Practical Volatility and Correlation Modeling for Financial Market Risk Management

Torben G. Andersen[a], Tim Bollerslev[b],
Peter F. Christoffersen[c] and Francis X. Diebold[d]

October 14, 2004

Abstract: What do academics have to offer market risk management practitioners in financial institutions? Current industry practice largely follows one of two extremely restrictive approaches: historical simulation or RiskMetrics. In contrast, we favor flexible methods based on recent developments in financial econometrics, which are likely to produce more accurate assessments of market risk. Clearly, the demands of real-world risk management in financial institutions – in particular, real-time risk tracking in very high-dimensional situations – impose strict limits on model complexity. Hence we stress parsimonious models that are easily estimated, and we discuss a variety of practical approaches for high-dimensional covariance matrix modeling. We thus aim to stimulate dialog between the academic and practitioner communities, hopefully stimulating the development of improved market risk management technologies that draw on the best of both worlds.

[*] This paper is prepared for Mark Carey and René Stulz (eds.), *Risks of Financial Institutions*, NBER. For research support, Andersen, Bollerslev and Diebold thank the U.S. National Science Foundation, and Christoffersen thanks FQRSC, SSHRC and IFM2.

[a] Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208, and NBER
phone: 847-467-1285, e-mail: t-andersen@kellogg.northwestern.edu

[b] Department of Economics, Duke University, Durham, NC 27708, and NBER
phone: 919-660-1846, e-mail: boller@econ.duke.edu

[c] Faculty of Management, McGill University, Montreal, Quebec, H3A 1G5, and CIRANO
phone: 514-398-2869, e-mail: peter.christoffersen@mcgill.ca

[d] Department of Economics, University of Pennsylvania, Philadelphia, PA 19104, and NBER
phone: 215-898-1507, e-mail: fdiebold@sas.upenn.edu

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.

Here we survey, unify and extend recent developments in the financial econometrics of time-varying volatility, focusing exclusively on practical applications 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.

Several themes appear repeatedly and are so important as to merit highlighting. 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 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 is the issue of low-frequency vs. high-frequency data, and the associated issue of parametric vs. 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 consistently and intentionally parametric.

The third is the issue of unconditional vs. 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 2, we consider portfolio level analysis, directly modeling portfolio volatility using historical simulation, exponential smoothing, and GARCH methods. In section 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 4, we explore the use of high-frequency data for improved covariance matrix modeling, treating realized variance and covariance, and again discussing dimensionality reduction methods. In section 5 we treat the construction of complete conditional density forecasts via simulation methods. We conclude in section 6.

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-based methods. In contrast, active portfolio risk management, including portfolio allocation, requires a multivariate model, as we discuss subsequently in section 3.

Portfolio level analysis is rarely done other than via historical simulation (defined below). 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 easy outperformed by a simple univariate GARCH model which is defined below. 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 bank VaRs on average underestimate risk when comparing ex post P/Ls with ex ante VaR forecasts. This finding could however simply be due to the reported P/Ls being "dirty" in that they contain non-risky income from fees, commissions and intraday trading profits. 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 LTCM debacle sets 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 below.

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.

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:

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

In practice, however, historical asset prices for the assets held today may not be available. Examples where difficulties arise include derivatives, individual bonds with various maturities, private equity, new public companies, merger companies and so on.

For these cases 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 can then be constructed using the historical prices on these substitute assets.

2.2 Volatility via Historical Simulation

Banks often rely on Value-at-Risk (VaR) from historical simulation VaR (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

$$HS\text{-}VaR^{p}_{T+1|T} \equiv r_w((T+1)p), \tag{2}$$

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

Historical simulation has some serious problems, which have been well-documented. Perhaps most importantly, it does not build in conditionality into the VaR forecast. The only source of dynamics in the HS-VaR is the fact that the sample window in (1) is updated over time. However, regulators do not require that the data window is updated every day so this source of conditionality is minor in practice.[1]

Figure 1 illustrates the hidden dangers of HS as discussed by Pritsker (2001). We plot the daily percentage loss on an S&P500 portfolio along with the 1% HS-VaR calculated from a 250 day moving window. The crash on October 19, 1987 dramatically increased market volatility; however, the HS-VaR barely moves. 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% 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.

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% 1-day HS-VaR

---

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

may be possible on a window of 250 observations but calculating a 10-day 1% VaR on 250 daily returns is not. Often the 1-day VaR is scaled by the square root of 10 but this extrapolation is only valid under the assumption of i.i.d. returns. A redeeming feature of the daily HS-VaR is exactly that it does not rely on an assumption of i.i.d., 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 2. We first simulate daily portfolio returns from a mean-reverting volatility model and then calculate the nominal 1% 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% HS-VaR can be calculated using the return generating model. Figure 2 shows the conditional coverage probability of the 1% HS-VaR over time. Notice from the figure how an HS-VaR with a nominal coverage probability of 1% can have a true conditional probability as high as 10%. On any given day the risk manager thinks that there is a 1% chance of getting a return worse than the HS-VaR but in actuality there may as much as a 10% chance of exceeding the VaR. Figure 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 which will try to keep the conditional coverage rate at 1% on any given day creating a horizontal line in Figure 2.

The above 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 below can equally well be used to calculate expected shortfall and other measures.

2.3 Volatility via Exponential Smoothing

Although the HS-VaR methodology discussed above makes no explicit assumptions about the distributional model generating the returns, the RiskMetrics (RM) model instead assumes a very tight parametric model. 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

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

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 $\hat{\sigma}^2$.

Note that back substitution in (3) yields an expression for the current smoothed value as an exponentially weighted moving average of past squared returns:

$$\sigma_t^2 = \sum_{j=0}^{\infty} w_j r_{t-1-j}^2 , \quad \text{where} \quad w_j = (1-\lambda)\lambda^j .$$

Hence the name "exponential smoothing."

Following RM we calculate the VaR simply as

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

where Φ^{-1} is the standard normal inverse cumulative density. Although the smoothing parameter λ can be estimated using standard techniques, it is often simply calibrated to 0.94 for daily returns. The implicit assumption of a zero mean and standard normal innovations implies that no other parameters need to be estimated.

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

$$Var(r_{w,t+k} + r_{w,t+k-1} + \ldots + r_{w,t+1} \,|\, \mathscr{F}_t) \equiv \sigma_{t:t+k|t}^2 = k\sigma_{t+1}^2 . \tag{5}$$

The RM model can thus be thought of as a random walk model in variance. The lack of mean-version in the RM variance model implies that the term structure of volatility is flat. Figure 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 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 extrapolating the daily variance by k as is done in (5) are discussed further in Diebold, Hickman, Inoue, and Schuermann (1998).

2.4 Volatility via GARCH

The implausible temporal aggregation properties of the RM model which we discussed above 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

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

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 using any methodology, mean-dynamics could in principle easily be incorporated into the models we discuss below.

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

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

Extensions to higher order models are in principle straightforward, but for notational simplicity we will concentrate on the (1,1) case here and throughout the chapter.

Note that back substitution in (7) 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 GARCH(1,1) volatility measurement is, perhaps surprisingly, 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. We estimates the vector of GARCH parameters θ by maximizing the log likelihood function,

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

Note that the assumption of conditional normality underlying the (quasi) likelihood function in (8) is merely a matter of convenience, allowing for the calculation of consistent and asymptotic normal parameter estimates. The conditional return distribution will generally not be normal. The log-likelihood optimization in (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. This optimization can be performed in a matter of seconds on a standard desktop computer using standard MBA-ware, as discussed by Christoffersen (2003).

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 "random walk plus noise" model (a "local level" model in the terminology of Harvey, 1989), in which case the minimum MSE 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, the volatilities are unlikely to follow random walks, and hence that the flat forecast function associated with exponential smoothing is unrealistic and undesirable for producing volatility forecasts.

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

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

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 write

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

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

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

which shows how the GARCH(1,1) forecasts by making adjustments around the long run variance with variance persistence governed by (α+β) 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 important 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

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

The variance of the k-day cumulative returns, which we use to calculate the volatility term structure, is then

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

Compare this mean-reverting expression with the RM forecast in (5). Note that the speed of mean-reversion in the GARCH(1,1) model is governed by α+β. The mean-reverting line in Figure 3 above is calculated from (12), normalizing by k and taking the square root.

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 conditions it implies a component structure obtained by allowing for time variation in the long-run variance in (9):

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

with the long-run component, q_t, modeled as an autoregressive process:

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

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.

To take a second example of the extensibility of GARCH models, note that all models considered thus far imply symmetric response to positive vs. 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; the "asymmetric GARCH(1,1)" model of Glosten, Jagannathan and Runkle (1993) is

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

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

3. Asset Level Analysis: Modeling Asset Return Covariance Matrices

The discussion above focused on the specification of dynamic volatility models for the aggregate portfolio return. These methods are well-suited for providing forecasts of portfolio level risk measures such as the aggregate VaR. However they are less well-suited for providing input into the active portfolio and risk management process. If, for example, the portfolio manager wants to know the optimal portfolio weights to minimize portfolio variance, then a multivariate model is needed, which provides a forecast for the entire covariance matrix.

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 many traders on each desk, so any sub-portfolio analysis is not possible with the portfolio-based, aggregate approach.

In this section we therefore consider the specification of models for the full N-dimensional conditional distribution of asset returns. In general we write the multivariate model as

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

where we have again set the mean to zero and where I denotes the identity matrix. $\Omega_t^{1/2}$ can be thought of as the Cholesky decomposition of the covariance matrix Ω_t.

This section will focus on specifying a dynamic NxN covariance matrix, and section 5 will suggest methods for specifying the distribution of the innovation vector Z.

Constructing positive semidefinite (psd) covariance matrix forecasts, which ensures that the portfolio variance is always non-negative, presents a key challenge below. The covariance matrix will have ½N(N+1) distinct elements, but structure needs to be imposed to guarantee psd. The practical issues involved in estimating the parameters guarding the dynamics for the ½N(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 attention to methods that are applicable even with N is relatively large.

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:

$$\Omega_t = \lambda \Omega_{t-1} + (1 - \lambda) R_{t-1} R_{t-1}^{\prime}. \tag{17}$$

The covariance matrix recursion may 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 RM has costs, it also has benefits; in particular, the simple structure (17) guarantees psd estimated covariance matrices, because 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 above, 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 positive semi-definite matrices is itself is positive semi-definite.

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 the multivariate GARCH model. The most general multivariate GARCH(1,1) model is

$$vech(\Omega_t) = vech(C) + B\,vech(\Omega_{t-1}) + A\,vech(R_{t-1} R_{t-1}^{\prime}), \tag{18}$$

where the *vech* (vector half") operator converts the upper triangular elements of a symmetric matrix into a column vector, and A and B are [½N(N+1)] x [½N(N+1)] matrices.[2] 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." In its most general form the GARCH(1,1) model (18) has a total of $1/2N^4 + N^3 + N^2 + 1/2N = O(N^4)$ parameters. Hence, for example, for N=100 the model has 51,010,050 parameters! Estimating this many parameters is obviously infeasible. Note also that without specifying more structure on the model there is no guarantee of positive definiteness of 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):

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

This is very useful in practice, because 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" pameterization 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 own lags and own lagged squared returns, and conditional covariance depend only on 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. We write

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

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 modification that it introduces, a non-degenerate long-run covariance matrix Ω to which Ω_t reverts. Notice also

---

[2] ½N(N+1) is the number of distinct elements in the covariance matrix.

though that all variance and covariances are assumed to have the same speed of mean reversion, because of common α and β parameters, which may be overly restrictive.

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 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 (20). Specific transformations of the parameter matrices from the bivariate models ensure that the resulting A, B and C matrices will be psd, which in turn ensures that the conditional covariance matrix forecast is psd. Flex-GARCH appears to be a viable modeling approach when N is larger than say five, where the general diagonal GARCH model becomes intractable. However, when N is of the order of thirty 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.

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,

$$\Omega_t \equiv D_t \, \Gamma_t D_t . \qquad (21)$$

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

Engle (2002) generalizes Bollerslev's (1990) constant correlation 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, which can differ from asset to asset, and then standardizes the returns by their estimated GARCH conditional standard deviations. Then one uses the standardized return vector e_t to model the correlation dynamics. A simple GARCH(1,1) correlation dynamic would be

$$Q_t = C + \beta Q_{t-1} + \alpha (e_{t-1} e_{t-1}^{\prime}), \qquad (23)$$

where the individual correlations in the matrix Γ_t are the corresponding normalized elements of Q_t:

$$\rho_{i,j,t} = q_{i,j,t} / (\sqrt{q_{i,i,t}} \sqrt{q_{j,j,t}}). \qquad (24)$$

The normalization in (24) ensures that all correlation forecasts fall in the [-1;1] interval, and the simple scalar structure of the dynamics of Q_t in (23) ensures that Γ_t is psd.

Notice that only two parameters need be estimated in (23), because C can be pre-estimated by correlation targeting as discussed earlier. Estimating variance dynamics asset-by-asset and then assuming a simple structure for the correlation dynamics 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. Estimation of the correlation parameters can be done via quasi-maximum likelihood under a normality assumption.

Although the DCC model offers a promising framework for exploring correlation dynamics in large systems, the simple dynamic structure in (23) may be too restrictive for many applications. For example, volatility and correlation responses may be asymmetric in the signs of shocks.[3] 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 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.

4. Exploiting High-Frequency Return Data for Improved Covariance Matrix Modeling

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

4.1 Realized Variances

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

[3] A second 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.

$$\sigma^2_{t,\Delta} \equiv \sum_{j=1}^{1/\Delta} r^2_{t-1+j\Delta,\Delta}, \tag{25}$$

where 1/   is, for example, 48 for 30-minute returns in 24-hour markets. Theoretically, letting   go to zero, which implies sampling continuously, we approach the true integrated volatility of the underlying continuous time process on day t.[4]

In practice, market microstructure noise will affect the RV estimate when   gets too small. Prices sampled at 15-30 minute intervals, depending on the market, are therefore often used. Notice also that, in markets that are not open 24 hours per day, the potential jump from the closing price on day t-1 to the opening price on day t must be accounted for. This can be done using the method in Hansen and Lunde (2004). As is the case for the daily GARCH models considered above, corrections may 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 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 to be the true daily variance, modeling and forecasting it using standard ARMA time series tools. Allowing for certain kinds of measurement error can be done in an ARMA framework as well. The upshot is that if the fundamental frequency of interest is daily, then using sufficiently high-quality intra-day price data enables the risk manager to treat volatility as essentially observed. This is vastly different from the GARCH style models discussed above, in which the daily variance is constructed recursively from past daily returns.

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

$$\log(\sigma_{t,\Delta}) \equiv c + \beta \log(\sigma_{t-1,\Delta}) + \nu_t, \tag{26}$$

which can be estimated using simple OLS. The log specification guarantees positivity of forecasted volatilities and induces normality, as first shown in Andersen, Bollerslev, Diebold and Labys (2001a, b). ABDL demonstrate empirically the superior forecasting properties of RV-based forecasts compared with GARCH forecasts. Rather than relying on simple short-memory ARMA models they specify a fractionally integrated models, which capture the long memory routinely found in realized volatility dynamics.

Figure 5 shows clear evidence of long-memory in foreign exchange RVs as captured by the sample autocorrelation function for lags of 1 through 100 days. We first construct the daily RVs from 30-minute FX returns and then construct the sample autocorrelations of the RVs. Note that the RV autocorrelations are significantly positive for all 100 lags when compared with the Bartlett bands.

---

[4] For a full treatment, see Andersen, Bollerslev and Diebold (2004).

Note that RV forecasts may be integrated into the GARCH framework. For example, 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 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.

<u>4.2 Realized Covariances</u>

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

$$\Omega_{t,\Delta} \equiv \sum_{j=1}^{1/\Delta} R_{t-1+j\Delta,\,\Delta} \, R'_{t-1+j\Delta,\,\Delta} \,. \tag{27}$$

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 GARCH models above. Using intra-day 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 N < 1/ , the realized covariance matrix will be positive definite. However 1/ is, for example, 48, and so in large portfolios the condition is likely to be violated unless high-quality price data is available at very high frequencies. 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. Non-synchronous trading, however, creates additional potential problems in the multivariate case. These are similar, but potentially more severe, than the non-synchronous trading issues that arise in the estimation of say, monthly covariances and CAPM betas with non-synchronous daily data. A possible fix involves inclusion of additional lead and lag terms in the realized covariance measure (27), 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 realized covariances. They are quite speculative, as little work has been done in terms of assessing the economic value of realized covariances for portfolio optimization problems.[5] Treating the realized covariances as an observable vector time series, and paralleling the tradition of the scalar diagonal GARCH model, suggests the model

---

[5] One notable exception is the work of Fleming, Kirby, Oestdiek (2003), which suggests dramatic improvements vis-a-vis the RM and multivariate GARCH frameworks.

$$vech\,(\Omega_{t,\Delta}) \;=\; vech\,(C) \;+\; \beta\, vech\,(\Omega_{t-1,\Delta}) \;+\; \nu_t, \tag{28}$$

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 non-scalar diagonal GARCH model could be developed as well, 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

$$\Omega_{t,\Delta} \;\equiv\; P_{t,\Delta} P_{t,\Delta}^{\prime}, \tag{29}$$

where $P_{t,\Delta}$ is a unique lower triangular matrix. The data vector is then vech($P_{t,\Delta}$), and we substitute the forecast of vech($P_{t,\Delta}$) back into (29) to construct a forecast of $\Omega_{t,\Delta}$.

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

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

$$\Omega_{t,\Delta} \;\equiv\; D_{t,\Delta}\, \Gamma_{t,\Delta} D_{t,\Delta}, \tag{30}$$

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

$$\rho_{i,j,t,\Delta} \;\equiv\; \sigma_{i,j,t,\Delta} / (\sigma_{i,i,t,\Delta}\sigma_{j,j,t,\Delta}). \tag{31}$$

Following the DCC idea, we model the standard deviations asset-by-asset in the first step, and we model the correlations in the second step. Keeping a simple structure as in (23), we have

$$vech\,(Q_{t,\Delta}) \;=\; vech\,(C) \;+\; \beta\, vech\,(Q_{t-1,\Delta}) \;+\; \nu_t, \tag{32}$$

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

$$\hat{\rho}_{i,j,t,\Delta} \;=\; \hat{q}_{i,j,t,\Delta} \,/\, (\sqrt{\hat{q}_{i,i,t,\Delta}}\,\sqrt{\hat{q}_{j,j,t,\Delta}}). \tag{33}$$

The advantages of this approach are twofold: first, high-frequency information is used to more precise forecasts of volatility and correlation. 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 intra-day 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.

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 often work from a set of, say, 30 observed base assets believed to be key drivers of risk. These of course depend on the portfolio at hand but can, for example, consist of equity market indices, FX rates, benchmark interest rates, and so on, which are believed to capture the key sources of uncertainty in the portfolio. The assumptions made on the multivariate distribution of base assets are naturally of key 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 with $N_F < 1/$ 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). We can write the factor model as

$$R_t = \beta R_{F,t} + \nu_t. \tag{34}$$

We obtain the factor loadings in the NxN_F matrix B from regression (if data exists), or via pricing model sensitivities (if a pricing model exists), and otherwise using 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.[6]

We now need a multivariate model for the base assets. We write

$$R_{F,t} = \Omega_{F,t}^{1/2} Z_{F,t} \qquad Z_{F,t} \sim i.i.d. \quad E(Z_{F,t}) = 0 \quad Var(Z_{F,t}) = I\,, \tag{35}$$

and we use the modeling strategies discussed above to construct the $N_F x N_F$ realized factor

[6] 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.

covariance matrix $\Omega_{F,t}$.[7]

5. Modeling Entire Conditional Distributions

Proper portfolio and risk management requires knowing the entire multivariate distribution of asset or base asset returns, not just the second moments. Conventional risk measures such as Value-at-Risk (VaR) and expected shortfall, however, capture only limited aspects of the distribution.

In this section we therefore explore various approaches to complete the model. Notice that above 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 below 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 such higher-ordered conditional moment dynamics, however, is 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 non-normality in the unconditional returns, conditional returns still exhibit non-normal features. Interestingly, however, the features vary from market to market. For example, mature market FX returns appear to be strongly conditionally kurtotic whereas equity index returns appear to be conditionally skewed.

As an example of the latter, we show in Figure 6 the daily QQ plots for S&P500 returns from January 2, 1990 to December 31, 2002, standardized using volatilities from an asymmetric GARCH(1,1) model. That is, we plot quantiles of standardized returns against quantiles of the standard normal distribution. Figure 6 makes clear that the left tail of S&P500 returns conforms much less well to the normal distribution than does the right tail: there are more large negative returns 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 6 suggests that a normal-GARCH VaR would work well for any coverage rate for a portfolio which is short the S&P500. It may also work well for a long portfolio but only if the coverage rate is relatively large, say 5-10%.

Consider now instead the distribution of returns standardized by realized volatility. In contrast to the poor fit in the left tail evident in Figure 6, the distribution in Figure 7 is strikingly close to normal, as first noticed by Zhou (1996) and Andersen, Bollerslev, Diebold and Labys

---

[7] $\Omega_{F,t}^{1/2}$ is the Cholesky factor of the covariance matrix $\Omega_{F,t}$.

(2000). Figures 6 and 7 rely on the same series of daily S&P500 returns but simply use two different volatility measures to standardize the raw returns. The conditional non-normality of daily returns has been a key stylized fact in market risk management. Finding a volatility measure which can generate standardized returns that are close to normal is therefore surprising and noteworthy.

Figure 7 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 / log-normal 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 log-normal. This idea is explored empirically in ABDL, who find that a log-normal / normal mixture VaR model performs very well in an application to foreign exchange returns.

The recent empirical results in Andersen, Bollerslev and Diebold (2003) 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 standard methods available first for univariate and then for multivariate models.

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 in the conditional distribution. In the case of VaR we are looking for ways to calculate the cut-off z_p^{-1} in

$$VaR^{p}_{T+1|T} \equiv \sigma_{T+1} z_p^{-1} . \qquad (36)$$

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, which relies on only one parameter and is able to generate symmetric fat tails. Recently, generalizations of the Student's t which allow of 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 non-normal distributions via Cornish-Fisher approximations. The only inputs needed are the sample estimates of skewness and kurtosis of standardized returns. Extreme value theory provides another approximation alternative, in which the tail of the conditional distribution is estimated using only the extreme observations, as suggested in Diebold, Schuermann, and Stroughair

(1998).

A common problem of 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 daily returns are GARCH(1,1)-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 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 allows for calculation of the term structure of VaR, for example via the methods of Albanese et al. (2004). But in general, simulation methods are needed, and we now discuss a viable approach which combines a parametric volatility model with a data-driven conditional distribution.

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 very easy to do. We simply calculate the standardized pseudo portfolio returns as

$$\hat{z}_{w,t} = r_{w,t} \,/\, \hat{\sigma}_t, \quad for\ t = 1, 2, \ldots, T, \tag{37}$$

using a variance model from section 2. For one-day VaR we simply use the order statistic for the standardized returns combined with the volatility forecast,

$$FHS\text{-}VaR^{p}_{T+1} \equiv \sigma_{T+1}\, \hat{z}_w((T+1)p). \tag{38}$$

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

5.3 Asset Level: Multivariate Analytic Methods

Just as a fully specified univariate distribution is needed for complete risk measurement, so too is a fully specified multivariate distribution needed for proper portfolio and risk management to be done. We therefore now need to 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. So long as the realized volatilities are available, a multivariate version of the log-normal mixture model could therefore be developed.

As noted earlier, however, construction and use of realized covariance matrices may be problematic in when liquidity is not high, in which case traditional parametric models may be used. As in the univariate case, however, multivariate standardization using covariance matrices estimated from traditional parametric models, although obviously convenient[8], does not generally provide an accurate picture of tail risk.

A few analytic alternatives to the multivariate normal paradigm do exist, such as the multivariate t, as in Glasserman, Heidelberger, and Shahbuddin (2002). Recently much attention has also been focused on construction of multivariate densities from the marginal densities via copulas, as in 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 Figure 7 we show the extreme correlations between the U.S. and German returns calculated for various tail thresholds. 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; as revealed in Figure 8, the correlation between negative extremes is much larger than the normal distribution would suggest.[9] 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 extreme dependencies. Once again, however, it is not yet clear whether such methods will be operational in large-scale systems.

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

5.4 Asset Level: Multivariate Simulation Methods

In the general multivariate case, we can in principle use FHS with dynamic correlations,

---

[8] In the multivariate case the normal distribution is even more tempting to use, because it implies that the aggregate portfolio distribution itself is normally distributed, so that the portfolio VaR is easily calculated.

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

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

$$\hat{Z}_t = \hat{\Omega}_t^{-1/2} R_t \quad for \; t = 1, 2, \ldots, T, \tag{39}$$

where we calculate $\hat{\Omega}_t^{-1/2}$ from the Cholesky decomposition of the inverse covariance matrix $\hat{\Omega}_t^{-1}$. Resampling with replacement vector-wise from the standardized returns will ensure that particular features of the marginal distributions, as well as, for example, the constant cross-sectional dependencies such as those shown in Figure 8, will be preserved in the simulated data.

The dimensionality of the system in (39) may render the necessary multivariate standardization infeasible, but in the base asset setup we can also do FHS. From (35) we can resample from the factor innovations calculated as

$$\hat{Z}_{F,t} = \hat{\Omega}_{F,t}^{-1/2} R_{F,t} \quad for \; t = 1, 2, \ldots, T, \tag{40}$$

where we again use the Cholesky decomposition to build up distribution of factor returns. Then from (34) we can construct idiosyncratic asset innovations

$$\hat{\nu}_t = R_t - \hat{B} R_{F,t} \quad for \; t = 1, 2, \ldots, T, \tag{41}$$

and we then resample from $\hat{Z}_t$ and $\hat{\nu}_t$ to build up the distribution of asset returns in the base asset model.

If one is willing to assume constant correlations, then the standardization can simply be done on an asset-by-asset basis using the GARCH volatilities. Resampling vector-wise from the standardized returns will preserve the constant correlation present in the historical data.

6. Summary and Directions for Future Research

We hope to have demonstrated in this survey the potential general usefulness of dynamic econometric asset return models for financial risk management. Needless to say, the rapidly developing literature on high-frequency volatility modeling has important implications for risk management practice. A conclusive survey on this topic, however, would be premature at this point, so we have simply tried to foster interaction between academics and practitioners, surveying and unified recent academic results on volatility and covariance modeling, with emphasis on implications for practical risk management.

6.1 Summary

Our key points are as follows:

- 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.

- For the purpose of risk *measurement*, specifying a univariate density model directly on the portfolio return is likely to lead to be most accurate. RiskMetrics offers one possible approach, but the temporal aggregation properties–including the volatility term structure–of RiskMetrics model appear to be counter-factual.

- The GARCH volatility models offer a convenient and parsimonious framework for modeling key dynamic features of returns, including volatility mean-reversion, long-memory, and asymmetric response.

- 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.

- 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.

- High-frequency volatility measures present a promising venue for risk management. Realized volatility and correlation measures give more accurate forecasts of future realizations. Because high-frequency information is only available for highly liquid assets, we suggest a base-asset factor approach.

- 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 non-normal returns.

- The near log-normality of realized volatility, together with the surprising near-normality of returns standardized by realized volatility, holds promise for log-normal / normal mixture models in financial risk management.

Several important topics have thus far been omitted from this survey. We conclude by outlining two key areas for future research, namely option valuation and the treatment of microstructure noise.

6.2 Future Research I: Option Valuation

The seminal Black-Scholes-Merton (BSM) model has produced indispensable insight into the valuation of derivative securities. However, many empirical studies have shown that the assumptions of constant volatility and normal returns cause the BSM model systematically to misprice options. Hence we have developed increasingly sophisticated models to deal with the shortcomings of the BSM model. Nevertheless, practitioners often rely on curves fitted to the implied BSM volatility surface to compute option prices. Thus there appears to be a disconnect between the academic and practitioner camps in this area.

One reason for the disconnect is clearly that the more sophisticated academic models involve an increasing number of unobserved parameters and volatility factors, rendering them much harder to implement than the simple BSM model. However, recent advances in estimation of continuous time models (e.g. Eraker, 2004) may help reduce that problem.

Another way to close the apparent gap is to develop potentially more tractable discrete-time analogues of the continuous time models. Heston and Nandi (2000) thus define a discrete time GARCH version of Heston's (1993) stochastic volatility model. Christoffersen, Heston and Jacobs (2004) extend the Heston and Nandi model to allow for conditionally non-normal returns.

We have already discussed the need to model the long-memory features of return volatility dynamics. Getting the volatility dynamics right is clearly key for option valuation. One way to approximate long memory is via component models. Bates (2000) studies continuous time volatility component models in derivative pricing contexts, and Christoffersen, Jacobs and Wang (2004) study discrete time analogues.

6.3 Future Research II: Microstructure Noise

The potential usefulness of intra-day returns for daily risk management has been a key theme in our discussion so far. The realized volatility literature is still quite new, and many important practical considerations are still being explored. The influence of market microstructure noise on realized volatility is one such consideration.

The theoretical result giving integrated volatility as the limit of realized volatility when the sampling frequency goes to infinity is derived under the assumption of no market microstructure noise. When allowing for market microstructure noise, the simple limit result no longer holds, and sampling infinitely often – or in practice, "as often as possible" – is not generally optimal. Bandi and Russel (2004a, 2004b) derive optimal finite sampling frequencies for given specifications of market microstructure noise. But the suggestion to sample "less often than possible," thus discarding information, may be inefficient, and Ait-Sahalia, Mykland and Zhang (2004) suggest methods for using all available price information rather than sampling at certain prespecified frequencies.

References

Ait-Sahalia, Y., P. Mykland, and L. Zhang, 2004, How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, Review of Financial Studies, forthcoming .

Alizadeh, S., M. Brandt, and F. Diebold, 2002, Range-Based Estimation of Stochastic Volatility Models, Journal of Finance, 57, 1047-1091.

Andersen, T. 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., T. Bollerslev, and F.X. Diebold, 2003, Some Like it Smooth, and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modeling, and Forecasting Asset Return Volatility, Manuscript, Northwestern University, Duke University and University of Pennsylvania.

Andersen, T.G., T. Bollerslev, and F.X. Diebold, 2004, "Parametric and Nonparametric Volatility Measurement," in L.P. Hansen and Y. Ait-Sahalia (eds.), Handbook of Financial Econometrics. Amsterdam: North-Holland, forthcoming

Andersen, T., T. Bollerslev, F.X. Diebold, and P. Labys, 2000, Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian, Multinational Finance Journal, 4, 159-179.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2001, "The Distribution of Realized Exchange Rate Volatility," Journal of the American Statistical Association, 96, 42-55.

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., T. Bollerslev, F.X. Diebold, and P. Labys, 2003, Modeling and Forecasting Realized Volatility, Econometrica , 71, 529-626.

Baillie, R.T. and T. Bollerslev, 1992, Prediction in Dynamic Models with Time-Dependent Conditional Variances, Journal of Econometrics, 51, 91-113.

Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical Performance of Alternative Option Pricing Models, Journal of Finance 52, 2003-2050.

Bandi, F. and J. Russell, 2004a, Separating Market Microstructure Noise from Volatility, Manuscript, University of Chicago.

Bandi, F. and J. Russell, 2004b, Microstructure Noise, Realized Volatility, and Optimal Sampling, Manuscript, University of Chicago.

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.

Barndorff-Nielsen, O. and N. Shephard, 2004, Power and Bipower Variation with Stochastic Volatility and Jumps, Journal of Financial Econometrics, 2, 1-48.

Barone-Adesi, G., Bourgoin, F. and K. Giannopoulos, 1998, Don't Look Back, Risk, 11, August, 100-104.

Bates, D., 2000, Post-87 Crash Fears in S&P 500 Futures Options, Journal of Econometrics 94, 181-238.

Benson, P. and P. Zangari, 1997, A General Approach to Calculating Var Without Volatilities and Correlations, RiskMetrics Monitor, 2nd Quarter, 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, May, 64-67.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

Bollerslev, T., 1987, A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, Review of Economics and Statistics, 69, 542-547.

Bollerslev, T., R.F. Engle, and J. Wooldridge, 1988, A Capital Asset Pricing Model with Time-Varying Covariances, Journal of Political Economy, 96, 116-131.

Bollerslev, T. and H.O. Mikkelsen, 1999, Long-Term Equity Anticipation Securities and Stock Market Volatility Dynamics, Journal of Econometrics, 92, 75-99.

Bollerslev, T. and H. Zhou, 2002, Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility, Journal of Econometrics, 109, 33-65.

Brandt, M.W. and F.X. Diebold, 2004, A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations, Journal of Business, forthcoming.

Brandt, M. and C. Jones, 2002, Volatility Forecasting with Range-Based EGARCH Models, Manuscript, University of Pennsylvania.

Cappiello, L., R.F. Engle, and K. Sheppard, 2004, Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns, 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, 1204-1221.

Christoffersen, P., K. Jacobs, and Y. Wang, 2004, Option Valuation with Volatility Components, Manuscript, McGill University.

Christoffersen, P., S. Heston, and K. Jacobs, 2004, Option Valuation with Conditional Skewness, Journal of Econometrics, forthcoming .

Christoffersen, P. and N. Meddahi, 2004, The Term Structure of Financial Risk Using Realized Volatility, work in progress, McGill University and the University of Montreal.

De Santis, G., R. Litterman, A. Vesval, and K. Winkelmann, 2003, Covariance Estimation, in Modern Investment Management: An Equilibrium Approach, edited by R. Littermann. London: Wiley.

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 97-34. Published in condensed form as Scale Models, Risk, 11, 104-107.

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 A.-P. N. Refenes, A.N. Burgess and J.D. Moody (eds.), Decision Technologies for Computational Finance, 3-12. Amsterdam: Kluwer Academic Publishers. (Reprinted in Journal of Risk Finance, 1 (Winter 2000), 30-36, and in P. Embrechts (ed.), Extremes and Integrated Risk Management. London: Risk Publications, 2000.)

Engle, R.F., 2002, Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, Journal of Business and Economic Statistics, 20, 339-350.

Engle, R.F. and G.G.J. Lee, 1999, A Permanent and Transitory Component Model of Stock Return Volatility, in R.F. Engle and H. White (eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger, 475-497. Oxford, UK: Oxford University Press.

Engle, R.F. and J. Mezrich, 1996, GARCH for Groups, Risk, 8, 36-40.

Eraker, B., 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, Journal of Finance 59, 1367-1403.

Fernandez, C. and M.F.J. Steel, 1998, On Bayesian Modeling of Fat Tails and Skewness, Journal of the American Statistical Association, 93, 359-371.

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, 617-631.

Glasserman, P., P. Heidelberger, and P. Shahbuddin, 2002, Portfolio Value-at-Risk with Heavy-tailed Risk Factors, Mathematical Finance, 239–269.

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, 2004, Term Structure of Risk under Alternative Econometric Specifications, Manuscript, University of Virginia and UCSD.

Hansen, B., 1994, Autoregressive Conditional Density Estimation, International Economic Review, 35, 705-730.

Hansen, P.R. and A. Lunde, 2004, A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data, Manuscript, Stanford University.

Harvey, A.C., 1989, Forecasting Structural Time Series Models and the Kalman Filter, Cambridge University Press.

Heston, S., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, Review of Financial Studies 6, 327-343.

Heston, S. and S. Nandi, 2000, A Closed-Form GARCH Option Pricing Model, Review of Financial Studies 13, 585-626.

Jorion, P., 2000, Value-at-Risk. New York: McGraw-Hill.

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, 603-621.

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

Longin, F. And B. Solnik, 2001, Extreme Correlation of International Equity Markets, Journal of Finance, 56, 649-676.

Martens, M., 2004, Estimating Unbiased and Precise Realized Covariances, Manuscript, Erasmus University Rotterdam.

Nerlove, M. and S. Wage, 1964, On the Optimality of Adaptive Forecasting, Management Science, 10, 207-229.

Patton, A.J., 2002, Modeling Time-varying Exchange Rate Dependence Using the Conditional Copula. Working paper, LSE.

Pelletier, D., 2004, Regime Switching for Dynamic Correlations, Journal of Econometrics, forthcoming.

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.

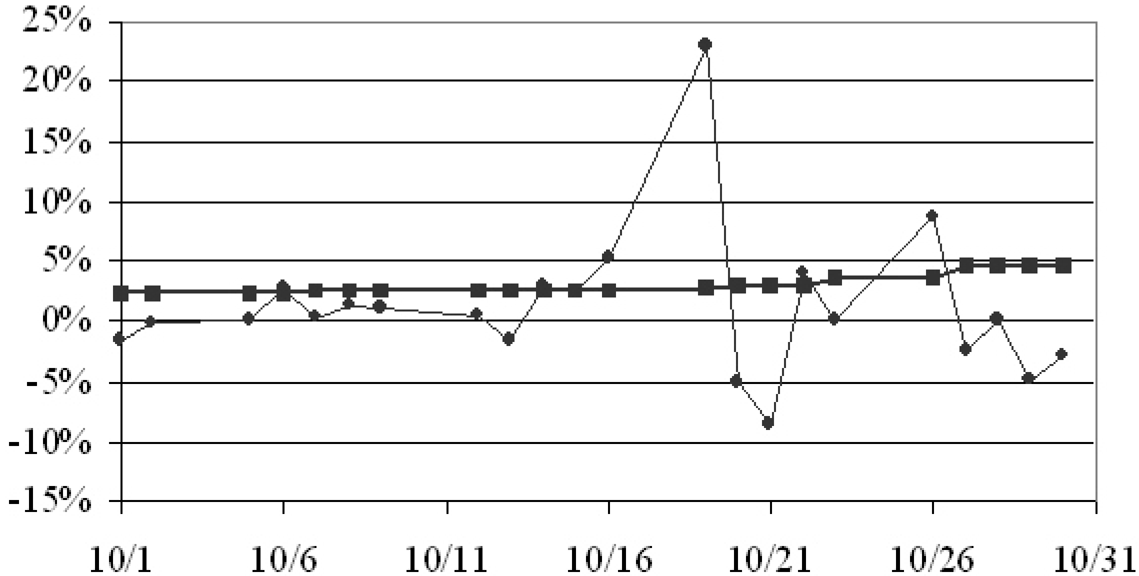
Pritsker, M., 2001, The Hidden Dangers of Historical Simulation, Manuscript, Federal Reserve Board.

Ray, B., 1993, Modeling Long-memory Processes for Optimal Long-Range Prediction, Journal of Time Series Analysis, 14, 511-525.

Theil, H. and S. Wage, 1964, Some Observations on Adaptive Forecasting, Management Science, 10, 198-206.

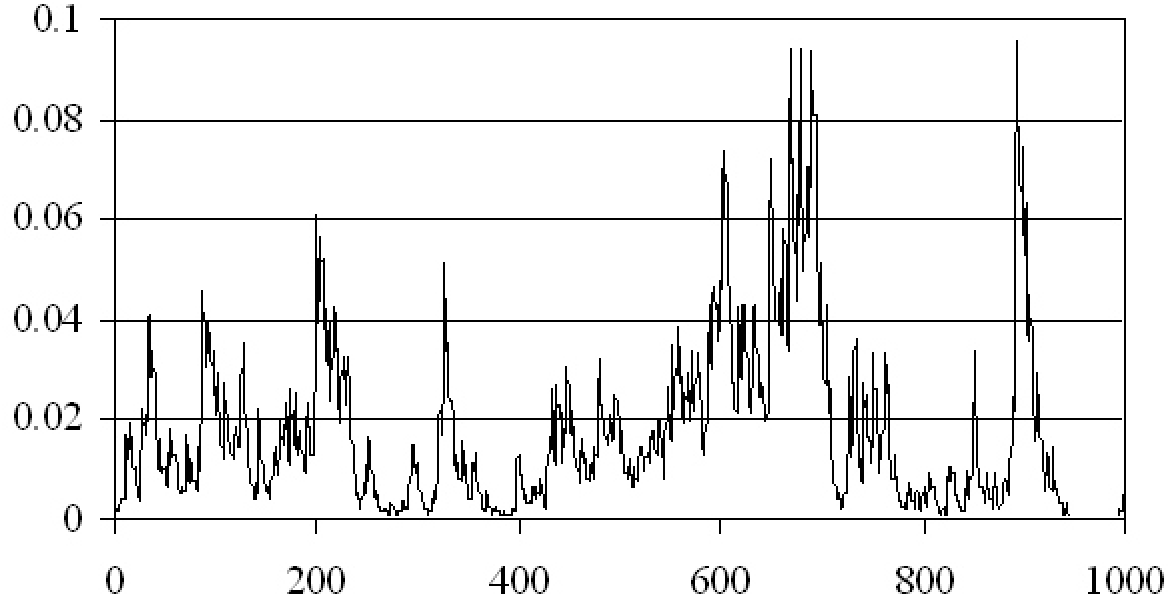
Zhou, B., 1996, “High-Frequency Data and Volatility in Foreign-exchange Rates, Journal of Business and Economic Statistics, 14, 45-52.

Figure 1. October 1987: Daily S&P500 Loss and 1% HS-VaR



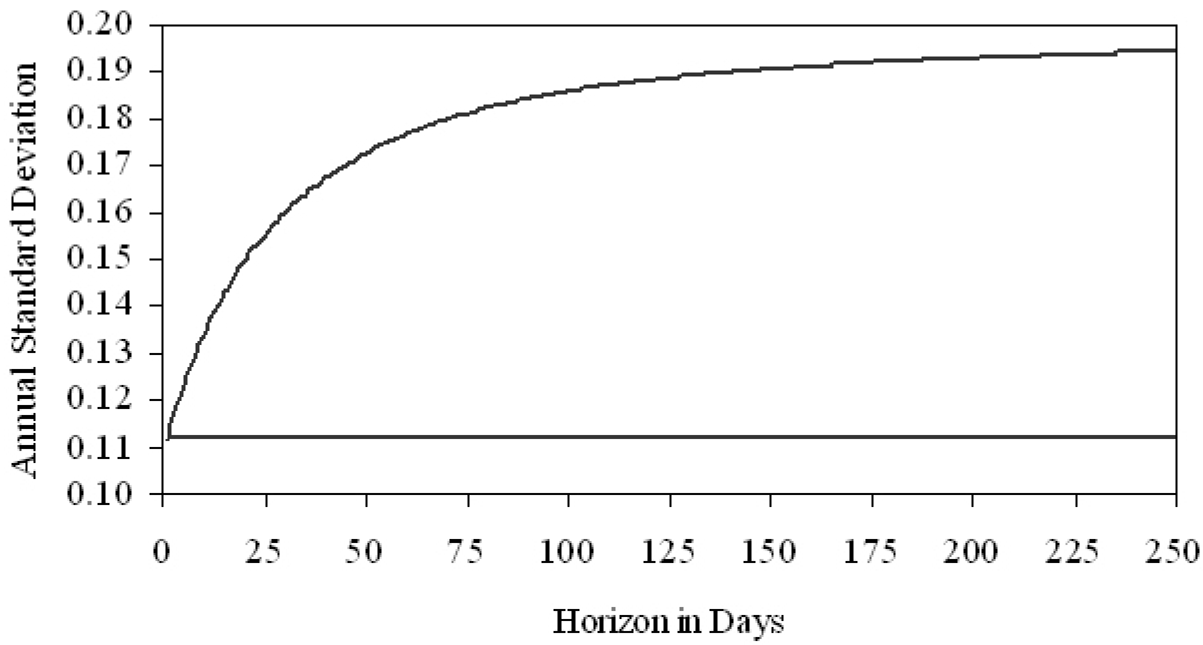
Notes to Figure: The thin line with diamonds shows the daily percentage loss on an S&P500 portfolio during October 1987. The thick line with squares shows the daily 1% VaR from historical simulation using a 250-day window.

Figure 2. True Conditional Coverage of 1% VaR from Historical Simulation



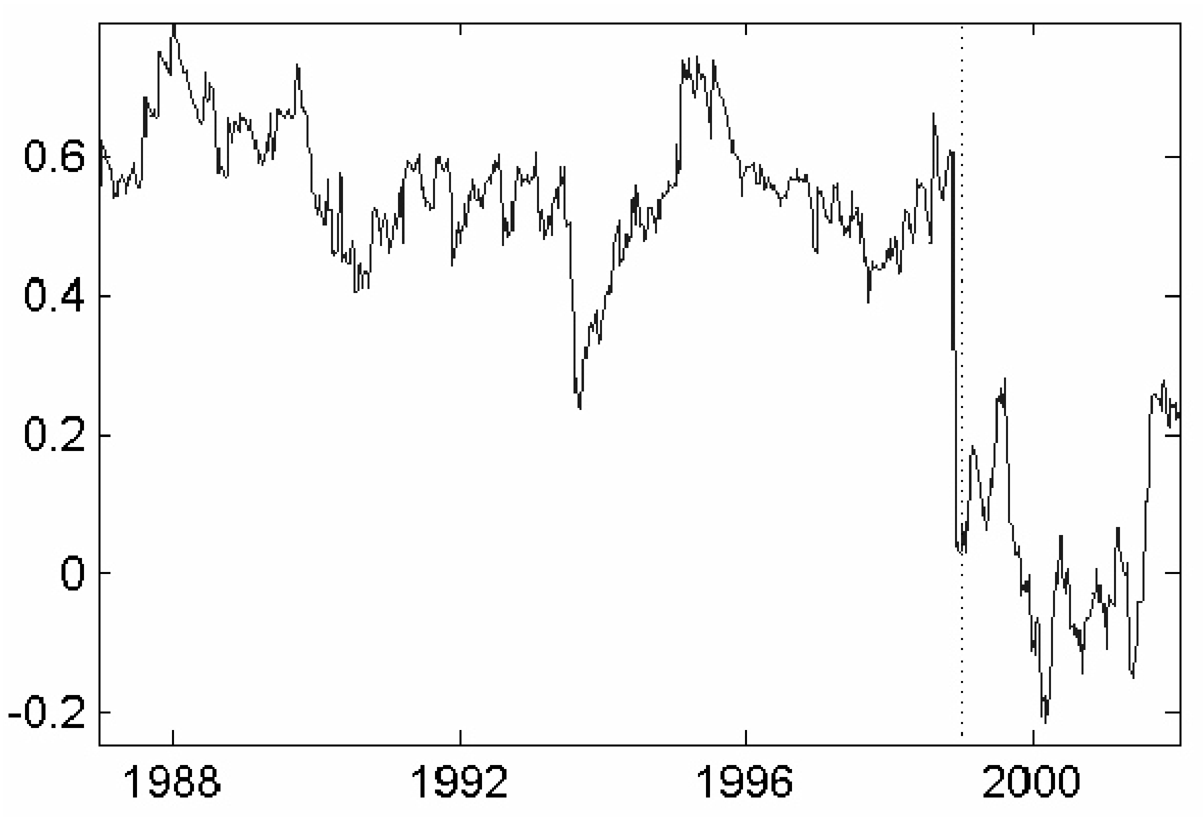
Notes to Figure: We simulate returns from a GARCH model with normal innovations, after which we compute the 1% HS-VaR using a rolling window of 250 observations, and then we plot the *true* conditional coverage probability of the HS-VaR, which we calculate using the GARCH structure.

Figure 3. Term Structure of Variance in GARCH and RiskMetrics Models



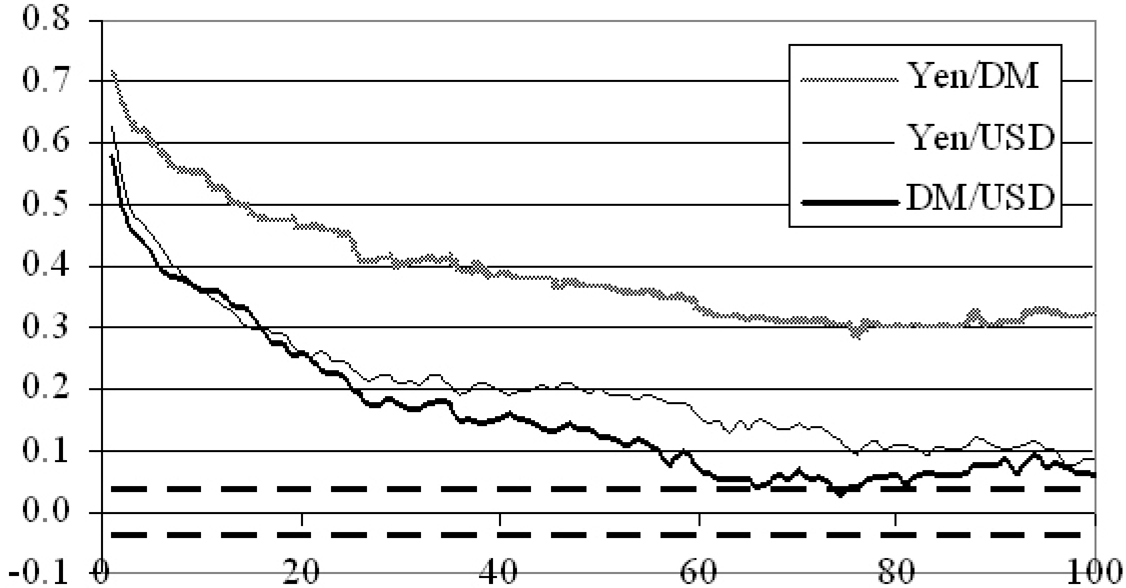
Notes to Figure: We plot the term structure of variance from a mean-reverting GARCH model (thick line) as well as the term structure from a RiskMetrics model (thin line). The current variance is assumed to be identical across models.

Figure 4. Time-Varying Bond Return Correlation: Germany and Japan



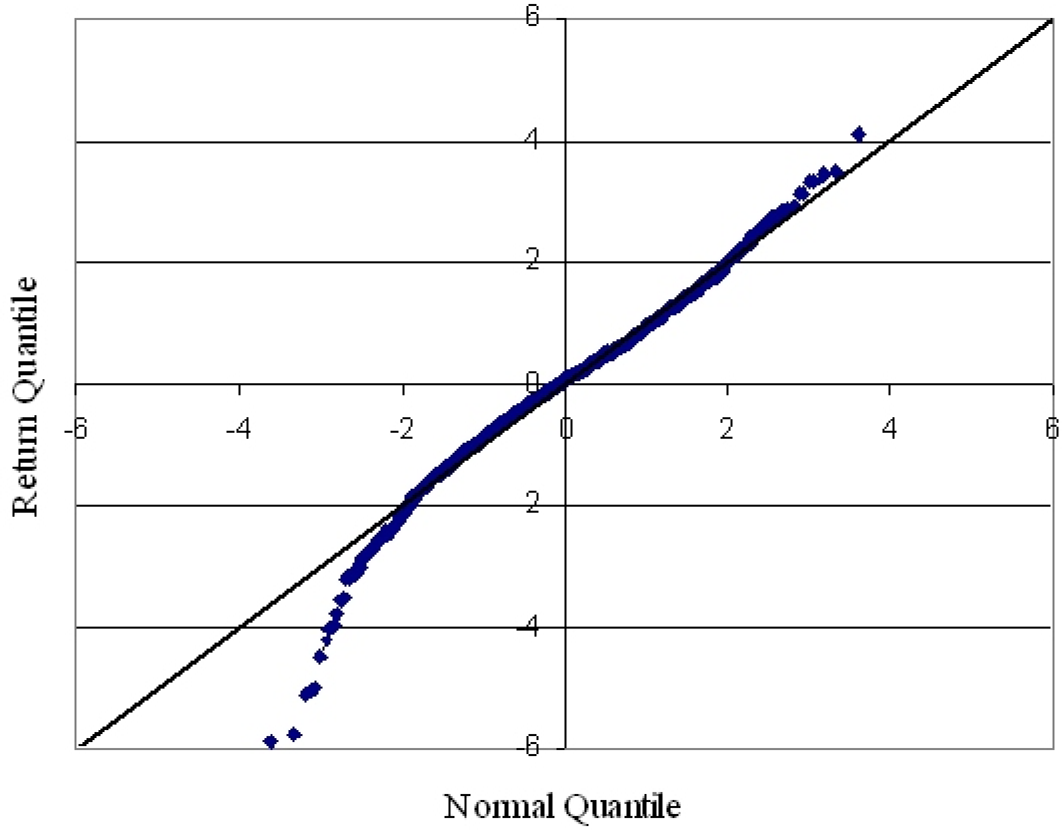
Notes to Figure: We reconstruct this figure from Capiello, Engle and Sheppard, 2004, plotting the correlation between German and Japanese government bond returns calculated from a DCC model allowing for asymmetric correlation responses to positive and negative returns. The vertical dashed line denotes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

Figure 5. Sample Autocorrelations of Realized Volatility: Three Currencies



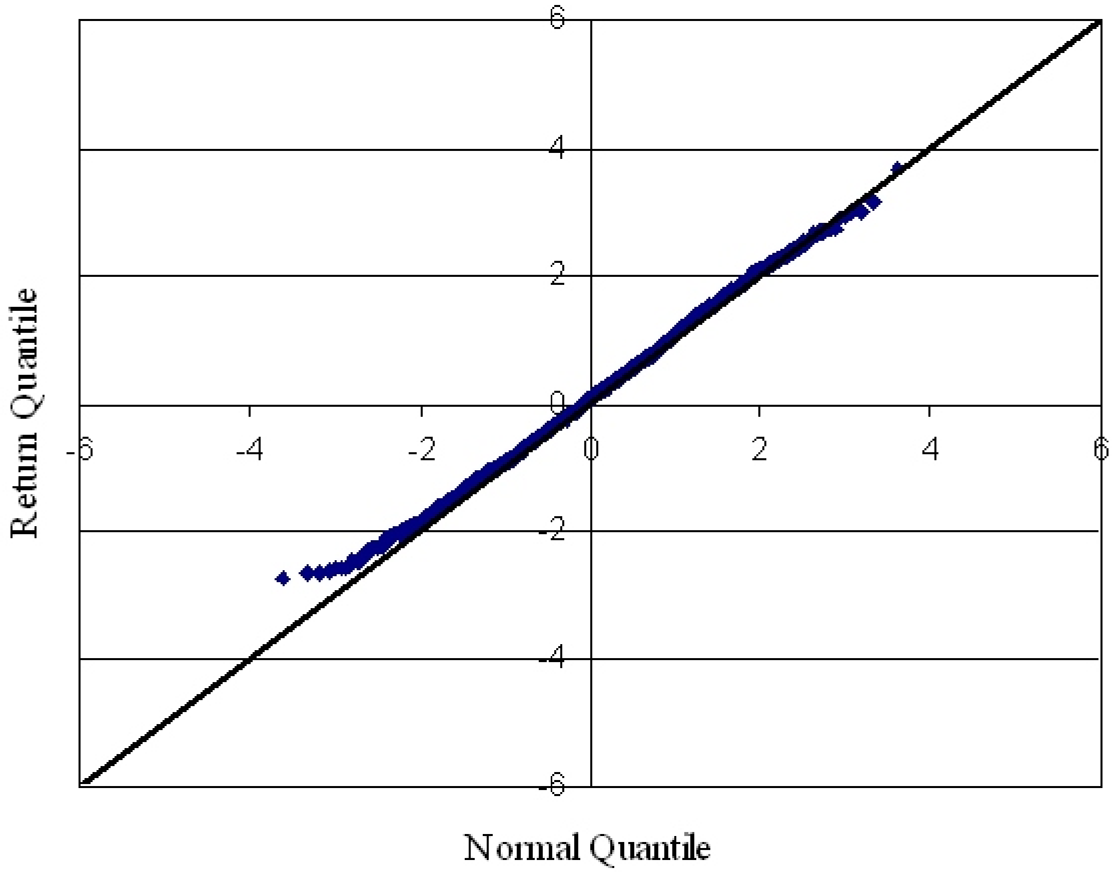
Notes to Figure: We plot the sample autocorrelations of daily realized log standard deviations for three FX rates, together with Bartlett's +/- 2 standard error bands for the sample autocorrealtions of white noise. We construct the underlying daily realized variance using 30-minute returns from December 1, 1986, through December 1, 1996.

Figure 6. QQ Plot of S&P500 Returns Standardized by GARCH Volatility



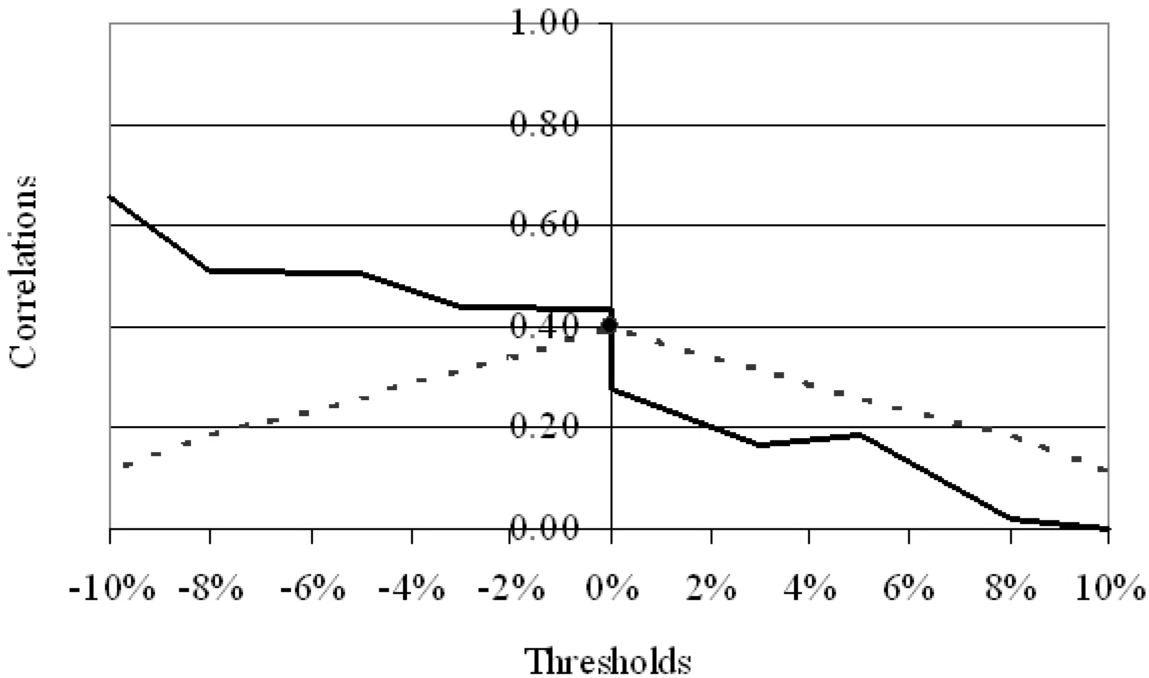
Notes to Figure: We show quantiles of daily S&P500 returns from January 2, 1990 to December 31, 2002, standardized by volatility from a fitted asymmetric GARCH(1,1) model, against the corresponding quantiles from a standard normal distribution.

Figure 7. QQ Plot of S&P500 Returns Standardized by Realized Volatility



Notes to Figure: We show quantiles of daily S&P500 returns from January 2, 1990 to December 31, 2002, standardized by realized volatility calculated from 5-minute futures returns, against the corresponding quantiles from a standard normal distribution.

Figure 8. Asymmetric Extreme Equity Correlations. U.S. versus Germany



Notes to Figure: We reconstruct this figure from Longin and Solnik, 2001, using monthly country index returns from MSCI (1959-1996) to construct correlations of returns truncated at various thresholds (solid line). We show in dashes the extreme correlations implied by the bivariate normal distribution.