

This PDF is a selection from a published volume from the
National Bureau of Economic Research

Volume Title: The Risks of Financial Institutions

Volume Author/Editor: Mark Carey and René M. Stulz,
editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-09285-2

Volume URL: <http://www.nber.org/books/care06-1>

Conference Date: October 22-23, 2004

Publication Date: January 2007

Title: Practical Volatility and Correlation Modeling for Financial
Market Risk Management

Author: Torben G. Andersen, Tim Bollerslev, Peter Christoffersen,
Francis X. Diebold

URL: <http://www.nber.org/chapters/c9618>

Practical Volatility and Correlation Modeling for Financial Market Risk Management

Torben G. Andersen, Tim Bollerslev,
Peter F. Christoffersen, and Francis X. Diebold

11.1 Introduction

It is now widely agreed that financial asset return volatilities and correlations (henceforth “volatilities”) are time varying, with persistent dynamics. This is true across assets, asset classes, time periods, and countries. Moreover, asset return volatilities are central to finance, whether in asset pricing, portfolio allocation, or market risk measurement. Hence the field of financial econometrics devotes considerable attention to time-varying volatility and associated tools for its measurement, modeling, and forecasting.

In this chapter we suggest practical applications of recent developments in financial econometrics dealing with time-varying volatility to the measurement and management of market risk, stressing parsimonious models that are easily estimated. Our ultimate goal is to stimulate dialog between the academic and practitioner communities, advancing best-practice market risk measurement and management technologies by drawing upon the best of both worlds. Three themes appear repeatedly, and so we highlight them here.

The first is the issue of aggregation level. We consider both aggregated (portfolio-level) and disaggregated (asset-level) modeling, emphasizing the related distinction between risk *measurement* and risk *management*, because risk measurement generally requires only a portfolio-level model, whereas

For helpful comments we would like to thank Ken Abbott, Casper de Vries, Philipp Hartmann, Patricia Jackson, Jim O’Brien, Hashem Pesaran, and Pedro Santa-Clara. For research support, Andersen, Bollerslev, and Diebold thank the U.S. National Science Foundation, and Christoffersen thanks Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC), Social Sciences and Humanities Research Council of Canada (SSHRC), and Institut de Finance Mathématique de Montréal (IFM2).

risk management requires an asset-level model. At the asset level, the issue of dimensionality and dimensionality reduction arises repeatedly, and we devote considerable attention to methods for tractable modeling of the very high-dimensional covariance matrices of practical relevance.

The second theme concerns the use of low-frequency versus high-frequency data, and the associated issue of parametric versus nonparametric volatility measurement. We treat all cases, but we emphasize the appeal of volatility *measurement* using nonparametric methods in conjunction with high-frequency data, followed by *modeling* that is intentionally parametric.

The third theme relates to the issue of unconditional versus conditional risk measurement. We argue that, for most financial risk management purposes, the *conditional* perspective is exclusively relevant, notwithstanding, for example, the fact that popular approaches based on historical simulation and extreme-value theory typically adopt an unconditional perspective. We advocate, moreover, moving beyond a conditional *volatility* perspective to a full conditional *density* perspective, and we discuss methods for constructing and evaluating full conditional density forecasts.

We proceed systematically in several steps. In section 11.2, we consider portfolio-level analysis, directly modeling portfolio volatility using historical simulation, exponential smoothing, and generalized autoregressive conditional heteroskedastic (GARCH) methods. In section 11.3, we consider asset-level analysis, modeling asset covariance matrices using exponential smoothing and multivariate GARCH methods, paying special attention to dimensionality-reduction methods. In section 11.4, we explore the use of high-frequency data for improved covariance matrix measurement and modeling, treating realized variance and covariance, and again discussing procedures for dimensionality reduction. In section 11.5 we treat the construction of complete conditional density forecasts via simulation methods. We conclude in section 11.6.

11.2 Portfolio Level Analysis: Modeling Portfolio Volatility

Portfolio risk measurement requires only a univariate portfolio-level model (e.g., Benson and Zangari 1997). In this section, we discuss such univariate portfolio methods. In contrast, active portfolio risk management, including value-at-risk (VaR) minimization and sensitivity analysis, requires a multivariate model, as we discuss subsequently in section 11.3.

In particular, portfolio level analysis is rarely done other than via historical simulation (defined subsequently). But we will argue that there is no reason why one cannot estimate a parsimonious dynamic model for portfolio-level returns. If interest centers on the distribution of the portfolio returns, then this distribution can be modeled directly rather than via aggregation based on a larger and almost inevitably less-well-specified multivariate model.

Berkowitz and O'Brien (2002) find evidence that existing bank risk models perform poorly and are easily outperformed by a simple univariate GARCH model (defined subsequently). Their result is remarkable in that they estimate a GARCH model fit to the time series of actual historical portfolio returns where the underlying asset weights are changing over time. Berkowitz and O'Brien find that banks' reported ex ante VaR forecasts are exceeded by the ex post profits and losses (P/Ls) on less than the predicted 1 percent of days. This apparent finding of risk underestimation could, however, simply be due to the reported P/Ls being "dirty" in that they contain nonrisky income from fees, commissions, and intraday trading profits.¹ More seriously, though, Berkowitz and O'Brien find that the VaR violations which do occur tend to cluster in time. Episodes such as the fall 1998 Russia default and Long-term Capital Management (LTCM) debacle set off a dramatic and persistent increase in market volatility which bank models appear to largely ignore, or at least react to with considerable delay. Such VaR violation clustering is evidence of a lack of conditionality in bank VaR systems, which in turn is a key theme in our discussion that follows.²

We first discuss the construction of historical portfolio values, which is a necessary precursor to any portfolio-level VaR analysis. We then discuss direct computation of portfolio VaR via historical simulation, exponential smoothing, and GARCH modeling.³

11.2.1 Constructing Historical Pseudo-Portfolio Values

In principle it is easy to construct a time series of historical portfolio returns using current portfolio holdings and historical asset returns:

$$(1) \quad r_{w,t} = \sum_{i=1}^N w_{i,T} r_{i,t} \equiv W'_T R_t, \quad t = 1, 2, \dots, T.$$

In practice, however, historical prices for the assets held today may not be available. Examples of such difficulties include derivatives, individual bonds with various maturities, private equity, new public companies, merger companies, and so on. For these cases, "pseudo historical" prices must be constructed using either pricing models, factor models, or some ad hoc considerations. The current assets without historical prices can, for example, be matched to similar assets by capitalization, industry, leverage, and duration. Historical pseudo asset prices and returns can then be constructed using the historical prices on these substitute assets.

1. Although the Basel Accord calls for banks to report 1 percent VaRs, for various reasons most banks tend to actually report more conservative VaRs. Rather than simply scaling up a 1 percent VaR based on some arbitrary multiplication factor, the procedures that we subsequently discuss are readily adapted to achieve any desired, more conservative, VaR.

2. See also Jackson, Maude, and Perraudin (1997).

3. Duffie and Pan (1997) provide an earlier incisive discussion of related VaR procedures and corresponding practical empirical problems.

11.2.2 Volatility via Historical Simulation

Banks often rely on VaRs from historical simulations (HS-VaR). In this case, the VaR is calculated as the $100p$ 'th percentile or the $(T + 1)p$ 'th order statistic of the set of pseudo returns calculated in (1). We can write

$$(2) \quad HS\text{-}VaR_{T+1|T}^p \equiv r_w([T + 1]p),$$

where $r_w([T + 1]p)$ is taken from the set of ordered pseudo returns ($r_w[1]$, $r_w[2]$, \dots , $r_w[T]$). If $[T + 1]p$ is not an integer value then the two adjacent observations can be interpolated to calculate the VaR.

Historical simulation has some serious problems, which have been well documented. Perhaps most importantly, it does not properly incorporate conditionality into the VaR forecast. The only source of dynamics in the HS-VaR is the fact that the sample window in equation (1) is updated over time. However, this source of conditionality is minor in practice.⁴

Figure 11.1 illustrates the hidden dangers of HS as discussed by Pritsker (2001). We plot the daily percentage loss on an S&P 500 portfolio along with the 1 percent HS-VaR calculated from a 250-day moving window. The crash on October 19, 1987, dramatically increased market volatility; however, the HS-VaR barely moved. Only after the second large drop, which occurred on October 26, does the HS-VaR increase noticeably.

This admittedly extreme example illustrates a key problem with the HS-VaR. Mechanically, from equation (2) we see that HS-VaR changes significantly only if the observations around the order statistic $r_w([T + 1]p)$ change significantly. When using a 250-day moving window for a 1 percent HS-VaR, only the second and third smallest returns will matter for the calculation. Including a crash in the sample, which now becomes the smallest return, may therefore not change the HS-VaR very much if the new second smallest return is similar to the previous one.

Moreover, the lack of a properly defined conditional model in the HS methodology implies that it does not allow for the construction of a term structure of VaR. Calculating a 1 percent one-day HS-VaR may be possible on a window of 250 observations, but calculating a ten-day 1 percent VaR on 250 daily returns is not. Often the one-day VaR is simply scaled by the square root of 10, but this extrapolation is only valid under the assumption of i.i.d. normal daily returns. A redeeming feature of the daily HS-VaR is exactly that it does not rely on an assumption of normal returns, and the square root scaling therefore seems curious at best.

In order to further illustrate the lack of conditionality in the HS-VaR method, consider figure 11.2. We first simulate daily portfolio returns from

4. Bodoukh, Richardson, and Whitelaw (1998) introduce updating into the historical simulation method. Note, however, the concerns in Pritsker (2001).

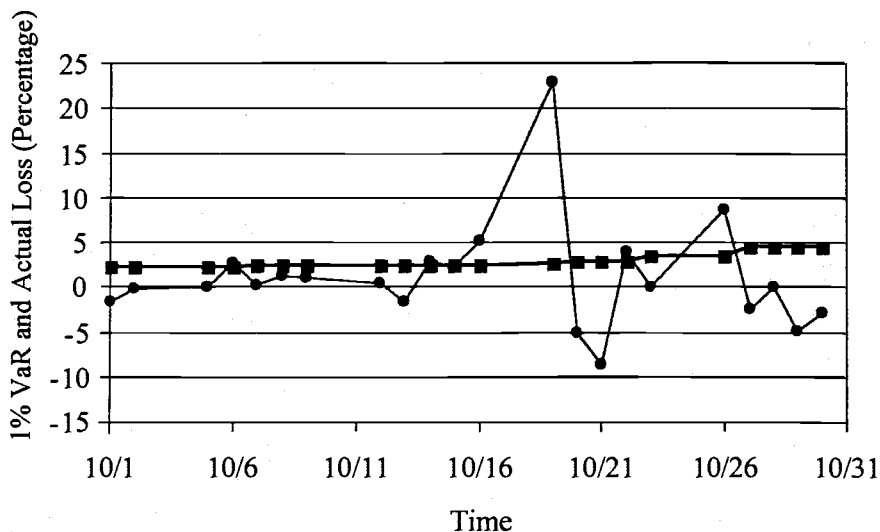


Fig. 11.1 October 1987: Daily S&P 500 Loss and 1 percent HS-VaR

Notes: The thin line with dots shows the daily percentage loss on an S&P 500 portfolio during October 1987. The thick line with squares shows the daily 1 percent VaR from historical simulation using a 250-day window.

a mean-reverting volatility model and then calculate the nominal 1 percent HS-VaR on these returns using a moving window of 250 observations. As the true portfolio return distribution is known, the true daily coverage of the nominal 1 percent HS-VaR can be calculated using the return-generating model. Figure 11.2 shows the conditional coverage probability of the 1 percent HS-VaR over time. Notice from the figure how an HS-VaR with a nominal coverage probability of 1 percent can have a true *conditional* probability as high as 10 percent, even though the *unconditional* coverage is correctly calibrated at 1 percent. On any given day the risk manager thinks that there is a 1 percent chance of getting a return worse than the HS-VaR, but in actuality there may be as much as a 10 percent chance of exceeding the VaR. Figure 11.2 highlights the potential benefit of conditional density modeling: the HS-VaR computes an essentially unconditional VaR, which on any given day can be terribly wrong. A conditional density model will generate a dynamic VaR in an attempt to keep the conditional coverage rate at 1 percent on any given day, thus creating a horizontal line in figure 11.2.

The preceding discussion also hints at a problem with the VaR risk measures itself. It does not say anything about how large the expected loss will be on the days where the VaR is exceeded. Other measures, such as expected shortfall, do, but VaR has emerged as the industry risk measurement standard and we will focus on it here. The methods we will suggest

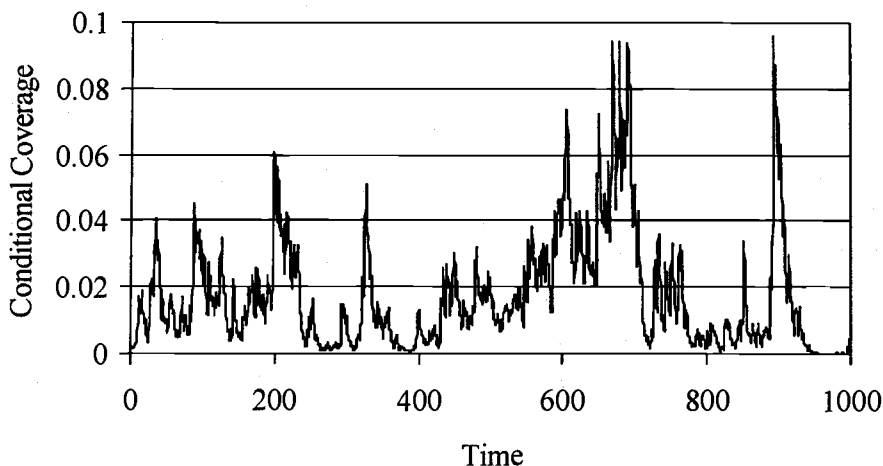


Fig. 11.2 True conditional coverage of 1 percent VaR from historical simulation

Notes: We simulate returns from a GARCH model with normal innovations, after which we compute the 1 percent HS-VaR using a rolling window of 250 observations, and then we plot the *true* conditional coverage probability of the HS-VaR, which we calculate using the GARCH structure. The true conditional coverage probability plotted thus denotes the likelihood each day of getting a VaR violation when using a misspecified 1 percent HS-VaR when the returns are simulated using GARCH.

can, however, equally well be used to calculate expected shortfall and other related risk measures.

11.2.3 Volatility via Exponential Smoothing

Although the HS-VaR methodology discussed previously makes no explicit assumptions about the distributional model generating the returns, the RiskMetrics (RM) filter/model instead assumes a very tight parametric specification. One can begin to incorporate conditionality via univariate portfolio-level exponential smoothing of squared portfolio returns, in precise parallel to the exponential smoothing of individual return squares and cross products that underlies RM.

Still taking the portfolio-level pseudo returns from (1) as the data series of interest, we can define the portfolio-level RM variance as

$$(3) \quad \sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{w,t-1}^2,$$

where the variance forecast for day t is constructed at the end of day $t - 1$ using the square of the return observed at the end of day $t - 1$ as well as the variance on day $t - 1$. In practice, this recursion can be initialized by setting the initial σ_0^2 equal to the unconditional sample standard deviation, for example, $\hat{\sigma}^2$.

Note that back substitution in equation (3) yields an expression for the

current smoothed value as an exponentially weighted moving average of past squared returns:

$$\sigma_t^2 = \sum_{j=0}^{\infty} \phi_j r_{w,t-1-j}^2,$$

where $\phi_j = (1 - \lambda)^j \lambda^j$. Hence the name “exponential smoothing.”

Following RM, the VaR is simply calculated as

$$(4) \quad RM\text{-}VaR_{T+1|T}^p \equiv \sigma_{T+1} \Phi_p^{-1},$$

where Φ_p^{-1} denotes the p th quantile in the standard normal distribution. Although the smoothing parameter λ may in principle be calibrated to best fit the specific historical returns at hand, following RM it is often simply fixed at 0.94 with daily returns. The implicit assumption of zero mean and standard normal innovations therefore implies that no parameters need to be estimated.

The conditional variance for the k -day aggregate return in RM is simply

$$(5) \quad \text{Var}(r_{w,t+k} + r_{w,t+k-1} + \dots + r_{w,t+1} \mid \mathcal{F}_t) \equiv \sigma_{t,t+k|t}^2 = k\sigma_{t+1}^2.$$

The RM model can thus be thought of as a random-walk model in variance. The lack of mean-reversion in the RM variance model implies that the term structure of volatility is flat. Figure 11.3 illustrates the difference between the volatility term structure for the random-walk RM model versus a mean-reverting volatility model. Assuming a low current volatility,

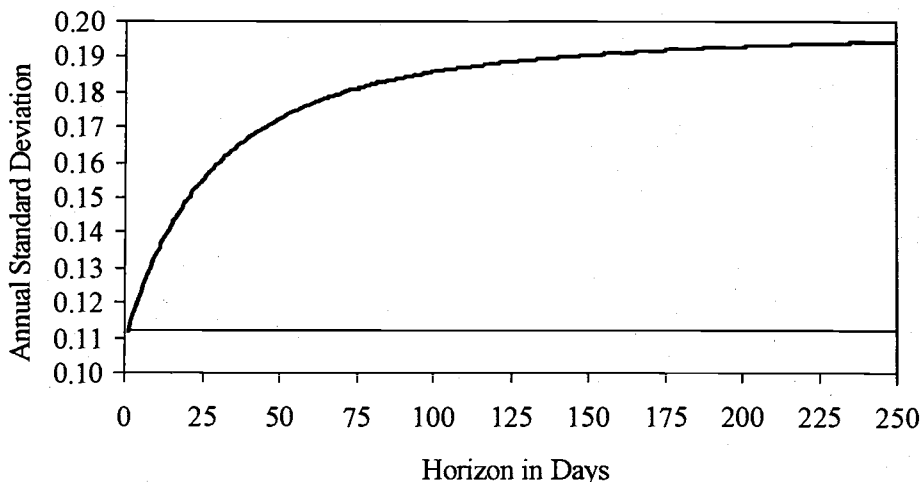


Fig. 11.3 Term Structure of Variance in GARCH and RiskMetrics Models

Notes: We plot the term structure of variance from a mean-reverting GARCH model (thick line) as well as the term structure from a RiskMetrics model (thin line). The current variance is assumed to be identical across models.

which is identical across models, the mean-reverting model will display an upward sloping term structure of volatility, whereas the RM model will extrapolate the low current volatility across all horizons. When taken this literally, the RM model does not appear to be a prudent approach to volatility modeling. The dangers of scaling the daily variance by k , as done in equation (5), are discussed further in Diebold, Hickman, Inoue, and Schuermann (1998).

11.2.4 Volatility via GARCH

The implausible temporal aggregation properties of the RM model, which we discussed earlier, motivates us to introduce the general class of GARCH models, which imply mean-reversion and which contain the RM model as a special case.

First we specify the general univariate portfolio return process

$$(6) \quad r_{w,t} = \mu_t + \sigma_t z_t \quad z_t \sim \text{i.i.d.} \quad E(z_t) = 0 \quad \text{Var}(z_t) = 1.$$

In the following, we will assume that the mean is zero, which is common in risk management, at least when short horizons are considered. Although difficult to estimate with much accuracy in practice, mean-dynamics could in principle easily be incorporated into the models discussed in the following.

The simple symmetric GARCH(1,1) model introduced by Bollerslev (1986) is written as

$$(7) \quad \sigma_t^2 = \omega + \alpha r_{w,t-1}^2 + \beta \sigma_{t-1}^2.$$

Extensions to higher-order models are straightforward, but for notational simplicity we will concentrate on the (1,1) case here and throughout the chapter. Repeated substitution in (7) readily yields

$$\sigma_t^2 = \frac{\omega}{1 - \beta} + \alpha \sum \beta^{j-1} r_{t-j}^2,$$

so that the GARCH(1,1) process implies that current volatility is an exponentially weighted moving average of past squared returns. Hence the GARCH(1,1) volatility measurement is seemingly very similar to RM volatility measurement. There are crucial differences, however.

First, GARCH parameters, and hence ultimately GARCH volatility, are estimated using rigorous statistical methods that facilitate probabilistic inference, in contrast to exponential smoothing, in which the parameter is set in an ad hoc fashion. Typically we estimate the vector of GARCH parameters θ by maximizing the log likelihood function,

$$(8) \quad \log L(\theta; r_{w,T}, \dots, r_{w,1}) \propto - \sum_{t=1}^T [\log \sigma_t^2(\theta) - \sigma_t^{-2}(\theta) r_{w,t}^2].$$

Note that the assumption of conditional normality underlying the (quasi) likelihood function in equation (8) is merely a matter of convenience. The

conditional return distribution will generally be nonnormal, but it does not need to be: quasi maximum likelihood estimation still produces consistent and asymptotically normal parameter estimates. The log-likelihood optimization in equation (9) can only be done numerically. However, GARCH models are parsimonious and specified directly in terms of univariate portfolio returns, so that only a single numerical optimization needs to be performed.⁵

Second, the covariance stationary GARCH(1,1) process has dynamics that eventually produce reversion in volatility to a constant long-run value, which enables interesting and realistic forecasts. This contrasts sharply with the RM exponential smoothing approach. As is well-known (e.g., Nerlove and Wage 1964, Theil and Wage 1964), exponential smoothing is optimal if and only if squared returns follow a “random walk plus noise” model (a “local level” model in the terminology of Harvey 1989), in which case the minimum mean squared error forecast at any horizon is simply the current smoothed value. The historical records of volatilities of numerous assets (not to mention the fact that volatilities are bounded below by zero) suggest, however, that volatilities are unlikely to follow random walks, and hence that the flat forecast function associated with exponential smoothing is unrealistic and undesirable for volatility forecasting purposes.

Let us elaborate. We can rewrite the GARCH(1,1) model in equation (7) as

$$(9) \quad \sigma_t^2 = (1 - \alpha - \beta)\sigma^2 + \alpha r_{w,t-1}^2 + \beta\sigma_{t-1}^2,$$

where $\sigma^2 \equiv \omega/(1 - \alpha - \beta)$ denotes the long-run, or unconditional daily variance. This representation shows that the GARCH forecast is constructed as an average of three elements. Equivalently, we can also write the model as

$$(10) \quad \sigma_t^2 = \sigma^2 + \alpha(r_{w,t-1}^2 - \sigma^2) + \beta(\sigma_{t-1}^2 - \sigma^2),$$

which explicitly shows how the GARCH(1,1) model forecasts by making adjustments to the current variance and the influence of the squared return around the long-run, or unconditional variance. Finally, we can also write

$$\sigma_t^2 = \sigma^2 + (\alpha + \beta)(\sigma_{t-1}^2 - \sigma^2) + \alpha\sigma_{t-1}^2(z_{t-1}^2 - 1),$$

where the last term on the right-hand side, on average, is equal to zero. Hence, this shows how the GARCH(1,1) forecasts by making adjustments around the long-run variance, with variance persistence governed by $(\alpha + \beta)$ and the (contemporaneous) volatility-of-volatility linked to the level of volatility as well as the size of α .

The mean-reverting property of GARCH volatility forecasts has impor-

5. This optimization can be performed in a matter of seconds on a standard desktop computer using standard software such as Excel, as discussed by Christoffersen (2003). For further discussion of inference in GARCH models, see also Andersen, Bollerslev, Christoffersen, and Diebold (2005).

tant implications for the volatility term structure. To construct the volatility term structure corresponding to a GARCH(1,1) model, we need the k -day ahead variance forecast, which is

$$(11) \quad \sigma_{t+k|t}^2 = \sigma^2 + (\alpha + \beta)^{k-1}(\sigma_{t+1}^2 - \sigma^2).$$

Assuming that the daily returns are serially uncorrelated, the variance of the k -day cumulative returns, which we use to calculate the volatility term structure, is then

$$(12) \quad \sigma_{t+k|t}^2 = k\sigma^2 + (\sigma_{t+1}^2 - \sigma^2)[1 - (\alpha + \beta)^k](1 - \alpha - \beta)^{-1}.$$

Compare this mean-reverting expression with the RM forecast in equation (5). In particular, note that the speed of mean reversion in the GARCH(1,1) model is governed by $\alpha + \beta$. The mean-reverting line in figure 11.3 is calculated from equation (12), normalizing by k and taking the square root to display the graph in daily standard deviation units.

Third, the dynamics associated with the GARCH(1,1) model afford rich and intuitive interpretations, and they are readily generalized to even richer specifications. To take one important example, note that the dynamics may be enriched via higher-ordered specifications, such as GARCH(2,2). Indeed, Engle and Lee (1999) show that the GARCH(2,2) is of particular interest, because under certain parameter restrictions it implies a component structure obtained by allowing for time variation in the long-run variance in (10),

$$(13) \quad \sigma_t^2 = q_t + \alpha(r_{w,t-1}^2 - q_{t-1}) + \beta(\sigma_{t-1}^2 - q_{t-1}),$$

with the long-run component, q_t , modeled as a separate autoregressive process,

$$(14) \quad q_t = \omega + \rho q_{t-1} + \phi(r_{w,t-1}^2 - \sigma_{t-1}^2).$$

Many authors, including Gallant, Hsu, and Tauchen (1999) and Alizadeh, Brandt, and Diebold (2002) have found evidence of component structure in volatility, suitable generalizations of which can be shown to approximate long memory (e.g., Andersen and Bollerslev 1997, and Barndorff-Nielsen and Shephard 2001), which is routinely found in asset return volatilities (e.g., Bollerslev and Mikkelsen 1999).

To take a second example of the extensibility of GARCH models, note that all models considered thus far imply symmetric response to positive versus negative return shocks. However, equity markets, and particularly equity indexes, often seem to display a strong asymmetry, whereby a negative return boosts volatility by more than a positive return of the same absolute magnitude. The GARCH model is readily generalized to capture this effect. In particular, the asymmetric GJR GARCH(1,1) model of Glosten, Jagannathan, and Runkle (1993) is simply defined by

$$(15) \quad \sigma_t^2 = \omega + \alpha r_{w,t-1}^2 + \gamma r_{w,t-1}^2 \mathbf{1}(r_{w,t-1} < 0) + \beta \sigma_{t-1}^2.$$

Asymmetric response in the conventional direction thus occurs when $\gamma > 0$.⁶

11.3 Asset Level Analysis: Modeling Asset Return Covariance Matrices

The preceding discussion focused on the specification of dynamic volatility models for the aggregate portfolio return. These methods are well suited to providing forecasts of portfolio-level risk measures such as aggregate VaR. However they are less well suited for providing input into the active risk management process. If, for example, the risk manager wants to know the sensitivity of the portfolio VaR to increases in stock market volatility and asset correlations, which typically occur in times of market stress, then a multivariate model is needed. Active risk management such as portfolio VaR minimization also requires a multivariate model, which provides a forecast for the entire covariance matrix.⁷

Multivariate models are also better suited for calculating sensitivity risk measures to answer questions such as: “If I add an additional 1,000 shares of IBM to my portfolio, how much will my VaR increase?” Moreover, bank-wide VaR is made up of many desks with multiple traders on each desk, and any subportfolio analysis is not possible with the aggregate portfolio-based approach.⁸

In this section we therefore consider the specification of models for the full N -dimensional conditional distribution of asset returns. Generalizing the expression in equation (6), we write the multivariate model as

$$(16) \quad R_t = \Omega_t^{1/2} Z_t \quad Z_t \sim \text{i.i.d.} \quad E(Z_t) = 0 \quad \text{Var}(Z_t) = I,$$

where we have again set the mean to zero and where I denotes the identity matrix. The $N \times N$ $\Omega_t^{1/2}$ matrix can be thought of as the square root, or Cholesky decomposition, of the covariance matrix Ω_t . This section will focus on specifying a dynamic model for this matrix, whereas section 11.5 will suggest methods for specifying the distribution of the innovation vector Z_t .

Constructing positive semidefinite (psd) covariance matrix forecasts, which ensures that the portfolio variance is always nonnegative, subsequently presents a key challenge. The covariance matrix will have $(1/2)N(N + 1)$ distinct elements, but structure needs to be imposed to guarantee psd.

6. Engle (2001, 2004) demonstrates empirically that allowing for asymmetries in the conditional variance can materially affect GARCH-based VaR calculations.

7. Brandt, Santa-Clara, and Valkanov (2004) provide an alternative and intriguing new approach for dimension reduction by explicitly parameterizing the portfolio weights as a function of observable state variables, thereby sidestepping the need to estimate the full covariance matrix. See also Pesaran and Zaffaroni (2004).

8. See Manganelli (2004) for an interesting new low-dimensional approach to this problem.

The practical issues involved in estimating the parameters guarding the dynamics for the $(1/2)N(N + 1)$ elements are related and equally important. Although much of the academic literature focuses on relatively small multivariate examples, in this section we will confine our attention to methods that are applicable even with N (relatively) large.

11.3.1 Covariance Matrices via Exponential Smoothing

The natural analogue to the RM variance dynamics in (3) assumes that the covariance matrix dynamics are driven by the single parameter λ for all variances and covariance in Ω_t :

$$(17) \quad \Omega_t = \lambda\Omega_{t-1} + (1 - \lambda)R_{t-1}R'_{t-1}.$$

The covariance matrix recursion may again be initialized by setting Ω_0 equal to the sample average coverage matrix.

The RM approach is clearly very restrictive, imposing the same degree of smoothness on all elements of the estimated covariance matrix. Moreover, covariance matrix forecasts generated by RM are in general suboptimal, for precisely the same reason as with the univariate RM variance forecasts discussed earlier. If the multivariate RM approach has costs, it also has benefits. In particular, the simple structure in (17) immediately guarantees that the estimated covariance matrices are psd, as the outer product of the return vector must be psd unless some assets are trivial linear combinations of others. Moreover, as long as the initial covariance matrix is psd (which will necessarily be the case when we set Ω_0 equal to the sample average coverage matrix as suggested earlier, so long as the sample size T is larger than the number of assets N), RM covariance matrix forecasts will also be psd, because a sum of psd matrices is itself psd.

11.3.2 Covariance Matrices via Multivariate GARCH

Although easily implemented, the RM approach (17) may be much too restrictive in many cases. Hence we now consider multivariate GARCH models. The most general multivariate GARCH(1,1) model is

$$(18) \quad \text{vech}(\Omega_t) = \text{vech}(C) + B \text{vech}(\Omega_{t-1}) + A \text{vech}(R_{t-1}R'_{t-1}),$$

where the vech (“vector half”) operator converts the unique upper triangular elements of a symmetric matrix into a $(1/2)N(N + 1) \times 1$ column vector, and A and B are $(1/2)N(N + 1) \times (1/2)N(N + 1)$ matrices. Notice that in this general specification, each element of Ω_{t-1} may potentially affect each element of Ω_t , and similarly for the outer product of past returns, producing a serious “curse-of-dimensionality” problem. In its most general form, the GARCH(1,1) model (18) has a total of $(1/2)N^4 + N^3 + N^2 + (1/2)N = O(N^4)$ parameters. Hence, for example, for $N = 100$ the model has 51,010,050 parameters! Estimating this many free parameters is obviously infeasible. Note also that without specifying more structure on the model

there is no guarantee of positive definiteness of the fitted or forecasted covariance matrices.

The dimensionality problem can be alleviated somewhat by replacing the constant term via “variance targeting,” as suggested by Engle and Mezrich (1996). Variance targeting forces the model-implied unconditional covariance matrix to equal a precalculated estimate from the simple sample average. This, in turn, avoids the cumbersome nonlinear estimation of the matrix of constant terms, which instead is computed from the other parameters as follows:

$$(19) \quad \text{vech}(C) = (I - A - B)\text{vech}\left(\frac{1}{T}\sum_{t=1}^T R_t R_t'\right).$$

This is also very useful from a forecasting perspective, as small perturbations in A and B sometimes result in large changes in the implied unconditional variance to which the long-run forecasts converge. However, there are still too many parameters to be estimated simultaneously in A and B in the general multivariate model when N is large.

More severe (and hence less palatable) restrictions may be imposed to achieve additional parsimony, as, for example, with the “diagonal GARCH” parameterization proposed by Bollerslev, Engle, and Wooldridge (1988). In a diagonal GARCH model, the matrices A and B have zeros in all off-diagonal elements, which in turn implies that each element of the covariance matrix follows a simple dynamic with univariate flavor: conditional variances depend only on their own lags and own lagged squared returns, and conditional covariances depend only on their own lags and own lagged cross products of returns. Even the diagonal GARCH framework, however, results in $O(N^2)$ parameters to be jointly estimated, which is computationally infeasible in systems of medium and large size.

One approach is to move to the most draconian version of the diagonal GARCH model, in which the matrices B and A are simply scalar matrices. Specifically,

$$(20) \quad \Omega_t = C + \beta\Omega_{t-1} + \alpha(R_{t-1}R_{t-1}'),$$

where the value of each diagonal element of B is β , and each diagonal element of A is α . Rearrangement yields

$$\Omega_t = \Omega + \beta(\Omega_{t-1} - \Omega) + \alpha(R_{t-1}R_{t-1}' - \Omega),$$

which is closely related to the multivariate RM approach, with the important difference that it introduces a nondegenerate long-run covariance matrix Ω , to which Ω_t reverts (provided that $\alpha + \beta < 1$). Notice also, though, that all variances and covariances are assumed to have the same speed of mean reversion, because of common α and β parameters, which may be overly restrictive.

11.3.3 Dimensionality Reduction I: Covariance Matrices via Flex-GARCH

Ledoit, Santa-Clara, and Wolf (2003) suggest an attractive Flex-GARCH method for reducing the computational burden in the estimation of the diagonal GARCH model without moving to the scalar version. Intuitively, Flex-GARCH decentralizes the estimation procedure by estimating $N(N + 1)/2$ bivariate GARCH models with certain parameter constraints, and then pasting them together to form the matrices A , B , and C in equation (18). Specific transformations of the parameter matrices from the bivariate models ensure that the resulting conditional covariance matrix forecast is psd. Flex-GARCH appears to be a viable modeling approach when N is larger than, say, 5, where estimation of the general diagonal GARCH model becomes intractable. However, when N is of the order of 30 and above, which is often the case in practical risk management applications, it becomes cumbersome to estimate $N(N + 1)/2$ bivariate models, and alternative dimensionality reduction methods are necessary. One such method is the dynamic conditional correlation framework, to which we now turn.

11.3.4 Dimensionality Reduction II: Covariance Matrices via Dynamic Conditional Correlation

Recall the simple but useful decomposition of the covariance matrix into the correlation matrix pre- and post-multiplied by the diagonal standard deviation matrix,

$$(21) \quad \Omega_t \equiv D_t \Gamma_t D_t.$$

Bollerslev (1990) uses this decomposition, along with an assumption of constant conditional correlations ($\Gamma_t = \Gamma$) to develop his Constant Conditional Correlation (CCC) GARCH model. The assumption of constant conditional correlation, however, is arguably too restrictive over long time periods.

Engle (2002) generalizes Bollerslev's (1990) CCC model to obtain a Dynamic Conditional Correlation (DCC) model. Crucially, he also provides a decentralized estimation procedure. First, one fits to each asset return an appropriate univariate GARCH model (the models can differ from asset to asset) and then standardizes the returns by the estimated GARCH conditional standard deviations. Then one uses the standardized return vector, say $e_t \equiv R_t \hat{D}_t^{-1}$, to model the correlation dynamics. For instance, a simple scalar diagonal GARCH(1,1) correlation dynamic would be

$$(22) \quad Q_t = C + \beta Q_{t-1} + \alpha(e_{t-1}e'_{t-1}),$$

with the individual correlations in the Γ_t matrix defined by the corresponding normalized elements of Q_t ,

$$(23) \quad \rho_{i,j,t} = q_{i,j,t} / (\sqrt{q_{i,i,t}} \sqrt{q_{j,j,t}}).$$

The normalization in (23) ensures that all correlation forecasts fall in the $[-1; 1]$ interval, while the simple scalar structure for the dynamics of Q_t in equation (22) ensures that Γ_t is psd.

If C is preestimated by correlation targeting, as discussed earlier, only two parameters need to be estimated in equation (22). Estimating variance dynamics asset by asset and then assuming a simple structure for the correlation dynamics thus ensures that the DCC model can be implemented in large systems: $N + 1$ numerical optimizations must be performed, but each involves only a few parameters, regardless of the size of N .

Although the DCC model offers a promising framework for exploring correlation dynamics in large systems, the simple dynamic structure in (22) may be too restrictive for many applications. For example, volatility and correlation responses may be asymmetric in the signs of past shocks.⁹ Researchers are therefore currently working to extend the DCC model to more general dynamic correlation specifications. Relevant work includes Franses and Hafner (2003), Pelletier (2004), and Cappiello, Engle, and Sheppard (2004).

To convey a feel for the importance of allowing for time-varying conditional correlation, we show in figure 11.4 the bond return correlation between Germany and Japan estimated using a DCC model allowing for asymmetric correlation responses to positive versus negative returns, reproduced from Cappiello, Engle, and Sheppard (2004). The conditional correlation clearly varies a great deal. Note in particular the dramatic change in the conditional correlation around the time of the euro's introduction in 1999. Such large movements in conditional correlation are not rare, and they underscore the desirability of allowing for different dynamics in volatility versus correlation.¹⁰

11.4 Exploiting High-Frequency Return Data for Improved Covariance Matrix Measurement

Thus far our discussion has implicitly focused on models tailored to capturing the dynamics in returns by relying only on daily return information. For many assets, however, high-frequency price data are available and should be useful for the estimation of asset return variances and covari-

9. A related example is the often-found positive relationship between volatility changes and correlation changes. If present but ignored, this effect can have serious consequences for portfolio hedging effectiveness.

10. As another example, cross-market stock-bond return correlations are often found to be close to zero or slightly positive during bad economic times (recessions), but negative in good economic times (expansions); see, for example, the discussion in Andersen, Bollerslev, Diebold, and Vega (2004).

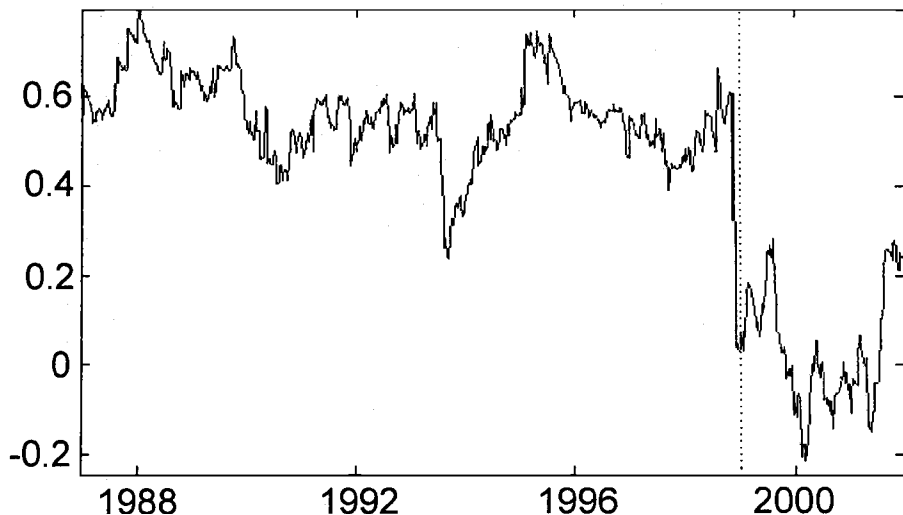


Fig. 11.4 Time-Varying Bond Return Correlation: Germany and Japan

Notes: We reconstruct this figure from Capiello, Engle, and Sheppard (2004), plotting the correlation between Germany and Japanese government bond returns calculated from a DCC model allowing for asymmetric correlation responses to positive and negative returns. The vertical dashed line denotes the euro's introduction in 1999.

ances. Here we review recent work in this area and speculate on its usefulness for constructing large-scale models of market risk.

11.4.1 Realized Variances

Following Andersen, Bollerslev, Diebold, and Labys (2003; henceforth ABDL), define the realized variance (RV) on day t using returns constructed at the Δ intraday frequency as

$$(24) \quad \sigma_{t,\Delta}^2 \equiv \sum_{j=1}^{1/\Delta} r_{t-1+j\Delta,\Delta}^2$$

where $1/\Delta$ is, for example, 48 for thirty-minute returns in twenty-four-hour markets. Theoretically, letting Δ go to zero, which implies sampling continuously, we approach the true *integrated* volatility of the underlying continuous time process on day t .¹¹

In practice, market microstructure noise will affect the RV estimate when Δ gets too small. Prices sampled at fifteen to thirty minute intervals, depending on the market, are therefore often used. Notice also that, in markets that are not open twenty-four hours per day, the potential jump from the closing price on day $t - 1$ to the opening price on day t must be ac-

11. For a full treatment, see Andersen, Bollerslev, and Diebold (forthcoming).

counted for. This can be done using the method in Hansen and Lunde (2005). As is the case for the daily GARCH models considered earlier, corrections may also have to be made for the fact that days following weekends and holidays tend to have higher-than-average volatility.

Although the daily realized variance is just an estimate of the underlying integrated variance and is likely measured with some error, it presents an intriguing opportunity: it is potentially highly accurate, and indeed accurate enough such that we might take the realized daily variance as an observation of the true daily variance, modeling and forecasting it using standard autoregressive moving average (ARMA) time series tools. Allowing for certain kinds of measurement error can also easily be done in this framework. The upshot is that if the fundamental frequency of interest is daily, then using sufficiently high-quality intraday price data enables the risk manager to treat volatility as essentially observed. This is vastly different from the GARCH style models discussed earlier, in which the daily variance is constructed recursively from past daily returns.

As an example of the direct modeling of realized volatility, one can specify a simple first-order autoregressive model for the log realized volatility,

$$(25) \quad \log(\sigma_{t,\Delta}) \equiv c + \beta \log(\sigma_{t-1,\Delta}) + \nu_t,$$

which can be estimated using simple ordinary least squares (OLS). The log specification guarantees positivity of forecasted volatilities and induces (approximate) normality, as demonstrated empirically in Andersen, Bollerslev, Diebold, and Labys (2000, 2001). ABDL show the superior forecasting properties of RV-based forecasts compared with GARCH forecasts. Rather than relying on a simple short-memory ARMA model as in equation (25), they specify a fractionally integrated model to better account for the apparent long-memory routinely found in volatility dynamics.

Along these lines, figure 11.5 shows clear evidence of long-memory in foreign exchange RVs as evidenced by the sample autocorrelation function for lags of 1 through 100 days. We first construct the daily RVs from thirty-minute FX returns and then calculate the corresponding daily sample autocorrelations of the RVs. Note that the RV autocorrelations are significantly positive for all 100 lags when compared with the conventional 95 percent Bartlett confidence bands.

The RV forecasts may also be integrated into the standard GARCH modeling framework, as explored in Engle and Gallo (2004).¹² Similarly, rather than relying on GARCH variance models to standardize returns in the first step of the DCC model, RVs can be used instead. Doing so would result in a more accurate standardization and would require only a single

12. Intriguing new procedures for combining high-frequency data and RV-type measures with lower-frequency daily returns in volatility forecasting models have also recently been developed by Ghysels, Santa-Clara, and Valkanov (2005).

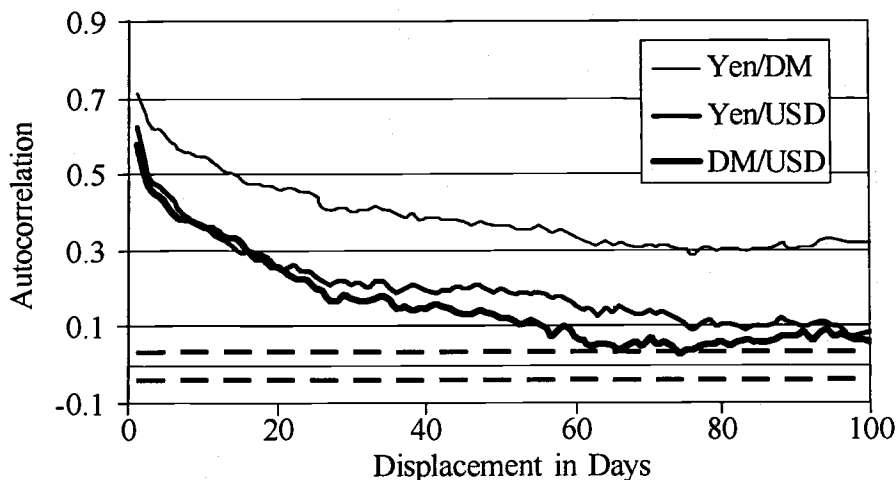


Fig. 11.5 Sample autocorrelations of realized volatility: Three currencies

Notes: We plot the sample autocorrelations of daily realized log standard deviations for three FX rates, together with Bartlett's ± 2 standard error bands for the sample autocorrelations of white noise. We construct the underlying daily realized variances using thirty-minute returns from December 1, 1986, through December 1, 1996.

numerical optimization step—estimation of correlation dynamics—thereby rendering the computational burden in DCC nearly negligible.

We next discuss how realized variances and their natural multivariate counterparts, realized covariances, can be used in a more systematic fashion in risk management.

11.4.2 Realized Covariances

Generalizing the realized variance idea to the multivariate case, we can define the daily realized covariance matrix as

$$(26) \quad \Omega_{t,\Delta} \equiv \sum_{j=1}^{1/\Delta} R_{t-1+j\Delta,\Delta} R'_{t-1+j\Delta,\Delta}.$$

The upshot again is that variances and covariances no longer have to be extracted from a nonlinear model estimated via treacherous maximum-likelihood procedures, as was the case for the preceding GARCH models. Using intraday price observations, we essentially observe the daily covariances and can model them as if they were observed. ABDL show that, as long as the asset returns are linearly independent and the number of assets, N , is less than $1/\Delta$, the realized covariance matrix will be positive definite. However, for a sampling interval of, for example, thirty minutes in twenty-four-hour markets, $1/\Delta$ is 48, so in large portfolios the condition is likely to be violated. We return to this important issue at the end of this section.

Microstructure noise may plague realized covariances, just as it may plague realized variances. Nonsynchronous trading, however, creates additional complications in the multivariate case. These are similar, but potentially more severe, than the nonsynchronous trading issues that arise in the estimation of, say, monthly covariances and CAPM betas with nonsynchronous daily data. A possible fix involves the inclusion of additional lead and lag terms in the realized covariance measure (26), along the lines of the Scholes and Williams (1977) beta-correction technique. Work on this is still in its infancy, and we will not discuss it any further here, but an important recent contribution is Martens (2004).

We now consider various strategies for modeling and forecasting realized covariances, treating them as directly observable vector time series. These all are quite speculative, as little work has been done to date in terms of actually assessing the economic value of using realized covariances for practical risk measurement and management problems.¹³

Paralleling the tradition of the scalar diagonal GARCH model, directly suggests the following model

$$(27) \quad \text{vech}(\Omega_{t,\Delta}) = \text{vech}(C) + \beta \text{vech}(\Omega_{t-1,\Delta}) + v_t,$$

which requires nothing but simple OLS to implement, while guaranteeing positive definiteness of the corresponding covariance matrix forecasts for any positive definite matrix C and positive values of β . This does again, however, impose a common mean-reversion parameter across variances and covariances, which may be overly restrictive. Realized covariance versions of the nonscalar diagonal GARCH model could be developed in a similar manner, keeping in mind the restrictions required for positive definiteness.

Positive definiteness may also be imposed by modeling the Cholesky decomposition of the realized covariance matrix rather than the matrix itself, as suggested by ABDL. We have

$$(28) \quad \Omega_{t,\Delta} \equiv P_{t,\Delta} P'_{t,\Delta},$$

where $P_{t,\Delta}$ is a unique lower triangular matrix. The data vector is then $\text{vech}(P_{t,\Delta})$, and we substitute the forecast of $\text{vech}(P_{t+k,\Delta})$ back into equation (28) to construct a forecast of $\Omega_{t+k,\Delta}$.

Alternatively, in the tradition of Ledoit and Wolf (2003), one may induce positive definiteness of high-dimensional realized covariance matrices by shrinking toward the covariance matrix implied by a single-factor structure, in which the optimal shrinkage parameter is estimated directly from the data.

13. One notable exception is the work of Fleming, Kirby, and Oestdiek (2003), which suggests dramatic improvements vis-à-vis the RM and multivariate GARCH frameworks for standard mean-variance efficient asset allocation problems.

We can also use a DCC-type framework for realized correlation modeling. In parallel to equation (21) we write

$$(29) \quad \Omega_{t,\Delta} \equiv D_{t,\Delta} \Gamma_{t,\Delta} D_{t,\Delta},$$

where the typical element in the diagonal matrix $D_{t,\Delta}$ is the realized standard deviation, and the typical element in $\Gamma_{t,\Delta}$ is constructed from the elements in $\Omega_{t,\Delta}$ as

$$(30) \quad \rho_{i,j,t,\Delta} \equiv \sigma_{i,j,t,\Delta} / (\sigma_{i,i,t,\Delta} \sigma_{j,j,t,\Delta}).$$

Following the DCC idea, we model the standard deviations asset by asset in the first step, and the correlations in a second step. Keeping a simple structure, as in equation (22), we have

$$(31) \quad \text{vech}(Q_{t,\Delta}) = \text{vech}(C) + \beta \text{vech}(Q_{t-1,\Delta}) + v_t,$$

where simple OLS again is all that is required for estimation. Once again, a normalization is needed to ensure that the correlation forecasts fall in the $[-1;1]$ interval. Specifically,

$$(32) \quad \hat{\rho}_{i,j,t,\Delta} = \hat{q}_{i,j,t,\Delta} / (\sqrt{\hat{q}_{i,i,t,\Delta}} \sqrt{\hat{q}_{j,j,t,\Delta}}).$$

The advantages of this approach are twofold: first, high-frequency information is used to obtain more precise forecasts of variances and correlations. Second, numerical optimization is not needed at all. Long-memory dynamics or regime switching could, of course, be incorporated as well.

Although there appear to be several avenues for exploiting intraday price information in daily risk management, two key problems remain. First, many assets in typical portfolios are not liquid enough for intraday information to be available and useful. Second, even in highly liquid environments, when N is very large the positive definiteness problem remains. We now explore a potential solution to these problems.

11.4.3 Dimensionality Reduction III: (Realized) Covariance Matrices via Mapping to Liquid Base Assets

Multivariate market risk management systems for portfolios of thousands of assets in many cases work from a set of, say, thirty observed base assets believed to be key drivers of risk. Such a base asset factor structure is, of course, more justified for a relatively specialized application such as a U.S. equity portfolio than for a large diversified entity such as a major international bank. The choice of factors depends on the portfolio at hand but can, for example, consist of equity market indexes, FX rates, benchmark interest rates, and so on, which are believed to capture the main sources of uncertainty in the portfolio. The assumptions made on the multivariate distribution of base assets are naturally of crucial importance for the accuracy of the risk management system.

Note that base assets typically correspond to the most liquid assets in the market. The upshot here is that we *can* credibly rely on realized volatility and covariances in this case. Using the result from ABDL, a base asset system of dimension $N_F < 1/\Delta$ will ensure that the realized covariance matrix is psd and therefore useful for forecasting.

The mapping from base assets to the full set of assets is discussed in Jorion (2000). In particular, the factor model is naturally expressed as¹⁴

$$(33) \quad R_t = \mathbf{B}R_{F,t} + v_t,$$

where v_t denotes the idiosyncratic risk. The factor loadings in the $N \times N_F$ matrix \mathbf{B} may be obtained from regression (if data exists), or via pricing model sensitivities (if a pricing model exists). Otherwise the loadings may be determined by ad hoc considerations, such as matching a security without a well-defined factor loading to another similar security which has a well-defined factor loading.

We now need a multivariate model for the N_F base assets. However, assuming that

$$(34) \quad R_{F,t} = \Omega_{F,t}^{1/2} Z_{F,t} \quad Z_{F,t} \sim \text{i.i.d.} \quad E(Z_{F,t}) = 0 \quad \text{Var}(Z_{F,t}) = I,$$

we can use the modeling strategies discussed earlier to construct the $N_F \times N_F$ realized factor covariance matrix $\Omega_{F,t}$ and the resulting systematic covariance matrix measurements and forecast.

11.5 Modeling Entire Conditional Return Distributions

Best-practice risk measurement and management often requires knowing the entire distribution of asset or base asset returns, not just the second moments. Conventional risk measures such as VaR and expected shortfall, however, capture only limited aspects of the distribution. They collapse a two-dimensional object, the return distribution function, into a one-dimensional object, the risk measure. Clearly information is lost in this dimension reduction in all but certain counterfactual special cases such as the normal distribution with a zero mean, which only depends on one parameter (the variance).

In this section we explore various approaches to complete the model. Notice that in equation (34) we deliberately left the distributional assumption on the standardized returns unspecified. We simply assumed that the standardized returns were i.i.d. We will keep the assumption of i.i.d. standardized returns and focus on ways to estimate the constant conditional density. This is, of course, with some loss of generality, as dynamics in moments beyond second order could be operative. The empirical evidence for

14. Diebold and Nerlove (1989) construct a multivariate ARCH factor model in which the latent time-varying volatility factors can be viewed as the base assets.

such higher-ordered conditional moment dynamics is, however, much less conclusive at this stage.

The evidence that daily standardized returns are not normally distributed is, however, quite conclusive. Although GARCH and other dynamic volatility models do remove some of the nonnormality in the unconditional returns, conditional returns still exhibit nonnormal features. Interestingly, these features vary systematically from market to market. For example, mature FX market returns are generally strongly conditionally kurtotic, but approximately symmetric. Meanwhile, most aggregate index equity returns appear to be both conditionally skewed and fat tailed.

As an example of the latter, we show in figure 11.6 the daily quantile-quantile (QQ) plot for S&P 500 returns from January 2, 1990, to December 31, 2002, standardized using the (constant) average daily volatility across the sample. That is, we plot quantiles of standardized returns against quantiles of the standard normal distribution. Clearly the daily returns are not unconditionally normally distributed.

Consider now figure 11.7, in which the daily returns are instead stan-

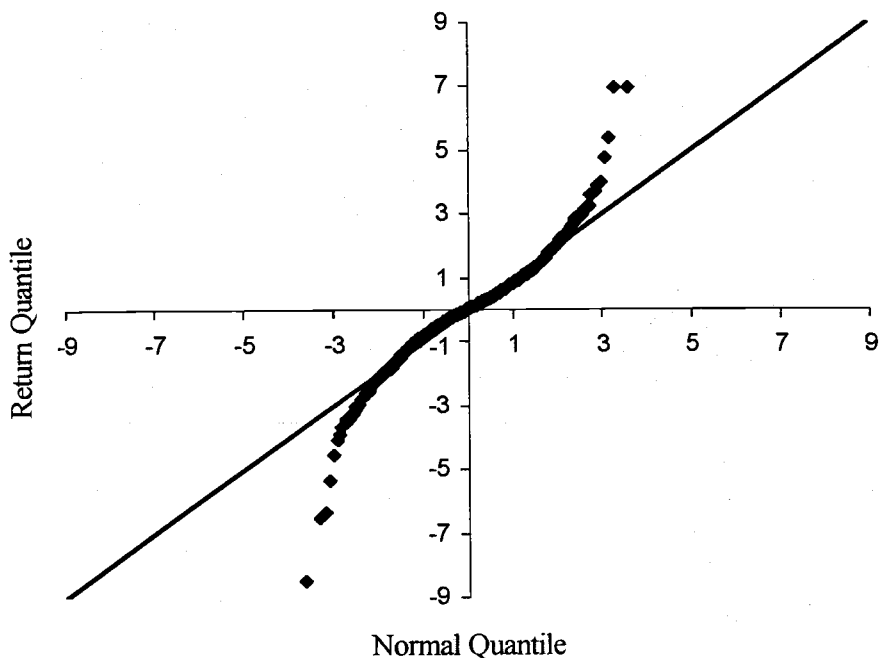


Fig. 11.6 QQ plot of S&P 500 returns standardized by the average volatility

Notes: We show quantiles of daily S&P 500 returns from January 2, 1990, to December 31, 2002, standardized by the average daily volatility during the sample, against the corresponding quantiles from a standard normal distribution.

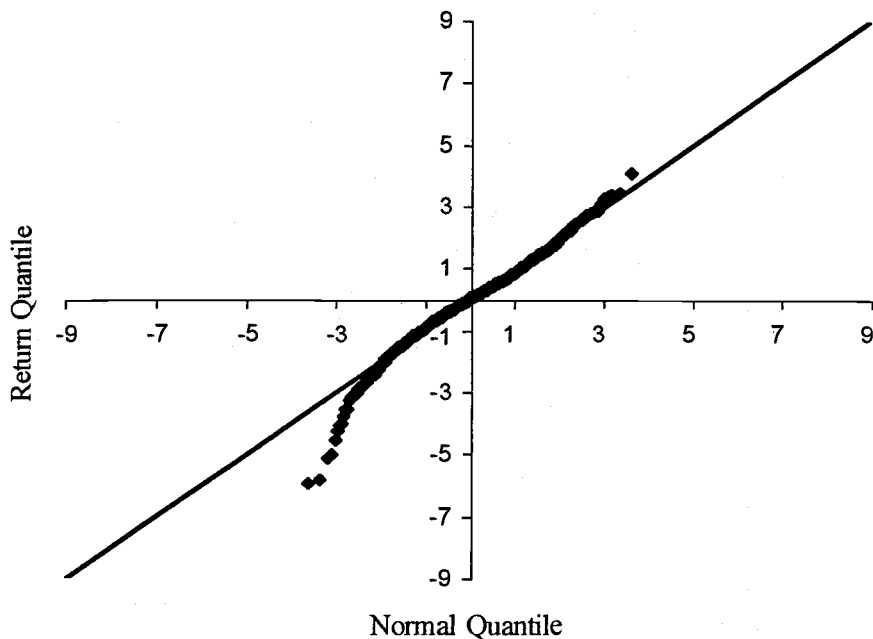


Fig. 11.7 QQ plot of S&P 500 returns standardized by GARCH volatility

Notes: We show quantiles of daily S&P 500 returns from January 2, 1990, to December 31, 2002, standardized by volatility from an estimated asymmetric GJR GARCH(1,1) model, against the corresponding quantiles from a standard normal distribution. The units on each axis are standard deviations.

standardized by the time-varying volatilities from an asymmetric GJR GARCH(1,1) model. The QQ plot in figure 11.7 makes clear that although the GARCH innovations conform more closely to the normal distribution than do the raw returns, the left tail of the S&P 500 returns conforms much less well to the normal distribution than does the right tail: there are more large innovations than one would expect under normality.

As the VaR itself is a quantile, the QQ plot also gives an assessment of the accuracy of the normal-GARCH VaR for different coverage rates. Figure 11.7 suggests that a normal-GARCH VaR would work well for any coverage rate for a portfolio which is short the S&P 500. It may also work well for a long portfolio, but only if the coverage rate is relatively large, say in excess of 5 percent.

Consider now instead the distribution of returns standardized by realized volatility. In contrast to the poor fit in the left tail evident in figure 11.7, the distribution in figure 11.8 is strikingly close to normal, as first noticed by Zhou (1996) and Andersen, Bollerslev, Diebold, and Labys (2000). Figures 11.7 and 11.8 rely on the same series of daily S&P 500 returns but

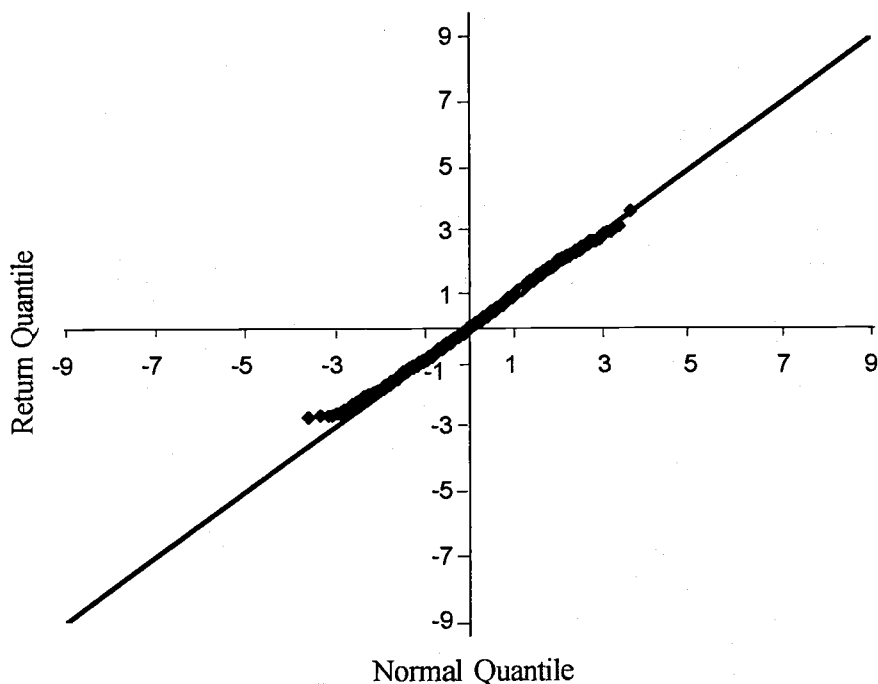


Fig. 11.8 QQ plot of S&P 500 returns standardized by realized volatility

Notes: We show quantiles of daily S&P 500 returns from January 2, 1990, to December 31, 2002, standardized by realized volatility calculated from five-minute futures returns, against the corresponding quantiles from a standard normal distribution. The units on each axis are standard deviations.

simply use two different volatility measures to standardize the raw returns. The conditional nonnormality of daily returns has been a key stylized fact in market risk management. Finding a volatility measure that can generate standardized returns that are close to normal is therefore surprising and noteworthy.

Figure 11.8 and the frequently found lognormality of realized volatility itself suggest that a good approximation to the distribution of returns may be obtained using a normal/lognormal mixture model. In this model, the standardized return is normal and the distribution of realized volatility at time t conditional on time $t - 1$ information is lognormal. This idea is explored empirically in ABDL, who find that a lognormal/normal mixture VaR model performs very well in an application to foreign exchange returns.

The recent empirical results in Andersen, Bollerslev, and Diebold (2006) suggest that even better results may be obtained by separately measuring

and modeling the part of the realized volatility attributable to jumps in the price process through so-called realized bipower variation measures, as formally developed by Barndorff-Nielsen and Shephard (2004). These results have great potential for application in financial risk management, and their practical implications are topics of current research.

Although realized volatility measures may be available for highly liquid assets, it is often not possible to construct realized volatility-based portfolio risk measures. We therefore now survey some of the more conventional methods, first for univariate and then for multivariate models.

11.5.1 Portfolio Level: Univariate Analytic Methods

Although the normal assumption works well in certain cases, we want to consider alternatives that allow for fat tails and asymmetry in the conditional distribution, as depicted in figure 11.7. In the case of VaR we are looking for ways to calculate the cutoff z_p^{-1} in

$$(35) \quad VaR_{T+1|T}^p \equiv \sigma_{T+1} z_p^{-1}.$$

Perhaps the most obvious approach is simply to look for a parametric distribution more flexible than the normal while still tightly parameterized. One such example is the (standardized) Student's t distribution suggested by Bollerslev (1987), which relies on only one additional parameter in generating symmetric fat tails. Recently, generalizations of the Student's t that allow for asymmetry have also been suggested, as in Fernandez and Steel (1998) and Hansen (1994).

Rather than assuming a particular parametric density, one can approximate the quantiles of nonnormal distributions via Cornish-Fisher approximations. Baillie and Bollerslev (1992) first advocated this approach in the context of GARCH modeling and forecasting. The only inputs needed are the sample estimates of skewness and kurtosis of the standardized returns. Extreme value theory provides another approximation alternative, in which the tail(s) of the conditional distribution is estimated using only the extreme observations, as suggested in Diebold, Schuermann, and Stroughair (1998), Longin (2000), and McNeil and Frey (2000).

A common problem with most GARCH models, regardless of the innovation distribution, is that the conditional distribution of returns is not preserved under temporal aggregation. Hence even if the standardized daily returns from a GARCH(1,1) model were normal, the implied weekly returns will not be. This in turn implies that the term structure of VaR or expected shortfall needs to be calculated via Monte Carlo simulation, as in, for example, Guidolin and Timmermann (2004). But Monte Carlo simulation requires a properly specified probability distribution, which would rule out the Cornish-Fisher and extreme-value-theory approximations.

Heston and Nandi (2000) suggest a specific affine GARCH-normal

model, which may work well for certain portfolios, and which, combined with the methods of Albanese, Jackson, and Wiberg (2004), allows for relatively easy calculation of the term structure of VaRs. In general, however, simulation methods are needed; we now discuss a viable approach that combines a parametric volatility model with a data-driven conditional distribution.

11.5.2 Portfolio Level: Univariate Simulation Methods

Bootstrapping, or Filtered Historical Simulation (FHS) assumes a parametric model for the second-moment dynamics but bootstraps from standardized returns to construct the distribution. At the portfolio level this is easy to do. Calculate the standardized pseudo portfolio returns as

$$(36) \quad \hat{z}_{w,t} = r_{w,t} / \hat{\sigma}_t, \quad \text{for } t = 1, 2, \dots, T,$$

using one of the variance models from section 11.2. For the one-day-ahead VaR, we then simply use the order statistic for the standardized returns combined with the volatility forecast to construct

$$(37) \quad FHS\text{-}VaR_{T+1}^p \equiv \sigma_{T+1} \hat{z}_w([T+1]p).$$

Multiday VaR requires simulating paths from the volatility model using the standardized returns sampled with replacement as innovations. This approach has been suggested by Diebold, Schuermann, and Stroughair (1998), Hull and White (1998), and Barone-Adesi, Bourgoin, and Gianopoulos (1998), who coined the term FHS. Pritsker (2001) also provides evidence on its effectiveness.

11.5.3 Asset Level: Multivariate Analytic Methods

Just as a fully specified univariate distribution is needed for risk measurement, so too is a fully specified multivariate distribution often needed for risk management. For example, a fully specified multivariate distribution allows for the computation of VaR sensitivities and VaR-minimizing portfolio weights. The cost, of course, is that we must make an assumption about the multivariate (but constant) distribution of Z_t in (16).

The results of Andersen, Bollerslev, Diebold, and Labys (2000) suggest that, at least in the FX market, the multivariate distribution of returns standardized by the realized covariance matrix is again closely approximated by a normal distribution. As long as the realized volatilities are available, a multivariate version of the lognormal mixture model discussed in connection with figure 11.8 could therefore be developed.

As noted earlier, however, construction and use of realized covariance matrices may be problematic in situations when liquidity is not high, in which case traditional parametric models may be used. As in the univariate case, however, the multivariate normal distribution, coupled with multivariate standardization using covariance matrices estimated from

traditional parametric models, although obviously convenient, does not generally provide an accurate picture of tail risk.¹⁵

A few analytic alternatives to the multivariate normal paradigm do exist, such as the multivariate Student's t distribution first considered by Harvey, Ruiz, and Sentana (1992), along with the more recent related work by Glasserman, Heidelberger, and Shahbuddin (2002). Recently much attention has also been focused on the construction of multivariate densities from the marginal densities via copulas, as in Jondeau and Rockinger (2004) and Patton (2002), although the viability of the methods in very high-dimensional systems remains to be established.

Multivariate extreme value theory offers a tool for exploring cross-asset tail dependencies, which are not captured by standard correlation measures. For example, Longin and Solnik (2001) define and compute extreme correlations between monthly U.S. index returns and a number of foreign country indexes. In the case of the bivariate normal distribution, correlations between extremes taper off to zero as the thresholds defining the extremes get larger in absolute value. The actual equity data, however, behave quite differently. The correlation between negative extremes is much larger than the normal distribution would suggest.¹⁶ Such strong correlation between negative extremes is clearly a key risk management concern. Poon, Rockinger, and Tawn (2004) explore the portfolio risk management implications of extremal dependencies, while Hartmann, Straetmans, and de Vries (chapter 4, this volume) consider their effect on banking system stability. Once again, however, it is not yet clear whether such methods will be operational in large-dimensional systems.

Issues of scalability, as well as cross-sectional and temporal aggregation problems in parametric approaches, thus once again lead us to consider simulation-based solutions.

11.5.4 Asset Level: Multivariate Simulation Methods

In the general multivariate case, we can in principle use FHS with dynamic correlations, but a multivariate standardization is needed. Using the Cholesky decomposition, we first create vectors of standardized returns from (16). We write the standardized returns from an estimated multivariate dynamic covariance matrix as

$$(38) \quad \hat{Z}_t = \hat{\Omega}_t^{-1/2} R_t \text{ for } t = 1, 2, \dots, T,$$

where we calculate $\hat{\Omega}_t^{-1/2}$ from the Cholesky decomposition of the inverse covariance matrix $\hat{\Omega}_t^{-1}$. Now, resampling with replacement vectorwise from the standardized returns will ensure that the marginal distributions as well

15. In the multivariate case the normal distribution is even more tempting to use, because it implies that the aggregate portfolio distribution itself is also normally distributed.

16. In contrast, and interestingly, the correlations of positive extremes appear to approach zero in accordance with the normal distribution.

as particular features of the multivariate distribution, as for example, the cross-sectional dependencies suggested by Longin and Solnik (2001), will be preserved in the simulated data.

The dimensionality of the system in equation (38) may render the necessary multivariate standardization practically infeasible. However, the same FHS approach can be applied with the base asset setup in equation (34), resampling from the factor innovations calculated as

$$(39) \quad \hat{Z}_{F,t} = \hat{\Omega}_{F,t}^{-1/2} R_{F,t} \text{ for } t = 1, 2, \dots, T,$$

where we again use the Cholesky decomposition to build up the distribution of the factor returns. From equation (33) we can then construct the corresponding idiosyncratic asset innovations as

$$(40) \quad \hat{v}_t = R_t - \hat{B}R_{F,t} \text{ for } t = 1, 2, \dots, T,$$

in turn resampling from \hat{Z}_t and \hat{v}_t to build up the required distribution of the individual asset returns in the base asset model.

Alternatively, if one is willing to assume constant conditional correlations, then the standardization can simply be done on an individual asset-by-asset basis using the univariate GARCH volatilities. Resampling vectorwise from the standardized returns will preserve the cross-sectional dependencies in the historical data.

11.6 Summary and Directions for Future Research

We have attempted to demonstrate the power and potential of dynamic financial econometric methods for practical financial risk management, surveying the large literature on high-frequency volatility measurement and modeling, interpreting and unifying the most important and intriguing results for practical risk management. The paper complements the more general and technical survey of volatility and covariance forecasting in Andersen, Bollerslev, Christoffersen, and Diebold (2005).

Our discussion has many implications for practical financial risk management; some point toward desirable extensions of existing approaches, and some suggest new directions. Key points include:

1. Standard model-free methods, such as historical simulation, rely on false assumptions of independent returns. Reliable risk measurement requires a *conditional* density model that allows for time-varying volatility.
2. For the purpose of risk *measurement*, specifying a univariate density model directly on the portfolio return is likely to be most accurate. RiskMetrics offers one possible approach, but the temporal aggregation properties—including the volatility term structure—of RiskMetrics appear to be counterfactual.
3. GARCH volatility models offer a convenient and parsimonious

framework for modeling key dynamic features of returns, including volatility mean-reversion, long-memory, and asymmetries.

4. Although risk measurement can be done from a univariate model for a given set of portfolio weights, risk *management* requires a fully specified multivariate density model. Unfortunately, standard multivariate GARCH models are too heavily parameterized to be useful in realistic large-scale problems.

5. Recent advances in multivariate GARCH modeling are likely to be useful for medium-scale models, but very large scale modeling requires decoupling variance and correlation dynamics, as in the dynamic conditional correlation model.

6. Volatility measures based on high-frequency return data hold great promise for practical risk management. Realized volatility and correlation measures give more accurate forecasts of future realizations than their conventional competitors. Because high-frequency information is only available for highly liquid assets, we suggest a base-asset factor approach.

7. Risk management requires fully specified conditional *density* models, not just conditional covariance models. Resampling returns standardized by the conditional covariance matrix presents an attractive strategy for accommodating conditionally nonnormal returns.

8. The near lognormality of realized volatility, together with the near normality of returns standardized by realized volatility, holds promise for relatively simple-to-implement lognormal/normal mixture models in financial risk management.

References

- Albanese, C., K. Jackson, and P. Wiberg. 2004. A new Fourier transform algorithm for Value-at-Risk. *Quantitative Finance* 4:328–38.
- Alizadeh, S., M. Brandt, and F. X. Diebold. 2002. Range-based estimation of stochastic volatility models. *Journal of Finance* 57:1047–91.
- Andersen, T. G., and T. Bollerslev. 1997. Heterogeneous information arrivals and return volatility dynamics: Uncovering the long run in high frequency returns. *Journal of Finance* 52:975–1005.
- Andersen, T. G., Bollerslev, P. Christoffersen, and F. X. Diebold. 2006. Volatility forecasting. In *Handbook of economic forecasting*, ed. G. Elliott, C. Granger, and A. Timmermann, 778–878. Amsterdam: North-Holland.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold. Forthcoming. Parametric and nonparametric volatility measurement. In *Handbook of financial econometrics*, ed. L. P. Hansen and Y. Ait-Sahalia, Amsterdam: North-Holland.
- . 2006. Forthcoming. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens. 2001. The distribution of realized stock return volatility. *Journal of Financial Economics* 61:43–76.

- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2000. Exchange rate returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal* 4:159–79.
- . 2001. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96:42–55.
- . 2003. Modeling and forecasting realized volatility. *Econometrica* 71:529–626.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega. 2004. Real-time price discovery in stock, bond, and foreign exchange markets. Unpublished manuscript, Northwestern University, Duke University, University of Pennsylvania, and University of Rochester.
- Baillie, R. T., and T. Bollerslev. 1992. Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics* 51:91–113.
- Barndorff-Nielsen, O. E., and N. Shephard. 2001. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, B* 63:167–241.
- . 2004. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2:1–48.
- Barone-Adesi, G., F. Bourgoign, and K. Giannopoulos. 1998. Don't look back. *Risk* 11:100–104.
- Benson, P., and P. Zangari. 1997. A general approach to calculating VaR without volatilities and correlations. *RiskMetrics Monitor* (Summer): 19–23.
- Berkowitz, J., and J. O'Brien. 2002. How accurate are Value-at-Risk models at commercial banks? *Journal of Finance* 57:1093–1112.
- Bodoukh, J., M. Richardson, and R. Whitelaw. 1998. The best of both worlds. *Risk* 11:64–67.
- Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31:307–27.
- . 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics* 69:542–47.
- Bollerslev, T., R. F. Engle, and J. Wooldridge. 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96:116–31.
- Bollerslev, T., and H. O. Mikkelsen. 1999. Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics* 92:75–99.
- Brandt, M., P. Santa-Clara, and R. Valkanov. 2004. Optimal portfolios with parametric weights. Unpublished manuscript, Duke University and University of California (Los Angeles).
- Cappiello, L., R. F. Engle, and K. Sheppard. 2004. Asymmetric dynamics in the correlations of global equity and bond returns. Unpublished manuscript, New York University.
- Christoffersen, P. 2003. *Elements of financial risk management*. San Diego: Academic Press.
- Christoffersen, P., and K. Jacobs. 2004. Which GARCH model for option valuation? *Management Science* 50:1204–21.
- Diebold, F. X., A. Hickman, A. Inoue, and T. Schuermann. 1998. Converting 1-day volatility to h-day volatility: Scaling by root-h is worse than you think. Wharton Financial Institutions Center, Working Paper no. 97-34. Philadelphia: Wharton School of Finance.
- Diebold, F. X., and M. Nerlove. 1989. The dynamics of exchange rate volatility: A multivariate latent-factor ARCH model. *Journal of Applied Econometrics* 4: 1–22.
- Diebold, F. X., T. Schuermann, and J. Stroughair. 1998. Pitfalls and opportunities in the use of extreme value theory in risk management. In *Decision technologies*

- for *computational finance*, ed. A.-P. N. Refenes, A. N. Burgess, and J. D. Moody, 3–12. Amsterdam: Kluwer Academic.
- Duffie, D., and J. Pan. 1997. An overview of Value at Risk. *Journal of Derivatives* (Spring): 7–49.
- Engle, R. F. 2001. GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives* 15:157–68.
- . 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20:339–50.
- . 2004. Risk and volatility: Econometric models and financial practice. *American Economic Review* 94:405–20.
- Engle, R. F., and G. M. Gallo. 2004. A multiple indicators model for volatility using intra-daily data. Unpublished manuscript. New York University and University of Firenze.
- Engle, R. F., and G. G. J. Lee. 1999. A permanent and transitory component model of stock return volatility. In *Cointegration, causality, and forecasting: A Festschrift in honor of Clive W. J. Granger*, ed. R. F. Engle and H. White, 475–97. Oxford, UK: Oxford University Press.
- Engle, R. F., and J. Mezrich. 1996. GARCH for groups. *Risk* 8:36–40.
- Fernandez, C., and M. F. J. Steel. 1998. On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* 93:359–71.
- Fleming, J., C. Kirby, and B. Ostdiek. 2003. The economic value of volatility timing using realized volatility. *Journal of Financial Economics* 67:473–509.
- Franses, P. H., and C. Hafner. 2003. A generalized dynamic conditional correlation model for many asset returns. Manuscript. Erasmus University Rotterdam.
- Gallant, A. R., C. T. Hsu, and G. E. Tauchen. 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics* 81:617–31.
- Ghysels, E., P. Santa-Clara, and R. Valkanov. 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131:59–95.
- Glasserman, P., P. Heidelberger, and P. Shahbuddin. 2002. Portfolio Value-at-Risk with heavy-tailed risk factors. *Mathematical Finance* 12 (3): 239–69.
- Glosten, L. R., R. Jagannathan, and D. Runkle. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48:1779–1801.
- Guidolin, M., and A. Timmermann. 2006. Term structure of risk alternative econometric specifications. *Journal of Econometrics* 131:285–308.
- Hansen, B. 1994. Autoregressive conditional density estimation. *International Economic Review* 35:705–30.
- Hansen, P. R., and A. Lunde. 2005. A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics* 3:525–54.
- Harvey, A. C. 1989. *Forecasting structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.
- Harvey, A. C., E. Ruiz, and E. Sentana. 1992. Unobserved component time series models with ARCH disturbances. *Journal of Econometrics* 52:129–57.
- Heston, S., and S. Nandi. 2000. A closed-form GARCH option pricing model. *Review of Financial Studies* 13:585–626.
- Hull, J., and A. White. 1998. Incorporating volatility updating into the historical simulation method for VaR. *Journal of Risk* 1:5–19.
- Jackson, P., D. Maude, and W. Perraudin. 1997. Bank capital and Value at Risk. *Journal of Derivatives* 4:73–89.
- Jondeau, E., and M. Rockinger. 2004. Forthcoming. The copula-GARCH model

- of conditional dependence: An international stock market application. *Journal of International Money and Finance*.
- Jorion, P. 2000. *Value-at-Risk*. New York: McGraw-Hill.
- 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* 85:735–47.
- Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10:603–21.
- Longin, F. 2000. From VaR to stress testing: The extreme value approach. *Journal of Banking and Finance* 24:1097–1130.
- Longin, F., and B. Solnik. 2001. Extreme correlation of international equity markets. *Journal of Finance* 56:649–76.
- Manganelli, S. 2004. Asset allocation by variance sensitivity analysis. *Journal of Financial Econometrics* 2:370–89.
- Martens, M. 2004. Estimating unbiased and precise realized covariances. Unpublished manuscript. Erasmus University, Rotterdam.
- McNeil, A. J., and R. Frey. 2000. Estimation of tail-related risk measures for heteroskedastic financial time series: An extreme value approach. *Journal of Empirical Finance* 7:271–300.
- Nerlove, M., and S. Wage. 1964. On the optimality of adaptive forecasting. *Management Science* 10:207–29.
- Patton, A. J. 2005. Modeling asymmetric exchange rate dependence. Forthcoming. *International Economic Review*.
- Pelletier, D. 2006. Regime switching for dynamic correlations. *Journal of Econometrics* 131:445–73.
- Pesaran, H., and P. Zaffaroni. 2004. Model averaging and Value-at-Risk based evaluation of large-multi asset volatility models for risk management. Unpublished manuscript, University of Cambridge.
- Poon, S.-H., M. Rockinger, and J. Tawn. 2004. Extreme value dependence in financial markets: Diagnostics, models and financial implications. *Review of Financial Studies* 17:581–610.
- Pritsker, M. 2001. The hidden dangers of historical simulation. Unpublished manuscript. Washington, DC: Federal Reserve Board.
- Scholes, M., and J. T. Williams. 1977. Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5:309–27.
- Theil, H., and S. Wage. 1964. Some observations on adaptive forecasting. *Management Science* 10:198–206.
- Zhou, B. 1996. High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics* 14:45–52.

Comment Pedro Santa-Clara

Andersen, Bollerslev, Christoffersen, and Diebold (henceforth ABCD) provide a comprehensive overview of financial risk management from the point of view of both Wall Street and the ivory tower. Most usefully, ABCD discuss a number of recent developments in the econometrics of time-

varying risk that hold vast promise for risk management applications: the dynamic conditional correlation model of Engle (2002), which permits large-scale, flexible modeling of conditional covariance matrices, the use of high-frequency data to measure realized variances and covariances that has been developed largely by the authors, and the modeling of the full distribution of conditional returns. In this discussion I will just offer a couple of comments and extensions to ABCD's very well-organized survey.

Unconditional Versus Conditional Risk

ABCD discuss extensively the pros and cons of both unconditional and conditional (dynamic) measures of risk. There is, however, an additional source of risk dynamics that is ignored in the paper and that, in fact, has not been studied much in the literature. Most financial assets are managed over time, and it is therefore more important to study the risks of dynamic investment strategies rather than the risks of static portfolios. Especially for supervision and regulation purposes, it matters more to forecast the risk of a portfolio taking into account the likely variation in its weights than to forecast the risk of the current positions that are unlikely to remain in place for long.

Assume that there exist some state variables that forecast both risk and return. A trader that adjusts the portfolio according to those state variables, for instance to maximize the conditional Sharpe ratio, will produce a portfolio with time-varying risk. Many authors have shown that the level of interest rates, the term spread, and the default spread have forecasting power for both first and second moments of returns of stocks and bonds. Brandt and Santa-Clara (2005) show that the optimal asset allocation for a mean-variance investor that recognizes the forecasting power of these state variables displays considerable time variation in portfolio weights and conditional moments.

As another example, investment strategies are typically conditioned on the level of risk in the markets. Either formally, through Value-at-Risk (VaR) constraints, or informally, according to the trader's feelings, the level of exposure is adjusted when risks change. Consider a trader with a VaR limit that manages the exposure of the portfolio to always be at that limit. When market risk is high, the exposure is reduced, and when risk is low, the exposure is increased. Interestingly, the result of this dynamic strategy is a series of returns that have constant conditional VaR. That is, in this case, a dynamic strategy produces a series of returns with static risk.

This example explains why the realized risk of a managed portfolio may not display GARCH characteristics even though the assets in the portfolio have them. Ex ante, if the portfolio were to remain constant, its risk would be changing. Ex post, given that the portfolio changes with the ex ante risk assessment, the realized risk is not time varying. This distinction between

ex ante and ex post risk of an investment strategy has been the basis of much confusion relating to the need of unconditional versus conditional risk models. It justifies the use of unconditional VaR by regulators, since they care only about ex post risk. On the other hand, traders need the more sophisticated models of conditional risk to be able to manage the exposures in a timely manner.

Modeling the Entire Distribution of Returns

ABCD explain that the common use of summary statistics such as volatility, VaR, or expected shortfall is likely to give a partial view of the true risk of a portfolio. Only the full (conditional) distribution of returns, including skewness and fat tails, will correctly capture the likelihood of different levels of losses.

Santa-Clara and Schwartz (2005) offer a simple alternative that captures the impact of the full distribution of returns on the risk of a portfolio. Their approach can be summarized briefly. The idea is that the investor (or the regulator) analyzes the distribution of returns through the lens of a utility function of returns that is concave (reflecting risk aversion). A simple example is the well-known power utility function, $u(r) = (1 + r)^{-\gamma}/(1 - \gamma)$, with relative risk aversion γ .

Given portfolio weights w , simulate the history of portfolio returns:

$$r_{p,t+1} = \sum_{i=1}^N w_i r_{i,t+1} \text{ for } t = 1, \dots, T - 1,$$

and evaluate the corresponding time series of realized utilities of the portfolio $u(r_{p,t+1})$. Then, regress the realized utilities on state variables z that condition the joint return distribution:¹

$$u(r_{p,t+1}) = \phi z_t + \varepsilon_{t+1}.$$

The fitted values of this regression are estimates of the conditional expected utility $E_t(u[r_{p,t+1}])$. At the current time T , the regression is estimated with historic data, and the fitted value $E_T(u[r_{p,T+1}]) = \phi z_T$ is a forecast of the risk of the portfolio in the next period $T + 1$. Actually, a more easily interpreted measure of risk is the conditional certainty equivalent $c_t = u^{-1}(E_t\{u\{r_{p,t+1}\}\})$, which is expressed in units of returns.

We can run similar regressions for the partial derivatives of the expected utility relative to portfolio weights. These derivatives can be used for risk management as they quantify how much the utility (or certainty equivalent) changes when the weight of each asset changes marginally.

Santa-Clara and Schwartz's measure of risk takes into account the full

1. The variable z may contain basis functions of a more fundamental set of state variables y . In this way the specification can accommodate a nonlinear relation between y and the expected utility. Also, the returns may be demeaned prior to running the regression in order to concentrate on risk and discard the effect of the average return on the investor's utility.

distribution of returns. The investor cares about the expected value of the utility, which in turn depends on all the moments of the distribution of the portfolio returns:

$$E_t[u(r_{p,t+1})] \approx u[E_t(r_{p,t+1})] + u''[E_t(r_{p,t+1})]\text{Var}_t(r_{p,t+1})/2 + u'''[E_t(r_{p,t+1})]\text{Skew}_t(r_{p,t+1})/6 + \dots,$$

which depend implicitly on the full joint distribution of the assets' returns. We have therefore a measure of risk that combines all the features of the distribution of returns weighted in an optimal manner according to the risk preferences of the investor.

Finally, this approach can easily accommodate dynamic investment strategies. Simply model the portfolio weights as a function of state variables x_t (which may or may not be different from z_t):

$$r_{p,t+1} = \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1} = \sum_{i=1}^N (\theta x_t) r_{i,t+1},$$

compute the realized utilities, and perform the above regression. Going a step further, the coefficients of the portfolio policy can be optimized to maximize the conditional expected utility of the portfolio along the lines of Brandt and Santa-Clara (2005) and Brandt, Santa-Clara, and Valkanov (2005):

$$\max_{\theta} \frac{1}{T} \sum_{t=1}^T u(r_{p,t+1}) = \frac{1}{T} \sum_{t=1}^T u \left[\sum_{i=1}^N (\theta x_t) r_{i,t+1} \right]$$

Conclusion

The econometrics of risk is an exciting area right now. ABCD's paper is a precious guide to recent developments and points to interesting directions for future research.

References

Brandt, Michael, and Pedro Santa-Clara. 2005. Forthcoming. Dynamic portfolio selection by augmenting the asset space. *Journal of Finance*.
 Brandt, Michael, Pedro Santa-Clara, and Rossen Valkanov. 2005. Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. UCLA Working Paper. Los Angeles: University of California.
 Engle, Robert F. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20:339–50.
 Santa-Clara, Pedro, and Eduardo Schwartz. 2005. Certainty equivalent Value at Risk. UCLA Working Paper. Los Angeles: University of California.

Discussion Summary

Ken Abbott opened the general discussion by suggesting that the methods suggested by the authors may be more applicable to modeling of credit risk, where correlation skew is a concern and copula methods are coming into favor, than in traditional market-risk applications. In his experience, historical simulation methods work well in practice and are relatively easy for bank staff and management to understand. The dynamic issues raised by the authors, which are particularly dramatic in cases like the 1987 crash, are handled in practice by stress-test exercises, which are done along with VaR modeling.

Patricia Jackson observed that the use to which a VaR model is put is a key consideration in its design. Where the purpose is estimating the capital required by the financial institution, including dynamic volatility is undesirable because volatility falls during safe periods and thus implied capital requirements fall. The change to the high volatility characteristic of periods of stress may occur quickly, leaving the institution with little time to increase its capital. Historical simulation methods are less subject to this problem. However, the methods suggested by the authors may be preferable for other uses.

On the other hand, *Jim O'Brien* noted that although historical simulation may tend to give the "correct" number of violations of a VaR quantile, violations tend to be bunched in time, which appears to be a sign of worrisome historical dependence.

The discussion turned to technical considerations; *Philipp Hartmann* noted that some of the methods suggested by the authors implicitly use linear measures in the tails of the return distribution, but tail events tend to occur during crisis periods and may require a more complex specification. *Hayne Leland* noted that bid-ask bounce and infrequent-trading problems can be an issue in the high-frequency data that the authors suggest be used for volatility estimation, and *Hashem Pesaran* noted that such data also are often rather dirty. *Peter Christoffersen* agreed that such problems exist, but suggested that they might be relatively easy to overcome for instruments traded in very liquid markets.