

Financial Risk Measurement for Financial Risk Management*

Torben G. Andersen^{a1}, Tim Bollerslev^{b2}, Peter F. Christoffersen^{c3}, and Francis X. Diebold^{d4}

^aKellogg School of Management, Northwestern University, Evanston, IL 60208, USA

^bDepartment of Economics, Duke University, Durham, NC 27708, USA

^cRotman School of Management, University of Toronto, Ontario, Canada M5S 3E6

^dDepartment of Economics, University of Pennsylvania, Philadelphia, PA 19104, USA

Contents

1. Introduction	1128
1.1 Six Emergent Themes	1129
1.2 Conditional Risk Measures	1130
1.3 Plan of the Chapter	1133
2. Conditional Portfolio-Level Risk Analysis	1133
2.1 Modeling Time-Varying Volatilities Using Daily Data and GARCH	1134
2.1.1 Exponential Smoothing and RiskMetrics	1134
2.1.2 The GARCH(1, 1) Model	1136
2.1.3 Extensions of the Basic GARCH Model	1139
2.2 Intraday Data and Realized Volatility	1142
2.2.1 Dynamic Modeling of Realized Volatility	1147
2.2.2 Realized Volatilities and Jumps	1151
2.2.3 Combining GARCH and RV	1154
2.3 Modeling Return Distributions	1156
2.3.1 Procedures Based on GARCH	1160

* For helpful comments we thank Hal Cole and Dongho Song. For research support, Andersen, Bollerslev and Diebold thank the National Science Foundation (US), and Christoffersen thanks the Social Sciences and Humanities Research Council (Canada).

¹ Torben G. Andersen is Nathan and Mary Sharp Distinguished Professor of Finance at the Kellogg School of Management, Northwestern University, Research Associate at the NBER, and International Fellow of CREATES, University of Aarhus, Denmark.

² Tim Bollerslev is Juanita and Clifton Kreps Professor of Economics, Duke University, Professor of Finance at its Fuqua School of Business, Research Associate at the NBER, and an International Fellow of CREATES, University of Aarhus, Denmark.

³ Peter F. Christoffersen is Professor of Finance at the Rotman School of Management, University of Toronto and affiliated with Copenhagen Business School and CREATES, University of Aarhus, Denmark.

⁴ Francis X. Diebold is Paul F. and Warren S. Miller Professor of Economics at the University of Pennsylvania, Professor of Finance and Statistics and Co-Director of the Financial Institutions Center at its Wharton School, and Research Associate at the NBER.

2.3.2	<i>Procedures Based on Realized Volatility</i>	1162
2.3.3	<i>Combining GARCH and RV</i>	1164
2.3.4	<i>Simulation Methods</i>	1165
2.3.5	<i>Extreme Value Theory</i>	1166
3.	Conditional Asset-Level Risk Analysis	1167
3.1	Modeling Time-Varying Covariances Using Daily Data and GARCH	1168
3.1.1	<i>Dynamic Conditional Correlation Models</i>	1171
3.1.2	<i>Factor Structures and Base Assets</i>	1174
3.2	Intraday Data and Realized Covariances	1176
3.2.1	<i>Regularizing Techniques for RCov Estimation</i>	1179
3.2.2	<i>Dynamic Modeling of Realized Covariance Matrices</i>	1184
3.2.3	<i>Combining GARCH and RCov</i>	1188
3.3	Modeling Multivariate Return Distributions	1190
3.3.1	<i>Multivariate Parametric Distributions</i>	1191
3.3.2	<i>Copula Methods</i>	1192
3.3.3	<i>Combining GARCH and RCov</i>	1194
3.3.4	<i>Multivariate Simulation Methods</i>	1196
3.3.5	<i>Multivariate Extreme Value Theory</i>	1197
3.4	Systemic Risk and Measurement	1199
3.4.1	<i>Marginal Expected Shortfall and Expected Capital Shortfall</i>	1199
3.4.2	<i>CoVaR and ΔCoVaR</i>	1200
3.4.3	<i>Network Perspectives</i>	1201
4.	Conditioning on Macroeconomic Fundamentals	1203
4.1	The Macroeconomy and Return Volatility	1204
4.2	The Macroeconomy and Fundamental Volatility	1205
4.3	Fundamental Volatility and Return Volatility	1207
4.4	Other Links	1207
4.5	Factors as Fundamentals	1209
5.	Concluding Remarks	1211
	References	1212

1. INTRODUCTION

Financial risk management is a huge field with diverse and evolving components, as evidenced by both its historical development (e.g. Diebold, 2012) and current best practice (e.g. Stulz, 2002). One such component—probably *the* key component—is risk *measurement*, in particular the measurement of financial asset–return volatilities and correlations (henceforth “volatilities”). Crucially, asset–return volatilities are time varying, with persistent dynamics. This is true across assets, asset classes, time periods, and countries, as vividly brought to the fore during numerous crisis events, most recently and prominently the 2007–2008 financial crisis and its long-lasting aftermath. 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 the new “volatility econometrics” 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.

1.1 Six Emergent Themes

Six key themes emerge, and we highlight them here. We treat some of them directly in explicitly focused sections, while we treat others indirectly, touching upon them in various places throughout the chapter, and from various angles.

The first theme concerns aggregation level. We consider both portfolio-level (aggregated, “top-down”) and asset-level (disaggregated, “bottom-up”) modeling, emphasizing the related distinction between risk measurement and risk management. Risk measurement generally requires only a portfolio-level model, whereas risk management requires an asset-level model.

The second theme concerns the frequency of data observations. We consider both low-frequency and high-frequency data, and the associated issue of parametric vs. non-parametric volatility measurement. We treat all cases, but we emphasize the appeal of volatility measurement using non-parametric methods used with high-frequency data, followed by modeling that is intentionally parametric.

The third theme concerns modeling and monitoring entire time-varying conditional densities rather than just conditional volatilities. We argue that a full conditional density perspective is necessary for thorough risk assessment, and that best-practice risk management should move—and indeed *is* moving—in that direction. We discuss methods for constructing full conditional density forecasts.

The fourth theme concerns dimensionality reduction in multivariate “vast data” environments, a crucial issue in asset-level analysis. We devote considerable attention to frameworks that facilitate tractable modeling of the very high-dimensional covariance matrices of practical relevance. Shrinkage methods and factor structure (and their interface) feature prominently.

The fifth theme concerns the links between market risk and macroeconomic fundamentals. Recent work is starting to uncover the links between asset-market volatility and macroeconomic fundamentals. We discuss those links, focusing in particular on links among equity return volatilities, real growth, and real growth volatilities.

The sixth theme, the desirability of conditional as opposed to unconditional risk measurement, is so important that we dedicate the following subsection to an extended discussion of the topic. We argue throughout the chapter that, for most financial risk management purposes, the conditional perspective is distinctly more relevant for monitoring daily market risk.

1.2 Conditional Risk Measures

Our emphasis on conditional risk measurement is perhaps surprising, given that many popular approaches adopt an unconditional perspective. However, consider, for example, the canonical Value-at-Risk (VaR) quantile risk measure,

$$p = Pr_T(r_{T+1} \leq -VaR_{T+1|T}^p) = \int_{-\infty}^{-VaR_{T+1|T}^p} f_T(r_{T+1}) dr_{T+1}, \quad (1)$$

where $f_T(r_{T+1})$ denotes the density of future returns r_{T+1} conditional on time- T information. As the formal definition makes clear, VaR is distinctly a conditional measure. Nonetheless, banks often rely on VaR from “historical simulation” (HS- VaR). The HS- VaR simply approximates the VaR as the 100 p th percentile or the T pth order statistic of a set of T historical pseudo portfolio returns constructed using historical asset prices but today’s portfolio weights.

Pritsker (2006) discusses several serious problems with historical simulation. Perhaps most importantly, it does not properly incorporate conditionality, effectively replacing the conditional return distribution in (1) with its unconditional counterpart. This deficiency of the conventional HS approach is forcefully highlighted by banks’ proprietary P/L as reported in Berkowitz and O’Brien (2002) and the clustering in time of the corresponding VaR violations, reflecting a failure by the banks to properly account for persistent changes in market volatility.⁵ The only source of dynamics in HS- VaR is the evolving window used to construct historical pseudo portfolio returns, which is of minor consequence in practice.⁶

Figure 1 directly illustrates this hidden danger of HS. We plot on the left axis the cumulative daily loss (cumulative negative return) on an S&P500 portfolio, and on the right axis the 1% HS- VaR calculated using a 500 day moving window, for a sample period encompassing the recent financial crisis (July 1, 2008–December 31, 2009). Notice that HS- VaR reacts only slowly to the dramatically increased risk in the fall of 2008. Perhaps even more strikingly, HS- VaR reacts very slowly to the decreased risk following the market trough in March 2009, remaining at its peak through the end of 2009. This happens because the early-sample extreme events that caused the increase in HS- VaR remain in the late-sample 500-day estimation window.

More generally, the sluggishness of HS- VaR dynamics implies that traders who base their positions on HS will reduce their exposure too slowly when volatility increases, and then increase exposure too slowly when volatility subsequently begins to subside.

⁵ See also Perignon and Smith (2010a).

⁶ Boudoukh, Richardson, and Whitelaw (1998) incorporate more aggressive updating into historical simulation, but the basic concerns expressed by Pritsker (2006) remain.

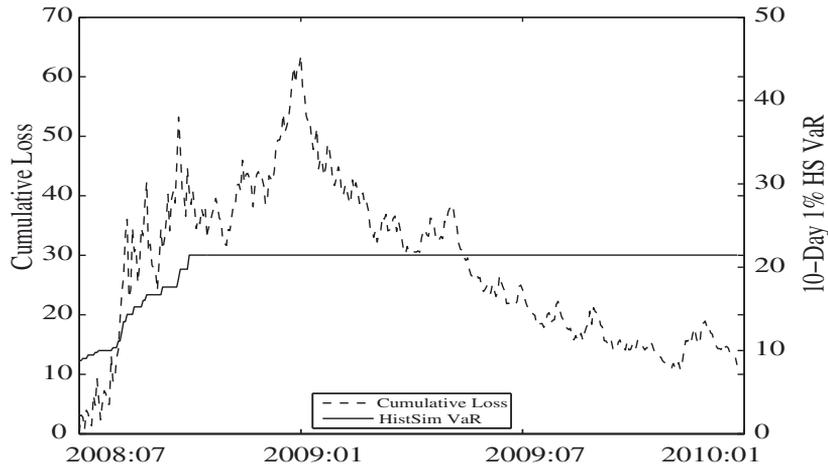


Figure 1 Cumulative S&P500 loss (left scale, dashed) and 1% 10-day HS- VaR (right scale, solid), July 1, 2008–December 31, 2009. The dashed line shows the cumulative percentage loss on an S&P500 portfolio from July 2008 through December 2009. The solid line shows the daily 10-day 1% HS- VaR based on a 500-day moving window of historical returns.

The sluggish reaction to current market conditions is only one shortcoming of HS- VaR . Another is the lack of a properly defined conditional model, which implies that it does not allow for the construction of a term structure of VaR . Calculating a 1% 1-day HS- VaR may be sensible on a window of 500 observations, but calculating a 10-day 1% VaR on 500 daily returns is not. Often the 1-day VaR is simply scaled by the square root of 10, but this extrapolation is typically not valid unless daily returns are i.i.d. and normally distributed, which they are not.⁷

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 500 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%, even though the unconditional coverage is correctly calibrated at 1%. 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 be as much as a 10% chance of exceeding the VaR . Figure 2 highlights the potential benefit of conditional density modeling: The HS- VaR may assess risk correctly

⁷ The i.i.d. return assumption alone is generally not enough because the distribution of returns, for the non-Gaussian case, will vary with the VaR horizon of interest; see, e.g. Bakshi and Panayotov (2010).

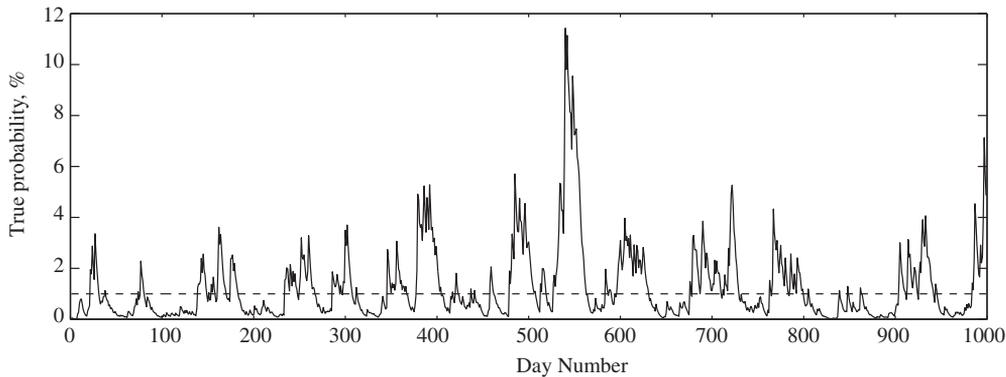


Figure 2 True exceedance probabilities of nominal 1% HS-*VaR* when volatility is persistent. We simulate returns from a realistically-calibrated dynamic volatility model, after which we compute 1-day 1% HS-*VaR* using a rolling window of 500 observations. We plot the daily series of true conditional exceedance probabilities, which we infer from the model. For visual reference we include a horizontal line at the desired 1% probability level.

on average (i.e. unconditionally) while still being terribly wrong at any given time (i.e. conditionally). A conditional density model will generate a dynamic *VaR* that attempts to keep the conditional coverage rate at 1% on any given day.

The above discussion also hints at a problem with the *VaR* risk measure itself. It does not say anything about how large the expected loss will be on days when *VaR* is exceeded. Other risk measures, such as expected shortfall (*ES*), attempt to remedy that defect. We define *ES* as

$$ES_{T+1|T}^p = p^{-1} \int_0^p VaR_{T+1|T}^\gamma d\gamma. \quad (2)$$

Because it integrates over the left tail, *ES* is sensitive to the shape of the *entire* left tail of the distribution.⁸ By averaging all of the *VaR*s below a prespecified coverage rate, the magnitude of the loss across all relevant scenarios matters. Thus, even if the *VaR* might be correctly calibrated at, say, the 5% level, this does not ensure that the 5% *ES* is also correct. Conversely, even if the 5% *ES* is estimated with precision, this does not imply that the 5% *VaR* is valid. Only if the return distribution is characterized appropriately throughout the entire tail region can we guarantee that the different risk measures all provide accurate answers.

Our main point of critique still applies, however. Any risk measure, whether *VaR*, *ES*, or anything else, that neglects conditionality, will inevitably miss important aspects

⁸ In contrast to *VaR*, the expected shortfall is a *coherent* risk measure in the sense of Artzner, Delbaen, Eber, and Heath (1999) as demonstrated by, e.g. Föllmer and Schied (2002). Among other things, this ensures that it captures the beneficial effects of portfolio diversification, unlike *VaR*.

of the dynamic evolution of risk. In the conditional analyses of subsequent sections, we focus mostly on conditional VaR , but we also treat conditional ES .⁹

1.3 Plan of the Chapter

We proceed systematically in several steps. In Section 2 we consider portfolio level analysis, directly modeling conditional portfolio volatility using exponential smoothing and GARCH models, along with more recent “realized volatility” procedures that effectively incorporate the information in high-frequency intraday data.

In Section 3 we consider asset-level analysis, modeling asset conditional covariance matrices, again using GARCH and realized volatility techniques. The relevant cross-sectional dimension is often huge, so we devote special attention to dimensionality-reduction methods.

In Section 4 we consider links between return volatilities and macroeconomic fundamentals, with special attention to interactions across the business cycle.

We conclude in Section 5.

2. CONDITIONAL PORTFOLIO-LEVEL RISK ANALYSIS

The portfolio risk measurements that we discuss in this section require only a univariate portfolio-level model. In contrast, active portfolio risk management, including VaR minimization and sensitivity analysis, as well as system-wide risk measurements, all require a multivariate model, as we discuss subsequently in Section 3.

In practice, portfolio-level analysis is often done via historical simulation, as detailed above. We argue, however, 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.

The construction of historical returns on the portfolio in place is a necessary precursor to any portfolio-level risk analysis. In principle it is easy to construct a time series of historical portfolio returns using current portfolio holdings, $W_T = (w_{1,T}, \dots, w_{N,T})'$ and historical asset returns,¹⁰ $R_t = (r_{1,t}, \dots, r_{N,t})'$:

$$r_{w,t} = \sum_{i=1}^N w_{i,T} r_{i,t} \equiv W_T' R_t, \quad t = 1, 2, \dots, T. \quad (3)$$

⁹ ES is increasingly used in financial institutions, but it has not been incorporated into the international regulatory framework for risk control, likely because it is harder than VaR to estimate reliably in practice.

¹⁰ The portfolio return is a linear combination of asset returns when simple rates of returns are used. When log returns are used the portfolio return is only approximately linear in asset returns.

In practice, however, historical 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 “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 the substitute assets.

We focus our discussion on *VaR*.¹¹ We begin with a discussion of the direct computation of portfolio *VaR* via exponential smoothing, followed by GARCH modeling, and more recent realized volatility-based procedures. Notwithstanding a number of well-known drawbacks, see, e.g. Stulz (2008), *VaR* remains by far the most prominent and commonly used quantitative risk measure. The main techniques that we discuss are, however, easily adapted to allow for the calculation of other portfolio-level risk measures, and we will briefly discuss how to do so as well.

2.1 Modeling Time-Varying Volatilities Using Daily Data and GARCH

The lack of conditionality in the HS-*VaR* and related HS approaches discussed above is a serious concern. Several procedures are available for remedying this deficiency. Chief among these are RiskMetrics (RM) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, both of which are easy to implement on a portfolio basis. We discuss each approach in turn.

2.1.1 Exponential Smoothing and RiskMetrics

Whereas the HS-*VaR* methodology makes no explicit assumptions about the distributional model generating the returns, the RM filter/model implicitly assumes a very tight parametric specification by incorporating conditionality via univariate portfolio-level exponential smoothing of squared portfolio returns. This directly parallels the exponential smoothing of individual return squares and cross products that underlies the basic RM approach at the individual asset level.¹²

¹¹ Although the Basel Accord calls for banks to report 1% *VaRs*, for various reasons banks tend to report more conservative *VaRs*; see, e.g. the results in Berkowitz and O’Brien (2002), Perignon, Deng, and Wang (2008), and Perignon and Smith (2010a, 2010b). Rather than simply scaling up a 1% *VaR* based on some “arbitrary” multiplication factor, the procedures that we discuss below may readily be used to achieve any desired, more conservative, *VaR*.

¹² Empirically more realistic long-memory hyperbolic decay structures, similar to the long-memory type GARCH models briefly discussed below, have also been explored by RM more recently; see, e.g. Zumbach (2006). However, following standard practice we will continue to refer to exponential smoothing simply as the RM approach.

Again, taking the portfolio-level pseudo returns from (3) as the data series of interest we can define the portfolio-level RM variance as

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{w,t-1}^2, \quad (4)$$

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 variance, say $\hat{\sigma}^2$. Note that repeated substitution in (4) yields an expression for the current smoothed value as an exponentially weighted moving average of past squared returns:

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

where:

$$\varphi_j = (1 - \lambda) \lambda^j.$$

Hence the name “exponential smoothing”.

In the RM framework, VaR is then simply obtained as

$$\text{RM-} VaR_{T+1|T}^p \equiv \sigma_{T+1} \Phi_p^{-1}, \quad (5)$$

where Φ_p^{-1} is the p th quantile of the standard normal distribution. Although other distributions and quantiles could be used in place of the normal—and sometimes are—the assumption of conditional normality remains dominant. Similarly, the smoothing parameter λ may in principle be calibrated to best fit the specific historical returns at hand although, following RM, it is typically fixed at a preset value of 0.94 with daily returns. Altogether, the implicit assumption of zero mean returns, a fixed smoothing parameter, and conditional normality therefore implies that no parameters and/or distributions need to be estimated.

Extending the approach to longer return horizons, the conditional variance for the k -day return in RM is

$$\text{Var}(r_{w,t+k} + r_{w,t+k-1} + \cdots + r_{w,t+1} | F_t) \equiv \sigma_{t:t+k|t}^2 = k \sigma_{t+1}^2. \quad (6)$$

Hence the RM model can be thought of as a random walk model in variance, insofar as the variance scales with the return horizon. More precisely, 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 MSE forecast at any horizon is simply the current smoothed value.¹³

¹³ See Nerlove and Wage (1964).

Unfortunately, however, the historical record of volatility across numerous asset classes suggests that volatilities are unlikely to follow random walks, and hence that the flat forecast function associated with exponential smoothing is inappropriate for volatility. In particular, the lack of mean-reversion in the RM variance calculations implies that the term structure of volatility is always flat, which violates both intuition and historical experience. Suppose, for example, that current volatility is high by historical standards, as was the case during the height of the financial crisis and the earlier part of the sample in Figure 1. The RM model will then simply extrapolate the high current volatility across all future horizons. By contrast, an empirically more realistic mean-reverting volatility model would correctly predict that the high volatility observed during the crisis would eventually subside.

The dangers of simply scaling the daily variance by the horizon k , as done in (6), are discussed further in Diebold, Hickman, Inoue, and Schuermann (1998a). Of course, the one-day RM volatility does adjust much more quickly to changing market conditions than the HS approach, but the flat volatility term structure is unrealistic and, when taken literally, RM does not appear to be a prudent approach to volatility modeling and measurement. Furthermore, it is only valid as a volatility filter and not as a data generating process for simulating future returns. Hence we now turn to GARCH models, which allow for much richer term structures of volatility and which can be used to simulate the return process forward in time.

2.1.2 The GARCH(1, 1) Model

To allow for time variation in both the conditional mean and variance of univariate portfolio returns, we write

$$r_{w,t} = \mu_t + \sigma_t z_t, \quad z_t \sim \text{i.i.d.}, \quad E(z_t) = 0, \quad \text{Var}(z_t) = 1. \quad (7)$$

For simplicity we will henceforth assume a zero conditional mean, $\mu_t \equiv 0$. This directly parallels the RM approach, and it is a common assumption in risk management when short (e.g. daily or weekly) return horizons are considered. It is readily justified by the fact that the magnitude of the daily volatility (conditional standard deviation) σ_t easily dominates that of μ_t for most portfolios of practical interest. This is also indirectly manifest by the fact that, in practice, accurate estimation of the mean is typically much more difficult than accurate estimation of volatility. Still, conditional mean dynamics could easily be incorporated into any of the GARCH models discussed below by considering de-measured returns $r_{w,t} - \mu_t$ in place of $r_{w,t}$.

The key object of interest is the conditional standard deviation, σ_t . If it depends non-trivially on the currently observed conditioning information, we say that $r_{w,t}$ follows a GARCH process. Numerous competing parameterizations for σ_t have been proposed in the literature for best capturing the temporal dependencies in the conditional variance of portfolio returns; see, e.g. the list of models and corresponding acronyms

in Bollerslev (2010). However, the simple symmetric GARCH(1, 1) introduced by Bollerslev (1986) remains by far the most commonly used formulation in practice. The GARCH(1, 1) model is defined by

$$\sigma_t^2 = \omega + \alpha r_{w,t-1}^2 + \beta \sigma_{t-1}^2. \quad (8)$$

Extensions to higher-order GARCH models are straightforward but usually unnecessary empirically, so we concentrate on the GARCH(1,1) throughout most of the chapter, while discussing some important generalizations in the following section.

Perhaps surprisingly, GARCH is closely related to exponential smoothing of squared returns. Repeated substitution in (8) yields

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

so 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 related to RM volatility measurement.

There are, however, crucial differences between GARCH and RM. First, the GARCH parameters, and hence ultimately the GARCH volatility, are estimated using rigorous statistical methods that facilitate probabilistic inference. By contrast, the parameters used in exponential smoothing are set in an ad hoc fashion. More specifically, the vector of GARCH parameters, $\theta = (\omega, \alpha, \beta)$, is typically estimated by maximizing the log-likelihood function,

$$\ln L(\theta; r_{w,T}, \dots, r_{w,1}) \propto - \sum_{t=1}^T [\ln(\sigma_t^2(\theta)) - \sigma_t^{-2}(\theta) r_{w,t}^2]. \quad (9)$$

This likelihood function is based on the assumption that z_t in (7) is i.i.d. $N(0, 1)$. However, the assumption of conditional normality underlying the (quasi-) likelihood function in (9) is merely a matter of convenience. If the conditional return distribution is non-normal, the resulting quasi MLE generally still produces consistent and asymptotically normal, albeit not fully efficient, parameter estimates, see, e.g. Bollerslev and Wooldridge (1992). 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 is needed.¹⁴

¹⁴This 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 (2006a).

Second, and crucially from the vantage point of financial market risk measurement, the covariance stationary GARCH(1, 1) process has dynamics that eventually produce reversion in volatility to a constant long-run value. This enables interesting and realistic forecasts and contrasts sharply with the RM exponential smoothing approach in which, as discussed earlier, the term structure of volatility is forced to be flat. To see the mean reversion that GARCH enables, rewrite the GARCH(1, 1) model in (8) as

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

where $\sigma^2 \equiv \omega/(1 - \alpha - \beta)$ denotes the long-run, or unconditional daily variance, or equivalently as

$$(\sigma_t^2 - \sigma^2) = \alpha(r_{w,t-1}^2 - \sigma^2) + \beta(\sigma_{t-1}^2 - \sigma^2). \quad (11)$$

Hence the forecasted deviation of the conditional variance from the long-run variance is a weighted average of the deviation of the current conditional variance from the long-run variance, and the deviation of the squared return from the long-run variance. RM's exponential smoothing creates a parallel weighted average, with the key difference that exponential smoothing imposes $\alpha + \beta = 1$, whereas covariance stationary GARCH(1,1) imposes $\alpha + \beta < 1$. Finally, we can rearrange (11) to write

$$(\sigma_t^2 - \sigma^2) = (\alpha + \beta)(\sigma_{t-1}^2 - \sigma^2) + \alpha \sigma_{t-1}^2 (z_{t-1}^2 - 1), \quad (12)$$

where the last term on the right has zero mean. Hence, the mean reversion of the conditional variance (or lack thereof) is governed by $(\alpha + \beta)$. So long as $(\alpha + \beta) < 1$, which must hold for the covariance stationary GARCH(1, 1) processes of empirical relevance, the conditional variance is mean reverting, with the speed of mean reversion governed by $(\alpha + \beta)$.

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 conditional variance forecast. By repeated substitution in (12), we obtain

$$\sigma_{t+k|t}^2 = \sigma^2 + (\alpha + \beta)^{k-1}(\sigma_{t+1}^2 - \sigma^2). \quad (13)$$

Under our maintained assumption that returns have conditional mean zero, the variance of the k -day cumulative return is simply the sum of the corresponding 1- through k -day ahead variance forecasts. Simplifying this sum, it may be informatively expressed as

$$\sigma_{t:t+k|t}^2 = k\sigma^2 + (\sigma_{t+1}^2 - \sigma^2) \left(\frac{1 - (\alpha + \beta)^k}{1 - \alpha - \beta} \right). \quad (14)$$

Hence, in contrast to the flat volatility term structure associated with the RM forecast in (6), the GARCH volatility term structure is upward or downward sloping depending on the level of current conditional variance compared to long-run variance.

To summarize the discussion thus far, we have seen that GARCH is attractive relative to RM because it moves from ad hoc exponential smoothing to rigorous yet simple likelihood-based probabilistic modeling, and because it allows for the mean reversion routinely observed in actual financial market volatilities. In addition, and crucially, the basic GARCH(1, 1) model is readily extended in a variety of important and empirically useful directions, to which we now turn.

2.1.3 Extensions of the Basic GARCH Model

One important generalization of the basic GARCH(1, 1) model involves the enrichment of the dynamics via higher-order specifications to obtain GARCH(p , q) models with $p \geq 1$, $q \geq 1$. Indeed, Engle and Lee (1999) show that the GARCH(2, 2) is of particular interest because, under certain parameter restrictions, it implies that conditional variance dynamics may be decomposed into long-run and short-run components,

$$(\sigma_t^2 - q_t) = \alpha(r_{w,t-1}^2 - q_{t-1}) + \beta(\sigma_{t-1}^2 - q_{t-1}), \quad (15)$$

where the long-run component, q_t , is a separate autoregressive process,

$$q_t = \omega + \rho q_{t-1} + \phi(r_{w,t-1}^2 - \sigma_{t-1}^2). \quad (16)$$

Of course, this “component GARCH” model is a very special version of a component model, and one may argue that it is not a component model at all, but rather just a restricted GARCH(2, 2).

More general component modeling is easily undertaken, however, allowing for additive superposition of independent autoregressive-type components, as in Gallant, Hsu, and Tauchen (1999), Alizadeh, Brandt, and Diebold (2002), Christoffersen, Jacobs, Ornathanalai, and Wang (2008), all of whom find evidence of component structure in volatility. Under appropriate conditions, such structures may be shown to approximate very strong dependence, i.e. “long-memory”, in which shocks to the conditional variance decay at a slow hyperbolic rate, see, e.g. Granger (1980), Cox (1981), Andersen and Bollerslev (1997), and Barndorff-Nielsen and Shephard (2001).

Exact long-memory behavior can also easily be incorporated into the GARCH modeling framework to more closely mimic the dependencies observed with most financial assets and/or portfolios; see, e.g. Bollerslev and Mikkelsen (1999).¹⁵

¹⁵ The basic RiskMetrics approach has also recently been extended to allow the smoothing parameters ψ_j used in filtering the returns to exhibit a fixed pre-specified hyperbolic slow long-memory type decay; see Zumbach (2006). However, the same general set of drawbacks pertaining to the basic RM filter remain.

As discussed further below, properly incorporating these types of long-memory dependencies generally also results in more accurate volatility forecasts over long horizons.

To take a second example of the extensibility of GARCH models, note that all of the models considered so far, including the RM filter, imply symmetric response to positive and 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 standard GARCH model is readily extended to capture this effect by simply including a separate term for the past negative return shocks, as in the so-called threshold-GARCH model proposed by Glosten, Jagannathan, and Runkle (1993),

$$\sigma_t^2 = \omega + \alpha r_{w,t-1}^2 + \gamma r_{w,t-1}^2 I(r_{w,t-1} < 0) + \beta \sigma_{t-1}^2, \quad (17)$$

where $I(\cdot)$ denotes the indicator function. For well-diversified equity portfolios γ is typically estimated to be positive and highly statistically significant. In fact, the asymmetry in the volatility appears to have increased over time and the estimate for the conventional ARCH coefficient, α in (17) is often insignificant with recent data, so that the dynamics appear to be driven exclusively by the negative shocks.

Other popular asymmetric GARCH models include the EGARCH model of Nelson (1991), in which the logarithmic conditional variance is a function of the “raw” and absolute standardized return shocks, and the NGARCH model of Engle and Ng (1993). In the NGARCH(1,1) model,

$$\sigma_t^2 = \omega + \alpha (r_{w,t-1} - \gamma \sigma_{t-1})^2 + \beta \sigma_{t-1}^2, \quad (18)$$

where asymmetric response in the conventional direction occurs for $\gamma > 0$.

In parallel to the RM- VaR defined in (5), a GARCH-based one-day VaR may correspondingly be calculated by simply multiplying the one-day volatility forecast from any GARCH model by the requisite quantile in the standard normal distribution,

$$\text{GARCH-}VaR_{T+1|T}^p \equiv \sigma_{T+1} \Phi_p^{-1}. \quad (19)$$

This GARCH- VaR , of course, implicitly assumes that the returns are conditionally normally distributed. This is a much better approximation than assuming the returns are unconditionally normally distributed, and it is entirely consistent with the fat tails routinely observed in unconditional return distributions.

As noted earlier, however, standardized innovations z_t from GARCH models sometimes have fatter tails than the normal distribution, indicating that conditional normality is not acceptable. The GARCH-based approach explicitly allows us to remedy this problem, by using other conditional distributions and corresponding quantiles in place

of Φ_p^{-1} , and we will discuss various ways for doing so in Section 2.3 below to further enhance the performance of the simple GARCH-*VaR* approach. Note also that in contrast to the RM-based *VaRs*, which simply scale with the square-root of the return horizon, the multi-day GARCH-based *VaRs* explicitly incorporate mean reversion in the forecasts. They cannot be obtained simply by scaling the *VaRs* in (19). Again, we will discuss this in more detail in Section 2.3.

For now, to illustrate the conditionality afforded by the GARCH-*VaR*, and to contrast it with HS-*VaR*, we plot in Figure 3 the *VaRs* from an NGARCH model and RiskMetrics (RM). The figure clearly shows that allowing for GARCH (or RM) conditionally makes the *VaRs* move up and, equally importantly, come down much faster than the HS-*VaRs*. Moreover, contrasting the two curves, it is evident that allowing for asymmetry in a rising market desirably allows GARCH-*VaR* to drop more quickly than RM-*VaR*. Conversely, the GARCH-*VaR* rises more quickly than RM-*VaR* (and *VaRs* based on symmetric GARCH models) in falling markets. Several studies by Engle (2001, 2004, 2009b, 2011) have shown that allowing for asymmetries in the conditional variance can materially affect GARCH-based *VaRs*.

The procedures discussed in this section were originally developed for daily or coarser frequency returns. However, high-frequency intraday price data are now readily

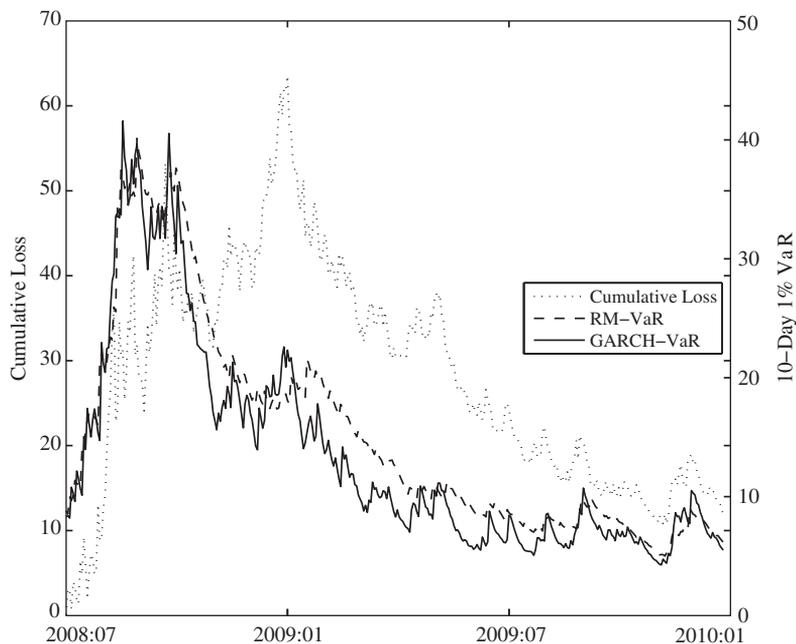


Figure 3 Cumulative S&P500 loss (dots, left scale) and 1% 10-day RM-*VaR* and GARCH-*VaR* (solid and dashed, right scale), July 1, 2008–December 31, 2009.

available for a host of different assets and markets. We next review recent research on so-called realized volatilities constructed from such high-frequency data, and show how to use them to provide even more accurate assessment and modeling of daily market risks.

2.2 Intraday Data and Realized Volatility

Higher frequency data add little to the estimation of expected returns. At the same time, however, the theoretical results in Merton (1980) and Nelson (1992) suggest that higher frequency data should be very useful in the construction of more accurate volatility models, and in turn expected risks. In practice, however, the statistical modeling of high-frequency data is notoriously difficult, and the daily GARCH and related volatility forecasting procedures discussed in the previous section have been shown to work poorly when applied directly to high-frequency intraday returns; see, e.g. Andersen and Bollerslev (1997) and Andersen, Bollerslev, and Lange (1999). Fortunately, extensive research efforts over the past decade have shown how the rich information inherent in the now readily available high-frequency data may be effectively harnessed through the use of so-called realized volatility measures.

To formally define the realized volatility concepts, imagine that the instantaneous returns, or logarithmic price increments, evolve continuously through time according to the stochastic volatility diffusion

$$dp(t) = \mu(t) dt + \sigma(t) dW(t), \quad (20)$$

where $\mu(t)$ and $\sigma(t)$ denote the instantaneous drift and volatility, respectively, and $W(t)$ is a standard Brownian motion.¹⁶ This directly parallels the general discrete-time return representation in Eqn (7), with $r_{w,t} \equiv p(t) - p(t-1)$ and the unit time interval normalized to a day. Just as the conditional mean in (7) can be safely set to zero, so too can the drift term in Eqn (20). Hence, in what follows, we set $\mu(t) = 0$.

Following Andersen and Bollerslev (1998b), Andersen, Bollerslev, Diebold, and Labys (2001b), and Barndorff-Nielsen and Shephard (2002), the realized variation (RV) on day t based on returns at the Δ intra-day frequency is then formally defined by

$$RV_t(\Delta) \equiv \sum_{j=1}^{N(\Delta)} \left(p_{t-1+j\Delta} - p_{t-1+(j-1)\Delta} \right)^2, \quad (21)$$

where $p_{t-1+j\Delta} \equiv p(t-1+j\Delta)$ denotes the intraday log-price at the end of the j th interval on day t , and $N(\Delta) \equiv 1/\Delta$. For example, $N(\Delta) = 288$ for 5-min returns in a

¹⁶ The notion of a continuously evolving around-the-clock price process is, of course, fictitious. Most financial markets are only open for part of the day, and prices are not continuously updated and sometimes jump. The specific procedures discussed below have all been adapted to accommodate these features and other types of market microstructure frictions, or “noise”, in the actually observed high-frequency prices.

24-h market, corresponding to $\Delta = 5/(24 \cdot 60) \approx 0.00347$, while 5-min returns in a market that is open for six-and-half hours per day, like the US equity markets, would correspond to $N(\Delta) = 78$ and $\Delta = 5/(6.5 \cdot 60) \approx 0.01282$. The expression in (21) looks exactly like a sample variance for the high-frequency returns, except that we do not divide the sum by the number of observations, $N(\Delta)$, and the returns are not centered around the sample mean.

Assume for the time being that the prices defined by the process in (20) are continuously observable. In this case, letting Δ go to zero, corresponding to progressively finer sampled returns, the RV estimator approaches the integrated variance of the underlying continuous-time stochastic volatility process on day t , formally defined by:¹⁷

$$IV_t = \int_{t-1}^t \sigma^2(\tau) d\tau. \quad (22)$$

Hence, in contrast to the RM- and GARCH-based volatility estimates discussed above, the *true* ex-post volatility for the day effectively becomes observable. And it does so in an entirely model-free fashion regardless of the underlying process that actually describes $\sigma(t)$.

In practice, of course, prices are not available on a continuous basis. However, with prices for many assets recorded, say, every minute, a daily RV could easily be computed from 1-min squared returns. Still, returns at the 1-min frequency are likely affected by various market microstructure frictions, or *noise*, arising from bid-ask bounces, a discrete price grid, and the like.¹⁸ Of course, even with 1-min price observations on hand, we may decide to construct the RV measures from 5-min returns, as these coarser sampled data are less susceptible to contamination from market frictions. Clearly, this involves a loss of information as the majority of the recorded prices are ignored. Expressed differently, it is feasible to construct five different sets of (overlapping) 5-min intraday return sequences from the given data, but in computing the regular 5-min based RV measure we exploit only one of these series—a theme we return to below.

The optimal choice of high-frequency grid over which to measure the returns obviously depends on the specific market conditions. The “volatility signature plot” of Andersen, Bollerslev, Diebold, and Labys (2000b) is useful for guiding this selection. It often indicates the adequacy of 5-min sampling across a variety of assets and markets, as originally advocated by Andersen and Bollerslev (1998a).¹⁹ Meanwhile, as many markets

¹⁷ More precisely, $\Delta^{-1/2}(\text{RV}_t(\Delta) - IV_t) \rightarrow N(0, 2IQ_t)$, where $IQ_t \equiv \int_0^1 \sigma^4(t-1+\tau) d\tau$ and the convergence is stable in law; for a full theoretical treatment, see, e.g. Andersen, Bollerslev, and Diebold (2010a).

¹⁸ Brownlees and Gallo (2006) contain a useful discussion of the relevant effects and some of the practical issues involved in high-frequency data cleaning.

¹⁹ See also Hansen and Lunde (2006) and the references therein.

have become increasingly liquid it would seem reasonable to resort to even finer sampling intervals with more recent data although, as noted below, the gains from doing so in terms of the accuracy of realized volatility based forecast appear to be fairly minor.

One way to exploit all the high-frequency returns, even if the RV measure is based on returns sampled at a lower frequency, is to compute an alternative RV estimator using a different offset relative to the first return of the trading day, and then combine them. For example, if 1-min returns are given, one may construct a new RV estimator using an equal-weighted average of the five alternative regular 5-min RV estimators available each day. We will denote this estimator $AvgRV$ below. The upshot is that the $AvgRV$ estimator based on 5-min returns is much more robust to microstructure noise than the single RV based on 1-min returns.

In markets that are not open 24h per day, the change from the closing price on day $t - 1$ to the opening price on day t should also be accounted for. This can be done by simply scaling up the trading day RV by the proportion corresponding to the missing over-night variation, or any of the other more complicated methods advocated in Hansen and Lunde (2005). As is the case for the daily GARCH models discussed above, corrections may also be made for the fact that days following weekends and holidays tend to have proportionally higher than average volatility.

Several other realized volatility estimators have been developed to guard against the influences of market microstructure frictions. In contrast to the simple $RV_t(\Delta)$ estimator, which formally deteriorates as the length of the sampling interval Δ approaches zero if the prices are observed with error, these other estimators are typically designed to be consistent for IV_t as $\Delta \rightarrow 0$, even in the presence of market microstructure noise. Especially prominent are the realized kernel estimator of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), the pre-averaging estimator of Jacod, Li, Mykland, Podolskij, and Vetter (2009), and the two-scale estimator of Aït-Sahalia, Mykland, and Zhang (2011). These alternative estimators are generally more complicated to implement than the $AvgRV$ estimator, requiring the choice of additional tuning parameters, smoothing kernels, and appropriate block sizes. Importantly, the results in Andersen, Bollerslev, and Meddahi (2011a) show that, when used for volatility forecasting, the simple-to-implement $AvgRV$ estimator performs on par with, and often better than, these more complex RV estimators.²⁰

To illustrate, we plot in Figure 4 the square root of daily $AvgRV$ s (in annualized percentage terms) as well as daily S&P 500 returns for January 1, 1990 through December 31, 2010. Following the discussion above, we construct $AvgRV$ from a 1-min grid of futures prices and the average of the corresponding five 5-min RVs.²¹

²⁰ Note, however, that while the $AvgRV$ estimator provides a very effective way of incorporating ultra high-frequency data into the estimation by averaging all of the possible squared price increments over the fixed non-trivial time interval $\Delta > 0$, the $AvgRV$ estimator is formally not consistent for IV as $\Delta \rightarrow 0$.

²¹ We have 1-min prices from 8:31 am to 3:15 pm each day. We do not adjust for the overnight return.

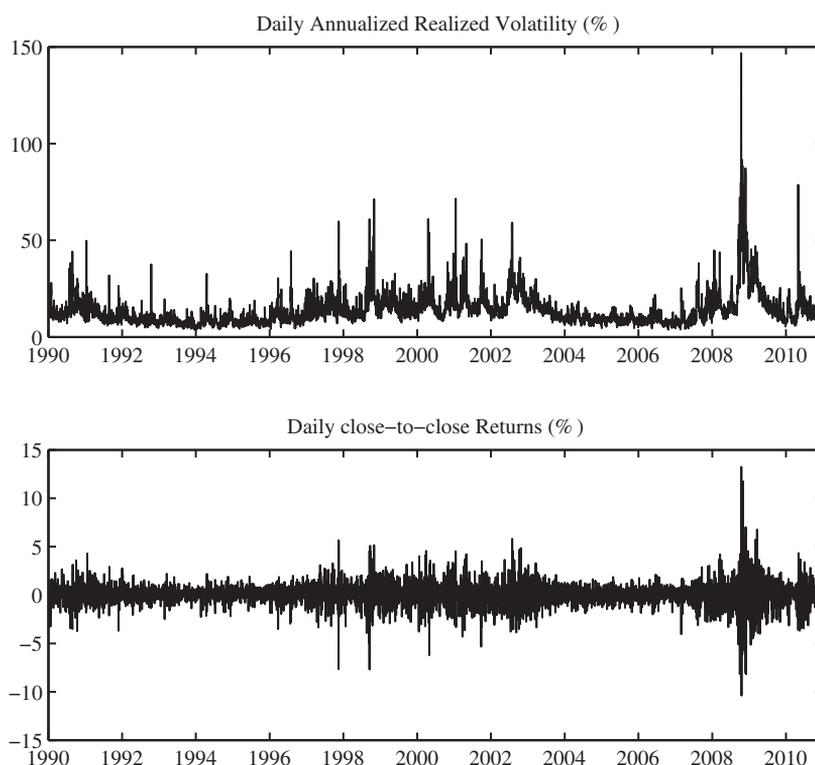


Figure 4 S&P500 daily returns and volatilities (Percent). The top panel shows daily S&P500 returns, and the bottom panel shows daily S&P500 realized volatility. We compute realized volatility as the square root of $AvgRV$, where $AvgRV$ is the average of five daily RVs each computed from 5-min squared returns on a 1-min grid of S&P500 futures prices.

Looking at the figure, the assumption of constant volatility is clearly untenable from a risk management perspective. The dramatic rise in the volatility in the fall of 2008 is also immediately evident, with the daily realized volatility reaching an unprecedented high of 146.2 on October 10, 2008, which is also the day with the largest ever recorded NYSE trading volume.

Time series plots such as that of Figure 4, of course, begin to inform us about aspects of the dynamics of realized volatility. We will shortly explore those dynamics in greater detail. But first we briefly highlight an important empirical aspect of the *distribution* of realized volatility, which has been documented in many contexts: realized volatility is highly right-skewed, whereas the natural logarithm of realized volatility is much closer to Gaussian. In Figure 5 we report two QQ (Quantile–Quantile) plots of different volatility transforms against the normal distribution. The top panel shows the QQ plot for daily $AvgRV$ in standard deviation form, while the bottom panel shows the QQ plot for

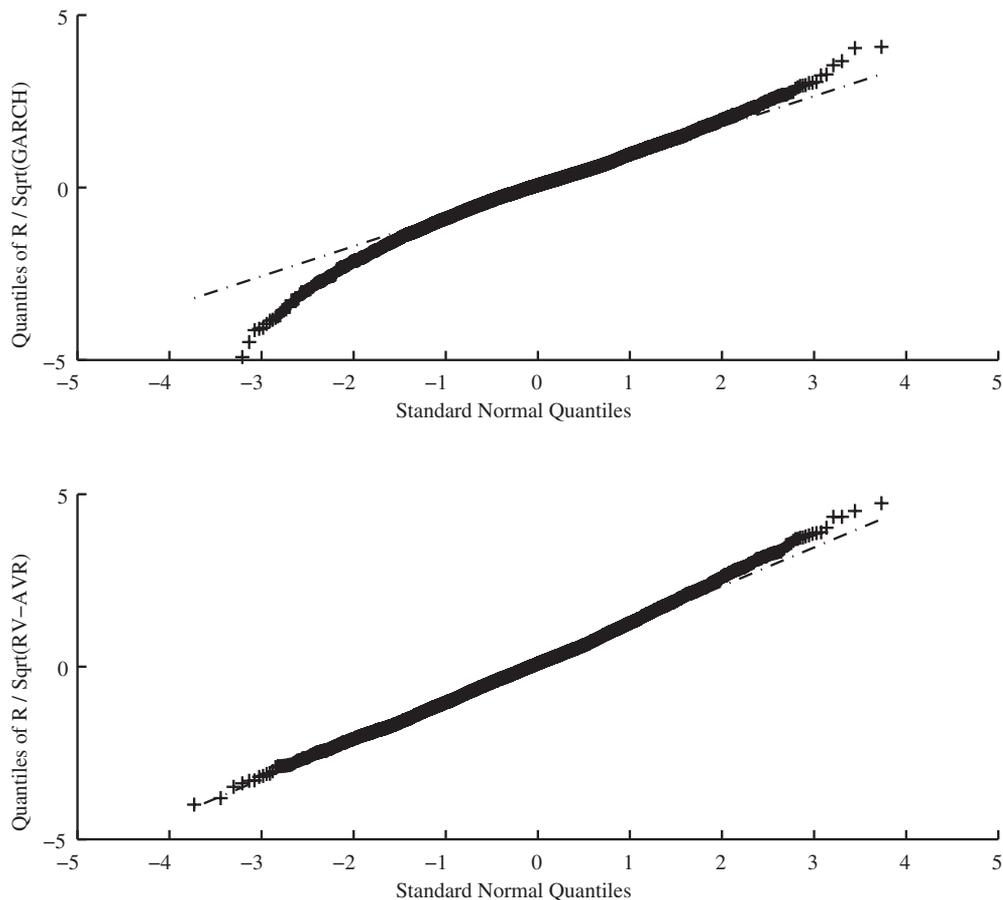


Figure 5 S&P500: QQ plots for realized volatility and log realized volatility. The top panel plots the quantiles of daily realized volatility against the corresponding normal quantiles. The bottom panel plots the quantiles of the natural logarithm of daily realized volatility against the corresponding normal quantiles. We compute realized volatility as the square root of $AvgRV$, where $AvgRV$ is the average of five daily RVs each computed from 5-min squared returns on a 1-min grid of S&P500 futures prices.

daily $AvgRV$ in logarithmic form. The right tail in the top panel is obviously much fatter than for a normal distribution, whereas the right tail in the bottom panel conforms more closely to normality. This approximate log-normality of realized volatility is often usefully exploited, even if it provides only a rough approximation, based on empirical observation rather than theoretical derivation.²²

²² Indeed, as noted by Forsberg and Bollerslev (2002), among others, RV cannot formally be log-normally distributed across all return horizons, because the log-normal distribution is not closed under temporal aggregation.

2.2.1 Dynamic Modeling of Realized Volatility

Although daily RV is ultimately only an estimate of the underlying true integrated variance, it is potentially highly accurate and thus presents an intriguing opportunity. By treating the daily RVs, or any of the other high-frequency-based RV measures, as direct ex post observations of the true daily integrated variances, the RV approach permits the construction of ex ante volatility forecasts using standard ARMA time series tools. Moreover, recognizing the fact that the measures are not perfect, certain kinds of measurement errors can easily be incorporated into this framework. The upshot is that if the frequency of interest is daily, then using sufficiently high-quality intraday price data enables the risk manager to treat volatility as effectively observed. This is fundamentally different from the RM filter and GARCH style models discussed above, in which the daily variances are inferred from past daily returns *conditional* on the specific structure of the filter or model.

To further help motivate such an approach, we plot in Figure 6 the autocorrelation function (ACF) of daily *AvgRV* and daily returns. The horizontal lines in each plot show

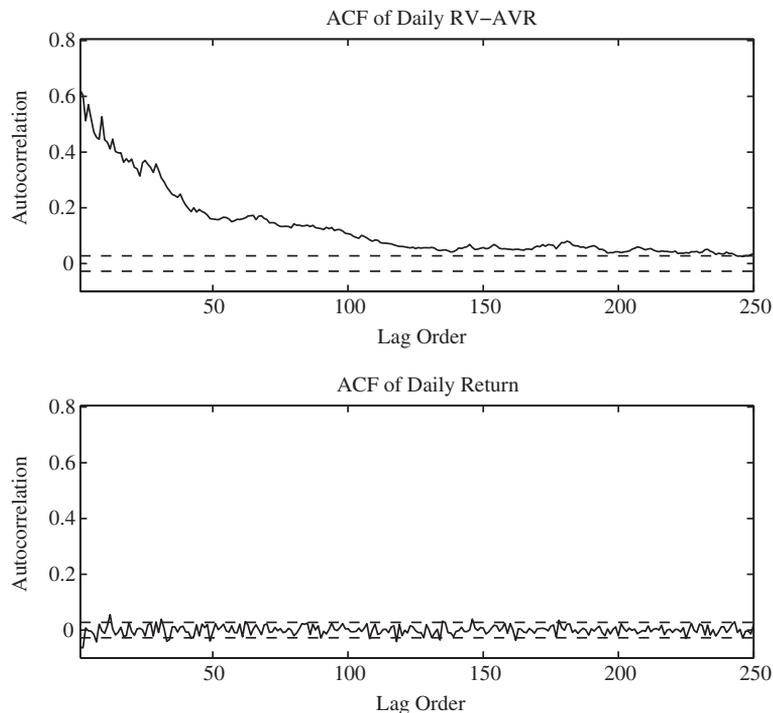


Figure 6 S&P500: sample autocorrelations of daily realized variance and daily return. The top panel shows realized variance autocorrelations, and the bottom panel shows return autocorrelations, for displacements from 1 through 250 days. Horizontal lines denote 95% Bartlett bands. Realized variance is *AvgRV*, the average of five daily RVs each computed from 5-min squared returns on a 1-min grid of S&P500 futures prices.

the Bartlett two-standard-deviation bands around zero. The ACFs are strikingly different; the realized variance ACF is always positive, highly statistically significant, and very slowly decaying, whereas the daily return ACF is insignificantly different from zero. The exceptionally slow decay of the realized variance ACF suggests long-memory dynamics, in turn implying that equity market volatility is highly forecastable. This long-memory property of RV is found across numerous asset classes; see, for example, Andersen et al. (2001b) for evidence on foreign exchange rates and Andersen, Bollerslev, Diebold, and Ebens (2001a) for comparable results pertaining to individual equities and equity-index returns.

Simple AR type models provide a natural starting point for capturing these dependencies. Let RV_t denote any of the high-frequency-based realized volatility measures introduced above. As an example, one could specify a simple first-order autoregressive model for the daily volatility series,

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + v_t. \quad (23)$$

This, and any higher-order AR models for RV_t can easily be estimated by a standard OLS regression package.

One could go farther and endow *integrated* variance with AR(1) dynamics, and recognize that RV_t contains some measurement error, since in real empirical work the underlying sampling cannot pass all the way to continuous time. Then RV_t would equal an AR(1) process plus a measurement error, which yields an ARMA(1,1) model if the two are independent:

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \alpha_1 v_{t-1} + v_t.$$

Estimation of this model formally requires use of nonlinear optimization techniques, but it is still very easy to do using standard statistical packages.

Although the simple short-memory AR(1) model above may be adequate for short-horizon risk forecasts, the autocorrelation function for $AvgRV$ shown in Figure 6 clearly suggests that when looking at longer, say monthly, forecast horizons, more accurate forecasts may be obtained by using richer dynamic models that better capture the long-range dependence associated with slowly decaying autocorrelations. Unfortunately, however, when $|\beta_1| < 1$ the AR(1) process has short memory, in the sense that its autocorrelations decay exponentially quickly. On the other hand, when $\beta_1 = 1$ the process becomes a random walk $(1 - L)RV_t = \beta_0 + v_t$, and has such strong memory that covariance stationarity and mean reversion are both lost. A useful middle ground may be obtained by allowing for fractional integration,²³

$$(1 - L)^d RV_t = \beta_0 + v_t. \quad (24)$$

²³The fractional differencing operator $(1 - L)^d$ is formally defined by its binomial expansion; see, e.g. Baillie, Bollerslev, and Mikkelsen (1996) and the discussion therein pertaining to the so-called fractional integrated GARCH (FIGARCH) model.

This long-memory model is mean reverting if $0 < d < 1$ and covariance stationary if $0 < d < 1/2$. Fractional integration contrasts to the extremely strong integer integration associated with the random walk ($d = 1$) or the covariance-stationary AR(1) case ($d = 0$). Crucially, it allows for long-memory dynamics in the sense that autocorrelations decay only hyperbolically, akin to the pattern seen in Figure 6.

Long-memory models can, however, be somewhat cumbersome to estimate and implement. Instead, a simpler approach may be pursued by directly exploiting longer run realized volatility regressors. Specifically, letting $RV_{t-4:t}$ and $RV_{t-20:t}$ denote the weekly and monthly realized volatilities, respectively, obtained by summing the corresponding daily volatilities. Many researchers, including Andersen, Bollerslev, and Diebold (2007a), have found that the so-called heterogenous autoregressive, or HAR-RV, model, originally introduced by Corsi (2009),

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-5:t-1} + \beta_3 RV_{t-21:t-1} + \nu_t, \quad (25)$$

provides a very good fit for most volatility series. As shown in Corsi (2009), the HAR model may be viewed as an approximate long-memory model. In contrast to the exact long-memory model above, however, the HAR model can easily be estimated by OLS. Even closer approximations to exact long-memory dependence can be obtained by including coarser, say quarterly, lagged realized volatilities on the right-hand side of the equation. A leverage effect, along the lines of the GJR-GARCH model discussed above, can also easily be incorporated into the HAR-RV modeling framework by including on the right-hand side additional volatility terms interacted with dummies indicating the sign of r_{t-1} , as in Corsi and Reno (2010).

The HAR regressions can, of course, also be written in logarithmic form

$$\log RV_t = \beta_0 + \beta_1 \log RV_{t-1} + \beta_2 \log RV_{t-5:t-1} + \beta_3 \log RV_{t-21:t-1} + \nu_t. \quad (26)$$

The log specification conveniently induces approximate normality, as demonstrated in Figure 5 above. It also ensures positivity of volatility fits and forecasts, by exponentiating to “undo” the logarithm.²⁴

Armed with a forecast for tomorrow’s volatility from any one of the HAR-RV or other time series models discussed above, say $\widehat{RV}_{T+1|T}$, a one-day VaR is easily computed as

$$RV\text{-VaR}_{T+1|T}^p = \widehat{RV}_{T+1|T} \Phi_p^{-1}, \quad (27)$$

²⁴ Note however that forecasts of RV_{t+1} obtained by exponentiating forecasts of $\log RV_{t+1}$ are generally biased, due to the nonlinearity of the $\exp(\cdot)$ transformation. Although we will not pursue it here, one could perform a bias correction, which would depend on the possibly time-varying variance of ν_t . A similar problem applies to the EGARCH model briefly discussed above.

where Φ_p^{-1} refers to the relevant quantile from the standard normal. Andersen, Bollerslev, Diebold, and Labys (2003a) use this observation to construct RV-based VaR s with properties superior to GARCH- VaR . We will discuss this approach in more detail in Section 2.3.2.

To illustrate, we show in Figure 7 the GARCH- VaR from Figure 3 together with the HAR- VaR based on (27) constructed using the simple linear HAR-RV specification in (25). The figure shows that HAR- VaR reaches its peak before GARCH- VaR . Equally important, the HAR- VaR drops back to a more normal level sooner than the GARCH- VaR after the trough in the market on March 2009. Intuitively, by using the more accurate RV measure of current volatility, the model is able to more quickly adjust to the changing market conditions and overall level of market risk. Of course, the commonly employed RM- VaR in Figure 3 is even slower to adjust than the GARCH- VaR , and the HS- VaR in Figure 1 adjusts so slowly that it remains at its maximum sample value at the end of 2009.

As discussed above, VaR and other risk measures are often computed for a two-week horizon. The risk manager is therefore interested in a 10-day volatility forecast.

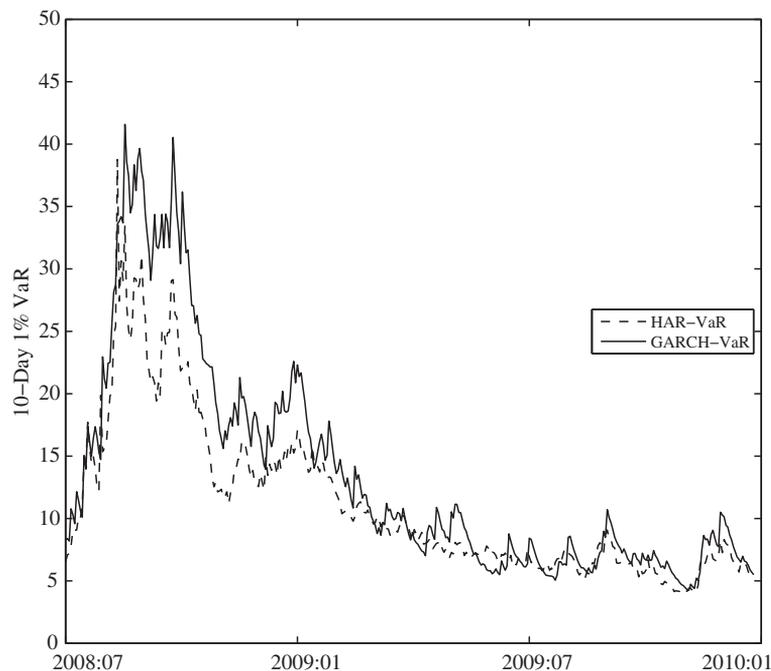


Figure 7 10-day 1% HAR- VaR and GARCH- VaR , July 1, 2008–December 31, 2009. The dashed line shows 10-day 1% HAR- VaR based on the HAR forecasting model for 10-day realized volatility. The solid line shows 10-day 1% GARCH- VaR . When computing VaR the 10-day returns divided by the expected volatility are assumed to be normally distributed.

Another advantage of the RV-based approach, and the HAR-RV model in particular, is that it can easily be adapted to deliver the required multi-period variance forecasts. Specifically, consider the modified HAR-RV regression,

$$RV_{t:t+9} = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-5:t-1} + \beta_3 RV_{t-21:t-1} + v_{t:t+9}. \quad (28)$$

An RV-based VaR can now easily be computed via

$$RV\text{-VaR}_{T+10|T}^p = \widehat{RV}_{T+1:T+10|T} \Phi_p^{-1},$$

where

$$\widehat{RV}_{T+1:T+10|T} = \hat{\beta}_0 + \hat{\beta}_1 RV_T + \hat{\beta}_2 RV_{T-4:T} + \hat{\beta}_3 RV_{T-20:T},$$

denotes the 10-day forecast obtained directly from the modified HAR-RV model in (28). Hence, in contrast to GARCH models, there is no need to resort to the use of complicated recursive expressions along the lines of the formula for $\sigma_{t:t+k|t}^2$ for the GARCH(1,1) model in (14). The modified HAR-RV model in (28) builds the appropriate mean reversion directly into the requisite variance forecasts.²⁵

2.2.2 Realized Volatilities and Jumps

The continuous-time process in (20) formally rules out discontinuities in the underlying price process. However, financial prices often exhibit “large” movements over short time intervals, or “jumps”. A number of these jumps are naturally associated with readily identifiable macroeconomic news announcements, see, e.g., Andersen, Bollerslev, Diebold, and Vega (2003b) and Andersen, Bollerslev, Diebold, and Vega (2007b), but many others appear idiosyncratic or asset specific in nature. Such large price moves are inherently more difficult to guard against, and the measurement and management of jump risk requires the use of different statistical distributions and risk management procedures from the ones needed to measure and manage the Gaussian diffusive price risks implied by the price process in (20).

In particular, taking into account the possibility of jumps in the underlying price process, the realized variation measures discussed above no longer converge to the integrated variance. Instead, the total ex-post variation is given by

$$QV_t = IV_t + JV_t, \quad (29)$$

where IV_t as before, in (22), accounts for the variation coming from the continuous, or smooth, price increments over the day, and

$$JV_t = \sum_{j=1}^{\mathcal{J}_t} J_{t,j}^2, \quad (30)$$

²⁵ Note however that a new HAR-RV model must be estimated for each forecast horizon of interest.

measures the variation due to the \mathcal{J}_t jumps that occurred on day t ; i.e., $J_{t,j}$, $j = 1, 2, \dots, \mathcal{J}_t$. This does not invalidate *AvgRV*, or any of the other RV estimators discussed above, as an ex post measure for the total daily quadratic variation, or QV_t . It does, however, suggest the use of more refined procedures for separately estimating QV_t and IV_t , and in turn JV_t .

Several alternative volatility estimators that are (asymptotically) immune to the impact of jumps have been proposed in the literature. The first was the bipower variation estimator of Barndorff-Nielsen and Shephard (2004b),

$$BPV_t(\Delta) = \frac{\pi}{2} \frac{N(\Delta)}{N(\Delta) - 1} \sum_{j=1}^{N(\Delta)-1} |\Delta p_{t-1+j\Delta}| |\Delta p_{t-1+(j+1)\Delta}|, \quad (31)$$

where $\Delta p_{t-1+j\Delta} \equiv p_{t-1+j\Delta} - p_{t-1+(j-1)\Delta}$. The idea behind the bipower variation estimator is intuitively simple. When Δ goes to zero the probability of jumps arriving both in time interval $j\Delta$ and $(j+1)\Delta$ goes to zero along with the absolute value of the non-jump returns. Jumps in the product $|\Delta p_{t-1+j\Delta}| |\Delta p_{t-1+(j+1)\Delta}|$ will therefore vanish asymptotically. Consequently, $BPV_t(\Delta)$ will converge to the integrated variance IV_t , as opposed to QV_t , for Δ approaching zero, even in the presence of jumps.²⁶ In contrast, the key terms in the realized variance estimator, namely the intraday squared returns $(\Delta p_{t-1+j\Delta})^2$, will include the price jumps as well as the “smooth” continuous price variation. The $RV_t(\Delta)$ estimator therefore always converges to QV_t for Δ approaching zero.

The $BPV_t(\Delta)$ estimator is subject to the same type of microstructure frictions that plague the $RV_t(\Delta)$ estimator at ultra-high sampling frequencies. Thus, even if a 1-min grid of prices is available, it might still be desirable to use coarser, say 5-min, returns in the calculation of $BPV_t(\Delta)$ to guard against market microstructure noise. A simple average of the five different $BPV_t(\Delta)$'s could then be used to compute an improved *AvgBPV* estimator.

Although the $BPV_t(\Delta)$ estimator is formally consistent for IV_t in the idealized setting without market microstructure noise, the presence of large jumps can result in non-trivial upward biases in practice. Motivated by this observation, Andersen, Dobrev, and Schaumburg (2012) recently proposed an alternative class of jump-robust estimators, the neighborhood truncation measures. The simplest version takes the form,

$$\text{MinRV}_t(\Delta) = \frac{\pi}{\pi - 2} \left(\frac{N(\Delta)}{N(\Delta) - 1} \right) \sum_{j=1}^{N(\Delta)-1} \min \{ |\Delta p_{t-1+j\Delta}|, |\Delta p_{t-1+(j+1)\Delta}| \}^2.$$

²⁶The $\pi/2$ normalization arises from the fact that the expected value of an absolute standard normal random variable equals $(\pi/2)^{1/2}$, while the ratio involving $N(\Delta)$ provides a finite-sample adjustment for the loss of one term in the summation.

The intuition behind the MinRV estimator is similar to that for the original *BPV* estimator. When Δ goes to zero, the probability of jumps arriving in two adjacent time intervals of length Δ goes to zero, so the minimum is unaffected by jumps. The main difference is that the jump is now fully neutralized, even at a given discrete sampling frequency, in the sense that the jump size has no direct impact on the estimator. Hence the finite sample distortion of the MinRV estimator is significantly less than that of *BPV* estimator.²⁷ By this same reasoning, a related jump-robust MedRV estimator may be constructed from the properly scaled square of the median of three adjacent absolute returns cumulated across the trading day, see Andersen, Dobrev, and Schaumburg (2012) for details.

Another intuitively simple approach for estimating IV_t , first explored empirically by Mancini (2001), is to use truncation, the idea being that the largest price increments are the ones associated with jumps. Specifically, by only summing the squared return below a certain threshold,

$$TV_t(\Delta) = \sum_{j=1}^{N(\Delta)} \Delta p_{t-1+j\Delta}^2 I(\Delta p_{t-1+j\Delta} < \mathcal{T}),$$

the resulting estimator again consistently estimates only the continuous variation provided that the threshold \mathcal{T} converges to zero at an appropriate rate as Δ goes to zero. Since the continuous variation changes over time, and in turn the likely magnitude of the corresponding continuous price increments, it is also important to allow the threshold to vary over time, both within and across days. This choice of time-varying threshold can be somewhat delicate to implement in practice; see, e.g. Bollerslev and Todorov (2011b) and the discussion therein.

Regardless of which of these different IV_t estimators is used, we obtain an empirically feasible decomposition of the total daily variation into the part associated with the “small”, or continuous, price moves, and the part associated with the “large”, and generally more difficult to hedge, price moves, or jumps. Even if the risk manager is not interested in this separation per se, this decomposition can still be very useful for the construction of improved *VaRs* and other related risk measures.

In particular, it is often the case that the variation associated with jumps tends to be much more erratic and less predictable than the variation associated with the continuous price component. As such, the simple HAR-RV type forecasting models discussed above may be improved by allowing for different dynamics for the two different sources

²⁷ This is true as long as there are no adjacent jumps at the sampling frequency used. Both estimators suffer from significant upward biases if adjacent jumps are present. This has led to additional procedures that enhance the robustness properties even further; see the discussion in Andersen, Dobrev, and Schaumburg (2011b).

of daily variation. Such an approach was first pursued by Andersen, Bollerslev, and Diebold (2007a), who found that the HAR-RV-CJ model,

$$\begin{aligned} \text{RV}_t = & \beta_0 + \beta_1 IV_{t-1} + \beta_2 IV_{t-5:t-1} + \beta_3 IV_{t-21:t-1} \\ & + \alpha_1 JV_{t-1} + \alpha_2 JV_{t-5:t-1} + \alpha_3 JV_{t-21:t-1} + v_t, \end{aligned} \quad (32)$$

indeed produces even better RV forecasts than the HAR-RV model in (25), which implicitly restricts the α_i and β_i coefficients in (32) to be identical. Instead, by allowing for “ α effects” and “ β effects” in the HAR-RV-CJ model, we capture the fact that the variation associated with jumps is less persistent and predictable than the continuous variation.

Further refinements allowing for leverage effects and/or other asymmetries and nonlinearities could easily be incorporated into the same HAR-RV modeling framework by including additional explanatory variables on the right-hand side. But the simple-to-estimate HAR-RV-CJ model typically does a remarkably good job of effectively incorporating the empirically most relevant dynamic dependencies of the intraday price data into the daily and longer-run volatility forecasts of practical interest.

2.2.3 Combining GARCH and RV

So far we have presented GARCH and RV-based procedures as two distinct approaches. There are, however, good reasons to combine the two. The ability of RV to rapidly deliver precise information regarding the current level of volatility, along with the ability of GARCH to appropriately smooth noisy volatility proxies make such a combination appealing. Another advantage of combined models is the ability to integrate the RV process naturally within a complete characterization of the return distribution, thus allowing the RV dynamics to become a natural and direct determinant of the time variation in risk measures such as *VaR* and expected shortfall. The following section will elaborate on those features of the approach.

The simplest way of combining GARCH and RV is to include the RV measure as an additional explanatory variable on the right-hand side of the GARCH equation,

$$\sigma_t^2 = \omega + \alpha r_{w,t-1}^2 + \beta \sigma_{t-1}^2 + \gamma \text{RV}_{t-1}. \quad (33)$$

This is often referred to as a GARCH-X model.²⁸ Estimating this model typically results in a statistically insignificant (ARCH) coefficient, α so that the model effectively reduces to

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \gamma \text{RV}_{t-1}. \quad (34)$$

²⁸ Professor Robert F. Engle in his discussion of Andersen et al. (2003a) at the 2000 Western Finance Association meeting in Sun Valley, Idaho, was among the first to empirically explore this idea. Related analysis appears in Engle (2002b). Lu (2005) provides another early comprehensive empirical study of GARCH-X type models.

Intuitively, the high-frequency-based RV measure affords a superior estimate of the true ex-post daily variation compared to the daily (de-measured) squared returns, in turn driving out the latter as an explanatory variable for tomorrow's volatility. As such, whenever high-frequency-based RV measures are available, it is always a good idea to use the GARCH-X model instead of the conventional GARCH(1,1) model based solely on daily return observations.²⁹

The GARCH-X model defined by (7) and (33) or (34) directly provides one-day volatility forecasts. The calculation of longer-run k -day forecasts $\sigma_{t+k|t}^2$ necessitates a model for forecasting RV_{t+k} as well. This could be accomplished in an ad hoc fashion by simply augmenting the GARCH-X model with any one of the HAR-RV type models discussed in the previous sections. The so-called Realized GARCH class of models developed by Hansen, Huang, and Shek (in press) provides a more systematic approach for doing exactly that.

As an example, consider the specific Realized GARCH model defined by (7) and

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \gamma RV_{t-1}, \quad (35)$$

$$RV_t = \omega_X + \beta_X \sigma_t^2 + \tau(z_t) + v_t, \quad (36)$$

where v_t denotes a random error with the property that $E_t(v_t) = 0$, and the $\tau(z_t)$ function allows for a contemporaneous leverage effect via the return shock z_t in (7).³⁰ Substituting the equation for σ_t^2 into the equation for RV_t shows that the model implies an ARMA representation for the realized volatility, but other HAR-RV type structures could, of course, be used instead. Note also that unlike regular GARCH, the Realized GARCH model has two separate innovations. However, because RV_t is observed, estimation of the model can still be done using bivariate maximum likelihood estimation techniques that closely mirror the easily implemented procedures available for regular GARCH models.

The Multiplicative Error Model (MEM) of Engle (2002b) and Engle and Gallo (2006) constitutes another framework for combining different volatility proxies (e.g. daily absolute returns, daily high-low ranges, RVs, IVs, or option-implied volatilities) into the estimation of a coherent multivariate model for return variances.³¹ It is natural

²⁹ In a related context, Visser (2011) has recently shown how the accuracy of the coefficient estimates in conventional daily GARCH models may be improved through the use of intraday RV-based measures in the estimation.

³⁰ A closely related class of two-shock Realized GARCH models, in which the return volatility is a weighted average of the GARCH and RV volatilities, has recently been proposed by Christoffersen, Feunou, Jacobs, and Meddahi (2011b). Their affine formulation has the advantage that option valuation can be done via Fourier inversion of the conditional characteristic function. Non-affine approaches to option valuation using RV have also been pursued by Corsi, Fusari, and LaVecchia (2011) and Stentoft (2008).

³¹ This approach has also been used by Brownlees and Gallo (2011) to compare different volatility measures and their uses in risk management.

to use this same framework to extend the GARCH-X model to allow for the construction of multi-day volatility forecasts.

In particular, building on the MEM structure, Shephard and Sheppard (2010) propose an extension of the basic GARCH-X model in (33), in which the conditional mean of realized volatility, $\mu_{RV,t} \equiv E_{t-1}(RV_t)$, is defined recursively by the equation,

$$\mu_{RV,t} = \omega_{RV} + \alpha_{RV}RV_{t-1} + \beta_{RV}\mu_{RV,t-1}. \quad (37)$$

Shephard and Sheppard (2010) refer to this model as a High-frequency Based Volatility model, or “HEAVY” model. Like the Realized GARCH class of models, HEAVY models have the advantage that they adapt to new information and market conditions much more quickly than the regular daily GARCH models. In contrast to the simple GARCH(1,1) model, for which the k -period variance forecasts in (13) converge monotonically to their long-run average values, the HEAVY model defined by (33) and (37) also might show momentum effects, so that the convergence of the multi-period variance forecasts to the long-run unconditional variance is not necessarily monotonic. This point is nicely illustrated by the volatility forecasts during the recent financial crises reported in Shephard and Sheppard (2010), which show how the model sometimes predicts rising volatility even when the current volatility is exceptionally high by historical standards.

Risk managers, of course, typically care not only about the dynamics of volatility but also more generally about the dynamics of the entire conditional distribution of portfolio returns. Movement in conditional variance is a key driver of movement in the conditional distribution, but only in the unlikely case of conditional normality is the entire story. Hence we next discuss how GARCH and realized variance may be used in broader modeling of entire return distributions.

2.3 Modeling Return Distributions

We have emphasized—and continue to emphasize—the potentially seriously misleading nature of unconditional risk analyses. Here we stress the similarly potentially seriously misleading nature of *Gaussian* risk analyses. There are four cases to consider, corresponding to the reliance on unconditional/conditional information and the use of Gaussian/non-Gaussian distributions.

Risk measurement in an unconditional Gaussian framework would be doubly defective, first because of the deficiencies of the unconditional perspective, and second because financial returns are simply not unconditionally Gaussian, as has been well known at least since the classic contributions of Mandelbrot (1963) and Fama (1965). For that reason, even crude approaches like historical HS-*VaR*, although maintaining an unconditional perspective, dispense with normality by building an approximation to the unconditional distribution from historically observed returns.

Figure 8 serves to illustrate the strong unconditional non-normality in returns, as it displays a QQ plot for daily S&P500 returns from January 2, 1990 to December 31, 2010. That is, it plots quantiles of the standardized returns against quantiles of the standard normal distribution. If the returns were unconditionally normally distributed, the points would fall along the 45-degree line. Clearly, the daily returns are not normally distributed.

Now consider the conditional case. Note that in specifying the general conditional variance model (7) we made no assumptions as to the conditional distribution of returns. That is, we made no assumptions as to the distribution of returns standardized by their conditional variance; i.e., the distribution of z_t in (7). But in converting objects like GARCH conditional variances into GARCH-VaR, for example, we *did* invoke conditional normality. At least four points are worth making.

First, conditional normality *can* be, and sometimes is, an acceptable assumption. Conditional normality does not imply unconditional normality, and indeed volatility dynamics “fatten” the tails of unconditional distributions relative to their conditional counterparts, so that conditionally Gaussian models sometimes match the unconditional fat tails present in the data. Put differently, distributions of returns standardized by their conditional volatilities can be approximately Gaussian, even if returns are clearly unconditionally non-Gaussian.

Second, conditional normality is not *necessarily* an acceptable assumption. Sometimes, for example, the unconditional distribution of returns might be *so* fat-tailed that the

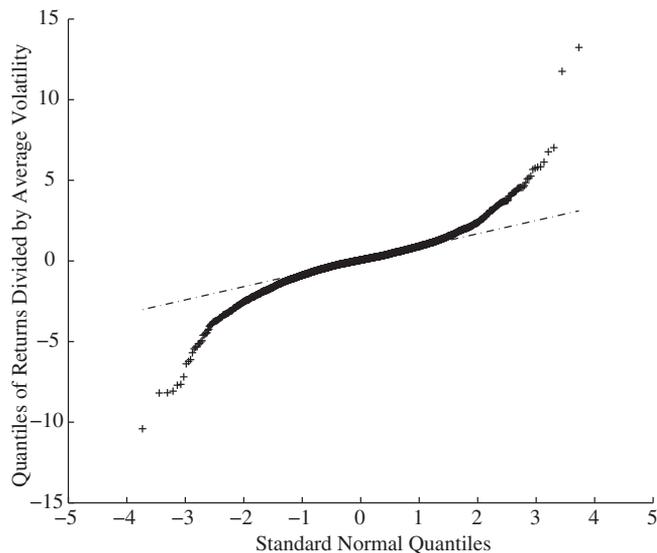


Figure 8 QQ plot of S&P500 returns. We show quantiles of daily S&P500 returns from January 2, 1990 to December 31, 2010, against the corresponding quantiles from a standard normal distribution.

volatility model cannot fatten conditionally Gaussian tails enough to match the unconditional distribution successfully.

Third, beyond fat unconditional tails, there may be *other* unconditional distributional features, such as skewness, that could *never* be captured under any symmetric conditional density assumption such as Gaussian, independent of the conditional variance model used. Matching the unconditional density in general requires flexible conditional variance *and* conditional density specifications.

Fourth, our goal in flexibly specifying the conditional density is not merely to replicate the unconditional density successfully. Rather, for risk measurement and management purposes the conditional density is *the object* of direct and intrinsic interest. That is, best-practice risk measurement and management often requires an estimate of the entire conditional distribution of returns, not just insufficient statistics like its conditional variance, conditional VaR, or conditionally expected shortfall. Hence we need a flexible specification of the conditional density.

Empirical analyses typically find that, although standardization by GARCH and related volatilities promotes normality, the standardized returns remain non-normal. The nature of the non-normality of standardized returns, moreover, varies systematically across asset classes. For example, standardized returns from mature foreign exchange markets are typically symmetric but leptokurtic, while standardized returns on aggregate equity indexes are typically skewed.

To illustrate we show in Figure 9 a Gaussian QQ plot for S&P500 returns standardized by the time-varying volatilities from the asymmetric NGARCH(1,1) model previously used in calculating the VaRs in Figure 3. The QQ plot reveals that the NGARCH-standardized returns conform more closely to normality than do the raw returns of Figure 8. It also reveals, however, that the left tail of the return distribution remains far from Gaussian. In particular, there are too many large negative returns relative to what one would expect if the standardized returns were Gaussian.

As the *VaR* itself refers to a specific quantile, this QQ plot in effect provides an assessment of the normal NGARCH-based *VaRs* defined in (19) across all possible coverage rates, p . In particular, judging by the coherence of the positive quantiles, the figure suggests that the normal-NGARCH-*VaR* approach works reasonably well at moderate coverage rates for a well-diversified portfolio representing a short position on the market index. On the other hand, for a diversified portfolio that is long on the market index, the approach only works if the desired coverage rate is relatively large, say in excess of about 15% or a value of around negative one in the figure. Moving further into the tail, the normal approximation deteriorates quite badly, rendering the corresponding normal-based *VaRs* unreliable. Of course, the corresponding conditional expected shortfall defined in (2) depends on the entire left tail, and will consequently be badly biased across all coverage rates due to the poor tail approximation.

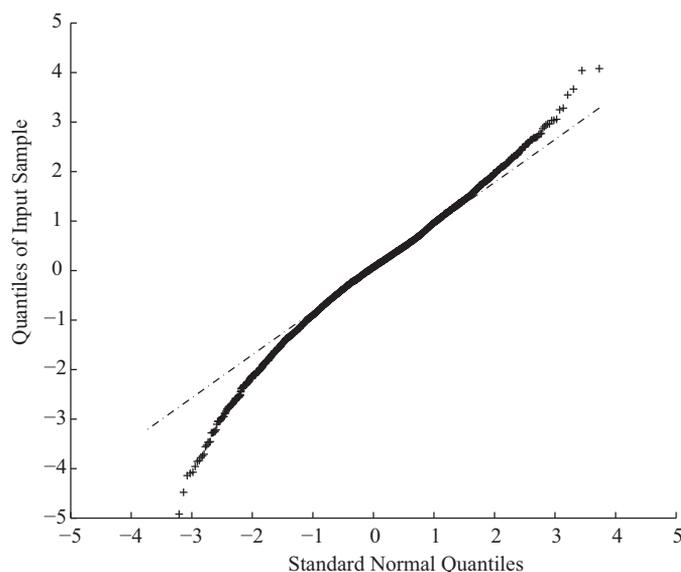


Figure 9 QQ plot of S&P500 returns standardized by NGARCH volatilities. We show quantiles of daily S&P500 returns standardized by the dynamic volatility from a NGARCH model against the corresponding quantiles of a standard normal distribution. The sample period is January 2, 1990 through December 31, 2010. The units on each axis are standard deviations.

Now consider standardizing the returns not by a GARCH or related model-based conditional volatility, but rather by *realized* volatility. Figure 10 shows a Gaussian QQ plot for daily S&P500 returns standardized by $AvgRV$. In contrast to the poor fit for the left tail evident in the QQ plot for the GARCH-standardized returns of Figure 9, the QQ plot for the $AvgRV$ -standardized returns in Figure 10 is remarkably close to normality throughout the support, including in the left tail. This striking empirical result was first systematically documented for exchange rates in Zhou (1996) and Andersen, Bollerslev, Diebold, and Labys (2000a), and extended to equity returns in Andersen et al. (2001a); see also the recent work by Andersen, Bollerslev, Frederiksen, and Nielsen (2010b) and the many references therein.³²

It is worth stressing that the QQ plots in Figures 9 and 10 rely on the *identical* daily S&P500 return series, but simply use two different volatility measures to standardize the raw returns: a GARCH-based estimate of σ_t and the realized volatility $AvgRV_t^{1/2}$. Putting things into perspective, the conditional non-normality of daily returns has long been seen as a key stylized fact in market risk management; see, e.g. Christoffersen (2003). Thus, identifying a volatility measure that produces approximately normally distributed

³² Andersen, Bollerslev, and Dobrev (2007c) explores the theoretical basis for this relationship and provides a detailed examination of the empirical fit for daily S&P500 returns.

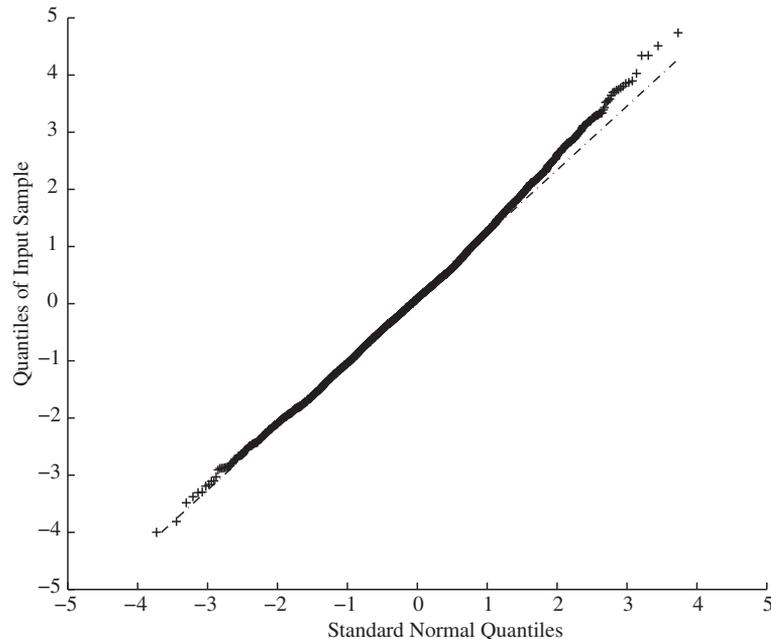


Figure 10 QQ plot of S&P500 returns standardized by realized volatilities. We show quantiles of daily S&P500 returns standardized by *AvgRV* against the corresponding quantiles of a standard normal distribution. The sample period is January 2, 1990 through December 31, 2010. The units on each axis are standard deviations.

standardized returns is both surprising and noteworthy. Of course, the realized volatility used in the standardization in Figure 10 is based on high-frequency data over the *same* daily time interval as the return, while the GARCH volatility used in Figure 9 is a true one-day-ahead prediction.

Against this background on the very different distributional properties of unstandardized, GARCH-standardized and RV-standardized returns, in this section we discuss how to use the different standardizations and resulting distributions to construct accurate predictive return distributions. An important part of that discussion, particularly in the GARCH-standardized case, involves specification of empirically realistic (i.e., non-Gaussian) conditional return distributions.

2.3.1 Procedures Based on GARCH

The GARCH dynamic directly delivers one-day ahead volatility forecasts. In order to complete the daily predictive return distribution, one simply needs to postulate a distribution for the z_t return shock in (7). Although the normal assumption may work well in certain cases, as Figure 9 makes clear, it often underestimates large downside risks. As such, it is important to consider alternatives that allow for fat tails and/or asymmetries

in the conditional distribution. Specifically, in the case of VaR we are looking for ways to more accurately assess the cut-off κ_p^{-1} in

$$VaR_{T+1|T}^p \equiv \sigma_{T+1} \kappa_p^{-1}, \quad (38)$$

instead of simply relying on Φ_p^{-1} from the standard normal distribution.³³ Of course, doing this for all values of $p \in [0, 1]$ essentially amounts to mapping out the entire conditional return distribution.

Perhaps the most obvious approach is to look for a parametric distribution that is more flexible than the normal. One example is the (standardized) Student- t distribution, which relies on only one additional degrees-of-freedom parameter in generating symmetric fat tails. Such an approach was first pursued by Bollerslev (1987), who showed how the likelihood function for the normal-GARCH model in (9) is readily extended to the GARCH- t case, thus allowing for the estimation of the degrees-of-freedom parameter (along with the other GARCH parameters) that best describes the return distribution, and in turn the requisite κ_p^{-1} for calculating the VaR in (38).

This approach works reasonably well when the conditional return distribution is close to symmetric. However, as illustrated by the QQ plots discussed above, equity portfolios are often severely left skewed. The Generalized Error Distribution (GED), first employed in this context by Nelson (1991), explicitly allows for asymmetries, as do some of the different generalizations of the Student- t distribution suggested by Hansen (1994) and Fernandez and Steel (1998), among others. Alternatively, following Engle and Gonzalez-Rivera (1991) the whole density for z_t may be approximated using more flexible semiparametric procedures.

Rather than postulating a particular parametric density, one can also simply approximate the quantiles of non-normal distributions via Cornish-Fisher type expansions. This approach was first advocated in the context of GARCH modeling and forecasting by Baillie and Bollerslev (1992). The only inputs needed for estimating κ_p^{-1} are the *unconditional* sample skewness and kurtosis statistics for the standardized returns.³⁴

Meanwhile, a common problem with most GARCH models, regardless of the innovation distribution, is that the specific distribution is not preserved under temporal aggregation; see, e.g. the discussion in Drost and Nijman (1993) and Meddahi and Renault (2004). For example, even if the standardized daily returns from a GARCH(1,1) model were normal, the implied weekly returns would not be. In turn, this implies that the term structure of VaR s is not closed under temporal aggregation either. Instead, the

³³ The 1996 amendment to the 1988 Basel Accord somewhat arbitrarily recommends the use of a multiplicative factor of at least -3.0 in the construction of a 1% VaR , relative to the $\Phi_{0.01}^{-1} = -2.33$ implied by the standard normal distribution; see also the discussion in Chan, Deng, Peng, and Xia (2007).

³⁴ More accurate approximations may in theory be obtained by including higher-order unconditional sample moments in the Cornish-Fisher expansion, but this does not always produce satisfactory results.

multi-period *VaRs* need to be computed via Monte Carlo simulations or other numerical methods, as exemplified by Guidolin and Timmermann (2006).³⁵ This also means that the Cornish–Fisher and related approximations, which only provide partial characterizations of the underlying daily return distribution in the form of specific quantiles, generally will not suffice for answering questions regarding the distribution of temporally aggregated returns. Below, we discuss a viable approach that effectively combines a parametric volatility model with a data-driven conditional distribution. First, however, we discuss how realized volatilities, if available, may be used in the calculation of even more accurate predictive return distributions by effectively incorporating the intraday information into the distributional forecasts.

2.3.2 Procedures Based on Realized Volatility

The basic idea underlying the construction of RV-based predictive return distributions is to treat the time series of RVs as stochastic. Hence, in contrast to the GARCH-based procedures, which seek to describe the predictive distribution through an appropriately specified univariate distribution for the standardized returns, the RV-based procedures necessitate, at a minimum, a bivariate random distribution for the returns *and* the realized volatilities.

This relatively new approach to risk measurement was first suggested by Andersen et al. (2003a). The approach is directly motivated by the empirical regularities pertaining to the RV measures highlighted above. First, as discussed in Section 2.2, simple time series models for the realized volatilities, like the HAR–RV specification, generally result in more accurate volatility forecasts than do the conventional GARCH models based on daily data only.³⁶ Second, as shown in Section 2.3, the distributions of daily returns standardized by the same-day RVs typically appear close to Gaussian. Taken together, this suggests a mixture-of-distributions type approach for characterizing the time $T + 1$ return distribution, in which the predictive distribution for RV_{T+1} serves as the mixture variable.³⁷

Specifically, assuming that the standardized return is normal, $r_{T+1}/RV_{T+1}^{1/2} \sim N(0, 1)$, and that the distribution of the time $T + 1$ realized volatility conditional on time T

³⁵ The affine GARCH models suggested by Heston and Nandi (2000) and Christoffersen, Heston, and Jacobs (2006), when combined with the methods of Albanese, Jackson, and Wiberg (2004), also allow for relatively easy-to-compute term structures for VaR, but some numerical calculations are still required.

³⁶ This empirical regularity may also be justified through more formal theoretical arguments, as to why the simple reduced form RV-based procedures often work better than structural model-based approaches in practice; see, Andersen, Bollerslev, and Meddahi (2004); Andersen, Bollerslev, and Meddahi (2011a); and Sizova (2011).

³⁷ There is a long history, dating back to Clark (1973), of using mixture-of-distributions to describe the unconditional distribution of returns. What is fundamentally different in the RV-based approach is to treat the mixing variable as directly observable and predictable.

information is log-normal, the resulting normal log-normal mixture distribution for the time $T + 1$ returns may be expressed as

$$f_T(r_{T+1}) = \frac{1}{2\pi\sigma_{\ell,T+1}^2} \int_0^\infty y^{-3/2} \exp\left\{-\frac{r_{T+1}^2}{2y} - \frac{1}{2\sigma_{\ell,T+1}^2} (\ln(y) - \mu_{\ell,T+1})^2\right\} dy,$$

where $\mu_{\ell,T+1}$ and $\sigma_{\ell,T+1}^2$ denote, respectively, the time T conditional mean and variance of $\log(RV_{T+1})$. For example, postulating a HAR-RV type model for logRV with homoskedastic errors, we obtain,

$$\mu_{\ell,T+1} = \beta_0 + \beta_1 \log(RV_T) + \beta_2 \log(RV_{T-4:T}) + \beta_3 \log(RV_{T-20:T}),$$

and $\sigma_{\ell,T+1}^2 = \sigma_v^2$, respectively.³⁸

The simple HAR-RV model for the conditional mean $\mu_{\ell,T+1}$ could, of course, be extended in several directions. For instance, as noted above, when modeling large equity portfolios, asymmetries, or “leverage effects”, are often statistically significant. Also, in their actual empirical implementation Andersen et al. (2003a) use a long-memory ARFIMA model for logRV in place of the HAR-RV formulation. This makes little difference for the maximum ten-day forecast horizons considered in their analysis, but it could be important to do so in the calculation of longer run, say quarterly (~ 66 days ahead) or annual (~ 252 days ahead), distributional forecasts.

The mixture distribution described above treats $\sigma_{\ell,t}$ as constant. However, it is natural to think about the volatility-of-volatility as being time varying with its own GARCH dynamics. Such an approach has been pursued empirically by Maheu and McCurdy (2011), who report that allowing for temporal variation in $\sigma_{\ell,t}$ does not actually result in materially different predictive return distributions. Going one step further, Bollerslev, Kretschmer, Pigorsch, and Tauchen (2009a) develop a joint conditional density model for the returns, the “smooth” volatility, and the variation due to jumps $\{r_t, \log(BPV_t), \log(RV_t = /BPV_t)\}$. In that model the predictive distribution for the returns is therefore obtained through a more complicated normal mixture involving two separate mixing variables the basic idea remains the same.

This continues to be an active area of research, and it is too early to say which of the different approaches will be the “winner”. It is evident, however, that any of the relatively simple RV-based procedures described above almost invariably generate more accurate predictive return distributions than the traditional GARCH-based distributional forecast, especially over relatively short one-day to one-week horizons.

³⁸ Although it is not possible to express the density function in closed form, it is easy to calculate numerically by repeated simulations from a normal distribution with a random variance drawn from a log-normal distribution with the requisite mean and variance.

2.3.3 Combining GARCH and RV

Just as the GARCH and RV concepts may be formally combined in the construction of volatility forecasts, they may be similarly combined to produce distributional forecasts. The procedures discussed in the previous section, of course, also utilize the realized volatility measures in the construction of the forecasts. However, they generally do not provide a direct link between the GARCH conditional variance σ_t and the realized volatility measures.

Forsberg and Bollerslev (2002) provides a first attempt at doing that. Their RV-GARCH style model is based on the assumption that RV is conditionally Inverse Gaussian distributed³⁹

$$f_T(RV_{T+1}) \sim IG(\sigma_{T+1}^2, \eta),$$

together with a GARCH-style process for the conditional expectation of RV,

$$E_T(RV_{T+1}) = \sigma_{T+1}^2 = \omega + \alpha r_{w,T}^2 + \beta \sigma_T^2.$$

Further assuming that the RV-standardized returns are normally distributed, results in the predictive normal inverse Gaussian (NIG) distribution with conditional variance σ_{T+1} ,

$$f_T(r_{w,T+1}) = \int f_T(r_{w,T+1}|RV_{T+1})f_T(RV_{T+1})dRV_{T+1} \sim NIG(\sigma_{T+1}^2, \eta).$$

Closely related RV-GARCH type models have also been developed and used in the context of option pricing by Christoffersen et al. (2011b), Corsi, Fusari, and LaVecchia (2011), and Stentoft (2008).

The more recent Realized GARCH and HEAVY models discussed in Section 2.2.3 take this approach one step further by providing a coherent joint modeling framework for $\{r_t, \sigma_t, RV_t\}$, where, importantly, the conditional variance of the returns, σ_t^2 , is not identical to the conditional expectation of RV_t . These models directly deliver one-day volatility and return distribution forecasts. In contrast to the GARCH-X style models and some of the RV-based procedures discussed above, multi-day distributional forecasts may also readily be computed using numerical simulation techniques.

These and other related GARCH-RV forecasting approaches are still being explored in the literature. Given the significant improvements afforded by incorporating the intraday information into the GARCH volatility forecasts through the RV measures,

³⁹The Inverse Gaussian distribution closely approximates the log-normal distribution for the realized volatility depicted in Figure 5 above.

especially during rapidly changing market conditions, we expect these procedures to play an increasingly important role as the field moves forward.

2.3.4 Simulation Methods

In the discussion above, we have often pointed to the use of numerical simulation techniques as a way to calculate quantiles or distributions that are not available in closed form. These techniques differ in terms of their underlying assumptions ranging from fully parametric to essentially non-parametric.

Bootstrapping, or Filtered Historical Simulation (FHS), assumes a parametric model for the second moment dynamics, and then bootstraps from the standardized returns to build up the required distribution. At the portfolio level this is easy to do. First calculate the standardized pseudo portfolio returns as,

$$\hat{z}_{w,t} = r_{w,t}/\hat{\sigma}_t, \quad t = 1, 2, \dots, T, \quad (39)$$

using one of the variance models discussed above. Then, in order to calculate a one-day-ahead VaR , one simply uses the order statistic for the standardized returns combined with the volatility forecast to construct,⁴⁰

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

This same idea could also be used to numerically calculate the VaR for parametric distributions where the quantiles are not readily available, by repeatedly drawing $z_{w,t}$ from the specific distribution.

The construction of multi-day VaR s is more time consuming, but conceptually straightforward. It requires simulating future paths from the volatility model using the standardized returns sampled with replacement as the innovations. This approach has been exploited by Diebold, Schuermann, and Stroughair (1998b), Hull and White (1998), and Barone-Adesi, Bourgoin, and Giannopoulos (1998), among others, and we refer to these studies for further details concerning its practical implementation.⁴¹

The FHS methodology was originally developed in a GARCH setting. However, for some of the RV-based procedures discussed above, one would naturally use RV or its expected value to standardize the portfolio returns. In these situations the standardized returns should be sampled from

$$\hat{z}_{w,t} = r_{w,t}/\sqrt{RV_t}, \quad t = 1, 2, \dots, T.$$

⁴⁰ For the expected shortfall in (2) one would simply average over the draws that exceed $z_w(T+1)p$.

⁴¹ Pritsker (2006) also provides additional evidence on the effectiveness of the FHS approach.

or

$$\hat{z}_{w,t} = r_{w,t} / \sqrt{E_{t-1}[\text{RV}_t]}, \quad t = 1, 2, \dots, T.$$

Of course, if the underlying model is based on a specific distributional assumption about the RV-standardized returns, that distribution should be used in lieu of the non-parametric bootstrap. Also, for RV-based GARCH models and related procedures, one might need to perform a bootstrap from the supposedly *i.i.d.* bivariate innovations for RV and returns, but the basic idea remains the same.

2.3.5 Extreme Value Theory

The different parametric and non-parametric procedures discussed above for characterizing the conditional return distribution, including the simulation based bootstrap procedures, are designed to work well for the center of the distribution and VaRs with relatively large coverage rates, say in excess of 5%. In many situations, however, one is primarily interested in the tails of the distributions and the risks associated with extremely large price changes. Extreme Value Theory (EVT) provides a formal statistical framework for meaningfully estimating the tails based on extrapolating from the available observations. McNeil, Frey, and Embrechts (2005) provide an excellent survey of these techniques and their application in quantitative risk management, and we merely highlight some of the key ideas here; early important work in this area also includes Diebold, Schuermann, and Stroughair (1998b), Longin (2000), and McNeil and Frey (2000).

Standard EVT is based on the assumption of i.i.d. observations. This may be a good approximation for many applications in actuarial science, but financial returns and large absolute price changes, in particular, are obviously not i.i.d. through time. However, in parallel to the FHS approach discussed immediately above, EVT may easily be combined with dynamic volatility models by applying the EVT-based approximations to the estimated return shocks $\hat{z}_{w,t} = r_{w,t} / \hat{\sigma}_t$ rather than the returns themselves. Since the return shocks are much closer to being i.i.d. than are the returns, this makes the application of EVT much more reasonable. Having estimated the tails for $\hat{z}_{w,t}$, these are easily transformed to tails or extreme quantiles of the raw returns by scaling with $\hat{\sigma}_t$.

EVT has the advantage that each tail of the distribution can be modeled separately. But it has the limitation that it only describes the tails of the distribution and not the entire distribution. It is therefore not possible to simulate data from an EVT distribution unless further assumptions are made. One way to proceed is to use EVT in the tails combined with FHS for characterizing the center of the distribution. Assume for example that EVT captures well the 2% most extreme positive shocks and the 3% most extreme negative shocks. Return shocks can then be simulated by first drawing a trinomial variable that comes up $\{-1, 0, +1\}$ with probabilities $\{.03, .95, .02\}$. When the trinomial comes up 0 then a shock is drawn randomly (with replacement) from the sample of $\hat{z}_{w,t}$

with the left 3% and right 2% extremes removed. When the trinomial comes up -1 then a shock is drawn from the left-tail EVT distribution. Similarly, a draw is made from the right-tail EVT distribution when the trinomial comes up $+1$. This same idea may also be used in “stress testing” the portfolio, by increasing the probabilities assigned to the tails, in turn generating a disproportionate number of draws from the extreme part of the distribution.

Portraying prices as evolving in continuous time, the extreme price increments are naturally thought of as “jumps”. The discussion in Section 2.2.2 above outlines several ways for disentangling the jumps on an ex-post basis with the help of high-frequency intraday data. Following the recent work of Bollerslev and Todorov (2011b), the high-frequency filtered jumps may in turn be used in the estimation of the corresponding jump tail distribution and the probability of observing an extreme price change. Work along these lines is still in its infancy. However, we conjecture that in parallel to the gains in predictive accuracy afforded by the use of realized volatility measures relative to GARCH type models estimated with daily data only, similar gains may be available through the proper use of the high-frequency data for more accurately estimating the jump tails and the extremes of the return distributions.

3. CONDITIONAL ASSET-LEVEL RISK ANALYSIS

Our discussion up until now has focused on dynamic volatility models for univariate returns. These methods are well suited for portfolio-level risk measures such as aggregate VaR and ES . 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 a simultaneous increase in stock market volatility and asset correlations, as typically occurs in times of market stress, then a multivariate model is needed. Active risk management, such as portfolio VaR minimization, also requires a multivariate model that provides a forecast for the entire covariance matrix.⁴² Bank-wide VaR is also made up of many desks with multiple traders on each desk, and any sub-portfolio analysis is not possible with the aggregate portfolio-based approach. Similarly, multivariate models are needed for calculating sensitivity risk measures and answering questions such as: “If I add an additional 1,000 shares of Apple to my portfolio, how much will my VaR increase?”

In this section we therefore consider the specification of models for the full N -dimensional conditional distribution of asset returns. To set out the notation, let Ω_t denote the $N \times N$ covariance matrix of the $N \times 1$ vector of asset returns R_t . The covariance matrix will have $\frac{1}{2}N(N + 1)$ distinct elements, but structure needs to be

⁴² Brandt, Santa-Clara, and Valkanov (2004) provide an alternative and intriguing 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.

imposed to guarantee that the covariance matrix forecasts are positive definite (pd), or even positive semi-definite (psd). A related, and equally important, practical issue involves the estimation of the parameters governing the dynamics for the $\frac{1}{2}N(N+1)$ individual elements.

We begin with a brief discussion of models and methods based on daily data. We then discuss how high-frequency data and realized variation measures may be incorporated into the construction of better covariance matrix and multivariate distributional forecasts. A notable aspect of our treatment is our inclusion and emphasis on methods that are applicable even when N is (relatively) large. This contrasts with much of the extant literature, which focuses on relatively low-dimensional models.⁴³

3.1 Modeling Time-Varying Covariances Using Daily Data and GARCH

The natural multivariate generalization of the RM variance dynamics in (4) provides a particularly simple approach to modeling large dimensional covariance matrices. It assumes that the dynamics of all the variances and covariances are driven by a single scalar parameter λ ,

$$\Omega_t = \lambda\Omega_{t-1} + (1 - \lambda)R_{t-1}R'_{t-1}. \quad (40)$$

In parallel to the univariate case, the recursion may be initialized by setting Ω_0 equal to the sample average coverage matrix.⁴⁴

The simple structure of (40) guarantees that the estimated covariance matrices are psd, and even pd if the initial covariance matrix, Ω_0 , is pd, as the sum of a psd and pd matrices is itself pd. Letting Ω_0 equal the sample coverage matrix, it will be pd as long as the sample size T exceeds the number of assets N and none of the assets are trivial linear combinations of others, thus rendering the RM covariance matrix forecasts pd as well.

At the same time, however, the RM approach is clearly very restrictive, imposing the same degree of smoothness on all elements of the covariance matrix. Moreover, covariance matrix forecasts generated by the multivariate RM approach inherit the implausible scaling properties of the univariate RM forecasts in Section 2.1, and will in general be suboptimal for the reasons discussed in the univariate context.

This, in turn, motivates a direct extension of the univariate GARCH approach to a multivariate setting. In particular, extending the expression in (7) to a vector setting, the

⁴³ See Bauwens, Laurent, and Rombouts (2006) for a survey of multivariate GARCH models, and Chib, Omori, and Asai (2009) for a survey of multivariate stochastic volatility models, involving daily data and moderate dimensions.

⁴⁴ As previously noted, empirically more realistic dependence structures have also been explored by RM, but following standard convention, we will continue to refer to exponential smoothing as the RM approach.

generic representation for a multivariate return process with time-varying conditional first- and second-order moments becomes

$$R_t = M_t + \Omega_t^{1/2} Z_t \quad Z_t \sim \text{i.i.d.}, \quad E(Z_t) = 0, \quad \text{Var}(Z_t) = \mathcal{I}, \quad (41)$$

where \mathcal{I} denotes the identity matrix, and the $N \times N$ matrix $\Omega_t^{1/2}$ is one of the “square-root” representations, e.g. the Cholesky decomposition, of the covariance matrix Ω_t . We refer to any specification in which Ω_t is a non-trivial function of the time $t - 1$ information set as a multivariate GARCH model. As with the univariate models discussed above, we will assume for simplicity that the daily means are all zero, or $M_t = 0$.⁴⁵

The most obvious extension of the popular univariate GARCH(1,1) model in (8) then takes the form

$$\text{vech}(\Omega_t) = \text{vech}(C) + B \text{vech}(\Omega_{t-1}) + A \text{vech}(R_{t-1}R'_{t-1}), \quad (42)$$

where the vech , or “vector-half”, operator converts the unique upper triangular elements of a symmetric matrix into a $\frac{1}{2}N(N+1) \times 1$ column vector, and the A and B matrices are both of dimension $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$. In parallel to the expression for the univariate model in (10), the long-run forecasts from the multivariate GARCH(1,1) model in (42) converge to $\text{vech}(\Omega) = (\mathcal{I} - A - B)^{-1} \text{vech}(C)$, provided the eigenvalues of $A + B$ are all less than unity and the inverse of the $(\mathcal{I} - A - B)^{-1} \text{vech}(C)$ matrix exists. This model-implied unconditional covariance matrix can be quite sensitive to small perturbations in A and B . As such, it is often desirable to restrict the matrix C to ensure that the long-run forecasts from the model are well behaved and converge to sensible values.

“Variance targeting” provides a powerful tool for doing that, in effect, “disciplining” multivariate volatility models. This idea was first suggested by Engle and Mezrich (1996), who proposed replacing the C matrix in the multivariate GARCH(1,1) model above with

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

This in turn ensures that the covariance matrix forecasts converge to their unconditional sample analog. Of course, if the risk manager has other information pertaining to

⁴⁵ This assumption is quite innocuous, and does not materially affect the inference over daily horizons. For models defined over longer return horizons, simply replace R_t with the de-meaned returns $R_t - M_t$ in all of the expressions below.

some of the elements in the covariance matrix, this may be used in a similar manner in fixing the relevant values in C .

Variance targeting also helps in the implementation of multivariate volatility models more generally, by reducing the number of parameters to be estimated. The most general version of the multivariate GARCH(1,1) model in Eqn (42), for example, has $O(N^4)$ parameters. More precisely, there are $N^4/2 + N^3 + N^2 + N/2$ parameters; hence, for example, for $N = 100$ there are 51,010,050 parameters! Estimating this many free parameters is obviously infeasible.⁴⁶ The “diagonal GARCH” parameterization, originally proposed by Bollerslev, Engle, and Wooldridge (1988), helps by restricting the A and B matrices to be diagonal. The number of parameters is still $O(N^2)$, however, and full-fledged estimation of the diagonal model is generally deemed computationally infeasible for systems much larger than $N = 5$.

Going one step farther, we obtain the most draconian version of the diagonal GARCH(1,1) model by restricting the A and B matrices to be scalar,

$$\Omega_t = C + \beta\Omega_{t-1} + \alpha(R_{t-1}R'_{t-1}). \quad (43)$$

This, of course, closely mirrors the RM approach discussed above, with the important difference that the long-run covariance matrix forecasts converge to the non-degenerate matrix $\Omega = (1 - \alpha - \beta)^{-1}C$ (provided that $\alpha + \beta < 1$). Estimation of this model may again be further simplified through the use of covariance targeting, replacing the C matrix by

$$C = (\mathcal{I} - \alpha - \beta) \frac{1}{T} \sum_{t=1}^T R_t R'_t,$$

leaving only the two scalar parameters, α and β , to be determined.⁴⁷

Even so, estimation can still be very cumbersome in large dimensions due to the need to invert the $N \times N$ covariance matrix Ω_t for every day in the sample in order to evaluate the likelihood function, which, of course, must be done numerous times during a numerical optimization. In an effort to circumvent this problem, Engle, Shephard, and Sheppard (2008) suggested replacing the regular likelihood function in the optimization of the model by a Composite Likelihood (CL) based on summing the log-likelihoods of pairs of assets,

$$\text{CL}(\alpha, \beta) = \sum_{t=1}^T \sum_{i=1}^N \sum_{j>i}^N \log f(\alpha, \beta; R_{i,t}, R_{j,t}), \quad (44)$$

⁴⁶ Without further restricting the structure of the model, there is also no guarantee that covariance matrix forecasts produced by the model are actually psd.

⁴⁷ This model also readily ensures that Ω_t and the corresponding forecasts are psd, as long as $\alpha > 0$ and $\beta > 0$.

where $\log f(\alpha, \beta; R_{i,t}, R_{j,t})$ denotes the bivariate normal density for asset pair $R_{i,t}$ and $R_{j,t}$. Each pair of assets yields a valid (but inefficient) likelihood for α and β , but by summing over all pairs the resulting CL-estimator becomes “relatively efficient”. In contrast to the standard likelihood function, the CL approach requires the inversion of 2×2 matrices only, albeit a total of $N(N + 1)/2$ for each day in the sample, but that, of course, is easy to do even in high-dimensional situations.

Still, the assumption that all of the variances and covariances have the same speed of mean reversion, as dictated by the α and β scalar parameters, is obviously very restrictive. As such, more flexible procedures may be needed in describing the temporal variation in Ω_t in an empirically realistic fashion, especially when considering disperse types of assets or asset classes. One approach that has proven especially useful is to focus on modeling the correlations rather than the covariances.

3.1.1 Dynamic Conditional Correlation Models

A conditional covariance matrix may always be decomposed into a conditional correlation matrix pre- and post-multiplied by the diagonal matrix of conditional standard deviations,

$$\Omega_t = D_t \Gamma_t D_t. \quad (45)$$

Motivated by this decomposition, Bollerslev (1990) first proposed treating the conditional correlations as constant, $\Gamma_t = \Gamma$, so that the dynamic dependencies in Ω_t are driven solely by the temporal variation in the conditional variances. The resulting Constant Conditional Correlation (CCC) GARCH model has the advantage that it is easy to estimate, even in large dimensions, in essence requiring only the estimation of N univariate models. Specifically, for each of the individual assets, one may first estimate an appropriate univariate GARCH model. These models may differ from asset to asset, thus allowing for much richer, possibly asymmetric and long-memory style, dependencies than in the multivariate diagonal GARCH models discussed above. Then, denoting the resulting vector of standardized returns by $\hat{e}_t = R_t \hat{D}_t^{-1}$, the conditional correlation matrix Γ is efficiently estimated by the sample mean of the outer product of these standardized returns.

Although the CCC GARCH model is easy to estimate, and may work well over relatively short time-spans, the underlying assumption of constant conditional correlation is arguably too restrictive in many situations.⁴⁸ In response to this, Engle (2002a) and Tse and Tsui (2002) independently suggested allowing for dynamically varying conditional

⁴⁸The literature is rife with examples of time-varying correlations. Cross-market stock–bond return correlations, for instance, are often found to be close to zero or slightly positive during bad economic times (recessions), but negative in good economic times (expansions); see, e.g. the discussion in Andersen et al. (2007b). Numerous studies, including Longin and Solnik (1995), have also demonstrated that the correlations among international equity markets change over time. Similarly, there is ample evidence from the recent financial crisis that default correlations can change quite dramatically over short periods of time.

correlations within a GARCH framework. Specifically, assuming a simple scalar diagonal GARCH(1,1) structure for the correlations, the Dynamic Conditional Correlation (DCC) GARCH model, first proposed by Engle (2002a), may be expressed as,

$$Q_t = C + \beta Q_{t-1} + \alpha(e_{t-1}e'_{t-1}), \quad (46)$$

where as before $e_t = R_t D_t^{-1}$, and the matrix of conditional correlations are defined by the normalized elements of Q_t , $\rho_{i,j,t} = q_{i,j,t} / (\sqrt{q_{i,i,t}}\sqrt{q_{j,j,t}})$, or in matrix format,

$$\Gamma_t = \text{diag}\{Q_t\}^{-1/2} Q_t \text{diag}\{Q_t\}^{-1/2}. \quad (47)$$

This latter normalization ensures that all of the correlations fall between -1 and 1 .

In parallel to the CCC model, estimation of the DCC model may proceed in two steps, by first estimating univariate GARCH models for each of the assets. In contrast to the CCC model, however, the second step estimation in the DCC model, involving the dynamics of the Γ_t matrix, requires the use of numerical optimization techniques. To help facilitate this step, and at the same time ensure that the forecasts from the model are well behaved, it is often desirable to rely on correlation targeting. The parametrization in (46) does not immediately lend itself to that, as the unconditional expectation of Q_t differs from the unconditional expectation of $e_t e'_t$. Instead, following Aielli (2006) and re-parameterizing the dynamics for Q_t as

$$Q_t = (1 - \alpha - \beta)C^* + \beta Q_{t-1} + \alpha(e_{t-1}^* e_{t-1}^{*'}), \quad (48)$$

where $e_t^* = \text{diag}\{Q_t\}^{1/2} e_t$, it follows that $E(Q_t) = E(e_t^* e_t^{*'})$. Correlation targeting is therefore readily implemented by replacing C^* with the sample mean of the $e_t^* e_t^{*'}$ matrix, or some other hypothesized value. This corrected DCC (cDCC) model is relatively easy to estimate in high dimensions when combined with the composite likelihood idea discussed earlier.⁴⁹

Another easy-to-implement DCC type model has recently been proposed by Engle and Kelly (2008). In this model, instead of assuming the same dynamic dependencies for all of the correlations, the time-varying correlations are assumed to be the same across all pairs of assets. Hence the name dynamic equicorrelation, or DECO, model. The assumption of identical correlations, of course, is only applicable when modeling similar types of assets, such as, e.g. a large cross-section of stock returns.⁵⁰ Following Engle and Kelly (2008), the DECO model may be conveniently expressed as

$$\Gamma_t = (1 - \rho_t)\mathcal{I} + \rho_t\mathcal{J}, \quad (49)$$

⁴⁹ The original DCC model defined by (46) and (47), and the cDCC version in (48), also both guarantee that Γ_t is psd, provided that $\alpha > 0$ and $\beta > 0$.

⁵⁰ If this assumption is valid, imposing identical correlations will also generally enhance estimation efficiency relative to a model that treats the pairwise correlations as unrelated.

where \mathcal{I} denotes the N dimensional identity matrix, and \mathcal{J} refers to the $N \times N$ matrix of ones. This representation for Γ_t has the advantage that the inverse is available in closed form,⁵¹

$$\Gamma_t^{-1} = \frac{1}{(1 - \rho_t)} \left[\mathcal{I} - \frac{\rho_t}{1 + (N - 1)\rho_t} \mathcal{J} \right],$$

thus rendering the likelihood function easy to evaluate. Implementation of the DECO model, of course, still requires an assumption about the dynamic dependencies in the common conditional correlation. In particular, assuming a GARCH(1,1) structure,

$$\rho_t = \omega_\rho + \alpha_\rho u_t + \beta_\rho \rho_{t-1},$$

with the updating rule naturally given by the average conditional correlation of the standardized returns,

$$u_t = \frac{2 \sum_{i=1}^N \sum_{j>i}^N e_{i,t} e_{j,t}}{N \sum_{i=1}^N e_{i,t}^2},$$

the model has only three parameters, ω_ρ , α_ρ , and β_ρ , to be estimated.

To convey a feel for the importance of allowing for time-varying conditional correlation, we plot in Figure 11 the estimated equicorrelations from a DECO model for the aggregate equity-index returns for 16 different developed markets from 1973 through 2009.⁵² As the figure shows, there has been a clear low-frequency upward fluctuation in the cross-country correlations, from a typical value of approximately 0.25 in the late 1970s to around 0.70 toward the end of the sample. The movement has not been entirely monotone, however, thus highlighting the flexibility of the DECO modeling approach also to account for important short-run fluctuations in the $1/2 \times 16 \times 15 = 120$ pairwise correlations.

The scalar DCC model defined by (46) and (47), the modified DCC model in (48), and the DECO model in (49) are all extremely parsimonious and readily implemented for large N . They do, however, impose severe restrictions on the correlations, and may thus be seen as overly simplistic in applications involving only a few assets. More elaborate DCC models, including asymmetric formulations (e.g. Cappiello, Engle, and Sheppard (2006)) and regime switching type representations (e.g. Pelletier, 2006), have been proposed to allow for more nuanced modeling when N is small, say $N \leq 5$. We will

⁵¹ The inverse exists if and only if $\rho_t \neq 1$ and $\rho_t \neq -1/(n - 1)$, while Γ_t is psd for $\rho_t \in (-1/(n - 1), 1)$.

⁵² Similar figures are displayed by Christoffersen, Errunza, Jacobs, and Langlois (2011a), and we refer to their study for additional details concerning the data and the methods of estimation.

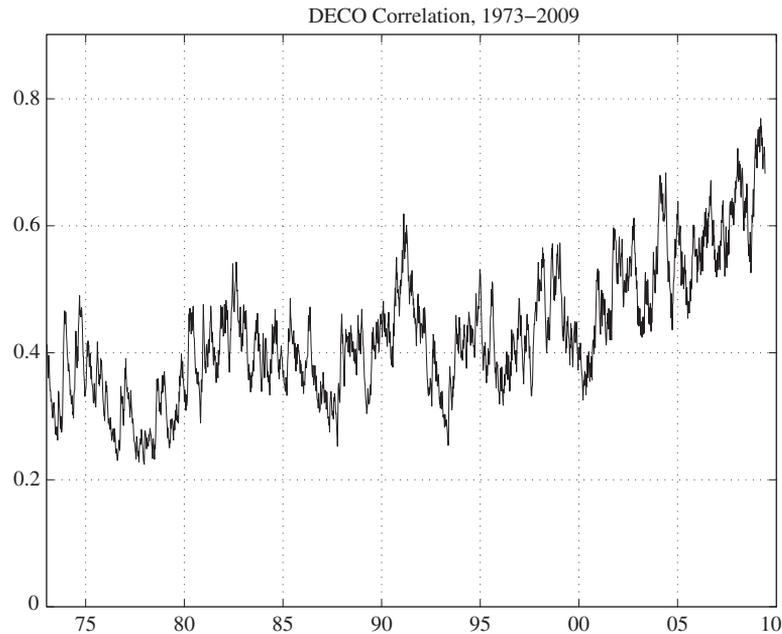


Figure 11 Time-varying international equity correlations. The figure shows the estimated equicorrelations from a DECO model for the aggregate equity-index returns for 16 different developed markets from 1973 through 2009.

not discuss these models here, but refer to the recent book by Engle (2009a) for a comprehensive survey of DCC models. Instead, we turn to an alternative way of disciplining the covariance matrix, namely factor structures.

3.1.2 Factor Structures and Base Assets

Factor structures are, of course, ubiquitous in finance. However, we will keep our discussion short and focused on their explicit use in simplifying the modeling and forecasting of large dimensional dynamic daily covariance matrices, as required for risk measurement and management purposes. More detailed discussions of the use of traditional factor models in the construction of *VaRs* and risk management more generally are available in Jorion (2007) and Connor, Goldberg, and Korajczyk (2010).

Market risk management systems for portfolios of thousands of assets often work from a set of smaller, say 30, observed base assets believed to be the key drivers of the underlying risks. The accuracy of the resulting risk management system, in turn, depends on the distributional assumptions for the base assets and the mapping from the base assets to the full set of assets. The specific choice of base assets depends importantly on the portfolio at hand but may, for example, consist of equity market indices, FX rates, benchmark interest rates, and so on, believed to capture the main

sources of uncertainty. These base assets will typically also be among the most liquid assets in the market. Such an approach is, of course, easier to contemplate for a relatively specialized application with readily identifiable risk factors, such as a US equity portfolio, than a very large diversified entity, such as a major international bank or conglomerate.

Specifically, let $R_{F,t}$ denote the $N_F \times 1$ vector of de-measured returns on the base assets, or systematic risk factors. The distribution of the factors may then generally be expressed as,

$$R_{F,t} = \Omega_{F,t}^{1/2} Z_{F,t}, \quad Z_{F,t} \sim \text{i.i.d.}, \quad E(Z_{F,t}) = 0, \quad \text{Var}(Z_{F,t}) = \mathcal{I}, \quad (50)$$

where the notation corresponds directly to the one in Eqn (41) above for the $N \times 1$ vector of returns R_t . The number of base assets may be considerably higher than usual for traditional factor models employed in finance, but the basic idea is to keep their number much lower than the total number of assets.

The mapping from the N_F base assets to the full set of N assets typically consists of a linear factor structure,

$$R_t = B_0 + B R_{F,t} + v_t, \quad (51)$$

where v_t denotes a $N \times 1$ vector of idiosyncratic risks, B_0 is an $N \times 1$ vector, and the factor loadings are contained in the $N \times N_F$ matrix B . The factor loadings may be obtained from regression, if sufficient historical data exists for the full cross-section of assets. Alternatively, one may exploit the implications from a specific pricing model, if such a model exists. Sometimes, the loadings are also determined in more of an ad hoc fashion, by matching a security without a factor loading to another similar security with a well-defined loading. Importantly, however, both B_0 and B are assumed to be constant.

Now, combining the distributional assumptions in (50) with the basic factor structure in (51), the resulting covariance matrix for R_t may be expressed as,

$$\Omega_t = B' \Omega_{F,t} B + \Omega_{v,t}, \quad (52)$$

where $\Omega_{v,t}$ denotes the $N \times N$ covariance matrix for v_t . Since Ω_t and $\Omega_{v,t}$ are both of the same dimension, this expression does not directly translate into any simplification in the estimation of the covariance matrix for the full set of N returns. However, assuming that the idiosyncratic risks are uncorrelated across assets and that their variances are constant, the expression for Ω_t simplifies to

$$\Omega_t = B' \Omega_{F,t} B + D_v, \quad (53)$$

where $D_v = \Omega_{v,t}$ is a time-invariant diagonal matrix. Moreover, the elements in D_v are readily estimated from the variances of the residuals in the factor model (51). This, of course, still leaves $\Omega_{F,t}$ to be determined. But, by keeping N_F moderately low, $\Omega_{F,t}$ is much easier to estimate than Ω_t . In fact, in addition to any of the techniques discussed in this section, some of the more advanced multivariate GARCH procedures alluded to above could be applied for estimating $\Omega_{F,t}$ when the number of base assets, or N_F , is kept sufficiently low.⁵³

Although convenient from a modeling perspective, the key assumption that $\Omega_{v,t}$ is diagonal and constant over time often appears at odds with the data. Just as variances (and covariances) of raw returns are clearly time varying, so are the variances (and covariances) of idiosyncratic risks. Related to this, the risk exposures of many assets, as encapsulated in the factor loadings, are also likely to change over time, rendering the key covariance matrix representation in (53) with B constant a poor approximation over long time periods. However, for applications exploiting high-frequency intraday data, it is often feasible to alleviate these drawbacks and, as we shall see below, factor structures are often invoked in such settings.

3.2 Intraday Data and Realized Covariances

Thus far our discussion has focused on models tailored toward capturing the dynamics in daily covariances based on daily data. As discussed in Section 2.2, however, for many assets intraday price data are now readily available, and just as this information is useful for the estimation of daily variances, it should be equally, if not more, useful for the estimation of daily asset covariances.

Generalizing the univariate setting in (20), and providing a continuous-time analog to the discrete-time representation in (41), we assume that the $N \times 1$ log-price vector, $P(t)$, is governed by the following multivariate diffusion process,

$$dP(t) = M(t)dt + \Omega(t)^{1/2} dW(t), \quad (54)$$

where $M(t)$ and $\Omega(t)^{1/2}$ denote the $N \times 1$ instantaneous drift vector and the $N \times N$ positive definite “square-root” of the covariance matrix, respectively, while $W(t)$ denotes a N -dimensional vector of independent Brownian motions. As before, without much loss of generality, we assume that $M(t) = 0$, although non-zero drifts, as relevant over longer return horizons, easily can be incorporated into the analysis by considering demeaned returns. We also assume that the asset returns are linearly independent, i.e., no

⁵³ This basic idea was pioneered by Diebold and Nerlove (1989) in their construction of a multivariate ARCH factor model, in which the latent time-varying volatility factors may be viewed as the base assets; see also Engle, Ng, and Rothschild (1990) and Alexander (2001).

redundant asset is included in the basic set of returns, implying that the covariance matrix $\Omega(t)$ is pd.⁵⁴

The natural multivariate extension of the realized variation measure, defined in (21), to the notion of a daily realized covariance matrix is simply

$$RCov_t(\Delta) \equiv \sum_{j=1}^{N(\Delta)} R_{t-1+j\Delta,\Delta} R'_{t-1+j\Delta,\Delta}, \quad (55)$$

where, as before, $N(\Delta) = 1/\Delta$. If, ideally, the price vector process in (54) is continuously observable, then letting Δ go to zero enables us to compute the realized covariance matrix in (55) at ever finer sampling intervals. In this scenario, the $RCov_t$ estimator converges to the integrated covariance matrix of the continuous-time stochastic volatility process on day t , given as,

$$ICov_t = \int_{t-1}^t \Omega(\tau) d\tau. \quad (56)$$

This expression, and the underlying limiting arguments, represent a direct extension of the notion of the integrated variance for $N = 1$ in Eqn (22).⁵⁵

Hence, as for the univariate case, the *true* ex-post covariance matrix becomes directly observable in this ideal setting, even in the absence of a model for $\Omega(t)$. The upshot is that, as before, variances *and* covariances no longer have to be extracted from a nonlinear model estimated via treacherous likelihood procedures, along the lines of the multivariate GARCH models discussed above. Instead, by treating the realized covariance matrices as realizations of the true underlying series of interest, we may apply standard time series techniques for their modeling and forecasting.

Of course, the idealized frictionless setting motivating the recipe for $RCov_t$ in (55), and its limit in (56), provide only an approximate description of reality. For instance, as discussed in Section 2.2, trades are not consummated continuously, imposing a strict upper bound on the highest possible sampling frequency. This presents important new implementation challenges compared to the univariate case, especially if the number of assets is large and the trading intensities of some assets are relatively low. In particular, while some of the techniques discussed earlier may be adapted for consistently estimating the individual elements of the covariance matrix in the presence of market micro-structure noise, none of these generally guarantee that the estimated covariance matrix is positive definite (pd), or even positive semi-definite (psd).

⁵⁴ As we discuss at length later, when the cross-section, N , is large, it can be difficult to generate unbiased estimates of the realized covariance matrix that satisfy this important constraint.

⁵⁵ For more formal development of the associated asymptotic distribution theory, see, e.g. Andersen et al. (2003a) and Barndorff-Nielsen and Shephard (2004a).

Along these lines, Andersen et al. (2003a) first noted that the simple realized covariance matrix in (55) will be pd by construction, as long as the asset returns are linearly independent and the trading (or quoting) activity is sufficiently high. The specific requirement is that price updates are available for the full cross-section of assets over small enough time increments, Δ , to ensure that the number of intraday observations, $N(\Delta) = 1/\Delta$, exceeds the number of assets, N . For example, if we sample individual US stocks every 5 min across the official trading day, the $RCov_t$ matrix is trivially singular if the number of stocks exceeds 78.

For a set of very actively traded securities, the above conditions may not appear unduly restrictive. After all, many assets trade multiple times each minute on average, often generating thousands of new trade prices per day. Unfortunately, this is deceptive. The key point is that *all* assets must have traded within each sampling interval. If not, this will generally result in a downward bias in the covariance estimates due to the presence of zero returns induced purely by the absence of trades (or quote changes)—a feature commonly labeled the Epps effect following the early characterization in Epps (1979). Since many assets periodically experience a trading lull, there will often be extended periods of no-trading for some of the assets, so that this can be a major concern. Hence, when using the basic realized covariance matrix estimator in (55), it is critical to sample fairly sparsely to alleviate this bias.⁵⁶ Of course, this then restricts the size of the cross-section that can be analyzed quite dramatically.

More generally, the price synchronicity requirement implies that the realized covariance matrix cannot be estimated consistently unless the sampling scheme is adapted to the trading intensity of the least active asset at any given time. This idea is encapsulated in the “refresh time” sampling procedure advocated by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011) as part of their multivariate realized kernel approach to covariance matrix estimation. The kernel consists of the inclusion of a suitably chosen weight function for the lead and lag returns in the computation of the covariance matrix. This ensures consistency in the presence of general classes of microstructure noise, while also guaranteeing that the estimate of the covariance matrix is psd.

Direct application of this approach is eminently feasible for a limited number of actively traded assets. However, when the number of assets is large, refresh time sampling results in a dramatic loss of data as intermediary prices for active assets are discarded until the last asset trades. For example, Hautsch, Kyj, and Oomen (2012) assess that, with realistic intra-stock differences in trade arrival rates, more than 90% of the data are discarded for a system of twenty actively traded assets, and the proportion

⁵⁶ It is generally also advantageous to follow the subsampling strategy previously outlined in Section 2.2, where one generates multiple subsamples of the intraday return series by initiating the sampling at the given frequency at different offsets relative to the opening trade, and then average the resulting covariance measures across the subsamples. For example, by initiating sampling at each of the first 5-min marks during the trading day, one could secure five distinct 5-min return series for each asset.

continues to rise as the cross-section of assets increases. This implies that, for N rising, the effective sampling frequency, $1/\Delta$, drops quite dramatically, in turn rendering it difficult to satisfy the positive definiteness bound. Equally problematic is the loss in estimation precision as each pairwise covariance term is computed from fewer and fewer intraday observations, ultimately producing a poorly estimated overall covariance matrix with many zeros among the eigenvalues. In sum, this strategy fails for very large cross-sections of assets.

Two main approaches have hitherto been proposed in the literature to accommodate large cross-sections, while avoiding dramatic Epps style biases. One avenue is to initially ignore the requirement of positive definiteness and apply the refresh sampling scheme on smaller blocks of assets, thus mitigating the problems associated with the loss of data, and then to apply a regularization procedure to restore the psd property. The second approach is to exploit covariance matrix factor structure to reduce the effective dimension of the problem, thereby allowing for more reliable estimates from a given set of intraday observations. We now discuss these techniques.

3.2.1 Regularizing Techniques for RCov Estimation

The simplest method for converting a “vast” $N \times N$ positive semi-definite covariance matrix estimator $RCov_t(\Delta)$ of less than full rank and possibly containing multiple zero eigenvalues, into a strictly positive definite matrix is shrinkage. The idea is to combine $RCov_t(\Delta)$ with an $N \times N$ shrinkage target matrix, Υ_t , which is positive definite and well conditioned. Ideally, the target should also provide a sensible benchmark covariance matrix to minimize the resulting bias. Formally,

$$\hat{\Omega}_t^S = \kappa RCov_t(\Delta) + (1 - \kappa)\Upsilon_t, \quad (57)$$

where the weight assigned to the realized covariance matrix satisfies $0 < \kappa < 1$, so the shrinkage estimator is a convex linear combination of a positive semi-definite and a positive definite matrix, implying it will be positive definite.

As an extraordinarily simple illustration of this basic principle, in a setting with daily data and time-varying covariance matrices, Ledoit and Wolf (2004) propose shrinkage toward the identity matrix, i.e., $\Upsilon_t = \mathcal{I}$, with the weight, κ , determined optimally according to an asymptotic quadratic loss function. While this will reduce the variance, it may, of course, induce a rather severe bias, as asset returns generally are highly correlated.

To counteract this bias, Ledoit and Wolf (2003) suggest shrinkage toward the covariance structure implied by a simple one-factor market model. Specifically, following the discussion in Section 3.1.2 above,

$$\Upsilon_t = \sigma_M^2 bb' + D_v, \quad (58)$$

where σ_M^2 refers to the variance of the market return, b denotes the $N \times 1$ vector of factor loadings for each of the assets with respect to the market portfolio, and D_v is a diagonal matrix composed of the corresponding idiosyncratic variances. Importantly, all of these parameters are easy to estimate from simple time series regressions.

In contrast to $\Upsilon_t = \mathcal{I}$, this procedure allows for non-trivial positive return correlation across assets, thus providing a more suitable shrinkage target for covariance estimation. However, it assumes that the relevant second-order return moments are time-invariant, so that a long time series of daily returns can be used for estimating b , along with the other parameters. This is counter to the spirit of high-frequency return-based estimation, where we seek to determine the time variation in the covariance matrix and, as an implication, the fluctuations in systematic market risk exposures, or factor loadings.⁵⁷ The extreme dichotomy between the realized covariance matrix, estimated without bias but with poor precision, and the shrinkage target, which may be strongly biased but is estimated with better precision, naturally suggest alternative approaches that better balance the two effects.

In this regard, Hautsch, Kyj, and Oomen (2012) have recently suggested breaking the covariance matrix into blocks according to the trading intensity of the underlying assets, thus minimizing the loss of data from refresh time sampling when using the multivariate realized kernels to estimate the different blocks. Of course, simply piecing the covariance matrix together from separate blocks generally produces an indefinite matrix with negative as well as positive eigenvalues. To circumvent this problem, Hautsch, Kyj, and Oomen (2012) adopt so-called eigenvalue cleaning to “regularize” the covariance matrix in a second step, by separating the set of large and significant eigenvalues from those that are statistically insignificant and may have been generated by random noise.⁵⁸

Specifically, denote the first stage realized kernel blocking estimator for the integrated covariance matrix on day t by $\hat{\Omega}_t$. Eigenvalue cleaning then consists of the following steps. First, define the realized correlation matrix by,

$$\hat{\Gamma}_t = \hat{D}_t^{-1} \hat{\Omega}_t \hat{D}_t^{-1}. \quad (59)$$

where, as for (45), $\hat{D}_t = \text{diag}(\hat{\Omega}_t)^{1/2}$ denotes the diagonal matrix of realized standard deviations. Using the conventional spectral decomposition, rewrite the correlation matrix as,

$$\hat{\Gamma}_t = \hat{P}_t \hat{\Lambda}_t \hat{P}_t', \quad (60)$$

⁵⁷ Again, Ledoit and Wolf (2003) envision their estimator to be applied for daily data but, as mentioned previously, there are recent attempts to adapt similar procedures to the high-frequency setting.

⁵⁸ This approach is motivated by random matrix theory; see, e.g. Mehta (1990) for an introduction to the theory and Tola, Lillo, Gallegati, and Mantegna (2008) for a recent application to portfolio choice.

where $\hat{\Lambda}_t$ is the diagonal matrix of eigenvalues, $\hat{\lambda}_i$, $i = 1, \dots, N$, sorted in descending order so that $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_N$, and \hat{P}_t denotes the orthonormal matrix of corresponding eigenvectors. Now, letting $\bar{\lambda}$ indicate the appropriate (positive) threshold for the significant eigenvalues, separate the first, say, k eigenvalues which exceed $\bar{\lambda}$ into one group. Next, equate all negative eigenvalues to zero and compute the average value, $\bar{\lambda}_t^B$, of the positive and (modified) zero eigenvalues that are less than $\bar{\lambda}$. The regularized covariance matrix is then constructed from the “cleaned” matrix of eigenvalues $\hat{\Lambda}_t^B$, with the original k eigenvalues as the first k diagonal elements and the remaining $N - k$ diagonal elements replaced by $\bar{\lambda}_t^B$, according to the formula

$$\hat{\Omega}_t^{RnB} = \hat{D}_t \hat{P}_t \hat{\Lambda}_t^B \hat{P}_t' \hat{D}_t. \quad (61)$$

Pursuing a similar approach, but taking the decomposition of the covariance matrix to a logical extreme, Lunde, Shephard, and Sheppard (2011) suggest estimating all covariance terms using only the corresponding bivariate realized kernel estimator. This minimizes the loss of information due to refresh time sampling, while permitting an optimal choice of kernel bandwidth for each pairwise return series. The first stage estimator is then obtained by assembling all the elements into a “composite realized kernel” covariance estimator. This heightens the quality of the estimate for each individual term, but it sacrifices the coherence of the overall matrix by not imposing the pd (or psd) property beyond the bivariate systems. Since the resulting composite covariance matrix typically will be “far” from pd, it requires a more substantial transformation of the entries in the covariance matrix to obtain a pd matrix than is the case for the RnB estimator of Hautsch, Kyj, and Oomen (2012), which usually operates with only 3–5 blocks.⁵⁹

Another closely related approach to the estimation of $RCov$, inspired by the idea of dimension reduction through the imposition of a factor structure, has also been suggested by Lunde, Shephard, and Sheppard (2011). The idea is to let the correlation structure be determined only by the eigenvectors associated with the largest and most significant eigenvalues. Again, the significant eigenvalues are identified day-by-day using the “i.i.d. noise threshold” prescribed by random matrix theory.⁶⁰ Formally, let

$$\tilde{\Gamma}_t = \tilde{P}_t \tilde{\Lambda}_t \tilde{P}_t', \quad (62)$$

⁵⁹ The notion of a distance between covariance matrices requires the adoption of a matrix norm. Since our discussion is heuristic, we abstain from any detailed account; see, e.g. Fan, Fan, and Lv (2008) for a discussion of alternative norms in the context of covariance matrix estimation.

⁶⁰ Alternatively, one may exploit an initial procedure to help decide on an appropriate fixed number of eigenvectors, or “factors”, in order to maintain a constant dimensionality of the correlation structure across days.

where $\tilde{\Lambda}_t$ denotes the $k \times k$ diagonal matrix containing the upper left $k \times k$ sub-matrix of $\hat{\Lambda}_t$, while \tilde{P}_t denotes the $N \times k$ matrix containing the first k columns of eigenvectors from \hat{P}_t associated with the largest k eigenvalues. The resulting $N \times N$ matrix, $\tilde{\Gamma}_t$, is of rank k and thus not strictly positive definite. It is also not a proper correlation matrix, as it generally fails to have unit entries along the diagonal. Nonetheless, it embodies the correlation structure implied by the k most important eigenvectors, or the first k principal components of the intraday returns. Hence, it is natural to modify this matrix to construct a proper correlation matrix,⁶¹

$$\tilde{\Gamma}_t^{PC} = \mathcal{I} + \left[\tilde{P}_t \tilde{\Lambda}_t \tilde{P}_t' - \text{diag} \left(\tilde{P}_t \tilde{\Lambda}_t \tilde{P}_t' \right) \right]. \quad (63)$$

The resulting principal component regularized realized covariance matrix estimator is then obtained by simply scaling up $\tilde{\Gamma}_t^{PC}$,

$$\hat{\Omega}_t^{PC} = \hat{D}_t \tilde{\Gamma}_t^{PC} \hat{D}_t \quad (64)$$

as in (61).

It remains a matter for future work to systematically characterize the performance of these approaches to $RCov_t$ estimation based on the spectral decomposition in (60) for empirically realistic situations involving different scenarios for the number of included assets and the trading (quoting) intensities.

Rather than extracting principal components day-by-day to obtain a factor structure for the realized covariance matrix, a number of authors propose using pre-specified observable factors, or returns on factor mimicking portfolios, as a way to reduce the dimensionality of the problem and the associated estimation errors.⁶²

Recall the basic linear factor structure in (51), where the parameters are assumed to be constant across days. Extending the corresponding expression for the discrete-time returns on the factors in (50) to a continuous-time setting, maintaining the same diffusion representation for the logarithmic factor price process as for the returns in (54), we may write,

$$dP_F(t) = \Omega_F(t)^{1/2} dW_F(t), \quad (65)$$

where $\Omega_F(t)^{1/2}$ denotes the $N_F \times N_F$ positive definite “square-root” of the instantaneous covariance matrix, and $W_F(t)$ is a N_F -dimensional vector of independent

⁶¹ Notice that for any square matrix A , the operation $A - \text{diag}(A)$ leaves the off-diagonal entries in A unchanged, while producing zeros along the diagonal. Hence, $\mathcal{I} + [A - \text{diag}(A)]$ yields a matrix with unit entries on the diagonal and off-diagonal entries inherited from A .

⁶² Fan, Fan, and Lv (2008) provide a formal theoretical analysis of the impact of dimensionality on the estimation of covariance matrices in the context of factor models.

Brownian motions. Denoting the resulting day t realized covariance matrix for the factors by $\hat{\Omega}_{F,t}$, an implied day-by-day realized covariance matrix estimator for the N -dimensional vector of returns may then be constructed as,

$$\hat{\Omega}_t^F = \hat{B} \hat{\Omega}_{F,t} \hat{B}' + \hat{D}_v, \quad (66)$$

where \hat{B} and \hat{D}_v refer to estimates of the factor loadings and the (diagonal) covariance matrix for the idiosyncratic variances, respectively.

This approach has been successfully implemented by Bannouh, Oomen, and van Dijk (2010) for the estimation of large dimensional daily covariance matrices for hundreds of individual stocks. Bannouh et al. (2010) rely on a set of highly liquid exchange traded funds (ETFs) as factors. Prices for these contracts are essentially free of microstructure noise at relatively high frequencies, allowing for accurate estimation of $\hat{\Omega}_{F,t}$. In contrast, they estimate the factor loadings from daily data to avoid biases due to microstructure and Epps type effects. An even simpler approach would be to rely on the market model, effectively setting $\kappa = 0$ in the earlier (57) and (58) for the shrinkage estimator, thereby only exploiting the realized return variation of the market index as the single dynamic factor driving the covariance matrix in accordance with (66).

Of course, as already noted in Section 3.1.2, the restriction that the covariance matrix of the idiosyncratic returns is diagonal is rather strong. For example, it precludes sector specific effects. In an effort to relax this assumption, Fan, Liao, and Mincheva (2011) allow for some correlation in the error covariance matrix by imposing the weaker requirement that the matrix is “sparse”. Their estimation procedure exploits random matrix theory as they achieve the requisite parsimony, or sparsity, in the idiosyncratic covariance matrix via so-called thresholding techniques.⁶³

The assumption that the factor loadings are constant may, of course, also be problematic in some situations. Just as high-frequency data for the factors may be used in accurately estimating $\hat{\Omega}_{F,t}$, high-frequency data for the factors *and* the returns could similarly be used in the estimation of day-by-day realized factor loadings, or “betas”. This idea for the estimation of daily realized factor loadings from intraday data was first pursued empirically by Bollerslev and Zhang (2003) and Andersen, Bollerslev, Diebold, and Wu (2006b) for the three Fama–French portfolios and the market, respectively.⁶⁴

From a practical perspective, however, the estimation of the loadings runs into the exact same market microstructure problems that plague the original $RCov_t$ estimator: it is difficult to implement with illiquid assets and the large dimensions typically required for asset-level risk analysis. These difficulties may, of course, be partly overcome by resorting

⁶³ Related banding and thresholding procedures for estimating daily realized covariance matrices are discussed in Wang and Zou (2010).

⁶⁴ Estimation and forecasting of betas based on high-frequency data have also been explored more recently within the Realized GARCH framework by Hansen, Lunde, and Voev (2010).

to some of the techniques already outlined above. This mainly involves suitably combining the different procedures, and we abstain from fleshing out the details. Hence, instead of further discussion of techniques for measuring the current realized covariance matrix, we now turn to different dynamic models for forecasting realized covariance matrices.

3.2.2 *Dynamic Modeling of Realized Covariance Matrices*

All of the different procedures discussed in the preceding section for estimating the realized covariance matrix may in principle be applied as short-term daily forecasts as well, when augmented with a martingale assumption for the realized covariance matrix, e.g. tomorrow's expected covariance matrix equals today's realization.⁶⁵ Of course, the martingale hypothesis is at best a short-term approximation, as both variances and covariances generally display mean reversion. Hence, for longer horizons explicit time series models must be developed as a basis for sensible forecasts.

Building on the univariate procedures discussed earlier, this section outlines various strategies for modeling and forecasting integrated covariance matrices, treating the realized covariance matrix as directly observable, albeit with some measurement error. Since the literature on the estimation of large realized covariance matrices is recent and remains limited, there are still no authoritative studies of the relative performance of different approaches.⁶⁶ Consequently, our review of existing techniques is invariably somewhat speculative. However, we anticipate this to be an area where substantial progress will be made over the coming years, and therefore summarizes what we see as some of the more promising new directions.

In parallel to the notation for the variance forecasts discussed earlier, we denote the $N \times N$ point forecast of the integrated return covariance matrix for period $t + k$ based on information through period t , by $\hat{\Omega}_{t+k|t}$, while the corresponding measures for the realized covariance matrix in period t is generically labeled $\hat{\Omega}_t$.⁶⁷ Just as many of the forecasting models for the realized volatilities discussed in Section 2.2 were directly

⁶⁵ Both Hautsch, Kyj, and Oomen (2012) and Lunde, Shephard, and Sheppard (2011) base their exploration of one-day-ahead covariance matrix forecasts on this hypothesis.

⁶⁶ The set of potential applications is literally unlimited, thus making it hard to settle on a simple metric for assessing the economic value of improved forecasts, even if one focuses on practical risk measurement and management problems. An early study inspiring this literature is Fleming, Kirby, and Ostdiek (2003), who suggest dramatic improvements vis-a-vis the RM and multivariate GARCH frameworks for standard mean-variance efficient asset allocation problems.

⁶⁷ Of course, as discussed in the previous section, there are many alternative proposals for estimating Ω_t and associated procedures for forecasting it, so $\hat{\Omega}_t$ and $\hat{\Omega}_{t+k|t}$ merely serve as generic indicators for the realized covariance measure and forecast being entertained at a given point in the exposition. We reserve the more specific notation, $RCov_t(\Delta)$, for the standard realized covariance estimator based on the cross-product of returns sampled at fixed frequency Δ . Also, as in the univariate case, the models will typically stipulate a specific dynamic evolution for Ω_t , whereas any empirical analysis will be based on the time series of observed $\hat{\Omega}_t$.

inspired by existing techniques for forecasting with daily or lower frequency data, so are many of the procedures for dynamic realized covariance matrix modeling.

In particular, directly emulating the RiskMetrics approach in (40), it is natural to postulate,

$$\hat{\Omega}_{t+1|t} = \lambda \hat{\Omega}_{t|t-1} + (1 - \lambda) \hat{\Omega}_t, \quad (67)$$

where $0 < \lambda < 1$. Thus, the integrated covariance matrix forecast is generated as an exponentially weighted average of past realized covariance matrix measures with λ controlling the relative weight ascribed to the more recent realizations.⁶⁸ Intuitively, this allows for persistent time-variation in the realized covariance matrices, while implicitly acknowledging that each realization is measured with error. Of course, this approach also inherits all of the problems with the conventional RM approach, including the lack of mean reversion, and as such may not be appropriate for longer forecast horizons.

Alternatively, mimicking the scalar diagonal AR(1) model in (23) suggests the following multivariate regression specification,

$$\text{vech}(\hat{\Omega}_{t+1}) = \text{vech}(C) + \beta \text{vech}(\hat{\Omega}_t) + \xi_{t+1}, \quad (68)$$

where the $N(N + 1)/2 \times 1$ vector ξ_t denotes an error term. This system requires nothing but OLS to implement, and conditional on the estimated parameters, \hat{C} and $\hat{\beta}$, the forecast for the integrated covariance matrix is readily obtained from,

$$\text{vech}(\hat{\Omega}_{t+1|t}) = \text{vech}(\hat{C}) + \hat{\beta} \text{vech}(\hat{\Omega}_t). \quad (69)$$

Strict positive definiteness of the covariance matrix forecast in (69) is guaranteed for any pd matrix \hat{C} and positive values of $\hat{\beta}$, as long as $\hat{\Omega}_t$ is psd.

Even though the above procedure generalizes the “martingale” hypothesis, corresponding to $C = 0$ and $\beta = 1$, it still assumes a common degree of mean reversion across all variances and covariances. As noted previously, this is likely overly restrictive, especially when considering a diverse set of assets, so it is worthwhile contemplating suitable generalizations.

Pushing the above approach one step further, any of the other procedures discussed in Section 3.1 could be similarly adapted to modeling realized covariances, keeping in mind the restrictions required for positive definiteness. For example, the DCC-type framework naturally suggests first modeling the realized standard deviations asset-by-asset using any of the procedures discussed in Section 2.2, and the corresponding

⁶⁸ This particular procedure is among the set of dynamic specifications explored by e.g. Fleming, Kirby, and Ostdiek (2003), Liu (2009), Bannouh et al. (2010), and Varneskov and Voev (2010).

realized correlations in a second step. Specifically, maintaining a simple dynamic structure as in (68), the dynamics of the correlations for the standardized returns could be modeled as,

$$\text{vech}(\hat{Q}_t) = \text{vech}(C) + \beta \text{vech}(\hat{Q}_{t-1}) + \xi_t, \quad (70)$$

where we have extended the notation for the conventional DCC model in the obvious way. Again, simple OLS is all that is required for estimation. As for the conventional DCC model, an additional normalization along the lines of (47) is needed to ensure that the resulting correlation matrix forecast is well defined, with ones along the diagonal and all of the off-diagonal elements falling between -1 and 1 .

The advantages of these approaches are twofold. First, high-frequency information is used to obtain more precise estimates of current variances and covariances, in turn resulting in better “initial conditions” for forecast calculations. Second, by treating the covariance matrices as directly observable no numerical optimization is needed for the estimation of the models.

Even though we have focused on simple first-order models and corresponding one-day-ahead forecasts, all the procedures discussed above could easily be iterated forward to generate multi-period forecasts $\hat{\Omega}_{t+k|t}$. More complicated long-memory dynamics, regime-switching, or asymmetries, could also be incorporated into the models, provided the dimensionality of the estimation problem is kept in check.

A major obstacle for adopting more realistic and complex representations for the realized covariance matrix dynamics than offered by, e.g. (68) is, as discussed at length previously, the requirement for positive definiteness. A possible solution consists of first applying a nonlinear transform to the $RCov_t$ matrix with the property that the inverse transform will ensure positive definiteness. One may then specify and estimate the dynamics of the transformed system without imposing any constraints. Once the future expected value of the transformed system is determined, the inversion back into a covariance representation automatically produces a pd matrix forecast. A popular example of this approach within the univariate setting is the specification of dynamic models for log volatility, as in the EGARCH and log-HAR-RV models discussed in Sections 2.1 and 2.2, respectively.

In this regard, Andersen et al. (2003a) proposed modeling the Cholesky decomposition of $RCov$ rather than the matrix itself. The Cholesky decomposition provides one possible definition of a unique square-root of a positive definite realized covariance matrix estimator,

$$\hat{\Omega}_t = L_t L_t', \quad (71)$$

where L_t is a unique lower triangular matrix. The data vector subjected to dynamic modeling is then $\text{vech}(L_t)$, and one simply substitute the forecast of $\text{vech}(L_{t+k})$ back

into (71) to construct a forecast of $\Omega_{t+k,\Delta}$.⁶⁹ One drawback to the use of Cholesky decompositions, and other nonlinear transformations, is that the estimated parameters can be difficult to interpret in terms of the marginal impacts of shocks to specific elements in the covariance matrix. Related to this, the dynamic Cholesky modeling strategy inevitable involves a bias, arising from modeling and forecasting a nonlinear transformation and then mapping the resulting point forecasts back into the covariance matrix.⁷⁰

Another strategy, proposed by Bauer and Vorkink (2011), is to exploit the matrix logarithmic function.⁷¹ Specifically, provided that $\hat{\Omega}_t$ is positive definite, then the $N \times N$ symmetric matrix,

$$A_t = \log m \left(\hat{\Omega}_t \right), \quad (72)$$

is implicitly defined by the inverse of the matrix exponential function,

$$\hat{\Omega}_t = \sum_{n=0}^{\infty} \frac{1}{n!} A_t^n. \quad (73)$$

One may then proceed as before by specifying the dynamics of $\text{vech}(A_t)$, estimating the system and constructing the implied $\hat{\Omega}_{t+k|t}$ forecasts. Of course, the dynamic specification for $\text{vech}(A_t)$ must be kept relatively simple to remain tractable in large dimensions.⁷² Also, the same general problems arising from the use of a nonlinear transformation in the Cholesky decomposition discussed above remain for the A_t to $\hat{\Omega}_t$ transformation.

In summary, while the literature on modeling the covariance matrix dynamics is progressing rapidly along many different directions, there is still no consensus on the relative merits of the approaches. It is clear, however, that the use of high-frequency intraday data and realized covariance measures hold the promise of substantially improving the accuracy of covariance matrix forecasting. Going one step further, in direct parallel to the approach taken in the univariate setting of Section 2.2.3, the realized covariance forecasts discussed above may also be embedded within a multivariate GARCH setting to provide a vehicle

⁶⁹ Building on this framework, Chiriac and Voev (2011) explore various dynamic specifications of the realized covariance matrix for six liquid US stocks, and find that a long-memory vector ARFIMA model performs well. The reliance on approximate maximum likelihood estimation renders their approach problematic for large scale systems, but it should be feasible to adopt simpler specifications that would enable estimation when N is large.

⁷⁰ The aforementioned study by Chiriac and Voev (2011) also provides approximate bias correction terms for this, but deem the extent of the bias to be relatively minor in their empirical application.

⁷¹ A related multivariate matrix Exponential GARCH model was proposed by Kawakatsu (2006).

⁷² The actual application in Bauer and Vorkink (2011) is relatively modest in terms of dimensionality, and too highly parameterized to be practical for high-dimensional applications.

for combining the realized covariance matrices with a multivariate distribution for the return innovations. We briefly discuss some recent ideas for implementing this next.

3.2.3 Combining GARCH and RCov

As with the univariate setting, it is tempting to combine the precision of high-frequency realized volatility based measures with the powerful and flexible econometric tools provided by (quasi) likelihood estimation of GARCH models in extracting the volatility dynamics for multivariate systems. This can be done in a variety of ways, especially if one breaks the approach down into multiple steps. Nonetheless, the literature dealing with this approach remains nascent and we have little evidence regarding the relative performance of alternative procedures, so we only briefly illustrate how these methods may be combined to construct candidate models with non-trivial dynamic covariance structures through a couple of examples.

First, it is natural to exploit the various techniques for estimation of the realized correlation matrix, discussed in the initial parts of Section 3, with the flexible dynamic modeling of the individual conditional variances afforded by GARCH style models. Recall the decomposition in (45), $\Omega_t = D_t \Gamma_t D_t$. The diagonal conditional standard deviation matrix, D_t , may be obtained from univariate models, each estimated in isolation using flexible dynamic specifications. When high-frequency data are available, the candidate univariate volatility models include the GARCH-X and Realized GARCH techniques reviewed in Section 2.3.3. These approaches ensure volatility dynamics that quickly respond to changes in the underlying realized volatility measures and provide a great deal of freedom in adapting the estimation to accommodate critical features of each specific series, including asymmetric return-volatility relations, long-memory dynamic dependencies, calendar effects, and the degree of heavy tails in the return distributions.

The conditional correlation matrix, Γ_t , also changes over time, but it is likely to evolve more slowly than the conditional variances. As such, one may exploit wider estimation windows to enhance the precision of estimation. Technically, one may simply stipulate a constant correlation matrix, $\Gamma_t = \Gamma$, for a period of one week or one month, say, but allow this constant matrix to be estimated over a rolling window so that it does evolve slowly over time. The longer time series allows for additional flexibility in estimating the realized correlation matrix, even for a very large set of assets, using the various techniques discussed in the previous sections. The candidate procedures for estimating Γ , include the basic $RCov_t$ estimator using appropriately sparse sampling frequencies, the shrinkage estimators, or the various techniques exploiting regularization via principal components, observable factor structures, thresholding, and blocking.⁷³ Clearly, the potential for developing alternative approaches along these lines is vast

⁷³ One example of applying such procedures is Rosenow (2008) although he only applies the procedures for daily data.

and we currently have only limited knowledge about the relevant empirical tradeoffs that will govern the success of the different techniques.

Second, we briefly discuss a proposal that directly combines realized covariance measures with GARCH style dynamics, namely the multivariate HEAVY model of Noureldin, Shephard, and Sheppard (in press), which extends the univariate specification in (37). In the general form, the model inherits the curse of dimensionality from multivariate GARCH representations, so the empirical work focuses on parsimonious, and restrictive, representations. The model is explicitly designed for the low-frequency (daily) realized return cross-product, but the information set is given by corresponding high-frequency observations. Denoting the realized daily return cross-product by U_t , the model may be defined as follows,

$$U_t = R_t R_t' = H_t^{1/2} \Xi_t \left(H_t^{1/2} \right)', \quad (74)$$

where the $N \times N$ matrix H_t denotes the covariance matrix of the daily return vector conditional on an information set including the high-frequency returns up to day t , while Ξ_t is a $N \times N$ symmetric innovation matrix with $E_{t-1}[\Xi_t] = \mathcal{I}$.

Forecasting the covariance matrix requires a dynamic model for H_t . One tractable option is the scalar HEAVY parametrization, which is well defined subject to regularity conditions resembling those from the scalar multivariate GARCH model,

$$H_{t+1} = C_H C_H' + b_H H_t + a_H V_t. \quad (75)$$

Here, a_H and b_H are positive scalars, C_H is a $N \times N$ matrix of constants, which may be fixed by covariance targeting, and V_t denotes a realized covariance measure, such as, e.g. the realized covariance matrix based on 5-min sampling.

Equation (75) allows for one-step-ahead forecasting, but multi-step forecasting requires an explicit representation of the dynamics for V_t as well. Letting $M_t = E_{t-1}[V_t]$, the evolution for V_t is stipulated to follow,

$$V_t = M_t^{1/2} \Psi_t M_t^{1/2}, \quad (76)$$

where the Ψ_t is a $N \times N$ symmetric innovation matrix with $E_{t-1}[\Psi_t] = \mathcal{I}$. The associated dynamic representation for M_t is analogous to the scalar GARCH style specification of (75), and directly generalizes (37),

$$M_{t+1} = C_M C_M' + b_M M_t + a_M V_t. \quad (77)$$

With covariance matrix targeting, the scalar HEAVY system may be estimated by standard likelihood techniques once we provide a conditional distribution for the stochastic

shocks to the system. In particular, if the return innovations are i.i.d. Gaussian, the innovation matrix, Ξ_t , in (74) will be Wishart distributed. Likewise, one may assume Ψ_t in (76) to be Wishart distributed.

In parallel to the univariate literature, Noureldin, Shephard, and Sheppard (in press) find the inclusion of the high-frequency return information to provide significant improvements over corresponding GARCH models utilizing only daily return observations. The upshot is that generalizations of multivariate GARCH models into settings that accommodate the inclusion of high-frequency data appear to provide a similar boost to the predictive performance that was observed in the univariate case. Obviously, the models still impose quite unsatisfactory constraints on the dynamic evolution of the system as well as the conditional return innovations, rendering further tractable extensions to the framework important objectives for future work.

In summary, the opportunities for combining factor structures, multiple components, GARCH modeling approaches, and realized covariance measures in distinct ways are nearly unlimited. The literature is progressing in different directions, but we lack consensus on how to assess and rank the performance of alternative procedures. Moreover, it is evident that the focus on the covariance matrix fails to explicitly incorporate features of the return distribution beyond the second moments, which are potentially critical for active risk management. We now turn to such issues.

3.3 Modeling Multivariate Return Distributions

Just as a fully specified and realistic univariate distribution is needed for risk measurement, so too is a fully specified and realistic multivariate (non-Gaussian) distribution needed for risk management. For example, a fully specified multivariate distribution allows for the computation of VaR sensitivities and VaR minimizing portfolio weights.

The results of Andersen et al. (2000a) 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 log-normal mixture model discussed in Section 2.3.2 could therefore be developed.

As discussed at length above, however, construction and use of realized covariance matrices may be problematic in situations when liquidity is not high. In that situation one of the more traditional parametric GARCH type models discussed in Section 3.1 may be used for modeling the temporal dependencies in the conditional covariance matrix and then combined with an explicit (and by assumption time-invariant) multivariate distribution for the standardized returns.

Specifically, assuming the mean to be zero, or $M_t = 0$, we have from (41),

$$Z_t = \Omega_t^{-1/2} R_t, \quad Z_t \sim \text{i.i.d.}, \quad E_{t-1}(Z_t) = 0 \quad \text{Var}_{t-1}(Z_t) = \mathcal{I}, \quad (78)$$

Alternatively, recalling the decomposition in (45), it is sometimes more convenient to consider the vector of standardized, but correlated asset shocks

$$e_t = D_t^{-1}R_t, \quad E_{t-1}(e_t) = 0, \quad \text{Var}_{t-1}(e_t) = \Gamma_t, \quad (79)$$

where D_t denotes the diagonal matrix of conditional standard deviations for each of the assets, and Γ_t refers to the potentially time-varying conditional correlation matrix.

For concreteness, we focus on the DCC type decomposition in (79) and express the return distributions below in terms of e_t . As discussed in Section 3.1.1, this is often more convenient in large dimensions, but the same general ideas apply for the basic decomposition in (78) and distributions expressed in terms of Z_t .

3.3.1 Multivariate Parametric Distributions

The normal distribution is convenient and tempting (but dangerous) to use. It implies that aggregate portfolio returns are also conditionally normally distributed. The multivariate normal density has the simple form

$$f(e_t) = C(\Gamma_t) \exp\left(-\frac{1}{2}e_t' \Gamma_t^{-1} e_t\right), \quad (80)$$

where the $C(\Gamma_t)$ normalization factor ensures that the density integrates to one. The multivariate normal distribution, however, typically does not provide an accurate picture of tail risk. In parallel to our earlier discussion of univariate return distributions, several multivariate distributions have been proposed to remedy this deficiency.

Especially prominent among these is the multivariate Student's t -distribution first employed in this context by Harvey, Ruiz, and Sentana (1992); see also the more recent work by Glasserman, Heidelberger, and Shahabuddin (2002). The multivariate standardized symmetric t -distribution with correlation matrix Γ_t has the following density

$$f(e_t) = C(d, \Gamma_t) \left(1 + \frac{e_t' \Gamma_t^{-1} e_t}{(d-2)}\right)^{-(d+N)/2}, \quad (81)$$

where $C(d, \Gamma_t)$ again ensures that the density integrates to one. The $d > 2$ scalar parameter determines the degree of leptokurtosis in the distribution. When d goes to infinity the power-form of the t -distribution converges to an exponential function and the multivariate normal distribution emerges in the limit. Unlike the normal distribution, the multivariate t -distribution allows for nonlinear tail dependence between assets. It does so in a symmetric fashion, however. It cannot accommodate two assets having a higher probability of a large joint down move than a joint up move of the same magnitude.

The asymmetric t -distribution employed by Demarta and McNeil (2005) allows for more flexibility. Let ξ denote an $N \times 1$ vector of “asymmetry parameters”. The density for the standardized asymmetric t -distribution may then be expressed as

$$f(e_t) = \frac{C(d, \dot{\Gamma}_t) K_{\frac{d+N}{2}} \left(\sqrt{(d + (e_t - \dot{\mu})' \dot{\Gamma}_t^{-1} (e_t - \dot{\mu})) \xi' \dot{\Gamma}_t^{-1} \xi} \right) \exp((e_t - \dot{\mu})' \dot{\Gamma}_t^{-1} \xi)}{\left(1 + \frac{(e_t - \dot{\mu})' \dot{\Gamma}_t^{-1} (e_t - \dot{\mu})}{d} \right)^{\frac{(d+N)}{2}} \left(\sqrt{(d + (e_t - \dot{\mu})' \dot{\Gamma}_t^{-1} (e_t - \dot{\mu})) \xi' \dot{\Gamma}_t^{-1} \xi} \right)^{-\frac{(d+N)}{2}}} \quad (82)$$

where $K_{\frac{d+N}{2}}(\cdot)$ denotes the modified Bessel function of the third kind,

$$\dot{\mu} = -\frac{d}{d-2} \xi, \quad \dot{\Gamma}_t = \frac{d-2}{d} \left(\Gamma_t - \frac{2d^2}{(d-2)^2 (d-4)} \xi \xi' \right),$$

and $C(d, \dot{\Gamma}_t)$ is another normalization factor. The definitions of $\dot{\mu}$ and $\dot{\Gamma}_t$ ensure that the vector of standardized return shocks, e_t , has mean zero and correlation matrix Γ_t . Note that for $\xi = 0$ and the absence of any asymmetries, we have $\dot{\mu} = 0$ and $\dot{\Gamma}_t = \Gamma_t$. The asymmetric t -distribution therefore nests the symmetric t -distribution as a special case.

While the asymmetric t -distribution is more flexible than the symmetric t , it requires that the N asymmetry parameters in ξ be estimated simultaneously with the other parameters of the model. This becomes quite challenging in large dimensions. Instead copula methods sometimes provide a more flexible approach by allowing the univariate and distinctly multivariate distributional aspects to be specified in two separate steps.

3.3.2 Copula Methods

Much attention in risk management has focused on the construction of multivariate densities from the marginal densities via copulas, as in, for example, Li (2000), Jondeau and Rockinger (2006), Patton (2006), Rosenberg and Schuermann (2006), Creal, Koopman, and Lucas (2011), and Hafner and Manner (2012). We will not attempt an exhaustive review of this extensive literature here, referring instead to the in-depth treatment in McNeil, Frey, and Embrechts (2005).

The central result in copula theory is Sklar’s theorem. The theorem states that for a very general class of multivariate distribution functions, say $F(e)$, with marginal distributions $F_1(e_1), \dots, F_N(e_N)$, there exists a unique copula $G(\cdot)$ linking the marginals to the joint distribution

$$F(e) = G(F_1(e_1), \dots, F_N(e_N)) \equiv G(u_1, \dots, u_N) \equiv G(u), \quad (83)$$

where the $N \times 1$ vector, u , is defined via the N marginals. In turn, this implies that the multivariate density may be expressed as

$$f(e) = \frac{\partial^N G(F_1(e_1), \dots, F_N(e_N))}{\partial e_1 \dots \partial e_N} = g(u) \times \prod_{i=1}^N f_i(e_i). \quad (84)$$

The resulting log-likelihood function for a sample of size T therefore naturally decomposes into two separate sums

$$\log L = \sum_{t=1}^T \log g(u_t) + \sum_{t=1}^T \sum_{i=1}^N \log f_i(e_{i,t}). \quad (85)$$

This offers a potentially powerful framework for risk model builders by allowing the modeling of the marginal densities, corresponding to the second double summation, to be separated from the modeling of the copula function appearing in the first summation.⁷⁴

Of course, in order to actually implement this approach, we need to specify the copula function $g(\cdot)$. The most commonly employed copula is constructed from the multivariate normal distribution. It may be succinctly expressed as

$$g(u_t; \Gamma_t^*) = |\Gamma_t^*|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \Phi^{-1}(u_t)' (\Gamma_t^{*-1} - I) \Phi^{-1}(u_t) \right\}, \quad (86)$$

where $\Phi^{-1}(u_t)$ refers to the $N \times 1$ vector of standard inverse univariate normals, and the correlation matrix Γ_t^* pertains to the $N \times 1$ vector e_t^* with typical element,

$$e_{i,t}^* = \Phi^{-1}(u_{i,t}) = \Phi^{-1}(F_i(e_{i,t})). \quad (87)$$

The normal copula has the advantage that it is relatively easy to work with. However, even though it is more flexible than the standard multivariate normal distribution, for many financial risk applications it does not allow for sufficient dependence between tail events.

To remedy this an alternative copula model can be built from the multivariate t -distribution. The resulting t -copula allows for tail dependence between the marginal probabilities $u_{i,t}$ but only in a symmetric fashion. Going one step further, an asymmetric t -copula may also be developed from the asymmetric multivariate t -distribution discussed above. From a practical modeling perspective, t -copula models have the potential

⁷⁴ Note, this implicitly assumes that the copula function $g(\cdot)$ is constant through time. Although fundamentally different, this parallels the assumption of a time-invariant multivariate distribution $f(\cdot)$ for the standardized returns underlying the discussion in Section 3.3.1.

to break the curse of dimensionality, which is otherwise unavoidable in multivariate t -distributions when N is large. In particular, while the asymmetric t distribution in (82) requires the simultaneous estimation of ξ and d , amounting to a total of $N + 1$ parameters, when using the asymmetric t -copula instead, it is possible to separately estimate each of the N marginal distributions allowing for asset specific distributional features.⁷⁵ The marginal distributions may then be “tied” together using an asymmetric t -copula with only two parameters: a scalar copula d_G and a scalar copula asymmetry parameter ξ_G . This approach has successfully been implemented by Christoffersen et al. (2011a).

Many other classes of copula functions exist as well. Most of these, however, including the popular Gumbel and Clayton classes, are not yet operational in high dimensions. An intriguing approach to overcoming this general dimensionality problem has recently been suggested by Oh and Patton (2011), who recommend relying on a latent factor structure for the copula. Fully efficient estimation of this new class of models is complicated by the lack of closed-form expression for the likelihood function but it is relatively easy to do via simulation-based procedures that match appropriate rank statistics. Oh and Patton (2011) find that this new approach works well in an application involving one hundred individual stocks.⁷⁶ It is too early to tell how widely applicable this copula-factor structure is.

3.3.3 Combining GARCH and RCov

Another approach for obtaining full-fledged multivariate conditional return distributions is to combine the realized covariance measures and GARCH style dynamic specifications with specific distributional assumptions, along the lines of the procedures discussed in Section 3.2.3 where the innovation distributions were specified mostly to ensure tractable (quasi-likelihood) estimation of the underlying dynamic model parameters.

For example, if the distributions adopted for each of the univariate return innovation series in the GARCH specifications for the individual components of D_t in the DCC-style decomposition in (45) are taken as exact representations of the data generating process, this in principle defines a conditional one-step-ahead return distribution given the estimated (and assumed to be constant) realized correlation matrix. However, this is only tractable if simple, and restrictive, distributional assumptions are imposed. Typically, this implies resorting to a multivariate normal or student t -distribution for the return innovation vector. This severely limits the complexity and realism in modeling the individual return innovations and volatilities. Short-term multi-horizon forecasts may be similarly obtained, if one stipulates that the correlation matrix remains constant. For longer

⁷⁵ Of course, the need to estimate the $N \times N$ correlation matrix Γ_t further confound the estimation problem.

⁷⁶ Their actual estimation results also suggest significant tail dependencies for most of the individual stocks in their sample, with the degree of tail dependence being stronger in crashes than booms.

horizons, however, the dynamics of the realized correlation matrix would need to be modeled separately. In that situation the system quickly becomes analytically untractable, and simulation techniques are required for obtaining the multi-horizon density forecasts.

Another possible route involves the HEAVY model introduced in (74)–(77). Assuming both multivariate innovation distributions are truly Wishart, as discussed in Section 3.2.3, the model naturally delivers a complete characterization of the one-step-ahead joint return distribution. The multi-horizon density forecasts must again rely on Monte Carlo procedures.

As an alternative to these GARCH representations, there has recently been an upsurge in work on related multivariate stochastic volatility models. These specifications generalize GARCH models in the sense that the dynamics of the volatility process is governed by independent random shocks rather than a deterministic function of the return innovations. The models tend to be heavily parametric but they may, under appropriate simplifying assumptions, be combined with realized covariance matrix measures.⁷⁷ These models typically exploit Gaussian assumptions for the return and volatility (square-root covariance matrix) innovations as they produce “squares” that are Wishart distributed and thus known in closed form.⁷⁸

The additive component Wishart-RCOV-A(K) model in Jin and Maheu (2010) provides an interesting example of combining such stochastic volatility representations with realized measures, by exploiting features akin to a multivariate HAR-RV model for the individual components of the realized covariance matrix. Although the empirical results appear promising, the parametric assumptions remain somewhat restrictive and estimation must be performed via Bayesian techniques using Markov Chain Monte Carlo (MCMC) procedures that are tractable only for moderately sized systems.

To summarize, the work on incorporating time-varying realized covariance measures within the multivariate GARCH and related stochastic volatility model setting is in its infancy. Given the need for tractability, the existing procedures invoke overly simplistic distributional assumptions, rendering the multi-horizon density forecasts unable to fully account for critical features such as pronounced return-volatility asymmetries, the possibility of jumps, long-memory style volatility dynamics, and extreme correlations in down markets. For the time being, such features are more readily portrayed through the design of appropriate simulation methods.

⁷⁷ Among the initial contributions in this area are Philipov and Glickman (2006), who specify a standard Wishart transition density for the inverse covariance matrix of daily returns, as well as Gouriéroux, Jasiak, and Sufana (2009) who introduce the Wishart autoregressive model for daily data. Extensions of these models that involve realized covariance measures have been developed by, e.g. Bonato Caporin and Rinaldo (2009); Golosnoy, Glibisch, and Liesenfeld (2010); and Asai and So (2010).

⁷⁸ The Wishart distribution provides the matrix generalization of a “squared” normal distribution, i.e., just as the sum of squared i.i.d. normal variates are χ^2 distributed, the sampling distribution of the sample covariance matrix for draws from the multivariate normal distribution is Wishart.

3.3.4 Multivariate Simulation Methods

The multivariate normal distribution implies normally distributed portfolio returns so that the *VaR*, *ES* and most other risk measures are easily computed analytically. When using non-normal distributions, or any kind of copula, portfolio *VaR*, and *ES* must instead be computed via Monte Carlo simulation, rendering purely simulation-based methods relatively more attractive.

In the general multivariate case, we can in principle use the Filtered Historical Simulation (FHS) approach discussed in Section 2.3.4, but a multivariate standardization is needed. Using for example the Cholesky or the spectral decomposition we first create vectors of standardized returns as in (78); i.e.,

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

where $\hat{\Omega}_t^{-1/2}$ denotes the relevant decomposition of the estimated covariance matrix.⁷⁹ Now, resampling with replacement vector-wise from the standardized returns will ensure that the marginal distributions, as well as particular features of the multivariate distribution, as for example, the contemporaneous cross-sectional dependencies suggested by Longin and Solnik (2001), will be preserved in the simulated data.

The dimensionality of the system may render the general multivariate standardization above practically infeasible. However, the same FHS approach can be applied with the base asset setup discussed in Section 3.1.2, resampling from the factor innovations,

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

where we again rely on the spectral or Cholesky decomposition to build up the distribution of the factor returns. Given the specification in Section 3.1.2, the corresponding idiosyncratic asset innovations may then be constructed from,

$$\hat{v}_t = R_t - \hat{B}R_{F,t}, \quad t = 1, 2, \dots, T.$$

Thus, by resampling sequentially from \hat{Z}_t and \hat{v}_t , we can easily build up the required distribution of the individual asset returns. This, of course, assumes that the base asset model provides a good description of the joint dependencies.

Alternatively, if one is willing to assume constant conditional correlations, as in (45) with $\Gamma_t = \Gamma$, then the standardization can simply be done on an individual asset-by-asset basis using the univariate GARCH or RV-based predictive volatilities. Resampling vector-wise from the standardized returns will naturally preserve the cross-sectional dependencies in the historical data.

⁷⁹ Patton and Sheppard (2009) recommend the spectral decomposition because unlike the Cholesky, it is invariant to a reordering of the variables.

3.3.5 Multivariate Extreme Value Theory

The simulation procedures discussed above work well for numerically describing correlations and related “central” features of the joint return distributions. Multivariate Extreme Value Theory (EVT) offers a tool for exploring cross-asset dependencies in the “tails” of distributions, which are not well captured by standard parametric distributions or correlation measures.

For example, Longin and Solnik (2001) define and compute extreme correlations between monthly US 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. Actual financial returns, however, behave quite differently. In particular, the correlation between the large (in an absolute sense) negative returns reported in Longin and Solnik (2001) tend to be much larger than the normal distribution would suggest (while interestingly, the correlations of large positive returns appear to approach zero in accordance with the normal distribution).⁸⁰ Such strong correlation between negative extremes is clearly a key risk management concern.⁸¹

To illustrate the important deviations from multivariate normality commonly found in financial markets, consider the threshold plots in Figure 12. The solid lines in Figure 12 show the empirical equity-index threshold correlations averaged across the 120 possible pairs of correlations based on the same 16 developed market returns used in the estimation of the DECO model in Figure 11. For comparison, the dashed lines indicate the threshold correlations implied by a multivariate standard normal distribution with constant correlation, while the lines with square markers are the threshold correlations computed via simulations from the previously estimated DECO model.

As the figure clearly shows, the down-market threshold correlations are much stronger than the up-market correlations. The multivariate normal distribution with constant correlation captures quite closely the up-market correlations but it cannot simultaneously account for the much larger, and increasing with the threshold, down-market correlations. The dynamic normal distribution driven by the basic Gaussian DECO model generates larger threshold correlations overall, but the model does not explain the strong multivariate asymmetry that actually exists in the returns. The specification of dynamic multivariate models and distributions to satisfactorily account for these important nonlinear asymmetric extreme dependencies is challenging. It remains the

⁸⁰ See also Ang and Bekaert (2002), Ang and Chen (2002), and Ang, Chen, and Xing (2006), among many others, for additional empirical evidence on similar nonlinear dependencies in equity returns.

⁸¹ It is generally unclear where these increased dependencies in the “tails” are coming from. Poon, Rockinger, and Tawn (2004), for instance, report that “devolatilizing” the daily returns for a set of international stock markets significantly reduces the joint tail dependence, while Bae, Karolyi, and Stulz (2003) find that time-varying volatility and GARCH effects cannot fully explain the counts of coincident “extreme” daily price moves observed across international equity markets.

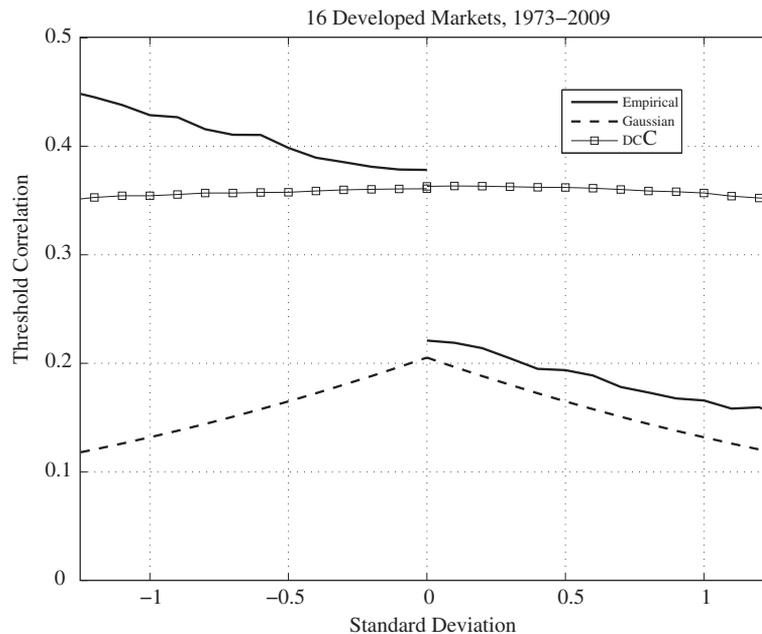


Figure 12 Average threshold correlations for 16 developed equity markets. The solid line shows the average empirical threshold correlation for GARCH residuals across sixteen developed equity markets. The dashed line shows the threshold correlations implied by a multivariate standard normal distribution with constant correlation. The line with square markers shows the threshold correlations from a DECO model estimated on the GARCH residuals from the 16 equity markets. The figure is based on weekly returns from 1973 to 2009.

focus of much ongoing work, much of which rely on the use of copulas and/or EVT type approximations.

A full treatment of this literature, and the extensive literature on multivariate EVT more generally, is well beyond the scope of the present chapter. Instead we refer to the books by Embrechts, Kluppelberg, and Mikosch (2002) and McNeil, Frey, and Embrechts (2005), along with the recent discussion in Embrechts (2009). Unfortunately, it is not yet clear whether multivariate EVT distributions will be operational in large-dimensional systems. Issues of scalability, as well as cross-sectional and temporal aggregation problems in parametric approaches, all present formidable challenges. Meanwhile, just as the newly available high-frequency data may be used in the construction of more accurate realized volatility measurements, and in turn covariance matrix forecasts, we conjecture that the intraday data may be constructively used in a similar manner for better measuring the “tails” of the return distributions, and in turn the joint extreme dependencies. The recent theoretical results in Bollerslev and Todorov (2011a) and related empirical findings in Bollerslev, Todorov, and Li (in press) are suggestive.

3.4 Systemic Risk and Measurement

We have emphasized using the univariate portfolio-level and multivariate asset-level risk models discussed in Sections 2 and 3, respectively, to construct real-time portfolio risk measures, such as VaR and ES , conditional on the history of returns. It is sometimes informative also to consider risk measures that condition not only on historical returns, but also on assumed *scenarios* for particular risk factors. We might, for example, be interested in the firm-specific effects of a market-wide shock, or the market-wide effects of a firm-specific shock.

Such scenario-based conditional risk measures are intrinsically related to measurement of *systemic* risk. There is no single definition of systemic risk, and we will shortly introduce several, but the defining characteristic is that systemic risk—one way or another—involves market-wide movements. Systemic risk measures can help firms to develop richer and more informative risk reports internally. They can also be used by supervisory authorities to measure and monitor the firm-specific effects of market-wide shocks, the market-wide effects of firm-specific shocks, and the corresponding totals across all firms. Those measurements may then impact policy decisions, such as which banks to support during crises, which bank mergers to approve, and so on.

3.4.1 Marginal Expected Shortfall and Expected Capital Shortfall

Marginal expected shortfall (MES) for firm j is

$$MES_{T+1|T}^{j|mkt} = E_T [r_{j,T+1} | \mathbb{C}(r_{mkt,T+1})], \quad (88)$$

where $r_{mkt,T+1}$ denotes the overall market return, and $\mathbb{C}(r_{mkt,T+1})$ denotes a systemic event, such as the market return falling below some threshold C . $MES_{T+1|T}^{j|mkt}$ tracks the sensitivity of firm j 's return to a market-wide extreme event, thereby providing a simple market-based measure of firm j 's fragility.

Ultimately, however, we are interested in assessing the likelihood of firm distress, and the fact that a firm's expected return is sensitive to market-wide extreme events—that is, the fact that its MES is large—does not necessarily mean that market-wide extreme events are likely to place it in financial distress. Instead, the distress likelihood should depend not only on MES , but also on how much capital the firm has on hand to buffer the effects of adverse market moves.

These distress considerations raise the idea of expected capital shortfall (ECS), which is closely related to, but distinct from, MES . ECS is the expected *additional capital* needed by firm j in case of a systemic market event. Clearly ECS should be related to MES , and Acharya, Pedersen, Philippon, and Richardson (2010) indeed show that in a simple model the two are linearly related,

$$ECS_{T+1|T}^{j|mkt} = a_{0j} + a_{1j}MES_{T+1|T}^{j|mkt}, \quad (89)$$

where a_{0j} depends on firm j 's "prudential ratio" of asset value to equity as well as its debt composition, and a_{1j} depends on firm j 's prudential ratio and initial capital. Note that the conditioning in *ECS* is closely related to the idea of *stress tests* popular in both best-practice private-sector and regulatory risk management. Indeed the thought experiment embodied in *ECS* is a stress test, with a very important and natural stress (a large market event) being tested.

Building on the theory of Acharya et al. (2010), Brownlees and Engle (2011) propose and empirically implement $ECS_{T+1|T}^{j|mkt}$ as a measure of firm j 's systemic risk exposure to the market at time T , with overall systemic risk then given by $\sum_{j=1}^N ECS_{T+1|T}^{j|mkt}$. Implementation of *MES* (and hence *ECS*) requires specification of the systemic market event $\mathbb{C}(r_{mkt,T+1})$, or more simply a market return threshold C . Values of $C = 2\%$ and $C = 40\%$ have, for example, been suggested for one-day and six-month returns, respectively. In addition, and of crucial importance, implementation of *MES* also requires a multivariate volatility model. That is, the conditioning on $\mathbb{C}(r_{j,T+1})$ in all of the measures above, from $MES_{T+1|T}^{j|mkt}$ through to $\sum_{j=1}^N ECS_{T+1|T}^{j|mkt}$, requires at least a bivariate volatility model for firm and market returns, and more generally a high-dimensional volatility model for all firms' returns. The models introduced in Sections 3.1, 3.2, 3.3 satisfy that need.⁸²

3.4.2 CoVaR and Δ CoVaR

In the previous section we introduced *MES* and *ECS*, which measure firm systemic risk *exposure* by conditioning firm events on market events. Here we introduce *CoVaR*, which works in the opposite direction, measuring firm systemic risk *contribution* by conditioning market events on firm events.

In (1) we defined firm j 's 1-step-ahead *VaR* at level p as the value of $VaR_{T+1|T}^{pj}$ that solves

$$p = \Pr_T \left(r_{j,T+1} < -VaR_{T+1|T}^{pj} \right).$$

Similarly, following Adrian and Brunnermeier (2011), one may define firm j 's 1-step-ahead "*CoVaR*" at level p conditional on a particular outcome for firm i , say $\mathbb{C}(r_{i,T+1})$, as the value of $CoVaR_{T+1|T}^{ji}$ that solves

$$p = \Pr_T \left(r_{j,T+1} < -CoVaR_{T+1|T}^{ji} | \mathbb{C}(r_{i,T+1}) \right). \quad (90)$$

⁸² Brownlees and Engle (2011), for example, use the daily GARCH-DCC modeling approach described in Section 3.1.1. Interestingly, they find that aggregate *MES* increased sharply starting in mid-2007, and that even by mid-2010 it was still much higher than in the pre-crisis period.

Because $\mathbb{C}(r_{i,T+1})$ is not in the time- T information set, $CoVaR$ will be different from the regular time- T conditional VaR . The leading choice of conditioning outcome, $\mathbb{C}(r_{i,T+1})$, is that firm i exceeds its VaR , or more precisely that $r_{i,T+1} < -VaR_{T+1|T}^{p,i}$. As such, $CoVaR$ is well suited to measure tail-event linkages between financial institutions.

A closely related measure, $\Delta CoVaR_{T+1|T}^{j|i}$ (read “Delta $CoVaR$ ”), is of particular interest. It measures the difference between firm- j VaR when firm- i is “heavily” stressed and firm- j VaR when firm i experiences “normal” times. More precisely,

$$\Delta CoVaR_{T+1|T}^{j|i} = CoVaR_{T+1|T}^{j|VaR(i)} - CoVaR_{T+1|T}^{j|Med(i)}, \quad (91)$$

where $CoVaR_{T+1|T}^{j|VaR(i)}$ denotes firm- j VaR when firm i 's return breaches its VaR , and $CoVaR_{T+1|T}^{j|Med(i)}$ denotes firm- j VaR when firm i 's return equals its median.

A direct extension lets us progress to the more interesting case of firm i 's overall systemic risk contribution, as opposed to just firm i 's contribution to firm j . We simply set $j = mkt$, so that $\Delta CoVaR_{T+1|T}^{mkt|i}$ then measures the difference between market VaR conditional on firm i experiencing an extreme return, and market VaR conditional on firm i experiencing a normal return.⁸³ Hence $\Delta CoVaR_{T+1|T}^{mkt|i}$ measures the contribution of firm i to overall systemic risk, $\sum_{i=1}^N \Delta CoVaR_{T+1|T}^{mkt|i}$.⁸⁴

The conditioning on $\mathbb{C}(r_{i,T+1})$ in all of the $CoVaR$ measures above, from $CoVaR_{T+1|T}^{j|i}$ through to $\sum_{i=1}^N \Delta CoVaR_{T+1|T}^{mkt|i}$, requires at least a bivariate volatility model for the returns on firms i and j , or i and mkt , and more generally a high-dimensional volatility model for all firms' returns. The models introduced in Sections 3.1,3.2,3.3 are again relevant.⁸⁵

3.4.3 Network Perspectives

Interestingly, modern network theory provides a powerful unifying framework for systemic risk measures, including measures like ECS and $CoVaR$ introduced above.⁸⁶

⁸³ Alternatively we might set $j = sys$, where sys denotes the financial system, as measured by the return on a portfolio of major financial institutions.

⁸⁴ The concept of $CoVaR$ also has interesting parallels to the conditioning of VaR in Garcia, Renault, and Tsafack (2007), who show that proper conditioning in VaR can eliminate the subadditivity problems raised by Artzner et al. (1999).

⁸⁵ Multivariate quantile models, such as those recently developed by White, Kim, and Manganello (2010), could also be used in this context.

⁸⁶ Here we provide a brief overview of key ideas. Extended discussion, references, and systemic risk measures based directly on network topology are contained in Diebold and Yilmaz (2009, 2011, 2012).

The simplest network is composed of N nodes, where any given pair of nodes may or may not be linked. We represent the network algebraically by an $N \times N$ symmetric adjacency matrix A of zeros and ones, $A = [a_{ij}]$, where $a_{ij} = 1$ if nodes i and j are linked, and $a_{ij} = 0$ otherwise. Because all network properties are embedded in A , any sensible connectedness measure must be based on A . The most important and popular, by far, are based on the idea of a node's *degree*, given by the number of its links to other nodes $\delta_i = \sum_j a_{ij}$, as well as aspects of the degree distribution across nodes. The total degree $\sum_i \delta_i$ (or mean degree $\frac{1}{N} \sum_i \delta_i$) is the key network connectedness measure.

The network structure sketched above is, however, rather too simple to describe the network connections of relevance in financial risk management. Generalization in two key directions is necessary. First, links may be of varying strength, not just 0–1. Second, links may be of different strength in different directions (e.g. firm i may impact firm j more than firm j impacts firm i). Note, for example, that the systemic risk measures introduced above are weighted and directional. For example, $CoVaR_{T+1|T}^{j|i}$ tracks effects from i to j , whereas $CoVaR_{T+1|T}^{i|j}$ tracks effects from j to i , and in general $CoVaR_{T+1|T}^{j|i} \neq CoVaR_{T+1|T}^{i|j}$.

It is a simple matter, however, to characterize directed, weighted networks in a parallel fashion. To allow for directionality, we allow the adjacency matrix A to be non-symmetric, and to allow for different relationship strengths we allow A to contain weights $a_{ij} \in [0, 1]$ rather than simply 0–1 entries. Node degrees are now obtained by summing weights in $[0, 1]$ rather than simply zeros and ones. In addition, and importantly, there are now “to-degrees” and “from-degrees”, corresponding to row sums and column sums, which generally differ since A is generally non-symmetric. The from-degree of node i is $\delta_i^{\text{from}} = \sum_j a_{ij}$, and the to-degree of node j is $\delta_j^{\text{to}} = \sum_i a_{ij}$. The total degree is $\delta = \sum_i \delta_i^{\text{from}} = \sum_j \delta_j^{\text{to}}$.

Crucially, the from- and to-degrees (and of course the total degree) measure aspects of systemic risk. The from- and to-degrees measure systemic risk with respect to particular firms. From-degrees measure exposures of individual firms to systemic shocks *from* the network, in a fashion analogous to $ECS_{T+1|T}^{j|mkt}$. To-degrees measure contributions of individual firms *to* systemic network events, in a fashion analogous to $\Delta CoVaR_{T+1|T}^{mkt|i}$. The total degree aggregates firm-specific systemic risk across firms, providing a measure of total system-wide systemic risk.

A key insight is that many approaches to systemic risk measurement fit naturally into the network framework. Consider, for example, $\Delta CoVaR$ measure. One can view the $\Delta CoVaR_{T+1|T}^{j|i}$, $i, j = 1, \dots, N$ (suitably scaled) as elements of an $N \times N$ adjacency matrix for a weighted directed network of firms. Then, for example, the systemic risk of firm i , $\Delta CoVaR_{T+1|T}^{sys|i}$, is the network to-degree of firm i , $\delta_i^{\text{to}} = \sum_j \Delta CoVaR_{T+1|T}^{j|i}$. And finally, the total systemic risk, $\sum_i \Delta CoVaR_{T+1|T}^{sys|i}$, is the network total degree δ .

4. CONDITIONING ON MACROECONOMIC FUNDAMENTALS

The risk models that we have discussed thus far are inherently “reduced form” in nature. They explain risk in an autoregressive fashion, as exemplified by the canonical GARCH family. Fortunately, even if the models fail to provide a deep structural understanding of volatility movements, they are nevertheless powerful and useful in a variety of contexts. We have obviously emphasized risk measurement and management, but other successful areas of application include portfolio allocation, spot and derivative asset pricing, active trading, and dynamic hedging.

Ultimately, however, we aspire to a deeper structural understanding. That is, we aspire to understand the connections between returns (especially, for our purposes, return volatilities) and macroeconomic fundamentals, say $r \leftrightarrow f$. Asset prices are risk-adjusted discounted claims on fundamental streams, so prices and their properties should ultimately depend on expected fundamentals and associated fundamental risks. Here we sketch emerging empirical aspects of those connections, through the lens of return and fundamental first and second moments, denoted μ_r , σ_r , μ_f , and σ_f , respectively.⁸⁷ Figure 13 provides a simple schematic diagram for all of the possible connections among σ_r , μ_r , σ_f , and μ_f . Each of the six connections represents a potentially important link, and a correspondingly important line of research inquiry.⁸⁸

Historically, however, it is well-known that σ_r , μ_r , σ_f , and μ_f have often appeared only weakly connected, or even disconnected. This observation is memorably enshrined in equity markets in the “excess volatility” puzzle of Shiller (1981), in foreign exchange markets in the “exchange rate disconnect” puzzle of Obstfeld and Rogoff (2000), in bond markets in Alan Greenspan’s long-maturity yield “conundrum,” and so on.⁸⁹

In contrast, we shall present and interpret a variety of accumulating evidences showing how returns—return volatilities in particular—are connected to fundamentals. Of course many of the links in Figure 13 remain incompletely understood, but they are receiving increased attention, and volatility features prominently throughout this

⁸⁷ In parallel to the models for returns emphasized so far in this chapter, we will content ourselves with means and variances, but one could, of course, also consider higher-order moments.

⁸⁸ Note that the links in Figure 13 are “undirected”, or “non-causal”, and as such more about correlation than causation. One could go even farther and consider directed, or causal, links, but that would require replacing each bi-directional arrow in Figure 13 with a *pair* of uni-directional arrows, thus doubling the number of links to be addressed.

⁸⁹ On the conundrum: “...the broadly unanticipated behavior of world bond markets remains a conundrum. Bond price movements may be a short-term aberration, but it will be some time before we are able to better judge the forces underlying recent experience” [Alan Greenspan, US congressional testimony, February 16, 2005]; see also Backus and Wright (2007).

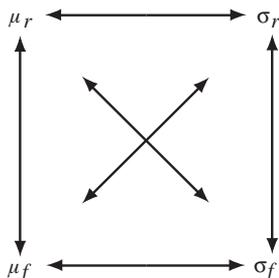


Figure 13 Return and fundamental mean and volatility linkages. Each link represents a distinct line of inquiry.

emerging research. Given the theme of the chapter, we will focus largely on three links directly involving σ_r and/or σ_f , namely $\mu_f \leftrightarrow \sigma_r$, $\mu_f \leftrightarrow \sigma_f$, and $\sigma_f \leftrightarrow \sigma_r$. We now address them in turn.

4.1 The Macroeconomy and Return Volatility

To begin, consider the link between macroeconomic fundamentals and return volatility, $\mu_f \leftrightarrow \sigma_r$. Officer (1973) was among the first to document and emphasize the very high stock market volatility during the very severe recession of the 1930s. The US stock market crash of 1987 spurred additional research into the fundamental determinants of volatility. In a well-known and exhaustive study in the wake of the 1987 crash, for example, Schwert (1989) went farther, showing that, surprisingly, the oft-suspected fundamentals (leverage, corporate profitability, etc.) have negligible impact on market volatility, while recessions *do*. In particular, return volatility is significantly higher in recessions, so that high volatility during bad times is not just a one-off Great Depression phenomenon, but rather a regularly recurring business cycle phenomenon.

These findings regarding the link between financial market volatility and the business cycle have since been echoed repeatedly. Hamilton and Lin (1996), for example, provide strong and sophisticated confirmation using regime-switching models of real growth and equity returns, allowing for both high and low real growth states and high and low equity-return volatility states. Their estimated regime transition probabilities indicate high positive steady-state coherence between low (high) real growth and high (low) equity return volatility.

More recent work, in particular Bloom, Floetotto, and Jaimovich (2009) as summarized in Table 1, also confirms and significantly amplifies Schwert's earlier result, showing, among other things, that it holds not only for stock returns at the aggregate level, but also for the cross-section of returns at the firm level. Table 1 makes clear not only the statistical significance of the "recession effect" on volatility, but also its sizable economic importance.

Table 1 Stock return volatility during recessions. Aggregate stock-return volatility is quarterly realized standard deviation based on daily return data. Firm-level stock-return volatility is the cross-sectional inter-quartile range of quarterly returns. *Source:* Adapted from Bloom, Floetotto, and Jaimovich (2009)

	Mean recession volatility increase (%)	Standard error (%)	Sample period
Aggregate returns	43.5	3.8	63Q1–09Q3
Firm-level returns	28.6	6.7	69Q1–09Q2

Although we have emphasized the links between macroeconomic fundamentals and equity market risk, one would expect related links in other market risk contexts. To take one example, consider foreign exchange. The expected real streams that underlie exchange rate determination are similar to those that underlie broad equity-market price determination, except that for exchange rates there are two streams, for two countries.

A second example is credit risk. In defaultable bond markets, for example, the celebrated Merton (1974) model directly links credit spreads to equity volatility, predicting that higher equity volatility should widen spreads, as emphasized empirically by Campbell and Tackler (2003). Hence the business cycle effects in equity volatility imply parallel business cycle effects in credit spreads, via the Merton model.

4.2 The Macroeconomy and Fundamental Volatility

The next link that we consider pertains to $\mu_f \leftrightarrow \sigma_f$; that is, real activity and its relationship to *real* (fundamental) volatility. It transpires that real fundamentals affect real volatility not only at business-cycle frequencies, but also at lower growth frequencies. Hence we treat both.

First consider fundamental volatility σ_f at business-cycle frequencies. Bloom, Floetotto, and Jaimovich (2009) show that σ_f is much higher in recessions (just as with σ_r), at both the aggregate level and at the cross-sectional firm level. We summarize their results in Table 2. Just as with the recession effect in stock return volatility, the recession effect in real growth volatility is notable not only for its statistical significance, but also for its sizable economic importance.⁹⁰

⁹⁰ Note that if stock return volatility and real growth volatility both increase during recessions, then they themselves must, of course, be positively related. We will return to this point below.

Table 2 Real growth volatility during recessions. Aggregate real growth volatility is quarterly conditional standard deviation. Firm-level real growth volatility is the cross-sectional inter-quartile range of quarterly real sales growth. *Source:* Adapted from Bloom, Floetotto, and Jaimovich (2009)

	Mean recession volatility increase (%)	Standard error (%)	Sample period
Aggregate growth	37.5	7.3	62Q1–09Q2
Firm-level growth	23.1	3.5	67Q1–08Q3

Observed links at business-cycle frequencies between real growth μ_f and real volatility σ_f are also well grounded in theory. Recent research, for example, explores dynamic stochastic general equilibrium models with heteroskedastic shocks (technology, preferences, policy, ...), as in Bloom (2009), Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (2011), and Basu and Bundick (2011).⁹¹

Now consider fundamental volatility σ_f at growth frequencies. Many have commented on the large reduction (roughly fifty percent) in US real GDP volatility beginning around 1985. Dubbed the “Great Moderation” by Stock and Watson (2002), it was originally documented by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000).

Perhaps the “Great Moderation” was just a long string of good luck, or perhaps it was a structural shift due to improved policy. In any event it seems likely that it is over, as the recession of 2007–2009 was very long and very deep. That is, even if a structural shift toward lower real volatility occurred in the mid-1980s, so too did a shift *back* around 2008. Hence it may be useful to think of the Great Moderation not as a one-off structural shift, but rather as a manifestation of a low-frequency real volatility dynamic driven by macroeconomic factors potentially very different from those that drive the earlier-discussed real volatility dynamics at business-cycle frequencies.

In intriguing recent work, Carvalho and Gabaix (2010) do precisely that, arguing that the Great Moderation was neither good policy nor good luck, but rather the natural outcome of the evolution of sectoral shares, which during the post-1984 period produced a better-diversified (and hence less volatile) GDP. In related work from an explicit network perspective, Acemoglu, Ozdaglar, and Tahbaz-Salehi (2010) make clear that the dynamic workings of “better diversification” are subtle and nuanced, depending not only on first-order connections among sectors, but also crucially on higher-ordered connections.

⁹¹ See also the insightful survey of Fernández-Villaverde and Rubio-Ramírez (in press).

4.3 Fundamental Volatility and Return Volatility

Now consider the links between fundamental volatility and return volatility, $\sigma_f \leftrightarrow \sigma_r$. Even with no additional work, our earlier discussion of $\mu_f \leftrightarrow \sigma_r$ and $\mu_f \leftrightarrow \sigma_f$ immediately implies that σ_r and σ_f must be positively related. This is so because σ_r and σ_f both covary negatively with the business cycle (μ_f), and hence they must covary positively with each other. Hence the case is closed as soon as it is opened; return volatility and real fundamental volatility are clearly related.

But one might want to go farther. First, one might want to complement our deduction of a $\sigma_f \leftrightarrow \sigma_r$ link with a direct exploration. Engle, Ghysels, and Sohn (2006) do just that, directly documenting the links between σ_f and σ_r after effectively removing high-frequency variation in returns and fundamentals using a persistent/transitory component model.

Second, one might want to explore cross-section and panel aspects. That can be useful because the precision with which relationships can be inferred depends on the amount of variation in the data, and there may be more variation over a broad cross section of countries than for a single country over time. Diebold and Yilmaz (2010) do this, showing that countries with higher fundamental volatility tend to have higher broad stock market volatility, even controlling for initial development level. In the most thorough study to date, Engle and Rangel (2008) explore time-series, cross-sections, and panels, clearly finding that the “long-term volatilities of macroeconomic fundamentals ... are primary causes of low-frequency market volatility”.

In closing this section we note that we have largely interpreted “market risk and macro fundamentals” as “market *volatility* and macro fundamentals”. As we have emphasized earlier in our discussion of portfolio-level risk measurement, however, one may naturally approach market volatility from a top-down (portfolio-level) or bottom-up (asset-level) perspective. In a bottom-up approach, not only conditional variances but also conditional *correlations* among individual returns are of central importance as they obviously impact portfolio (i.e., market) volatility. Hence the fundamental determinants of conditional correlations have also recently begun to receive attention, as in Rangel and Engle (2012).

4.4 Other Links

The links between volatility and fundamentals that we have discussed thus far do not involve μ_r . There are two main reasons. First, the horizons emphasized throughout most of the chapter tend to be fairly short—typically less than a month—and at such short horizons μ_r is small and arguably almost constant.⁹² Second, at longer horizons for which μ_r is larger and likely time varying in interesting ways, we can interpret μ_r as an excess return (“the equity premium”), which, of course, is the subject of an enormous

⁹² Indeed that is why we typically fix μ_r at zero in previous sections.

and distinguished literature that is treated extensively elsewhere in this volume. Hence we provide here only brief glimpses of aspects of the links $\mu_r \leftrightarrow \sigma_r$, $\mu_r \leftrightarrow \mu_f$ and $\mu_r \leftrightarrow \sigma_f$ as they relate most directly to our present concerns.

First, consider the equity premium and return-volatility relationship, $\mu_r \leftrightarrow \sigma_r$. Stimulated by the pioneering work of Markowitz (1959), an enormous amount of asset pricing research has focused on quantifying various aspects of this financial market “risk-return tradeoff”. Financial econometric research has followed suit, as exemplified by the GARCH-M model of Engle, Lilien, and Robbins (1987), defined by (7) and (8) above with $\mu_t = x_t' \beta + \delta \sigma_t$. In this model the conditional standard deviation enters directly as an explanatory variable for the conditional mean—together with other possible explanatory variables x_t —thus providing an econometric approximation to a time-varying risk premium.⁹³

Although intuitively appealing, a number of subtleties have emerged in both theory and empirics. Modern general equilibrium theory reveals that, in principle, positive contemporaneous risk-return correlation is not guaranteed, as subtle dynamic interactions may be operative; see, e.g. Abel (1988); Backus and Gregory (1993); Whitelaw (2000); and Bollerslev, Sizova, and Tauchen (2012) among others. In parallel, a wealth of recent empirical work reveals that, in practice, the contemporaneous risk-return correlation is often found to be *negative*; see, e.g. Bollerslev, Litvinova, and Tauchen (2006), Lettau and Ludvigson (2010), and Brandt and Wang (2010). Hence, rather ironically, we now realize that we know less than we thought about the most researched connection, $\mu_r \leftrightarrow \sigma_r$.

Second, consider the relationship between the equity premium and the business cycle, $\mu_r \leftrightarrow \mu_f$. Fama and French (1989) and Fama (1990) emphasize expected business conditions as a likely key driver of expected excess returns, with expected excess returns negative near business cycle peaks and positive near troughs. However, they, and the huge ensuing literature, use mostly *proxies* for expected business conditions, typically the dividend yield, the term premium, and the default premium; see, e.g. Campbell and Thompson (2008) and the literature cited therein.⁹⁴

Lettau and Ludvigson (2001) began a movement toward explicit incorporation of expected business condition variables with their celebrated generalized consumption-wealth ratio *cay*, or more precisely, the cointegrating residual between log consumption and log wealth. Campbell and Diebold (2009), and subsequently Goetzman, Watanabe, and Watanabe (2009), extended the movement with direct inclusion of expected real growth, or more precisely, Livingston survey expectations of real growth.⁹⁵ The results

⁹³ The conditional standard deviation is sometimes replaced by the conditional variance, $\mu_t = x_t' \beta + \delta \sigma_t^2$, or other monotone transformations of σ_t , in the estimation of the GARCH-M model.

⁹⁴ Note that, ironically, the standard proxies are financial rather than real.

⁹⁵ For details on the Livingston survey, see Croushore (1997).

suggest that expected growth is indeed a central determinant of expected excess returns, with the Livingston expectations generally the most stable and significant predictor across numerous competing specifications, including ones involving the “standard” financial predictor variables.

Having discussed a number of links involving fundamental volatility, we are now in a position to consider the final link, which also involves fundamental volatility, namely $\mu_r \leftrightarrow \sigma_f$. Modern asset-pricing theory emphasizes not only fundamental expectations, but also fundamental volatilities in the determination of the equity premium. An obvious example is the “long-run risk” model by Bansal and Yaron (2004), and its extension explicitly incorporating time-varying economic uncertainty in Bollerslev, Tauchen, and Hao (2009b). In this new class of models, which features Epstein and Zin (1989) preferences, variation in both consumption’s conditional mean *and* conditional variance contribute importantly to variation in the equity premium. Supporting empirical evidence is provided in Bansal, Khatacharian, and Yaron (2005) and Bollerslev, Sizova, and Tauchen (2012), among others.

4.5 Factors as Fundamentals

In our discussion of the links between market risk and macro fundamentals we have sometimes been casual in distinguishing returns from excess returns, realized from expected returns, realized from expected volatility, and related, in our treatment of timing. This is to some extent unavoidable, reflecting different conventions both within and among different and evolving literatures, as well as our desire to convey wide-ranging ideas in this broad survey. Nevertheless, a clearly emergent theme is that financial markets, as summarized by μ_r and σ_r , are very much linked to the business cycle, as summarized by μ_f and σ_f . Indeed it is not an exaggeration to claim that business-cycle risk may be the key driver of expected excess equity returns *and* return volatilities. Here we expand on that insight.

Although the business cycle may be a key risk factor, a long tradition, dating at least to Burns and Mitchell (1946) and actively extending to the present, recognizes that no single observed variable is “the business cycle” or “real activity”. Instead, we observe literally dozens of indicators (employment, industrial production, GDP, personal income, etc.), all of which contain information about the business cycle, which is not directly observable. Hence the key business cycle real activity fundamental underlying risk may be appropriately and productively viewed as a common factor to be extracted from many individual real activity indicators.

Expanding on this “factors as fundamentals” perspective, another likely relevant additional factor candidate is price/wage pressure, which may of course interact with real activity, as emphasized in Aruoba and Diebold (2010). In any event, the point is simply that, although we see hundreds of macroeconomic fundamentals, a drastically smaller set of underlying macroeconomic factors is likely relevant for tracking market risk.

This is useful not only for best-practice firm-level risk management, but also for regulators. In particular, the factors-as-fundamentals perspective has important implications for the design of stress tests that simulate financial market responses to fundamental shocks, suggesting that only a few key fundamentals (factors) need be stressed.

Not surprisingly, then, we advocate that risk managers pay closer attention to macroeconomic factors, as they are the ultimate drivers of market risk. We hasten to add, however, that due to the frequent “disconnect” problems mentioned earlier, we would never advocate conditioning risk assessments *only* on macroeconomic factors. Rather, macroeconomic factors complement, rather than substitute, for the methods discussed in earlier sections, by broadening the conditioning information set to include fundamentals in addition to past returns.

One might reasonably question the usefulness of conditioning on macroeconomic data for daily risk assessment, because macroeconomic data are typically available only quarterly (e.g. GDP and its components), or sometimes monthly (e.g. industrial production and the CPI). Recent developments that exploit state space methods and optimal filtering, however, facilitate high-frequency (e.g. daily) monitoring of latent macroeconomic fundamental factors. In particular, based on the high-frequency real activity monitoring approach of Aruoba, Diebold, and Scotti (2009), the Federal Reserve Bank of Philadelphia produces the “ADS index” of real activity, updated and written to the web in real time as new indicator data, released at different frequencies, are released or revised.⁹⁶

We have emphasized macroeconomic fundamentals for equity market risk, but the bond market is also closely linked to macroeconomic fundamentals. In particular, government bond yield curves are driven by just a few factors (level, slope, curvature), with the level factor closely linked to price/wage activity and the slope factor closely linked to real activity.⁹⁷ The same is true for yield curves of defaultable bonds, except that there is the additional complication of default risk, but that too is linked to the business cycle. Hence despite data on dozens of government bond yields, and dozens of macroeconomic indicators, the interesting reality is their much lower-dimensional “state vectors”—the level and slope factors beneath the yield curve, and the real and price/wage activity factors beneath the macroeconomy. One can easily imagine the usefulness for daily market and credit risk management (say) of systems linking yield curve factors (level, slope, curvature, ...), equity factors (market, HML, SMB, momentum, liquidity, ...), and macroeconomic factors (real, price/wage, ...). All of those factors are now readily available at daily frequency.

⁹⁶ The index and a variety of related materials are available at <http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index>.

⁹⁷ For background and references, see Diebold and Rudebusch (2012).

5. CONCLUDING REMARKS

We have attempted to demonstrate the power and potential of dynamic financial econometric methods for practical financial risk measurement and management. We have surveyed the large literature on high-frequency volatility modeling, interpreting and unifying the most important and intriguing results of practical relevance.⁹⁸ 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. Successful risk measurement may be achieved through the use of univariate density models directly for portfolio returns. GARCH volatility models offer a convenient and parsimonious framework for modeling key dynamic features of such portfolio returns, including volatility mean reversion, long memory, and asymmetries.
3. Successful risk management, in contrast, requires a fully specified multivariate density model. In that regard, standard multivariate models are too heavily parameterized to be useful in realistic medium- and large-scale financial market contexts. In medium-scale financial contexts, recently developed multivariate GARCH models are likely to be useful. In very large-scale financial contexts, more structure must be imposed, such as decoupling variance and correlation dynamics. In all cases, resampling methods applied to standardized returns are an attractive strategy for accommodating conditionally non-normal returns.
4. Volatility measures based on high-frequency return data hold great promise for practical risk management, as realized volatility and correlation measures produce more accurate risk assessments and forecasts than their conventional competitors. Because high-frequency information is only available for highly liquid assets, a base-asset factor approach may sometimes be useful. In addition, the near log-normality of realized volatility, together with the near-normality of returns standardized by realized volatility, holds promise for relatively simple-to-implement log-normal/normal mixture models in financial risk management.
5. The business cycle emerges as a key macroeconomic fundamental driving risk in a variety of markets, including equities and bond yields. Among other things, this means that our emphasis on conditioning applies not only at the short horizons

⁹⁸ We hasten to add that this chapter is a complement, not a substitute, for the more general and technical survey of volatility and covariance forecasting of Andersen et al. (2006a). In addition, space constraints and other considerations have invariably limited our choice of included topics. For instance, we have largely neglected stochastic volatility and other parameter-driven approaches to volatility modeling, as well as option-implied volatility.

(typically daily) stressed in Sections 2 and 3, but also at much longer horizons, once the information set is appropriately broadened to include macro fundamentals as opposed to just past returns.

REFERENCES

- Abel, A. (1988). Stock prices under time-varying dividend risk. *Journal of Monetary Economics*, 35, 375–393.
- Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi A. (2010). Cascades in networks and aggregate volatility. Manuscript, MIT
- Acharya, V., Pedersen, L., Philippon, T., & Richardson, M. (2010). Measuring systemic risk. Manuscript, NYU.
- Adrian, T., & Brunnermeier, M. (2011). CoVaR. Manuscript, Federal Reserve Bank of New York.
- Aielli, G. P. (2006). Consistent estimation of large scale dynamic conditional correlations. Manuscript, University of Florence.
- Ait-Sahalia, Y., Mykland, P. A., & Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160, 190–203.
- Albanese, C., Jackson, K., & Wiberg, P. (2004). A new Fourier transform algorithm for value-at-risk. *Quantitative Finance*, 4, 328–338.
- Alexander, C. (2001). *Market models: A guide to financial data analysis*. John Wiley and Sons.
- Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57, 1047–1091.
- Andersen, T. G., & Bollerslev, T. (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, T. (1998a). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39, 885–905.
- Andersen, T. G., & Bollerslev, T. (1998b). DM-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer-run dependencies. *Journal of Finance*, 53, 219–265.
- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006a). Volatility and correlation forecasting. In G. Elliot, C. W. J. Granger & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 777–878). North-Holland.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003a). Modeling and forecasting realized volatility. *Econometrica*, 71, 579–625.
- Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007a). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, 89, 701–720.
- Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2010a). Parametric and nonparametric volatility measurement. In L. P. Hansen & Y. Ait-Sahalia (Eds.), *Handbook of financial econometrics* (pp. 67–138). North-Holland.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2000a). Exchange rate returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal*, 4, 159–179.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2003b). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, 93, 38–62.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2007b). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73, 251–277.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Wu, G. (2006b). Realized beta: Persistence and predictability. In T. B. Fomby & D. Terrell (Eds.), *Advances in econometrics: Econometric analysis of economic and financial time series* (pp. 1–40). Elsevier.
- Andersen, T. G., Bollerslev, T., & Dobrev, D. (2007c). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional assumptions. *Journal of Econometrics*, 138, 125–180.
- Andersen, T. G., Bollerslev, T., Frederiksen, P. H., & Nielsen, M. (2010b). Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. *Journal of Applied Econometrics*, 25, 233–261.

- Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance*, 6, 457–477.
- Andersen, T. G., Bollerslev, T., & Meddahi, N. (2004). Analytical evaluation of volatility forecasts. *International Economic Review*, 45, 1079–1110.
- Andersen, T. G., Bollerslev, T., & Meddahi, N. (2011a). Market microstructure noise and realized volatility forecasting. *Journal of Econometrics*, 160, 220–234.
- Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169, 75–93.
- Andersen, T. G., Dobrev, D., & Schaumburg, E. (2011b). A functional filtering and neighborhood truncation approach to integrated quarticity estimation. Manuscript, Northwestern University.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61, 43–76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2000b). Great realizations. *Risk*, 13, 105–108.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96, 42–55.
- Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies*, 15, 1137–1187.
- Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63, 443–494.
- Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *Review of Financial Studies*, 19, 1191–1239.
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9, 203–228.
- Aruoba, S. B., & Diebold, F. X. (2010). Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review*, 100, 20–24.
- Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real time measurement of business conditions. *Journal of Business and Economic Statistics*, 27, 417–427.
- Asai, M., & So, M. K. P. (2010). Stochastic covariance models. Manuscript, Hong Kong University of Science and Technology.
- Backus, D., & Gregory, A. (1993). The relations between risk premia and conditional variances. *Journal of Business and Economic Statistics*, 11, 177–185.
- Backus, D., & Wright, J. (2007). Cracking the conundrum. *Brookings Papers on Economic Activity*, 38, 293–329.
- Bae, K. H., Karolyi, G. A., & Stulz, R. (2003). A new approach to measuring financial contagion. *Review of Financial Studies*, 16, 717–763.
- Baillie, R. T., & Bollerslev, T. (1992). Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics*, 52, 91–113.
- Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractional integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74, 3–30.
- Bakshi, G., & Panayotov, G. (2010). First-passage probability, jump models, and intra-horizon risk. *Journal of Financial Economics*, 95, 20–40.
- Bannouh, K., Martens M., Oomen, R., & van Dijk, D. (2010). Realized mixed-frequency factor models for vast dimensional covariance estimation. Manuscript, Erasmus University Rotterdam.
- Bansal, R., Khatacharian, V., & Yaron, A. (2005). Interpretable asset markets?. *European Economic Review*, 49, 531–560.
- Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59, 1481–1509.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76, 1481–1536.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162, 149–169.
- Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, B*, 63, 167–241.
- Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, B*, 64, 253–280.

- Barndorff-Nielsen, O. E., & Shephard, N. (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression and correlation in financial economics. *Econometrica*, 72, 885–925.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2, 1–48.
- Barone-Adesi, G., Bourgoin, F., & Giannopoulos, K. (1998). Don't look back. *Risk*, 11, 100–104.
- Basu, S., & Bundick, B. (2011). Uncertainty shocks in a model of effective demand. Manuscript, Boston College.
- Bauer, G. H., & Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160, 93–101.
- Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21, 79–109.
- Berkowitz, J., & O'Brien, J. (2002). How accurate are value-at-risk models at commercial banks?. *Journal of Finance*, 57, 1093–1111.
- Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77, 623–685.
- Bloom, N., Floetotto, M., & Jaimovich, N. (2009). Really uncertain business cycles. Manuscript, Stanford University.
- 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. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72, 498–505.
- Bollerslev, T. (2010). Glossary to ARCH (GARCH). In T. Bollerslev, J. R. Russell & M. W. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle* (pp.137–163). Oxford University Press.
- Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96, 116–131.
- Bollerslev, T., Kretschmer, U., Pigorsch, C., & Tauchen, G. (2009a). A discrete-time model for daily S&P 500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics*, 150, 151–166.
- Bollerslev, T., Litvinova, J., & Tauchen, G. (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, 4, 353–384.
- Bollerslev, T., & Mikkelsen, H. O. (1999). Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics*, 92, 75–99.
- Bollerslev, T., Sizova, N., & Tauchen, G. (2012). Volatility in equilibrium: Asymmetries and dynamic dependencies. *Review of Finance*, 16, 31–80.
- Bollerslev, T., Tauchen, G., & Hao, Z. (2009b). Expected stock returns and variance risk premia. *Review of Financial Studies*, 22, 4463–4492.
- Bollerslev, T., & Todorov, V. (2011a). Estimation of jump tails. *Econometrica*, 79, 1727–1783.
- Bollerslev, T., & Todorov, V. (2011b). Tails, fears and risk premia. *Journal of Finance*, 66, 2165–2211.
- Bollerslev, T., Todorov, V., & Li, S. Z. (in press). Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics*.
- Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11, 143–172.
- Bollerslev, T., & Zhang, Y. B. (2003). Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance*, 10, 533–558.
- Bonato Caporin, M., & Rinaldo, A. (2009). Forecasting realized (Co)Variances with a block structure Wishart autoregressive model. Manuscript, Swiss National Bank.
- Boudoukh, J., Richardson, M., & Whitelaw, R. (1998). The best of both worlds. *Risk*, 11, 64–67.
- Brandt, M., Santa-Clara, P., & Valkanov, R. (2004). Optimal portfolios with parametric weights. Manuscript, Duke University.
- Brandt, M. & Wang, L. (2010). Measuring the time-varying risk-return relation from the cross-section of equity returns. Manuscript, Duke University.
- Brownlees, C., & Engle, R.F. (2011). Volatility, correlation and tails for systemic risk measurement. Manuscript, NYU.

- Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics and Data Analysis*, 51, 2232–2245.
- Brownlees, C. T., & Gallo, G. M. (2011). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, 8, 29–56.
- Burns, A. F., & Mitchell, W. C. (1946). *Measuring business cycles*. National Bureau of Economic Research.
- Campbell, J. Y., & Tackler, G. B. (2003). Equity volatility and corporate bond yields. *Journal of Finance*, 58, 2321–2350.
- Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *Review of Financial Studies*, 21, 1509–1531.
- Campbell, S. D., & Diebold, F. X. (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business and Economic Statistics*, 27, 266–278.
- Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4, 537–572.
- Carvalho, V. M., & Gabaix, X. (2010). The great diversification and its undoing. Manuscript, NYU.
- Chan, N. H., Deng, S. J., Peng, L., & Xia, Z. (2007). Interval estimation for the conditional value-at-risk based on GARCH models with heavy tailed innovations. *Journal of Econometrics*, 137, 556–576.
- Chib, Y., Omori, S., & Asai, M. (2009). Multivariate stochastic volatility. In T. G. Andersen R. A. Davis J.-P. Kreiss & T. Mikosch (Eds.), *Handbook of Financial Time Series* (pp.365–400). Verlag: Springer.
- Chiriac, R., & Voev, V. (2011). Modeling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26, 922–947.
- Christoffersen, P. F. (2003). *Elements of financial risk management*. Academic Press.
- Christoffersen, P. F., Errunza, V., Jacobs, K., & Langlois, H. (2011a). Is the potential for international diversification disappearing? Manuscript, University of Toronto.
- Christoffersen, P. F., Feunou, B., Jacobs, K., & Meddahi, N. (2011b). The economic value of realized volatility. Manuscript, University of Toronto.
- Christoffersen, P. F., Heston, S., & Jacobs, K. (2006). Option valuation with conditional skewness. *Journal of Econometrics*, 131, 253–284.
- Christoffersen, P. F., Jacobs, K., Ornathanalai, C., & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, 90, 272–297.
- Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41, 135–155.
- Connor, G., Goldberg, L. R., & Korajczyk, R. A. (2010). *Portfolio risk analysis*. Princeton University Press.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7, 174–196.
- Corsi, F., Fusari, N., & LaVecchia, D. (2011). Realizing smiles: Options pricing with realized volatility. Swiss Finance Institute Research Paper No. 10–05.
- Corsi, F., & Reno, R. (2010). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Manuscript, University of Lugano.
- Cox, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics*, 8, 93–115.
- Creal, D., Koopman, S., & Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics*, 29, 552–563.
- Croushore, D. (1997). The Livingston survey: Still useful after all these years. *Federal Reserve Bank of Philadelphia Business Review*, March/April, 1–12.
- Demarta, S., & McNeil, A. (2005). The t copula and related copulas. *International Statistical Review*, 73, 111–129.
- Diebold, F. X. (2012). An historical perspective on nancial risk measurement and management. In: F. X. Diebold (Ed.), *Financial risk measurement and management. International library of critical writings in economics series*. Edward Elgar.
- Diebold, F. X., Hickman, A., Inoue, A., & Schuermann, T. (1998a). Scale models. *Risk*, 11, 104–107.
- Diebold, F. X., & Nerlove, M. (1989). The dynamics of exchange rate volatility: A multivariate latent-factor ARCH model. *Journal of Applied Econometrics*, 4, 1–22.
- Diebold, F. X., & Rudebusch, G. D. (2012). The dynamic Nelson–Siegel approach to yield curve modeling and forecasting. Manuscript, University of Pennsylvania.

- Diebold, F. X., Schuermann, T., & Stroughair, J. D. (1998b). Pitfalls and opportunities in the use of extreme value theory in risk management. In A.-P. N. Refenes, A. N. Burgess & J. D. Moody (Eds.), *Decision Technologies for Computational Finance* (pp.3–12). Kluwer Academic Publishers.
- Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, *119*, 158–171.
- Diebold, F. X., & Yilmaz, K. (2010). Macroeconomic volatility and stock market volatility, worldwide. In T. Bollerslev, J. Russell & M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle* (pp.97–116). Oxford University Press.
- Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*, 57–66.
- Diebold, F. X., & Yilmaz, K. (2011). On the network topology of variance decompositions: Measuring the connectedness of financial firms. Manuscript, University of Pennsylvania.
- Drost, F. C., & Nijman, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, *61*, 909–927.
- Embrechts, P. (2009). Linear correlation and EVT: Properties and caveats. *Journal of Financial Econometrics*, *7*, 30–39.
- Embrechts, P., Kluppelberg, C., & Mikosch, T. (2002). *Modelling extremal events*. Springer Verlag.
- Engle, R. F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, *15*, 157–168.
- Engle, R. F. (2002a). 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. (2002b). New frontiers in ARCH models. *Journal of Applied Econometrics*, *17*, 425–446.
- Engle, R. F. (2004). Risk and volatility: Econometric models and financial practice. *American Economic Review*, *94*, 405–420.
- Engle, R. F. (2009a). *Anticipating correlations: A new paradigm for risk management*. Princeton University Press.
- Engle, R. F. (2009b). The risk that risk will change. *Journal of Investment Management*, *7*, 24–28.
- Engle, R. F. (2011). Long-term skewness and systemic risk. *Journal of Financial Econometrics*, *9*, 437–468.
- Engle, R. F., & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, *131*, 3–27.
- Engle, R. F., Ghysels, E., & Sohn, B. (2006). On the economic sources of stock market volatility. Manuscript, NYU.
- Engle, R. F., & Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics*, *9*, 345–360.
- Engle, R. F., & Kelly, B. (2008). Dynamic equicorrelation. Manuscript, NYU.
- Engle, R. F., & Lee, G. G. J. (1999). A permanent and transitory component model of stock return volatility. In R. F. Engle & H. White (Eds.), *Cointegration, causality, and forecasting: A Festschrift in honor of Clive W.J. Granger* (pp.475–497). Oxford University Press.
- Engle, R. F., Lilien, D. M., & Robbins, R. P. (1987). Estimating time-varying risk premia in the term structure: The ARCH-M model. *Econometrica*, *55*, 391–407.
- Engle, R. F., & Mezrich, J. (1996). GARCH for groups. *Risk*, *9*, 36–40.
- Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, *48*, 1749–1778.
- Engle, R. F., Ng, V. K., & Rothschild, M. (1990). Asset pricing with a factor-ARCH covariance structure. *Journal of Econometrics*, *45*, 213–238.
- Engle, R. F., & Rangel, J. G. (2008). The spline GARCH model for unconditional volatility and its global macroeconomic causes. *Review of Financial Studies*, *21*, 1187–1222.
- Engle, R. F., Shephard, N., & Sheppard, K. (2008). Fitting vast dimensional time-varying covariance models. Manuscript, NYU.
- Epps, T. W. (1979). Comovement in stock prices in the very short run. *Journal of the American Statistical Association*, *74*, 291–298.
- Epstein, L., & Zin, S. E. (1989). Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*, 937–969.
- Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, *38*, 34–105.
- Fama, E. F. (1990). Stock returns, expected returns, and real activity. *Journal of Finance*, *45*, 1089–1108.

- Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25, 23–49.
- Fan, J., Fan, Y., & Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147, 186–197.
- Fan, J., Liao, Y., & Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. Manuscript, Princeton University.
- Fernandez, C., & Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 359–371.
- Fernández-Villaverde, J., Guerrón-Quintana, P. A., Rubio-Ramírez, J., & Uribe, M. (2011). Risk matters: The real effects of volatility shocks. *American Economic Review*, 101, 2530–2561.
- Fernández-Villaverde, J., & Rubio-Ramírez, J. F. (in press). Macroeconomics and volatility: Data, models, and estimation. In D. Acemoglu, M. Arellano, & E. Decker (Eds.), *Advances in economics and econometrics: Theory and applications, tenth world congress of the econometric society*, Cambridge University Press.
- Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using realized volatility. *Journal of Financial Economics*, 63, 473–509.
- Föllmer, H., & Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6, 429–447.
- Forsberg, L., & Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ECU) volatility and ARCH modeling (of the Euro): The GARCH-NIG model. *Journal Applied Econometrics*, 17, 535–548.
- Gallant, A. R., Hsu, C. T., & Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics*, 81, 617–631.
- Garcia, R., Renault, E., & Tsafack, G. (2007). Proper conditioning for coherent VaR in portfolio management. *Management Science*, 53, 483–494.
- Glasserman, P., Heidelberger, P., & Shahabuddin, P. (2002). Portfolio value-at-risk with heavy-tailed risk factors. *Mathematical Finance*, 12, 239–269.
- Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779–1801.
- Goetzman, W. N., Watanabe, A., & Watanabe, M. (2009). Investor expectations, business conditions, and the pricing of beta-instability risk. Manuscript, Yale University.
- Golosnoy, V., Glibisch, B., & Liesenfeld, R. (2010). The conditional autoregressive Wishart model for multivariate stock market volatility. Manuscript, Christian-Albrechts-University of Kiel.
- Gourieroux, C., Jasiak, J., & Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150, 167–181.
- Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14, 227–238.
- Guidolin, M., & Timmermann, A. (2006). Term structure of risk under alternative econometric specifications. *Journal of econometrics*, 131, 285–308.
- Hafner, C., & Manner, H. (2012). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, 27, 269–295.
- Hamilton, J. D., & Lin, G. (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics*, 11, 573–593.
- Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35, 705–730.
- Hansen, P. R., Huang, Z., & Shek, H. (in press). Realized GARCH: A complete model of returns and realized measures of volatility. *Journal of Applied Econometrics*.
- Hansen, P. R., & Lunde, A. (2005). A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics*, 3, 525–554.
- Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24, 127–161.
- Hansen, P. R., Lunde, A., & Voev, V. (2010). Realized beta GARCH: A multivariate GARCH model with realized measures of volatility and covolatility. Manuscript, University of Aarhus.
- Harvey, A., Ruiz, E., & Sentana, E. (1992). Unobserved component time series models with ARCH disturbances. *Journal of Econometrics*, 52, 129–157.

- Harvey, A. C. (1989). *Forecasting, Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Hautsch, N., Kyj, L. M., & Oomen, R. C. A. (2012). A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics*, 27, 625–645.
- Heston, S. L., & Nandi, S. (2000). A closed-form GARCH option valuation model. *Review of Financial Studies*, 13, 585–625.
- Hull, J., & White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1, 5–19.
- Jacod, J., Li, Y., Mykland, P., Podolskij, M., & Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications*, 119, 2249–2276.
- Jin, X., & Maheu, J. M. (2010). Modeling realized covariances and returns. Manuscript, University of Toronto.
- Jondeau, E., & Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25, 827–853.
- Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
- Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics*, 134, 95–128.
- Kim, C., & Nelson, C. R. (1999). Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics*, 81, 608–616.
- Ledoit, O., & Wolf, M. (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., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 68, 365–411.
- Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, 56, 815–849.
- Lettau, M., & Ludvigson, S. C. (2010). Measuring and modeling variation in the risk-return tradeoff. In L. P. Hansen & Y. Ait-Sahalia (Eds.), *Handbook of financial econometrics* (pp.617–690). Elsevier.
- Li, David X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income*, 9, 43–54.
- Liu, Q. (2009). On portfolio optimization: How and when do we benefit from high-frequency data. *Journal of Applied Econometrics*, 24, 560–582.
- Longin, F., & Solnik, B. (1995). Is the correlation in international equity returns constant: 1960–1990?. *Journal of International Money and Finance*, 14, 3–26.
- Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, 56, 649–676.
- Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance*, 24, 1097–1130.
- Lu, Y. (2005). *Modeling and forecasting daily stock return volatility with intraday price fluctuations information*. Ph.D. thesis, Department of Economics, University of Pennsylvania.
- Lunde, A., Shephard, N., & Sheppard, K. (2011). Econometric analysis of vast covariance matrices using composite realized kernels. Manuscript, University of Aarhus.
- Maheu, J. M., & McCurdy, T. (2011). Do high-frequency measures of volatility improve forecasts of return distributions?. *Journal of Econometrics*, 160, 69–76.
- Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric Brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36, 394–419.
- Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. John Wiley.
- McConnell, M. M., & Perez-Quiros, G. (2000). Output fluctuations in the United States: What has changed since the early 1980s. *American Economic Review*, 90, 1464–1476.
- McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7, 271–300.
- McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques, tools*. Princeton University Press.
- Meddahi, N., & Renault, E. (2004). Temporal aggregation of volatility models. *Journal of Econometrics*, 119, 355–379.
- Mehta, M. L. (1990). *Random matrices*. Academic Press.

- Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8, 323–361.
- Merton, Robert C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29, 449–470.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59, 347–370.
- Nelson, D. B. (1992). Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Financial Economics*, 52, 61–90.
- Nerlove, M., & Wage, S. (1964). On the optimality of adaptive forecasting. *Management Science*, 10, 207–224.
- Noureddin, D., Shephard, N., & Sheppard, K. (in press). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*.
- Obstfeld, M., & Rogoff, K. (2000). The six major puzzles in international macroeconomics: Is there a common cause?. In B. Bernanke & K. Rogoff (Eds.), *NBER Macroeconomics Annual 2000* (pp.339–412). MIT Press.
- Officer, R. R. (1973). The variability of the market factor of the New York stock exchange. *Journal of Business*, 46, 434–453.
- Oh, D. H., & Patton, A. (2011). Modelling dependence in high dimensions with factor copulas. Manuscript, Duke University.
- Patton, A., & Sheppard, K. (2009). Evaluating volatility and correlation forecasts. In T. G. Andersen R. Davis, J.-P. Kreiss & T. Mikosch (Eds.), *Handbook of Financial Time Series* (pp.801–838). Springer Verlag.
- Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47, 527–556.
- Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics*, 131, 445–473.
- Perignon, C., Deng, Z. Y., & Wang, Z. J. (2008). Do banks overstate their value-at-risk? *Journal of Banking and Finance*, 32, 783–794.
- Perignon, C., & Smith, D. R. (2010a). Diversification and value-at-risk. *Journal of Banking and Finance*, 34, 55–66.
- Perignon, C., & Smith, D. R. (2010b). The level and quality of value-at-risk disclosure by commercial banks. *Journal of Banking and Finance*, 34, 362–377.
- Philipov, A., & Glickman, M. E. (2006). Multivariate stochastic volatility via Wishart processes. *Journal of Business and Economic Statistics*, 24, 313–328.
- Poon, S. H., Rockinger, M., & Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Review of Financial Studies*, 17, 581–610.
- Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking and Finance*, 30, 561–582.
- Rangel, J. G., & Engle, R. F. (2012). The factor-spline-GARCH Model for high and low frequency correlations. *Journal of Business and Economic Statistics*, 30, 109–124.
- Rosenberg, J. V., & Schuermann, T. (2006). A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics*, 79, 569–614.
- Rosenow, B. (2008). Determining the optimal dimensionality of multivariate volatility models with tools from random matrix theory. *Journal of Economic Dynamics and Control*, 32, 279–302.
- Schwert, G. W. (1989). Why does stock market volatility change over time?. *Journal of Finance*, 44, 1115–1153.
- Shephard, N., & Sheppard, K. (2010). Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25, 197–231.
- Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends?. *American Economic Review*, 71, 421–436.
- Sizova, N. (2011). Integrated variance forecasting: Model based vs. reduced form. *Journal of Econometrics*, 162, 294–311.
- Stentoft, L. (2008). Option pricing using realized volatility. Manuscript, University of Aarhus.
- Stock, J. H., & Watson, M. W. (2002). Has the business cycle changed and why?. In M. Gertler & K. Rogoff (Eds.), *NBER Macroeconomics Annual* (pp.159–218). MIT Press.
- Stulz, R. M. (2002). Risk management and derivatives. South-Western.
- Stulz, R. M. (2008). Risk management failures: What are they and when do they happen?. *Journal of Applied Corporate Finance*, 20, 58–67.

- Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32, 235–258.
- Tse, Y. K., & Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroskedasticity model with time-varying correlations. *Journal of Business and Economic Statistics*, 20, 351–362.
- Varneskov, R. T., & Voev, V. (2010). The role of realized ex-post covariance measures and dynamic model choice on the quality of covariance forecasts. Manuscript, University of Aarhus.
- Visser, M. P. (2011). GARCH parameter estimation using high-frequency data. *Journal of Financial Econometrics*, 9, 162–197.
- Wang, Y., & Zou, J. (2010). Vast volatility matrix estimation for high-frequency financial data. *Annals of Statistics*, 38, 943–978.
- White, H., Kim, T.-H., & Manganelli, S. (2010). VAR for VaR: Measuring systemic risk using multivariate regression quantiles. Manuscript, UCSD.
- Whitelaw, R. F. (2000). Stock market risk and return: An equilibrium approach. *Review of Financial Studies*, 13, 521–547.
- Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics*, 14, 45–52.
- Zumbach, G. (2006). A gentle introduction to the RM 2006 methodology. Manuscript, RiskMetrics Group.

Non Print Items

ABSTRACT

Current practice largely follows restrictive approaches to market risk measurement, such as historical simulation or RiskMetrics. In contrast, we propose flexible methods that exploit recent developments in financial econometrics and are likely to produce more accurate risk assessments, treating both portfolio-level and asset-level analysis. Asset-level analysis is particularly challenging because 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 powerful yet parsimonious models that are easily estimated. In addition, we emphasize the need for deeper understanding of the links between market risk and macroeconomic fundamentals, focusing primarily on links among equity return volatilities, real growth, and real growth volatilities. Throughout, we strive not only to deepen our scientific understanding of market risk, but also cross-fertilize the academic and practitioner communities, promoting improved market risk measurement technologies that draw on the best of both.

KEYWORDS

Market risk, Volatility, GARCH

JEL CLASSIFICATIONS

C1, G1