same decay coefficient λ for all variance and covariance terms is equivalent to estimating directly the EWMA portfolio variance with coefficient λ. If the estimated coefficients of the GARCH(1,1) model applied to the minimum EWMA variance portfolio are not too different from the chosen λ, one would expect good performance from this model—whenever performance is measured by minimum variance. Notice, however, that this does not imply that EWMA produces reasonable estimates of the variance-covariance matrix. As we illustrated in Section 2, estimation of variance-covariance matrices is equivalent to finding the coefficients of a paraboloid with given vertex and curvature. Even if the vertex provided by EWMA is close to the true one, there is no guarantee that the same holds for the curvature.

A final aspect worth noticing is that the computation time of VSA does not seem to rise too much with the number of assets.[10] A 30-asset portfolio is optimized by VSA in less than 8 minutes on a standard Pentium IV computer. Computational feasibility seems to be a further attractive feature of the methodology proposed in this article.

4 CONCLUSION

Fitting variance models directly to the time series of portfolio returns has many advantages, such as the possibility of estimating parsimonious models and computational tractability. The problem of this strategy is that the multivariate dimension of the portfolio allocation is lost. This article suggests a strategy to overcome this problem, working within a GARCH framework. We assess the potential impact of a trade on the estimated variance by computing the sensitivity of the estimated variance with respect to the weight of the asset involved in the trade. This sensitivity measure is simply the derivative of the estimated variance with respect to portfolio weights. As a by-product of this analysis, we propose a new and simple method to estimate full variance-covariance matrices, which exploits the analytical relationship among variances, covariances, and sensitivity measures.

We illustrate the functioning and the performance of our methodology with two empirical applications. In the first one we estimate the variance sensitivity for a portfolio of two assets. We document how this sensitivity has been changing over time and stress its implications for risk management. We also compute the second derivative of the estimated variance with respect to portfolio weights. We

[10] Computation time of VSA actually decreases when moving from portfolios with 5 assets to portfolios with 10 assets. This result, however, is probably due to the quality of the initial conditions fed to the optimization algorithm.

argue that this measure gives an indication of the diversification opportunities at any given point in time: the higher the second derivative, the greater the gains (in terms of variance reduction) from a proper diversification strategy.

In the second application we implement the suggested methodology to find minimum variance portfolios at any given point in time. We test the model with five subsamples of the Dow Jones index. We evaluate the performance of our methodology against that of popular alternatives, including the DCC, the OGARCH, and the EWMA models. Our findings suggest that our methodology generally leads to considerable efficiency gains.

APPENDIX A

Table A.1 List of stocks used in the analysis

	Ticker	Name
1	AA	Alcoa Inc.
2	AXP	American Express Co.
3	T	AT&T Corp.
4	BA	Boeing Co.
5	CAT	Caterpillar Inc.
6	C	Citigroup Inc.
7	KO	Coca-Cola Co.
8	DIS	Walt Disney Co.
9	DD	E.I. DuPont de Nemours & Co.
10	EK	Eastman Kodak Co.
11	XOM	Exxon Mobil Corp.
12	GE	General Electric Co.
13	GM	General Motors Corp.
14	HPQ	Hewlett-Packard Co.
15	HD	Home Depot Inc.
16	HON	Honeywell International Inc.
17	INTC	Intel Corp.
18	IBM	International Business Machines Corp.
19	IP	International Paper Co.
20	JNJ	Johnson & Johnson
21	JPM	J.P. Morgan Chase & Co.
22	MCD	McDonald's Corp.
23	MRK	Merck & Co.
24	MSFT	Microsoft Corp.
25	MMM	3M Co.
26	MO	Altria Group Inc.
27	PG	Procter & Gamble
28	SBC	SBC Communications Inc.
29	UTX	United Technologies Corp.
30	WMT	Wal-Mart Stores Inc.

These were the 30 stocks composing the Dow Jones Industrial Average index as of March 11, 2002.

APPENDIX B: COMPUTATION OF FIRST AND SECOND DERIVATIVES UNDER NORMAL LIKELIHOOD FOR A GARCH(1,1) MODEL

We follow the conventions and rules on matrix differentiation, as in Appendix A.13 of Lütkepohl (1990). We indicate the explicit dependence of θ on a by $\theta(a)$. Whenever this dependence is not made explicit, it means that we treat θ as not depending on a. We derive the analytical expressions of Equations (3) and (4), under the assumption of a normal likelihood. Rewrite Equation (3), making explicit the dependence on a:

$$\frac{\partial h_t(\theta(a))}{\partial a} = \frac{\partial z_t(\theta(a))'}{\partial a}\theta + \frac{\partial \theta(a)'}{\partial a} z_t, \tag{A.1}$$

where $\frac{\partial z_t(\theta(a))'}{\partial a} = \begin{bmatrix} 0_{(n\times 1)} & \frac{\partial Y_{t-1}^2}{\partial a} & \frac{\partial h_{t-1}(\theta(a))}{\partial a} \end{bmatrix}$, $\frac{\partial h_{t-1}(\theta(a))}{\partial a}$ can be computed recursively, $\frac{\partial Y_{t-1}^2}{\partial a} = 2Y_{t-1}(y_{t-1\backslash n+1} - \iota y_{t-1,n+1})$, ι is an $(n, 1)$-vector of ones, and $y_{t\backslash n+1} = [y_{t,1}, \ldots, y_{t,n}]'$. To compute $\frac{\partial \theta(a)'}{\partial a}$ we apply Theorem 1, under the assumption that $l_t = -0.5[\ln(h_t) + Y_t^2 h_t^{-1}]$. The score is

$$\frac{\partial l_t}{\partial \theta} = -0.5\frac{\partial h_t}{\partial \theta} H, \tag{A.2}$$

where $H \equiv (h_t^{-1} - Y_t^2 h_t^{-2})$, $\frac{\partial h_t}{\partial \theta} = z_t + \frac{\partial z_t'}{\partial \theta}\theta$, and $\frac{\partial z_t'}{\partial \theta} = \begin{bmatrix} 0_{(3\times 2)} & \frac{\partial h_{t-1}}{\partial \theta} \end{bmatrix}$. Therefore

$$\frac{\partial^2 l_t}{\partial \theta \partial \theta'} = -0.5\left[\frac{\partial h_t}{\partial \theta}\frac{\partial H}{\partial \theta'} + H\frac{\partial^2 h_t}{\partial \theta \partial \theta'}\right], \tag{A.3}$$

where $\frac{\partial H}{\partial \theta'} = \frac{\partial h_t}{\partial \theta'}\tilde{H}$, $\quad \tilde{H} \equiv -h_t^{-2} + 2Y_t^2 h_t^{-3}$, $\frac{\partial^2 h_t}{\partial \theta \partial \theta'} = \frac{\partial z_t}{\partial \theta'} + \frac{\partial z_t'}{\partial \theta} + (\theta' \otimes I_3)\frac{\partial}{\partial \theta'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial \theta}\right)$ and $\frac{\partial}{\partial \theta'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial \theta}\right) = \begin{bmatrix} 0_{(6\times 3)} \\ \frac{\partial^2 h_{t-1}}{\partial \theta \partial \theta'} \end{bmatrix}$;

$$\frac{\partial^2 l_t}{\partial \theta \partial a'} = -0.5\left[\frac{\partial h_t}{\partial \theta}\frac{\partial H}{\partial a'} + H\frac{\partial^2 h_t}{\partial \theta \partial a'}\right], \tag{A.4}$$

where $\frac{\partial H}{\partial a'} = \frac{\partial h_t}{\partial a'}\tilde{H} - h_t^{-2}\frac{\partial Y_t^2}{\partial a'}$, $\quad \frac{\partial h_t}{\partial a'} = \theta'\frac{\partial z_t}{\partial a'}$, $\quad \frac{\partial z_t}{\partial a'} = \begin{bmatrix} 0_{(1\times n)} \\ \frac{\partial Y_{t-1}^2}{\partial a'} \\ \frac{\partial h_{t-1}}{\partial a'} \end{bmatrix}$, $\quad \frac{\partial^2 h_t}{\partial \theta \partial a'} = \frac{\partial z_t}{\partial a'} + (\theta' \otimes I_3)\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial \theta}\right)$, and $\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial \theta}\right) = \begin{bmatrix} 0_{(6\times n)} \\ \frac{\partial^2 h_{t-1}}{\partial \theta \partial a'} \end{bmatrix}$.

The second derivative is given by Equation (4):

$$\begin{aligned}\frac{\partial^2 h_t(\theta(a))}{\partial a \partial a'} &= \frac{\partial z_t(\theta(a))'}{\partial a}\frac{\partial \theta(a)}{\partial a'} + \frac{\partial \theta(a)'}{\partial a}\frac{\partial z_t(\theta(a))}{\partial a'} \\ &\quad + (\theta' \otimes I_n)\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t(\theta(a))'}{\partial a}\right) + (z_t(\theta(a))' \otimes I_n)\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial \theta(a)'}{\partial a}\right),\end{aligned} \tag{A.5}$$

where $\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t(\theta(a))'}{\partial a}\right) = \begin{bmatrix} 0_{(n\times n)} \\ \frac{\partial^2 Y_{t-1}^2}{\partial a \partial a'} \\ \frac{\partial^2 h_{t-1}(\theta(a))}{\partial a \partial a'} \end{bmatrix}$, $\frac{\partial Y_{t-1}^2}{\partial a \partial a'} = 2(y_{t-1\backslash n+1} - \iota y_{t-1,n+1}),(y_{t-1\backslash n+1} - \iota y_{t-1,n+1})'$, and $\frac{\partial^2 h_{t-1}(\theta(a))}{\partial a \partial a'}$ can be computed recursively. To compute $\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial \theta(a)'}{\partial a}\right)$, we again apply Theorem 1 and note that, if A is a (p, p) symmetric nonsingular matrix, $\frac{\partial \mathrm{vec}(A^{-1})}{\partial a'} = -(A^{-1}\otimes A^{-1})\frac{\partial \mathrm{vec}(A)}{\partial a'}$:

$$\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial \theta(a)'}{\partial a}\right) = -\left[(I_{\theta\theta})^{-1}\otimes \frac{\partial \theta(a)'}{\partial a}\right]\frac{\partial}{\partial a'}\mathrm{vec}[I_{\theta\theta}(\theta(a))] \\ - [(I_{\theta\theta})^{-1}\otimes I_n]\frac{\partial}{\partial a'}\mathrm{vec}[I'_{\theta a}(\theta(a))]. \quad (A.6)$$

We compute the individual components of $\frac{\partial}{\partial a'}\mathrm{vec}[I_{\theta\theta}(\theta(a))]$ and $\frac{\partial}{\partial a'}\mathrm{vec}[I'_{\theta a}(\theta(a))]$, in turn:

$$\frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 l_t}{\partial\theta\partial\theta'}(\theta(a))\right] = -0.5\left\{\left(I_3\otimes\frac{\partial h_t}{\partial\theta}\right)\frac{\partial}{\partial a'}\left(\frac{\partial H}{\partial\theta}(\theta(a))\right) + \left(\frac{\partial H}{\partial\theta}\otimes I_3\right)\frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial\theta}(\theta(a))\right)\right. \\ \left. + \mathrm{vec}\left(\frac{\partial^2 h_t}{\partial\theta\partial\theta'}\right)\frac{\partial H(\theta(a))}{\partial a'} + H\frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 h_t}{\partial\theta\partial\theta'}(\theta(a))\right]\right\}, \quad (A.7)$$

where

- $\frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial\theta}(\theta(a))\right) = \frac{\partial z_t(\theta(a))}{\partial a'} + \frac{\partial z_t'}{\partial\theta}\frac{\partial\theta(a)}{\partial a'} + (\theta'\otimes I_3)\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial\theta}(\theta(a))\right)$
- $\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial\theta}(\theta(a))\right) = \begin{bmatrix} 0_{(6\times n)} \\ \frac{\partial}{\partial a'}\left(\frac{\partial h_{t-1}}{\partial\theta}(\theta(a))\right) \end{bmatrix}$
- $\frac{\partial H(\theta(a))}{\partial a'} = \frac{\partial h_t(\theta(a))}{\partial a'}\tilde{H} - h_t^{-2}\frac{\partial Y_t^2}{\partial a'}$
- $\frac{\partial}{\partial a'}\left(\frac{\partial H}{\partial\theta}(\theta(a))\right) = \frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial\theta}(\theta(a))\right)\tilde{H} + \frac{\partial h_t}{\partial\theta}\frac{\partial\tilde{H}(\theta(a))}{\partial a'}$
- $\frac{\partial\tilde{H}(\theta(a))}{\partial a'} = (2h_t^{-3} - 6Y_t^2 h_t^{-4})\frac{\partial h_t(\theta(a))}{\partial a'} + 2h_t^{-3}\frac{\partial Y_t^2}{\partial a'}$
- $\frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial^2 h_t}{\partial\theta\partial\theta'}(\theta(a))] = \frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial z_t'}{\partial\theta}(\theta(a))] + \frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial z_t}{\partial\theta'}(\theta(a))] + \frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial}{\partial\theta}\mathrm{vec}(\frac{\partial z_t}{\partial\theta'})'(I_3\otimes\theta)]$
- $\frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial z_t}{\partial\theta'}(\theta(a))] = K_{3,3}[\frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial z_t'}{\partial\theta}(\theta(a))]]$, where $K_{3,3}$ is the commutation matrix [see Lütkepohl (1990: 466)]
- $\frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial}{\partial\theta}\mathrm{vec}(\frac{\partial z_t}{\partial\theta'})'(I_3\otimes\theta)] = (I_3\otimes G')\frac{\partial}{\partial a'}\mathrm{vec}(I_3\otimes\theta) + (I_3\otimes\theta'\otimes I_3)\frac{\partial}{\partial a'}\mathrm{vec}(G')$
- $G = \begin{bmatrix} 0_{(6\times 3)} \\ \frac{\partial^2 h_{t-1}}{\partial\theta\partial\theta'} \end{bmatrix}$
- $\frac{\partial}{\partial a'}\mathrm{vec}(I_3\otimes\theta) = (I_3\otimes K_{1,3}\otimes I_3)\left(\mathrm{vec}(I_3)\otimes\frac{\partial\theta(a)}{\partial a'}\right)$
- $\frac{\partial}{\partial a'}\mathrm{vec}(G') = \begin{bmatrix} 0_{(18\times n)} \\ \frac{\partial}{\partial a'}\mathrm{vec}[\frac{\partial^2 h_t}{\partial\theta\partial\theta'}(\theta(a))] \end{bmatrix}$.

It remains to compute $\frac{\partial}{\partial a'}\mathrm{vec}\lfloor I'_{\theta a}(\theta(a))\rfloor$ in Equation (A.6):

$$\frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 l_t}{\partial\theta'\partial a}(\theta(a))\right] = -0.5\left\{\left(I_3\otimes\frac{\partial H}{\partial a}\right)\frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial\theta}(\theta(a))\right)+\left(\frac{\partial h_t}{\partial\theta}\otimes I_n\right)\frac{\partial}{\partial a'}\left(\frac{\partial H}{\partial a}(\theta(a))\right)\right.$$
$$\left.+\mathrm{vec}\left(\frac{\partial^2 h_t}{\partial\theta'\partial a}\right)\frac{\partial H(\theta(a))}{\partial a'}+H\frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 h_t}{\partial\theta'\partial a}(\theta(a))\right]\right\}, \quad \text{(A.8)}$$

where

- $\frac{\partial}{\partial a'}\left(\frac{\partial H}{\partial a}(\theta(a))\right) = \tilde{H}\frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial a}(\theta(a))\right) + \frac{\partial h_t}{\partial a}\frac{\partial \tilde{H}(\theta(a))}{\partial a'} - h_t^{-2}\frac{\partial^2 Y_t^2}{\partial a\partial a'} + 2h_t^{-3}\frac{\partial Y_t^2}{\partial a}\frac{\partial h_t(\theta(a))}{\partial a'}$
- $\frac{\partial}{\partial a'}\left(\frac{\partial h_t}{\partial a}(\theta(a))\right) = \frac{\partial z_t'}{\partial a}\frac{\partial\theta(a)}{\partial a'} + (\theta'\otimes I_n)\frac{\partial}{\partial a'}\left(\frac{\partial z_t'}{\partial a}(\theta(a))\right)$
- $\frac{\partial}{\partial a'}\mathrm{vec}\left(\frac{\partial z_t'}{\partial a}(\theta(a))\right) = \begin{bmatrix} 0_{(n\times n)} \\ \frac{\partial^2 Y_{t-1}^2}{\partial a\partial a'} \\ \frac{\partial}{\partial a'}\left[\frac{\partial h_{t-1}}{\partial a}(\theta(a))\right] \end{bmatrix}$
- $\frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 h_t}{\partial\theta'\partial a}(\theta(a))\right] = \frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial z_t'}{\partial\theta}(\theta(a))\right] + \frac{\partial}{\partial a'}\mathrm{vec}[F(\theta(a))'(\theta(a)\otimes I_3)]$
- $F = \begin{bmatrix} 0_{(6\times n)} \\ \frac{\partial^2 h_{t-1}}{\partial\theta\partial a'} \end{bmatrix}$
- $\frac{\partial}{\partial a'}\mathrm{vec}\lfloor F(\theta(a))'(\theta(a)\otimes I_3)\rfloor = (I_3\otimes F')\frac{\partial}{\partial a'}\mathrm{vec}(\theta(a)\otimes I_3) + (\theta'\otimes I_3\otimes I_n)\times \frac{\partial}{\partial a'}\mathrm{vec}\lfloor F(\theta(a))'\rfloor$
- $\frac{\partial}{\partial a'}\mathrm{vec}(\theta(a)\otimes I_3) = (K_{3,3}\otimes I_3)\left(\frac{\partial\theta(a)}{\partial a'}\otimes\mathrm{vec}(I_3)\right)$
- $\frac{\partial}{\partial a'}\mathrm{vec}[F(\theta(a))'] = \begin{bmatrix} 0_{(6n\times n)} \\ \frac{\partial}{\partial a'}\mathrm{vec}\left[\frac{\partial^2 h_{t-1}}{\partial\theta'\partial a}(\theta(a))\right] \end{bmatrix}.$

Received December 18, 2002; revised March 9, 2004; accepted March 31, 2004

REFERENCES

Alexander, C. (2000), "Orthogonal Methods for Generating Large Positive Semi-Definite Covariance Matrices." Discussion Papers in Finance 2000-06, University of Reading.

Bollerslev, T. (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach." *Review of Economics and Statistics 72*, 498–505.

Brandt, M., and F. X. Diebold (2004). "A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations." Forthcoming in the *Journal of Business*.

Engle, R. F. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica 50*, 987–1007.

Engle, R. F. (2002). "Dynamic Conditional Correlation—A Simple Class of Multivariate GARCH Models." *Journal of Business and Economic Statistics 20*, 339–350.

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

Engle, R. F. and K. Sheppard. (2001). "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH." Discussion Paper 2001-15, University of California–San Diego.

Gourieroux, C., J. P. Laurent, and O. Scaillet. (2000). "Sensitivity Analysis of Values at Risk." *Journal of Empirical Finance 7*, 225–245.

Garman, M. (1996). "Improving on VaR." *Risk 9*, 61–63.

Ledoit, O., P. Santa-Clara, and M. Wolf. (2003). "Flexible Multivariate GARCH Modeling with an Application to International Stock Markets." *Review of Economics and Statistics 85*, 735–747.

Lütkepohl, H. (1990). *Introduction to Multiple Time Series Analysis*. New York: Springer-Verlag.

McNeil, A. J., and R. Frey. (2000). "Estimation of Tail Related Risk Measures for Heteroskedastic Financial Time Deries: An Extreme Value Approach." *Journal of Empirical Finance 7*, 271–300.

Nijman, T., and E. Sentana. (1996). "Marginalization and Contemporaneous Aggregation in Multivariate GARCH Processes." *Journal of Econometrics 71*, 71–87.

White, H. (1994). *Estimation, Inference and Specification Analysis*. Cambridge: Cambridge University Press.