Modeling Asymmetric Comovements of Asset Returns

Kenneth F. Kroner
Barclays Global Investors

Victor K. Ng
Goldman Sachs

Existing time-varying covariance models usually impose strong restrictions on how past shocks affect the forecasted covariance matrix. In this article we compare the restrictions imposed by the four most popular multivariate GARCH models, and introduce a set of robust conditional moment tests to detect misspecification. We demonstrate that the choice of a multivariate volatility model can lead to substantially different conclusions in any application that involves forecasting dynamic covariance matrices (like estimating the optimal hedge ratio or deriving the risk minimizing portfolio). We therefore introduce a general model which nests these four models and their natural "asymmetric" extensions. The new model is applied to study the dynamic relation between large and small firm returns.

The estimation of time-varying covariances between asset returns is crucial for asset pricing, portfolio selection, and risk management. Yet the development in this area is lagging significantly behind the development in the time-varying volatility area, as evidenced by the sparcity of the literature on modeling time-varying covariances compared to modeling time-varying volatility. There is also no study that compares the properties and relative performance of

This article was previously circulated under the title, "Multivariate ARCH Modeling of Asset Returns." We are grateful to Stephen Buser, Rob Engle, Richard Harjes, Andrew Karolyi, Gautam Kaul, Rene Stulz, Ravi Jagannathan, and an anonymous referee for helpful suggestions. We also thank Gautam Kaul, Jennifer Conrad, and Mustafa Gultekin for providing the data used in this article. Address correspondence to Kenneth F. Kroner, Advanced Strategies and Research Group, Barclays Global Investors, 45 Fremont St., San Francisco, CA 94105.

The Review of Financial Studies Winter 1998 Vol. 11, No. 4, pp. 817–844

the existing multivariate covariance models. In contrast, many univariate time-varying volatility models have been carefully examined and compared. For examples, see Pagan and Schwert (1990), Engle and Ng (1993), Amin and Ng (1994a, b), and Kim and Kon (1994).)

Moreover, none of the popular multivariate models capture the asymmetric volatility effect — a phenomenon that a negative return shock (unexpected price drop) will lead to a higher subsequent volatility than a positive return shock (unexpected price increase) of the same magnitude. In contrast, there are several univariate models that capture this property. For examples, see Nelson (1990), Glosten, Jagannathan, and Runkle (1993), and Engle and Ng (1993). In addition, none of the popular multivariate models allow for an asymmetric effect in the covariance. Such a phenomenon is likely if there is an asymmetric effect in the variance. For instance, if the asymmetric effect is caused by a leverage effect — an increase in the riskiness of the stock due to an increase in the debt:equity ratio of the firm following a price drop — then the change in financial leverage in the firm should also change the covariance between its stock return and the stock returns of other firms that have not experienced a change in financial leverage. Alternatively, if the asymmetric effect in volatility is caused by an increase in the information flow following bad news, then the covariance between stock returns should be affected because there will be a change in the relative rate of information flow across firms.

Furthermore, there are few specification tests for the multivariate models. Perhaps because of this, most multivariate time-varying covariance models are often chosen on an ad hoc basis. In many cases, the ease of estimation is the primary factor affecting the choice of model. Robustness checks that analyze the sensitivity of economic results to model specification are often omitted.

One objective of this article is to fill these gaps. Specifically, we illustrate how the existing time-varying covariance models differ from each other. We propose a way to evaluate the specification of these models, and demonstrate that the choice of a multivariate volatility model can affect estimated portfolio weights and hedge ratios. We also demonstrate that the existing models do not capture some important stylized facts about asymmetric volatility relationships, and in the spirit of Hentschel (1995) propose a model that encompasses the existing models while modeling these stylized facts.

A second objective of this article is to use a general approach to study the time-varying covariance between the stock returns of large and small firms. Conrad, Gultekin, and Kaul (1991) found that shocks to large firm returns are important to the future dynamics of their own volatility as well as the volatility of small firm returns. Conversely, shocks to small firms have no impact on the behavior of the volatility of large firms. Our application furthers this line of research by examining how robust the Conrad, Gultekin, and Kaul result is with respect to model specification by using our proposed encompassing model; extending the study to cover the differential effects of large and small firm shocks on the covariance; and extending the model to allow for the asymmetric effects of positive and negative shocks on both volatility and covariance.

The organization of this article is as follows: In Section 2, alternative approaches to modeling time-varying covariances are reviewed. In Section 3, these models are applied to a dataset containing a large-firm portfolio return series and a small-firm portfolio return series. Various summary statistics and graphical techniques are used to highlight the differences between the existing models on the assumed dynamics of large- and small-firm return volatilities. In Section 4, a formal testing approach is introduced. The tests are applied to the models to evaluate their ability to describe the dynamic behavior of the covariance between large- and small-firm returns. In Section 5, an encompassing modeling approach is introduced. The model, which nests many existing multivariate GARCH models as special cases, is further extended to allow for asymmetric effects in the variance and covariance. The general model is then applied to study the time-varying covariance between large- and small-firm returns. In Section 6 we illustrate the importance of our results in portfolio selection and dynamic hedging applications. Section 7 concludes the article.

1. Alternative Approaches to Modeling Time-Varying Covariances

Multivariate GARCH models are among the most widely used time-varying covariance models. These include the VECH model of Bollerslev, Engle, and Wooldridge (1988), the constant correlation (CCORR) model of Bollerslev (1990), the factor ARCH (FARCH) model of Engle, Ng, and Rothschild (1990), and the BEKK model of Engle and Kroner (1995). These models have been applied to many markets and many asset pricing and investment problems. For an extensive summary, see the survey by Bollerslev, Chou, and Kroner (1992).

To describe these models, we adopt the following notation:

R_{it} : the rate of return of asset i from time $t-1$ to time t.

μ_{it}: the expected return of asset i given all information at time $t-1$.

ϵ_{it}: the unexpected return of asset i ($\epsilon_{it} = R_{it} - \mu_{it}$).

h_{iit}: the conditional variance of R_{it} given all information at time $t-1$.

h_{ijt}: the conditional covariance between R_{it} and R_{jt} given all information at time $t-1$.

H_t: the conditional covariance matrix ($H_t = [h_{ijt}]$).

1.1 The VECH model

The VECH model is characterized by the following equation:[1]

$$h_{ijt} = \omega_{ij} + \beta_{ij} h_{ijt-1} + \alpha_{ij} \varepsilon_{it-1} \varepsilon_{jt-1} \qquad \text{for all } i, j = 1, \ldots, N, \quad (1)$$

where ω_{ij}, β_{ij}, and α_{ij}, $i = 1, \ldots, N$ and $j = 1, \ldots, N$ are parameters. An advantage of the VECH model is that it is easy to understand — it is simply an ARMA model for $\varepsilon_{it}\varepsilon_{jt}$. Provided that $\beta_{it} \in (0, 1)$ for all i and j, Equation (1) can also be rewritten as

$$h_{ijt} = \sigma_t^* + \alpha_{ij} \Sigma_{t=1,t} \beta_{ij}^{\tau-1} (R_{it-1} - \mu_{it-1})(R_{jt-1} - \mu_{jt-1}). \qquad (2)$$

where $\sigma_t^* = [\beta^t h_0 + \omega_{ij} \sigma_{t=0,t-1} \beta_{ij}^{\tau}]$. That is, except for an adjustment term σ_t^* which ensures that the expectation of h_{ijt} is the unconditional covariance between returns i and j, the VECH model estimates the covariance as a geometrically declining weighted average of past cross products of unexpected returns, with lower weights for older observations.

The VECH model has two practical shortcomings. First, it has $\frac{3}{2}N(N+1)$ parameters. So a 20-asset model will have 630 parameters.[2] A second implementation problem is that the model might not yield a positive definite covariance matrix unless nonlinear inequality restrictions are imposed that govern the rates at which the weights are reduced for older observations [see Kraft and Engle (1983)]. Without these restrictions, the weights for the covariance terms could decline too slowly relative to the weights for the variance terms, causing the off-diagonal terms of the estimated covariance matrix to become too big relative to the diagonal terms, thus causing the matrix to be nonpositive definite.

1.2 The BEKK model

The BEKK model represents a solution to the positive definiteness problem. It is characterized by the following equation:

$$H_t = \Omega + B' H_{t-1} B + A' \varepsilon_{t-1} \varepsilon_{t-1}' A, \qquad (3)$$

where Ω, A, and B are $N \times N$ matrices, with Ω symmetric and positive definite. In this model, the ijth covariance can be written as

$$h_{ijt} = \omega_{ij} + \text{cov}_{t-1}(\varepsilon_{rt}, \varepsilon_{st}) + \varepsilon_{p,t-1} \varepsilon_{q,t-1}$$

where ε_p, ε_q, ε_r, and ε_s are the unexpected shocks to portfolios p, q, r, and s, and ω_{ij} is the ijth element of Ω. The weights in portfolios p and q come from the ith and jth columns of the A matrix, and the weights in

[1] Some, including Bollerslev, Engle, and Wooldridge (1988), refer to this model as the "diagonal VECH" model.

[2] 210 of these are contained in the estimate of Ω.

portfolios r and s come from the ith and jth columns of the B matrix. If we restrict $B = \kappa A$ for some scalar κ, then we can interpret this model as one in which there are N factors (or portfolios of assets) driving the conditional covariance matrix.

This model assumes the conditional covariance matrix of asset returns is determined by the outer product matrices of the vector of past return shocks. Because the second and third terms on the right-hand side of Equation (3) are expressed in quadratic forms, the positive definiteness of the conditional covariance matrix of asset returns is guaranteed, provided that Ω is positive definite. While this model overcomes this major weakness of the VECH model, it still has $(5/2)N^2 + (N/2)$ parameters. So for $N = 20$, the BEKK model has 1010 parameters, seriously restricting the applicability of the BEKK model to many financial systems.

1.3 The factor ARCH (FARCH) model

The FARCH model was constructed to solve this large-system applicability problem, while retaining the benefits of positive definiteness. The model is characterized by the following equation:

$$H_t = \Omega + \lambda\lambda'[\beta w' H_{t-1} w + \alpha(w'\varepsilon_{t-1})^2], \tag{4}$$

where λ and w are $N \times 1$ vectors, α and β are scalars, and Ω is a symmetric positive $N \times N$ matrix. This model is a special case of the BEKK model in which the A and B matrices are rank one and equal except for a scale factor. More specifically, the BEKK model becomes the FARCH model if $A = \sqrt{\alpha} w\lambda'$ and $B = \sqrt{\beta} w\lambda'$. For an N variable system, the number of parameters in this model is $(1/2)N^2 + (5/2)N + 2$, which is significantly less than that of the general BEKK model. If $N = 20$, this model has 252 parameters, 210 of which are used for estimating Ω.

Let $R_{pt} = w'R_t$, where $R_t = (R_{1t}, \ldots, R_{Nt})'$ is the return to a portfolio formed with a vector of weights w. The return shock of this portfolio at time $t-1$ is $\varepsilon_{pt-1} = w'\varepsilon_{t-1}$, and the conditional variance of this portfolio at time $t-1$ is $h_{pt} = w'H_t w$. Using h_{pt} and ϵ_{pt-1} the FARCH model can be rewritten in the following alternative form:

$$h_{ijt} = \omega_{ij} + \lambda_i\lambda_j h_{pt} \qquad \text{for all } i, j = 1, \ldots, N \tag{5a}$$

$$h_{pt} = \omega_p + \beta H_{pt-1} + \alpha\varepsilon_{pt-1}^2, \tag{5b}$$

where $\omega_p = w'\Omega w$, $\sigma_{ij} = \Omega_{ij} - \lambda_i\lambda_j\omega_p$, ω_{ij} is the *(i,j)*th element of Ω. Intuitively this FARCH model assumes that there is a single portfolio whose variance is driving all the conditional variances and covariances of asset returns. This common portfolio, or factor, follows a GARCH process. In the one factor model in Ng, Engle, and Rothschild (1992), R_{pt} is taken to be the market return. Under this assumption, the entire conditional covariance matrix of stock returns is driven by the conditional variance of the market portfolio.

The key difference between the FARCH model and the BEKK model is the number of factors that are driving the conditional covariance matrix. If there are N factors driving the covariance matrix, the BEKK model is implied. If there is one factor, the FARCH model is implied. The extension to k factors, $1 < k < N$, is obvious, and is suggested in Ng, Engle, and Rothschild (1992).

1.4 The constant correlation (CCORR) model

The CCORR is another way to parsimoniously model a time-varying covariance matrix. It restricts the conditional covariance between two asset returns to be proportional to the product of the conditional standard deviations. In this model the conditional correlation coefficient of the two asset returns is time invariant. Specifically the model is

$$h_{iit} = \omega_{ii} + \beta_{ii} h_{iit-1} + \alpha_{ii}\varepsilon_{it-1}^2 \quad \text{for all } i = 1, \ldots, N \tag{6a}$$

$$h_{ijt} = \rho_{ij}(\sqrt{h_{iit}}\sqrt{h_{jjt}}) \quad \text{for all } i \neq j. \tag{6b}$$

The CCORR model is positive definite if and only if the correlation matrix $[\rho_{ij}]$ is positive definite. The number of parameters in this model is only $(1/2)N^2 + (7/2)N$. For $N = 20$, this is 270.

2. Properties of the Four Multivariate GARCH Models

Each of the four models presented in the previous section implicitly imposes a different set of restrictions on the variance and covariance processes. We illustrate these differences with a bivariate system of large-firm and small-firm portfolio returns obtained from the dataset used by Conrad, Gultekin, and Kaul (1991).[3] The sample period is from July 1962 to December 1988, for a total of 1371 weekly observations.

Since we are not interested in the behavior of the time-varying mean returns in this study, we simply model the mean of the return vector as a 10th order vector autoregression (VAR) with 10 lags of a threshold term. Specifically, the model is (for $i = 1, 2$)

$$R_{it} = \delta_{i0} + \Sigma_{j=1,2}\Sigma_{\tau=1,10}[\delta_{j_\tau} R_{jt_{-\tau}} + d_{j_t} \max(-R_{jt_{-\tau}}, 0)] + \varepsilon_{it}. \tag{7}$$

Throughout this article, $i = 1$ refers to the small-firm portfolio and $i = 2$ refers to the large-firm portfolio. The threshold terms are added to

[3] Since our main focus is on the effect of last period's return shocks on current volatility, we replaced the GARCH terms in all the models except CCORR with $[\beta_{ij}h_{ijt-1}]$, where $\beta_{ij} = \beta_{ji}$. We chose not to make this modification to the CCORR model because it would invalidate the essential feature of the CCORR model, which is that correlations are constant.

ensure that any asymmetric effects found in the variances and covariances are not caused by a misspecification in the mean. The estimation is done in two steps. First we estimate the mean equation to get the residuals ε_{1t} and ε_{2t}, then we estimate the conditional covariance matrix parameters using maximum likelihood, treating ε_{1t} and ε_{2t} as observable data.[4] The block diagonality of the information matrix under this setup guarantees that consistency and efficiency are not lost in such a procedure.

The four multivariate GARCH models give very different variance and covariance estimates. For evidence of this, consider first the summary statistics of the variance and covariance series obtained from the four models. These summary statistics, including the mean, standard deviation, minimum, and maximum, are reported in Table 1. The covariances obtained from the BEKK model and the FARCH model tend to be slightly higher and more volatile than those from the VECH model and the CCORR model. The BEKK model in particular produces a broad range of covariance estimates, as evidenced from the large maximum-minimum range. Focusing next on the variance estimates, the VECH and CCORR model estimates for the small-firm variance series are more volatile than those from the FARCH and the BEKK models. In contrast, the VECH and CCORR estimates of the large-firm variance series are less volatile than the FARCH and the BEKK estimates.

For further evidence that the four models can give very different variance and covariance estimates, consider the correlations between the estimates from the different models, reported in Table 2. The first panel of Table 2 gives the correlations between the small-firm variance estimates obtained from the four models. The second panel gives the correlations between the large-firm variance estimates. The third panel gives the correlations between the covariance estimates obtained from the four models. The correlations in panel 2 all exceed 0.99, suggesting that the four models give very similar large-firm variance estimates. Therefore, if we are only interested in estimating the large-firm variance in this dataset, the choice of models is unimportant. This conclusion does not hold for the small-firm variance and the covariance. The small-firm variance estimates obtained from the FARCH and BEKK models are not highly correlated with those obtained from the VECH and CCORR models. In fact, the correlation between the small-firm variance estimates obtained from the FARCH model and the CCORR model is only 0.366. This suggests that these models are producing substantially different small-firm variance estimates. Similarly, the correlations between the covariance estimates from all combinations of models are less than 0.89.

[4] This follows the two-step approach of Pagan and Schwert (1990), Gallant, Rossi, and Tauchen (1992), and Engle and Ng (1993).

Table 1
Summary statistics of estimated variance and covariance series

Model	Variable	Mean	SD	Minimum	Maximum
Small-firm variance					
	ε_{1t}^2	7.43	28.97	0.00	825.45
VECH	h_{11t}	7.73	6.31	3.82	104.98
CCORR	h_{11t}	7.91	7.08	3.54	117.60
FARCH	h_{11t}	7.51	4.14	2.87	33.24
BEKK	h_{11t}	7.53	4.91	2.64	50.67
Large-firm variance					
	ε_{2t}^2	3.93	8.12	0.00	147.87
VECH	h_{22t}	3.89	2.58	1.10	21.53
CCORR	h_{22t}	3.96	2.78	0.98	23.49
FARCH	h_{22t}	4.12	3.18	0.74	27.08
BEKK	h_{22t}	4.16	3.20	0.76	26.15
Covariance					
	$\varepsilon_{1t}\varepsilon_{2t}$	2.61	8.59	−45.44	167.55
VECH	h_{12t}	2.33	1.49	0.30	17.59
CCORR	h_{12t}	2.47	1.48	0.91	15.60
FARCH	h_{12t}	2.79	2.42	0.30	21.71
BEKK	h_{12t}	2.77	2.66	0.25	30.71

This table gives summary statistics for the variance and covariance series estimated from the four multivariate GARCH models discussed in the article. All four models were applied to the same dataset of large-firm and small-firm portfolio returns. ε_{1t} is the return shock to the small-firm portfolio and ε_{2t} is the return shock to the large-firm portfolio. h_{11t} is the estimated variance of the small-firm portfolio returns. h_{22t} is the estimated variance of the large-firm portfolio returns. h_{12t} is the estimated covariance between the small-firm and large-firm portfolio returns.

Judging from these results, it seems clear that the four models can produce substantially different covariance matrix estimates. These differences could affect the results of asset pricing exercises and portfolio management applications, making the choice of the model very important. Given this conclusion, the obvious next question is: What causes the differences in the variance and covariance estimates? To answer this, we introduce a multivariate generalization of the graphical "news impact curve" from Engle and Ng (1993). Univariate applications, which involve plotting the conditional variance against last period's shocks, appear in Engle and Ng (1993) and Hentschel (1995). The multivariate generalization plots the conditional variance and covariance against large- and small-firm shocks from the last period, holding the past conditional variances and covariances constant at their unconditional sample mean levels. We will call these "news impact surfaces."

Specifically, let z_{t-1} denote the vector of inputs (known at time $t-1$) for the determination of h_{ijt}, excluding ε_{it-1} and ε_{jt-1}. Also, let Z denote the unconditional mean of z_{t-1}. The news impact surface for h_{ijt} is the

Table 2
Correlations of estimated second moments from alternative models

Panel 1: Small-firm portfolio variance series

	VECH	CCORR	FARCH	BEKK
VECH	1.000			
CCORR	0.999	1.000		
FARCH	0.367	0.365	1.000	
BEKK	0.642	0.640	0.834	1.000

Panel 2: Large-firm portfolio variance series

	VECH	CCORR	FARCH	BEKK
VECH	1.000			
CCORR	0.999	1.000		
FARCH	0.999	0.999	1.000	
BEKK	0.954	0.954	0.955	1.000

Panel 3: Covariance series

	VECH	CCORR	FARCH	BEKK
VECH	1.000			
CCORR	0.876	1.000		
FARCH	0.748	0.785	1.000	
BEKK	0.752	0.719	0.885	1.000

Panel 1 of this table gives the correlation matrix of the small-firm variance series estimated from the four multivariate GARCH models discussed in the text. Panel 2 gives the correlation matrix of the large-firm portfolio variance series estimated from these same models. Panel 3 gives the correlation matrix of the covariance series estimated from these models.

three-dimensional graph of the function:

$$h_{ijt} = h_{ij}(\varepsilon_{it-1}, \varepsilon_{jt-1}; z_{t-1} = Z).$$

The news impact surfaces generated by the four models for the covariance h_{12t}, the small-firm variance h_{11t}, the large-firm variance h_{22t}, and the correlation $h_{12t}/\sqrt{h_{11t}}\sqrt{h_{22t}}$ are plotted in Figures 1, 2, 3, and 4, respectively.

Figure 1 shows that, in contrast to the univariate model comparisons in Hentschel (1995), the different models imply substantially different news impact surfaces for the covariance, even though they are fitted to the same dataset. These differences are caused by the different functional forms assumed by each of the specifications. For instance, under the VECH model, past return shocks to the large- and small-firm portfolios enter into the covariance equation in the cross-product form $(\varepsilon_{1t-1}\varepsilon_{2t-1})$. Hence when the shocks are both large but of opposite signs, the covariance can be small or even negative. This is evident from the saddle shape of the VECH news impact surface in Figure 1. On the other hand, under the CCORR model, the covariance is proportional to the product of the standard deviations. Thus when shocks are large, regardless of their signs, the standard deviations and hence the covariance will be large. Therefore the news impact surface for

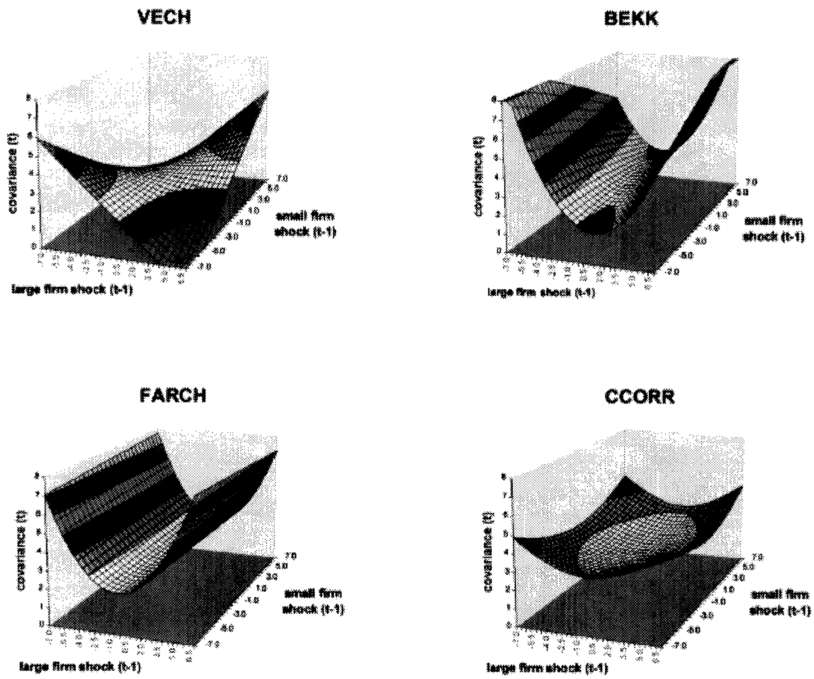

Figure 1
News impact surfaces for covariances
The figures give the news impact surfaces for the covariance between small-firm and large-firm return shocks under the VECH model, the BEKK model, the FARCH model, and the CCORR model.

the CCORR model is bowl-shaped. For the FARCH model, since the covariance is proportional to the variance of a factor which is loaded primarily on the large-firm portfolio, the FARCH covariance news impact surface is a U-shaped surface along the axis for the large-firm return shock.

These general shapes are not specific to the dataset analyzed here. For any application, the impact on covariances of opposite-signed shocks in the CCORR model will be substantially different than in the VECH model. Also, the news impact surface for covariances from a FARCH model will always be U-shaped, with the data determining the direction the parabola points. Ignoring these differences when choosing which multivariate GARCH model to employ could lead to a seriously misspecified model.

Figure 2 shows that there are also significant differences between the news impact surfaces for the variance of the small-firm portfolio obtained from the four different models. The VECH model and the CCORR model restrict the variance of the small-firm portfolio to depend only on the square of the return shock for the small-firm portfolio. This restriction forces the

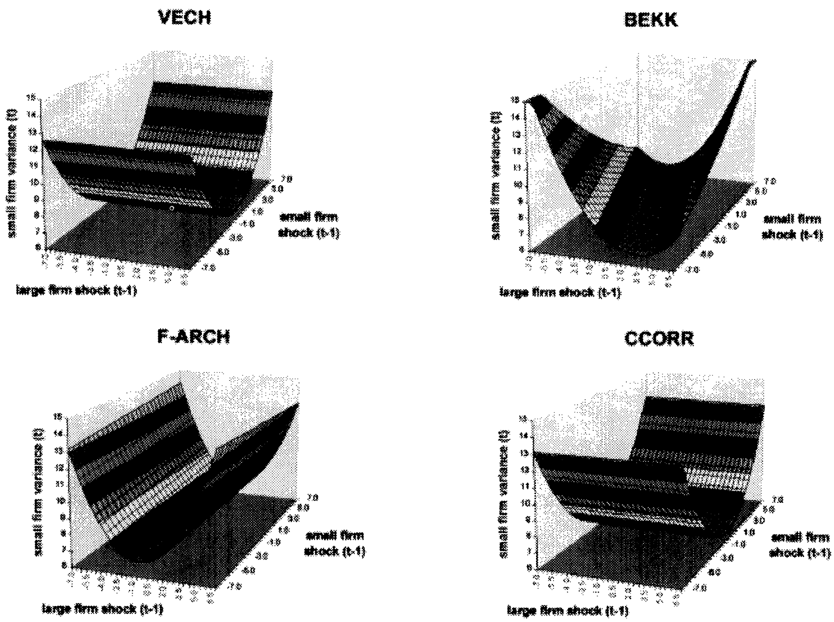

Figure 2
News impact surfaces for small-firm variances
The figures give the news impact surfaces for the small-firm variances under the VECH model, the BEKK model, the FARCH model, and the CCORR model.

large-firm return shock to have no effect on the small-firm variance, thereby restricting the news impact surfaces to be flat along any line parallel to the large-firm axis. On the other hand, the news impact surface implied by the FARCH model suggests that it is the large-firm return shock which has the biggest impact on the variance of the small firm portfolio, while the news impact surface from the BEKK model suggests that both large- and small-firm shocks matter.

Notice also that the news impact surface generated for the small-firm variance (Figure 2) and for the covariance (Figure 1) from the FARCH model are identical, except for a scale factor. This restriction must hold because the same portfolio drives all the elements of the FARCH covariance matrix.

Consider next the news impact surfaces for the large-firm variances in Figure 3. The four models give almost identical large-firm news impact surfaces, suggesting that when modeling large-firm variances, these four models are very similar. But notice also that the large-firm news impact surfaces from the CCORR and VECH models point along the opposite axes from the small-firm variance surfaces for these same models. This result

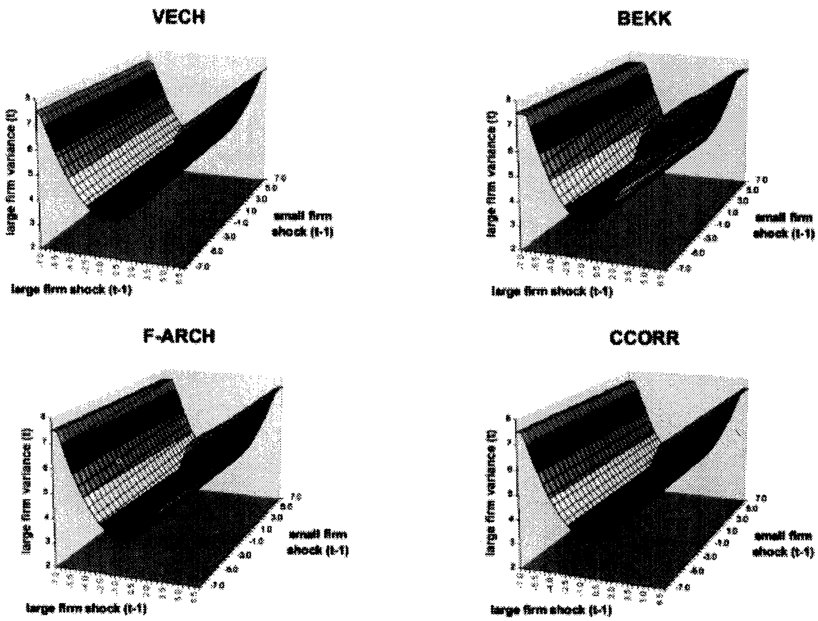

Figure 3
News impact surfaces for large-firm variances
The figures give the news impact surfaces for the large-firm variances under the VECH model, the BEKK model, the FARCH model, and the CCORR model.

is caused by restrictions in the underlying models. Similarly, restrictions in the FARCH model force the FARCH news impact surfaces to all point along the same axis.

Finally, we provide the news impact surface for the correlations in Figure 4. The surface from the CCORR model is flat, because correlations in this model are not a function of the information set. However, notice that the news impact surfaces for the correlations from the FARCH model are also quite flat. This is not surprising, because in a one-factor model the correlation is

$$\rho_t = \frac{\sigma_{ij} + \lambda_i \lambda_j h_{pt}}{\sqrt{(\sigma_{ij} + \lambda - i\lambda_i h_{pt})(\sigma_{jj} + \lambda_j \lambda_j h_{pt})}}.$$

The same dynamics are driving both the numerator and denominator of this expression, leaving only minimal dynamics in the ratio. In a two-factor model, one would expect more dynamics in the correlations.

Based on the above observations, we can conclude that (i) the different covariance matrix models impose significantly different restrictions on the dynamic behavior of the variances, covariances and correlations; and (ii) an

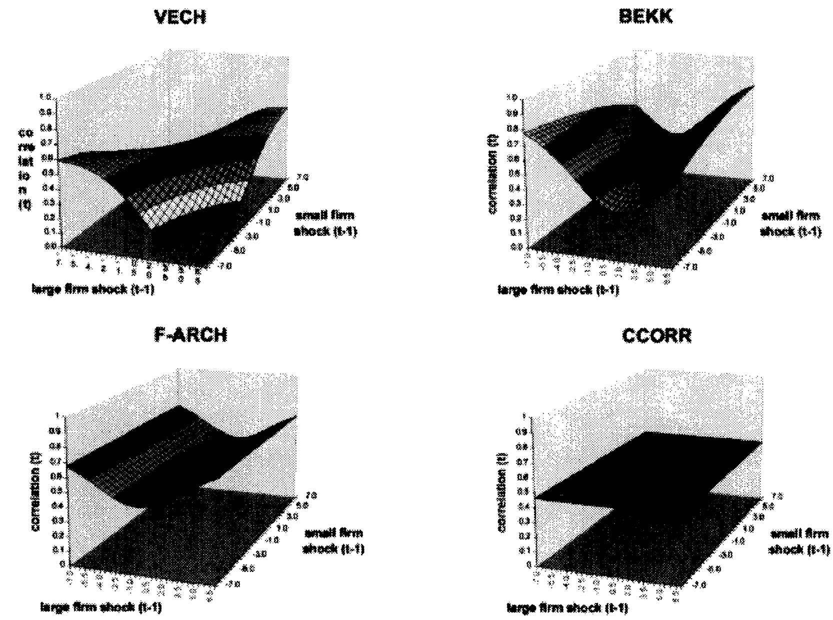

Figure 4
News impact surfaces for correlations
The figures give the news impact surfaces for the correlation between small-firm shocks and large-firm shocks under the VECH model, the BEKK model, the FARCH model, and the CCORR model.

important difference between the models is the way they allow past shocks of asset returns to affect the variances and covariances. These results cannot be overemphasized and are particularly important for applications that rely crucially on covariance estimates, such as portfolio choice problems, hedging problems, and asset pricing problems. Very careful consideration should go into the choice of a multivariate GARCH model before estimation is conducted, and thorough specification testing on the estimated model is essential before conclusions are made based on the model. It is therefore valuable to have a set of specification tests available, and it is to this that we now turn our attention.

3. Robust Conditional Moment Tests

To test the validity of a model, a natural approach is to compare the ex post cross-product matrix of the vector of residuals to the estimated covariance matrix. For the covariance case, this is like superimposing a scatter plot of the cross-product of residuals (a plot of $\varepsilon_{it}\varepsilon_{jt}$ against ε_{it-1} and ε_{jt-1}) on the three-dimensional graph for the covariance news impact surface. Recognizing that the covariance news impact surface is a graph of h_{ijt} against

ε_{it-1} and ε_{jt-1} and recognizing that the unconditional expectation of $\varepsilon_{it}\varepsilon_{jt}$ is h_{ijt}, we can test the models by measuring the vertical distance between $\varepsilon_{it}\varepsilon_{jt}$ and h_{ijt} and studying whether these distance measures follow some specific patterns. For instance, if the model gives a covariance news impact surface that is too low whenever ε_{it-1} is negative, then the vertical distance between $\varepsilon_{it}\varepsilon_{jt}$ and h_{ijt} will tend to be positive when ε_{it-1} is negative.

Based on this idea, we define a "generalized residual" u_{ijt} to be the (i,j)th element of $\varepsilon_t\varepsilon_t' - H_t$ so that $u_{ijt} = \varepsilon_{it}\varepsilon_{jt} - h_{ijt}$. A generalized residual is the distance between a point on the scatter plot of $\varepsilon_{it}\varepsilon_{jt}$ from a corresponding point on the news impact surface. If the model is correct, $E_{t-1}(u_{ijt}) = 0$, thus u_{ijt} should be uncorrelated with any variable known at time $t-1$. This observation gives us a natural way to identify misspecification by examining whether u_{ijt} is correlated with variables known at time $t-1$. These variables are called misspecification indicators. The choice of misspecification indicators is very important because different indicators can target different forms of misspecification. Inappropriate choice of misspecification indicators will reduce the ability of the test to detect misspecification.[5]

In this regard, the graphical representation of the news impact surface provided above has provided useful hints for finding suitable misspecification indicators. Knowing that a major difference between the models is their asymmetric property, a beneficial approach is to partition the $(\varepsilon_{it-1}, \varepsilon_{jt-1})$ space in a way that can highlight the asymmetric property. Misspecification indicators can then be built based on this partition.

A natural way is to partition the $(\varepsilon_{it-1}, \varepsilon_{jt-1})$ space into four quadrants corresponding to the following sign combinations of $(\varepsilon_{it-1}, \varepsilon_{jt-1})$: $(-,-)$, $(-,+)$, $(+,-)$, and $(+,+)$. Let $I(\cdot)$ be an indicator function that equals one if the argument is true and zero otherwise. The misspecification indicators corresponding to such a partition are

$$\begin{aligned}
x_{1t-1} &= I(\varepsilon_{it-1} < 0; \varepsilon_{jt-1} < 0) \\
x_{2t-1} &= I(\varepsilon_{it-1} < 0; \varepsilon_{jt-1} > 0) \\
x_{3t-1} &= I(\varepsilon_{it-1} > 0; \varepsilon_{jt-1} < 0) \\
x_{4t-1} &= I(\varepsilon_{it-1} > 0; \varepsilon_{jt-1} > 0).
\end{aligned}$$

Related to these, we can consider the "sign indicators,"

$$\begin{aligned}
x_{5t-1} &= I(\varepsilon_{it-1} < 0) \\
x_{6t-1} &= I(\varepsilon_{jt-1} < 0),
\end{aligned}$$

[5] See Brenner, Harjes, and Kroner (1996) for a detailed description of these types of tests and an illustration in a univariate GARCH framework.

which will allow us to test for traditional leverage-type asymmetries in the data.

As pointed out by Engle and Ng (1993), the magnitudes of the shocks can also play an important role. Furthermore, the effect of the size of a shock on the variances and covariances might also depend on the sign of the shock and possibly the sign of other shocks. To capture such possibilities, we scale the sign indicators by the size of the shocks. This yields another set of misspecification indicators:

$$\begin{aligned} x_{7t-1} &= \varepsilon_{it-1}^2 I(\varepsilon_{it-1} < 0) \\ x_{8t-1} &= \varepsilon_{it-1}^2 I(\varepsilon_{jt-1} < 0) \\ x_{9t-1} &= \varepsilon_{jt-1}^2 I(\varepsilon_{it-1} < 0) \\ x_{10t-1} &= \varepsilon_{jt-1}^2 I(\varepsilon_{jt-1} < 0). \end{aligned}$$

When $N = 1$, the entire set of 10 indicators reduces to indicators that match the sign and size-bias tests introduced in Engle and Ng (1993).

To complete the testing design, we borrow the robust conditional moment test framework of Wooldridge (1990). A test statistic that is robust to the conditional distribution used when estimating the multivariate GARCH model is constructed as

$$\mathbf{C}_{rcm} = [(1/T)\Sigma_{t=1,T} u_{ijt}\lambda_{gt-1}]^2[(1/T)\Sigma_{t=1,T} u_{ijt}^2\lambda_{gt-1}^2]^{-1}, \qquad (8)$$

where λ_{gt-1} is the residual from a regression of the misspecification indicator x_{gt-1} on the derivatives of h_{ijt} with respect to the parameters of the model. Under general regularity conditions, Wooldridge (1990) shows that $\mathbf{C}_{rcm}$ has an asymptotic $\chi^2(1)$ distribution.

The robust conditional moment test statistics can be computed easily from two auxiliary regressions. The first regression is x_{gt-1} on the derivatives of h_{ijt} with respect to all parameters of the null model. The second regression is a vector of ones on the product $u_{ijt}\lambda_{gt-1}$, where λ_{gt-1} is the residual from the first regression. The test statistic is T times the uncentered R^2 from the second regression.

The test statistics for the covariance between the small-firm and the large-firm returns are reported in panel A of Table 3.[6] To highlight the usefulness of the tests, the Ljung–Box tests for serial correlation in the normalized cross-product of residuals, $\varepsilon_{it}\varepsilon_{jt}/h_{ijt}$, are reported in panel B of Table 3. The Ljung–Box test is a popular diagnostic for models with time-varying conditional second moments because it addresses whether the model has adequately captured the serial correlation in the second moments.

As can be seen from Table 3, the Ljung–Box tests do not reject any of

[6] The final column of Table 3, labeled ADC, will be discussed shortly.

Table 3
Diagnostic tests for covariance specification

Panel A: Robust conditional moment tests

	VECH	CCORR	FARCH	BEKK	ADC
$I(\varepsilon_{1t-1} < 0)$	4.85	5.32	4.91	4.85	—
$I(\varepsilon_{2t-1} < 0)$	16.22	17.63	16.34	16.27	5.48
$I(\varepsilon_{1t-1} < 0; \varepsilon_{2t-1} < 0)$	5.88	10.95	6.40	6.45	—
$I(\varepsilon_{1t-1} < 0; \varepsilon_{2t-1} > 0)$	—	6.84	—	—	—
$I(\varepsilon_{1t-1} > 0; \varepsilon_{2t-1} < 0)$	—	—	—	4.10	—
$I(\varepsilon_{1t-1} > 0; \varepsilon_{2t-1} > 0)$	11.09	10.46	12.06	12.67	—
$\varepsilon^2_{1t-1} I(\varepsilon_{1t-1} < 0)$	—	—	—	—	—
$\varepsilon^2_{1t-1} I(\varepsilon_{2t-1} < 0)$	3.99	6.45	5.01	5.54	—
$\varepsilon^2_{2t-1} I(\varepsilon_{1t-1} < 0)$	—	4.03	4.24	5.40	—
$\varepsilon^2_{2t-1} I(\varepsilon_{2t-1} < 0)$	4.41	4.95	5.11	6.29	—

Panel B: Ljung–Box tests for serial correlation in $\varepsilon_{1t}\varepsilon_{2t}/h_{12t}$

	VECH	CCORR	FARCH	BEKK	ADC
$Q(6)$	3.81	6.18	1.18	2.50	5.63
$Q(12)$	6.42	10.39	4.28	5.58	7.90
$Q(18)$	9.39	13.37	9.07	9.88	10.43

Panel A gives the robust conditional moment test statistics for each of the five models estimated. The misspecification indicators are listed in the first column and the remaining five columns give the test statistics for each of the five models. This statistic is distributed χ^2_1 and has a 95% critical value of 3.84. Only those statistics that are significant at the 5% level are reported. ε_{1t-1} is the return shock to the small-firm portfolio, and ε_{2t-1} is the return shock to the large-firm portfolio. Panel B gives the Ljung–Box test statistic for serial correlation in the standardized cross-product of residuals from these five models. $Q(r)$ is the Ljung–Box statistic for rth order serial correlation. The 5% critical levels for $Q(6)$, $Q(12)$, and $Q(18)$ are 12.6, 21.0, and 36.4, respectively.

the models. This is hardly surprising given that we rarely see Ljung–Box tests rejecting any variations of GARCH models in the literature. However, looking at our robust conditional moment tests, the message is very different. Each model is strongly rejected. The test statistics indicate that all models fail to capture the asymmetric response of covariance to both large-firm and small-firm portfolio shocks. As evidenced by the strong test statistics arising from the $I(\varepsilon_{2t-1} < 0)$ indicator, the models are especially bad at capturing the asymmetric relationship between covariances and shocks to the large-firm portfolio. Also, judging from the test statistics, there are more rejections when the shocks are large than when the shocks are small, suggesting that the size of the shocks matters.

There are more rejections when both ε_{1t} and ε_{2t} are negative and when both are positive. There are two potential explanations for this. First, this might be due to the high correlation between small- and large-firm returns (their correlation is 48%), which leads to more observations and therefore higher power in the $(+, +)$ and $(-, -)$ quadrants.[7] And two, it could arise

[7] About one-third of the observations are in each of the $(+, +)$ and $(-, -)$ quadrants, and about one-sixth

because the asymmetric predictability is more pronounced when the common shock is large. Intuitively, any big shock (whether positive or negative) is likely to be shared by both the large firm and small firm portfolios, leading to large return shocks of the same sign. So the observations that drive the asymmetries are likely to be located in the $(+, +)$ and $(-, -)$ quadrants.

One important conclusion from these results is that we should not place too much confidence in statistically insignificant Ljung–Box statistics when evaluating GARCH models. Even badly misspecified models can capture the serial correlation in the second moments and give insignificant Ljung–Box test statistics.

4. A General Dynamic Covariance (GDC) Model

The robust conditional moment test results call for a more general model with an ability to capture the asymmetric effects explicitly. Instead of working on extensions of each of the four models and comparing the large number of possible extensions, we adopt a more structured approach, similar to Hentschel (1995). First, we introduce a general dynamic covariance matrix model that can nest many of the existing models. Then we generalize this model to include asymmetric effects. The resulting asymmetric covariance matrix model encompasses various asymmetric extensions of the four models. Model selection is much easier under this approach. The specification of the basic encompassing model is as follows.

4.1 A general dynamic covariance (GDC) matrix model — Definition

$$H_t = D_t R D_t + \Phi \circ \Theta_t,$$

where $\circ$ is the Hadamard product operator (element-by-element matrix multiplication) and

$$\begin{aligned}
D_t &= [d_{ijt}], \quad d_{iit} = \sqrt{\theta_{iit}} \quad \text{for all } i, \quad d_{ijt} = 0 \quad \text{for all } i \neq j \\
\Theta_t &= [\theta_i jt] \\
R &= [r_{ij}], \quad \phi_{ii} = 0 \quad \text{for all } i \\
\Phi &= [\phi_{ij}], \quad \phi_{ii} = 0 \quad \text{for all } i \\
\theta_{ijt} &= \omega_{ij} + b' H_{t-1} b_j + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_j \quad \text{for all } i, j
\end{aligned} \tag{9}$$

and

$a_i, b_i, i = 1, \ldots, N$ are $N \times 1$ vectors of parameters,
ω_{ij}, ρ_{ij}, and $\phi_{ij}, i, j = 1, \ldots, N$ are scalars with $\Omega \equiv [\omega_{ij}]$ positive definite.

---

in each of the other two quadrants.

The GDC model has two components, the first term, $D_t R D_t$, is like the constant correlation model, but with the variance functions given by that of the BEKK model. The second term, $\Phi \circ \Theta_t$, has zero diagonal elements, but has off-diagonal elements given by the BEKK-type covariance functions, scaled by the ϕ_{ij} parameters.

To further clarify the model structure, we can write down the expressions for the elements of H_t directly. Specifically, under the GDC model,

$$\begin{aligned} h_{iit} &= \theta_{iit} \quad \text{for all } i \\ h_{ijt} &= \rho_{ij}\sqrt{\theta_{iit}}\sqrt{\theta_{jjt}} + \phi_{ij}\theta_{ijt} \quad \text{for all } i \neq j, \end{aligned}$$

where, $\theta_{ijt} i, j = 1, \ldots, N$ are given by the above BEKK form (Equation 9).

The GDC model is therefore a hybrid of the CCORR model structure and the BEKK model structure. An interesting property of this model is that it encompasses the four multivariate GARCH models discussed above. This encompassing result is given in Proposition 1.

Proposition 1. *Consider the following set of conditions:*

(i) $\rho_{ij} = 0$ *for all* $i \neq j$.
(ii) $a_i = \alpha_i \iota_i$ *and* $b_i = \beta_i \iota_i$ *for all* i, *where* ι_i *is the* i*th column of an* $N \times N$ *identity matrix, and* α_i *and* β_i, $i = 1, \ldots, N$ *are scalars.*
(iii) $\phi_{ij} = 0$ *for all* $i \neq j$.
(iv) $\phi_{ij} = 1$ *for all* $i \neq j$.
(v) $A = \alpha(w\lambda')$ *and* $B = \beta(w\lambda')$, *where* $A = [a_1, \ldots, a_N]$, $B = [b_1, \ldots, b_n]$, w *and* λ *are* $N \times 1$ *vectors, and* α *and* β *are scalars.*

The GDC model will reduce to the different multivariate GARCH models under different combinations of these conditions. Specifically, the GDC model will become a restricted VECH model (with the restrictions $\beta_{ij} = \beta_{ii}\beta_{jj}$ *and* $\alpha_{ij} = \alpha_{ii}\alpha_{jj}$*) under conditions (i) and (ii), the CCORR model under conditions (ii) and (iii), the BEKK model under conditions (i) and (iv), and the FARCH model under conditions (i), (iv) and (v).*

Proof. See Appendix for all proofs.

This encompassing property makes the GDC model a useful and attractive framework for estimating time-varying covariance matrices and for comparing and testing existing multivariate GARCH models. The generality of the GDC model also provides a natural ground for an extension that permits asymmetric effects in both variances and covariances. An extension of the GDC model following the approach of Glosten, Jagannathan, and Runkle (1993) is given below.

4.2 Asymmetric Dynamic Covariance (ADC) Matrix Model — Definition

Let $\eta_{it} = \max[0, -\varepsilon_{it}]$ and $\eta_t = [\eta_{1t}, \ldots, \eta_{Nt}]'$, the ADC model is defined as

$$H_t = D_t R D_t + \Phi \circ \Theta_t,$$

where

$$\begin{aligned}
D_t &= [d_{ijt}], \quad d_{iit} = \sqrt{\theta_{iit}} \text{ for all } i, \quad d_{ijt} = 0 \text{ for all } i \neq j \\
\Theta_t &= [\theta_{ijt}] \\
R &= [r_{ij}], \quad r_{ii} = 1 \text{ for all } i, \quad r_{ij} = \rho_{ij} \text{ for all } i \neq j \\
\Phi &= [\phi_{ij}], \quad \phi_{ii} \text{ for all } i \\
\theta_{ijt} &= \omega_{ij} + b_i' H_{t-1} b_j + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_j + g_i' \eta_{t-1} \eta_{t-1}' g_j \text{ for all } i, j \quad (10)
\end{aligned}$$

and

$$\begin{aligned}
&a_i, b_i, \text{ and } g_i, i = 1, \ldots, N \text{ are } N \times 1 \text{ vectors of parameters,} \\
&\omega_{ij}, \rho_{ij}, \text{ and } \phi_{ij}, i, j = 1, \ldots, N \text{ are scalars.}
\end{aligned}$$

The essential difference between the ADC model and the GDC model is the addition of the term $g_i' \eta_{t-1} \eta_{t-1}' g_j$ in the equation for θ_{ijt} (Equation 10). The asymmetric dynamic covariance matrix model nests some natural extensions of the four multivariate GARCH models that allow for asymmetric effects in the variances and covariances. These are summarized in Proposition 2. The proof for Proposition 2 follows directly from the proof for Proposition 1 and hence is not given to conserve space.

Proposition 2. *Consider the following set of conditions:*

(i) $\rho_{ij} = 0$ *for all* $i \neq j$
(ii′) $a_i = \alpha_i \iota_i$, $b_i = \beta_i \iota_i$, *and* $g_i = \gamma_i \iota_i$ *for all* i, *where* ι_i *is the* i*th column of an* $N \times N$ *identity matrix, and* α_i, β_i, *and* γ_i, $i = 1, \ldots, N$ *are scalars*
(iii) $\phi_{ij} = 0$ *for* $i \neq j$
(iv) $\phi_{ij} = 1$ *for all* $i \neq j$
(v′) $A = \alpha(w\lambda')$, $B = \beta(w\lambda')$, *and* $G = \gamma(w\lambda')$ *where* $A = [a_1, \ldots, a_N]$, $B = [b_1, \ldots, b_N]$, *and* $G = [g_1, \ldots, g_N]$, w *and* λ *are* $N \times 1$ *vectors, and* α, β, *and* γ *are scalars.*

The ADC model will reduce to the different asymmetric multivariate GARCH models under different combinations of these conditions. Specifically the ADC model will become an asymmetric VECH model under conditions (i) and (ii′), an asymmetric CCORR model under conditions (ii′) and (iii), an asymmetric BEKK model under conditions (i) and (iv), and an asymmetric FARCH model under conditions (i), (iv) and (v′).

The exact form of these specialized multivariate asymmetric models are given below:

Asymmetric VECH:

$$\begin{aligned} h_{iit} &= \omega_{ii} + \beta_i^2 h_{iit-1} + \alpha_i^2 \varepsilon_{it-1}^2 + \gamma_i^2 \eta_{it-1}^2 \\ &\quad \text{for all } i \quad \text{(GJR asymmetric variance function)} \\ h_{ijt} &= \phi_{ij}\omega_{ij} + \phi_{ij}\beta_i\beta_j h_{ijt} + \phi_{ij}\alpha_i\alpha_j\varepsilon_{it-1}\varepsilon_{jt-1} + \phi_{ij}\gamma_i\gamma_j\eta_{it-1}\eta_{jt-1} \\ &\quad \text{for all } i \neq j. \end{aligned}$$

Asymmetric CCORR:

$$\begin{aligned} h_{iit} &= \omega_{ii} + \beta_i^2 h_{iit-1} + \alpha_i^2 \varepsilon_{it-1}^2 + \gamma_i^2 \eta_{it-1}^2 \\ &\quad \text{for all } i \quad \text{(GJR asymmetric variance function)} \\ h_{ijt} &= \rho_{ij}\sqrt{h_{iit}}\sqrt{h_{jjt}} \quad \text{for all } i \neq j. \end{aligned}$$

Asymmetric BEKK:

$$H_t = \Omega + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B + G'\eta_{t-1}\eta_{t-1}'G$$

Asymmetric FARCH:

$$\begin{aligned} h_{ijt} &= \sigma_{ij} + \lambda_i\lambda_j h_{pt} \quad \text{for all } i \\ h_{pt} &= \omega_p + \beta h_{pt-1} + \alpha\varepsilon_{pt-1}^2 + \gamma\eta_{pt-1}^2 \end{aligned}$$

where

$$h_{pt} \equiv w'H_t w, \quad \varepsilon_{pt} \equiv w'\varepsilon_t, \quad \eta_{pt} = w'\eta_t, \quad \text{and} \quad \sigma_{ij} \equiv \omega_{ij} - \lambda_i\lambda_j w'\Omega w$$

The above multivariate asymmetric GARCH models are natural extensions of their standard counterparts. The asymmetric VECH and asymmetric CCORR models have variance functions given by the Glosten, Jagannathan, and Runkle (GJR) model instead of the standard GARCH(1, 1) model. The asymmetric VECH also allows a cross-product term of the negative shocks to determine the covariance. An implication of this is that the covariance will be higher when there is bad news for both firms. The asymmetric BEKK extends the standard one by having an additional quadratic form that is dependent on the outer product of the vector of negative return shocks. Finally, the asymmetric FARCH model utilizes a bad news portfolio, η_p, produced by taking a weighted average of the individual asset bad news with the weights being the original factor weights.

To examine the performance of the ADC model and to further study the

Table 4
ADC model estimation results

Parameter	Estimate	Standard Error
ω_{11}	0.218	0.040
ω_{12}	-0.595	0.436
ω_{22}	0.027	0.019
a_{11}	0.217	0.015
a_{12}	-0.083	0.025
a_{21}	-0.070	0.057
a_{22}	0.254	0.033
g_{11}	0.075	0.041
g_{12}	-0.008	0.037
g_{21}	0.436	0.044
g_{22}	0.373	0.048
ρ_{12}	0.381	0.151
ϕ_{12}	0.626	0.163
B_{11}	0.868	0.014
B_{12}	0.495	0.241
B_{22}	0.884	0.015

This table gives the maximum likelihood estimates for the ADC model:

$$\begin{aligned}
h_{iit} &= \theta_{iit} \quad \text{for all } i = 1, 2 \\
h_{12t} &= \rho_{12}\sqrt{\theta_{11t}}\sqrt{\theta_{22t}} + \phi_{12}\theta_{12t} \\
\theta_{ijt} &= \omega_{ij} + B_{ij}h_{ijt-1} + a_i'\varepsilon_{t-1}\varepsilon_{t-1}'a_j \\
&\quad + g_i'\eta_{t-1}\eta_{t-1}'g_j \quad \text{for all } i, j \\
\text{where, } a_i &= [a_{i1}, a_{i2}]', \quad g_i = [g_{i1}, g_{i2}]', \\
i &= 1, 2
\end{aligned}$$

Heteroskedasticity-consistent standard errors are reported. $i = 1$ refers to the small-firm portfolio and $i = 2$ refers to the large-firm portfolio.

dynamic relation between large- and small-firm returns, we apply the ADC model to our large- and small-firm return series. The estimation results are reported in Table 4. With these results, the first question to ask is whether the estimated ADC model would reduce to one of the more specialized models. For this, we can test the conditions given in Proposition 2. The t statistic for the hypothesis $\rho_{12} = 0$ is 2.52. Thus condition (i) is rejected at the 5% level. The t statistic for $\phi_{12} = 0$ is 3.84. So condition (iii) is also rejected. The t statistic for $\phi_{12} = 1$ is 2.29. Thus condition (iv) is also rejected. These results indicate that the estimated ADC model is statistically different from any one of the specialized models. Also, the rejection of $\phi_{12} = 0$ implies that there exists an asymmetry in the covariances that is not driven by the asymmetry in the variances.

We next address whether the asymmetric effects are important for the variances and covariances. For this, the results in Table 4 show that both g_{21} and g_{22} are statistically significant. Since the g_2 vector captures the negative

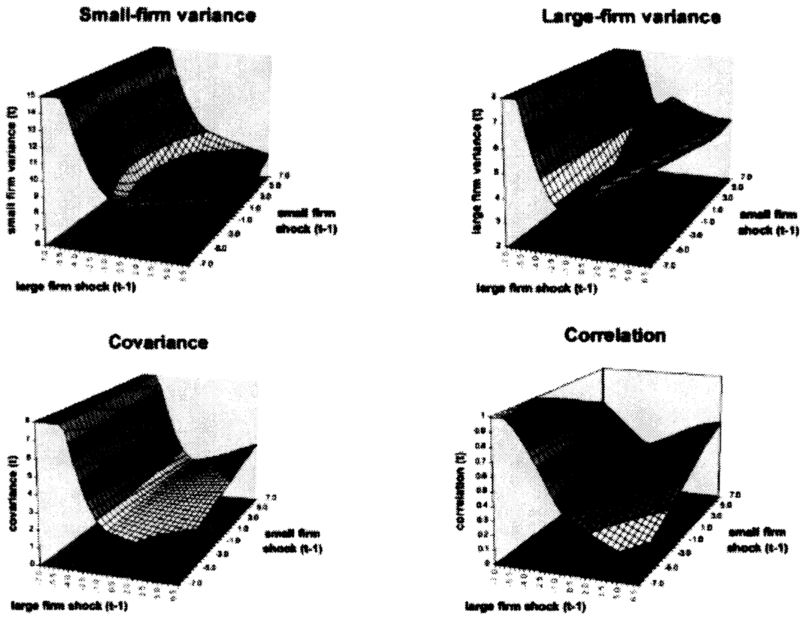

Figure 5
News impact surfaces for ADC model
The figures give the news impact surfaces for the small-firm variance, the large-firm variance, the covariance between small-firm and large-firm shocks, and the correlation between small-firm shocks and large-firm shocks under the ADC model.

shocks of the large-firm portfolio (as $i = 2$ is for the large firm), the results indicate that the sign of the large-firm return shocks is more important than the sign of the small-firm return shocks.

Such effects can also be seen more clearly by inspecting the variance and covariance news impact surfaces for the ADC model. These graphs are given in Figure 5. Panel 1 of Figure 5 indicates that the small-firm portfolio variance is only mildly affected by news to the small-firm portfolio. Instead, bad news to the large-firm portfolio has a dominant impact on small-firm variances. This should not be surprising, given existing results in the literature. For example, Nelson (1990) and Glosten, Jagannathan and Runkle (1993) demonstrate that bad news has a bigger impact on subsequent volatility than good news, and Conrad, Gultekin, and Kaul (1991) demonstrate that shocks to large-firm returns affect future small-firm return volatility. Putting these two results together (as our model does), we find that bad news to large firms affects small-firm volatility. The advantage the ADC model has over these other models is that it nests all these potential asymmetric relationships and spillovers in one model.

Consistent with Conrad, Gultekin, and Kaul (1991), panel 2 reveals that the variance of the large-firm portfolio is unaffected by small-firm shocks. In contrast, volatility of the large-firm portfolio is responsive to its own news, especially its own bad news. Panel 3 indicates that there is also an interesting asymmetric effect in the covariance which has not been documented before. Specifically, the covariance between large- and small-firm returns is higher following a negative shock to the large-firm portfolio, while it is almost unaffected by shocks to the small-firm portfolio. Panel 4 indicates that these asymmetries in the covariance are not driven entirely by the asymmetries in variances, because, for example, positive small-firm shocks have a different impact on correlations than negative small-firm shocks. Panel 4 also reveals that shared negative shocks have much stronger impacts on correlations than shared positive shocks.

Finally, to check for misspecification we apply our robust conditional moment tests to the model and report the results in the final column of Table 3. The only rejection observed is that the ADC model does not fully capture the asymmetric relation between large-firm shocks and covariances. In stark contrast to the other multivariate GARCH models we examined, the ADC model is well-specified along all other dimensions examined. The ADC model fits the data well.

5. Illustration of Economic Importance

Estimating the right time-varying covariance matrix is essential for asset pricing, portfolio selection, and risk management. To illustrate the importance of the covariance matrix to these types of financial problems, we applied our results to two problems.

First, consider the problem of calculating the optimal fully invested portfolio holdings subject to a no-shorting constraint. This application is illustrative of the kinds of problems faced by portfolio managers when deriving their optimal portfolio holdings. In order to avoid forecasting expected returns, we assume here that the expected returns are zero, making the problem equivalent to estimating the risk-minimizing portfolio weights. Define

$$\omega_t = \frac{h_{22t} - h_{12t}}{h_{11t} - 2h_{12t} + h_{22t}}.$$

Then it is easy to show that, assuming a mean-variance utility function, the optimal portfolio holdings of the small-firm portfolio are

$$\omega_t^* = \begin{cases} 0 & \text{if} \quad \omega_t < 0 \\ \omega_t & \text{if} \quad 0 \le \omega_t \le 1 \\ 1 & \text{if} \quad \omega_t > 1 \end{cases}$$

and the optimal holdings of the large-firm portfolio are $1 - \omega_t^*$. As shown

Table 5
Portfolio comparisons from the estimated models

Panel A: Optimal fully invested small-firm portfolio weights

	VECH	CCORR	FARCH	BEKK	ADC
Averages	0.222	0.232	0.209	0.222	0.181
VECH	1.000				
CCORR	0.977	1.000			
FARCH	0.698	0.702	1.000		
BEKK	0.796	0.721	0.361	1.000	
ADC	0.571	0.542	0.121	0.623	1.000

Panel B: Optimal risk-minimizing large-firm hedge ratios

	VECH	CCORR	FARCH	BEKK	ADC
Averages	0.666	0.692	0.639	0.640	0.753
VECH	1.000				
CCORR	0.728	1.000			
FARCH	−0.500	−0.553	1.000		
BEKK	0.393	0.047	0.403	1.000	
ADC	−0.010	−0.110	0.656	0.516	1.000

Panel A gives summary statistics for the optimal small-firm portfolio weights in a fully invested, no-shorting portfolio. The optimal large-firm weights are one minus the optimal small-firm weights. The first row gives the average weight over the sample period and the remaining rows give the correlation matrix of optimal weights. Panel B gives summary statistics for the risk-minimizing hedge ratio in a problem where the large-firm portfolio is used to hedge against small-firm return volatility. A hedge ratio of 0.67 means that the investor would short \$67 worth of the large-firm portfolio to hedge against a long position of \$100 in the small-firm portfolio. As in panel A, the first row gives the average hedge ratios and the remaining rows give the correlation matrix of optimal hedge ratios.

in panel A of Table 5, all the models result in very similar average optimal weights, with the averages ranging from 0.18 (ADC) to 0.22 (VECH and BEKK). However, the correlations between these portfolio weights are low, typically about 70%. So the optimal portfolio will depend on the covariance model chosen, meaning that portfolio managers would have to be very careful which covariance model they select, because the model matters.

Second, consider the problem of estimating a dynamic risk-minimizing hedge ratio using multivariate GARCH models. Several applications of this exist in the literature. For example, Kroner and Claessens (1991) and Kroner and Sultan (1993) use the CCORR model and Baillie and Myers (1991) use the VECH model. To minimize the risk of a portfolio that is long \$1 in small-firm portfolio, an investor should short $\$\beta$ of the large-firm portfolio, where the "risk minimizing hedge ratio" β is

$$\beta_t^* = -\frac{h_{12,t}}{h_{22,t}}$$

The summary statistics of the estimated hedge ratios from the different covariance models are given in panel B of Table 5. The average hedge ratio is about 0.70 for all the models, but the correlations between the hedge ratios are strikingly low. In fact, the correlation between the FARCH and CCORR hedge ratios is -0.55. Clearly the choice of models will seriously affect the estimated hedge ratios. Of importance, the formula for the optimal hedge ratio is the same as that for the market β if the second asset is the market. This suggests that any application that estimates time-varying βs must also pay careful attention to the model selection process.

6. Conclusion and Summary

Existing multivariate models allowing the covariance matrix to be time varying generally impose strong restrictions on how past shocks can affect the covariance matrix. Yet these restrictions are seldom compared and tested. Furthermore, asymmetric/leverage effects have been found in variances, but few studies have examined such effects in covariances, even though there are good reasons to believe that they exist and have important implications for portfolio management.

We filled these gaps by demonstrating the differences between several popular multivariate GARCH models; introducing a set of robust conditional moment tests to detect misspecification in the dynamics of the covariance matrix, with special emphasis on the asymmetric effects in the covariances; and introducing a general dynamic covariance matrix model which nests various existing models as special cases. More specifically, our model nests the constant correlation model of Bollerslev (1990), the FARCH model of Engle, Ng, and Rothschild (1990), the BEKK model of Engle and Kroner (1995), and the VECH model of Bollerslev, Engle, and Wooldridge (1988). We also introduced a generalization of the encompassing model that allows for asymmetric effects in the variances and covariances. This asymmetric dynamic covariance matrix model nests various asymmetric extensions of the four existing models.

We apply the asymmetric dynamic covariance matrix model to weekly returns from a large-firm portfolio and a small-firm portfolio to examine the dynamic relation between large- and small-firm returns. We found that all four existing models are misspecified, especially in the dynamics of the covariance. Our results confirm the general conclusion of Conrad, Gultekin, and Kaul (1991) in a more general setting. That is, large-firm returns can affect the volatility of small-firm returns, but small-firm returns do not have much effect on large-firm volatility. Moreover, we also show that there are significant asymmetric effects in both the variances and covariances which have not been documented before. In particular, bad news about large firms can cause volatility in both small-firm returns and large-firm returns. Furthermore, the conditional covariance between large-firm returns and small-firm returns tends to be higher following bad news about large firms than good news. In addition, news about small firms has minimal effect on the variances and covariance.

Perhaps the most important conclusion of this research is that the choice of a multivariate volatility model can substantially affect the conclusions of the analysis. This is especially important for portfolio selection, risk management, and asset pricing. For example, we showed that the correlations between the risk-minimizing hedge ratios derived from various popular multivariate volatility models are surprisingly low, and sometimes negative.

Appendix

Proof of Proposition 1.

The VECH model: If $\rho_{ij} = 0$ for all $i \neq j$, then the matrix R reduces to an $N \times N$ identity matrix. Hence

$$H_t = D_t' D_t + \Phi \circ \Theta_t,$$

or equivalently,

$$h_{iit} = \theta_{iit} \text{ for all } i, \quad \text{and } h_{ijt} = \phi_{ij}\theta_{ijt} \text{ for all } i \neq j.$$

With $a_i = \alpha_i \iota_i$ and $b_i = \beta_i \iota_i$ for all i, θ_{ijt} becomes

$$\theta_{ijt} = \omega_{ij} + \beta_i \beta_j (\iota_i' H_{t-1} \iota_j) + \alpha_i \alpha_j (\iota_i' \varepsilon_{t-1} \varepsilon_{t-1}' \iota_j) \quad \text{for all } i \text{ and } j.$$

Substituting this expression for θ_{ijt} back into the expressions for h_{iit} and h_{ijt}, we obtain the VECH model with the restriction that $\beta_{ij} = \beta_i \beta_j$:

$$\begin{aligned} h_{iit} &= \omega_{ii} + \beta_i^2 h_{iit-1} + \alpha_i^2 \varepsilon_{it-1}^2 \quad \text{for all } i, \text{ and} \\ h_{ijt} &= \phi_{ij}\omega_{ij} + \phi_{ij}\beta_i\beta_j h_{ijt} + \phi_{ij}\alpha_i\alpha_j \varepsilon_{it-1}\varepsilon_{jt-1} \quad \text{for all } i \neq j. \end{aligned}$$

The CCORR model: If $\phi_{ij} = 0$ for all $i \neq j$, then the matrix Φ becomes a null matrix. Hence

$$H_t = D_t' R D_t,$$

or equivalently,

$$h_{iit} = \theta_{iit} \text{ for all } i, \quad \text{and } h_{ijt} = \rho_{ij}\sqrt{\theta_{iit}}\sqrt{\theta_{jjt}} \text{ for all } i \neq j.$$

With $a_i = \alpha_i \iota_i$ and $b_i = \beta_i \iota_i$ for all i, θ_{ijt} becomes

$$\theta_{ijt} = \omega_{ij} + \beta_i \beta_j (\iota_i' H_{t-1} \iota_j) + \alpha_i \alpha_j (\iota_i' \varepsilon_{t-1} \varepsilon_{t-1}' \iota_j) \quad \text{for all } i \text{ and } j.$$

Substituting this expression for θ_{ijt} back into the expressions for h_{iit} and recognizing that $\theta_{iit} = h_{iit}$ and $\theta_{jjt} = h_{jjt}$, we obtain the constant correlation model:

$$
\begin{aligned}
h_{iit} &= \omega_{ii} + \beta_i^2 h_{iit-1} + \alpha_i^2 \varepsilon_{it-1}^2 \quad \text{for all } i \text{, and} \\
h_{ijt} &= \rho_i j \sqrt{h_{iit}} \sqrt{h_{jjt}} \quad \text{for all } i \neq j.
\end{aligned}
$$

The BEKK model: If $\rho_{ij} = 0$ and $\phi_{ij} = 1$ for all $i \neq j$, then the matrix R reduces to an $N \times N$ identity matrix and the matrix Φ reduces to a matrix with zero diagonal elements and unit off-diagonal elements. Hence

$$
H_t = D_t D_t + (\iota\iota' - I) \circ \Theta_t,
$$

where ι is a vector of ones. Expressing D_t and G_t in terms of θ, we have

$$
\begin{aligned}
h_{iit} &= \theta_{iit} = \omega_{ii} + b_i' H_{t-1} b_i + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_i \quad \text{for all } i \text{, and} \\
h_{ijt} &= \theta_{ijt} = \omega_{ij} + b_i' H_{t-1} b_j + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_j \quad \text{for all } i \neq j.
\end{aligned}
$$

In matrix notation, this is

$$
H_t = \Omega + A' \varepsilon_{t-1} \varepsilon_{t-1}' A + B' H_{t-1} B,
$$

where $A = [a_1, \ldots, a_N]$, $B = [b_1, \ldots, b_N]$, and $\Omega = [\omega_{ij}]$.

The FARCH model: As before, if $\rho_{ij} = 0$, $\phi_{ij} = 1$ for all $i \neq j$, then

$$
H_t = \Omega + A' \varepsilon_{t-1} \varepsilon_{t-1}' A + B' H_{t-1} B.
$$

If, in addition, $A = \alpha(w\lambda')$ and $B = \beta(w\lambda')$ then the expression for H_t can be rewritten as

$$
H_t = \omega + \lambda\lambda' [\beta(w' H_{t-1} w) + \alpha(w' \varepsilon_{t-1})^2],
$$

or equivalently,

$$
\begin{aligned}
h_{ijt} &= \sigma_{ij} + \lambda_i \lambda_j h_{pt} \quad \text{for all } i \\
h_{pt} &= \omega_p + \beta h_{pt-1} + \alpha \varepsilon_{pt-1}^2,
\end{aligned}
$$

where

$$
h_{pt} \equiv w' H_t w, \quad \varepsilon_{pt} \equiv w' \varepsilon_t, \quad \text{and } \sigma_{ij} \equiv \omega_{ij} - \lambda_i \lambda_j w' \Omega w. \qquad \blacksquare
$$

References

Amin, K., and V. Ng, 1994a, "A Comparison of Predictable Volatility Models Using Option Data," unpublished manuscript, International Monetary Fund.

Amin, K., and V. Ng, 1994b, "Heath, Jarrow and Morton Implied Volatility Function and Conditional Heteroskedasticity Models: Information in Eurodollar Futures Options," *Review of Financial Studies*, 10, 333–367.

Baillie, R. T., and R. J. Myers, 1991, "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge," *Journal of Applied Econometrics*, 6, 109–124.

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

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

Bollerslev, T., R. Y. Chou, and K. F. Kroner, 1992, "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5–59.

Brenner, R. J., R. H. Harjes, and K. F. Kroner, 1996, "Another Look at Alternative Models of the Short Term Interest Rate," *Journal of Financial and Quantitative Analysis*, 31, 85–108.

Conrad, J., M. Gultekin, and G. Kaul, 1991, "Asymmetric Predictability of Conditional Variances," *Review of Financial Studies*, 4, 597–622.

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

Engle, R., and V. Ng, 1993, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.

Engle, R., V. Ng, and M. Rothschild, 1990, "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, 213–238.

Gallant, R., P. Rossi, and G. Tauchen, 1992, "Stock Prices and Volume," *Review of Financial Studies*, 5, 199–242.

Glosten, L., R. Jagannathan, and D. Runkle, 1993, "Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.

Hentschel, L., 1995, "All in the Family: Nesting Symmetric and Asymmetric GARCH Models," *Journal of Financial Economics*, 39, 71–104.

Kim, D., and S. Kon, 1994, "Alternative Models of Conditional Heteroskedasticity of Stock Returns," *Journal of Business*, 67, 563–598.

Kraft, D., and R. Engle, 1983, "Autoregressive Conditional Heteroskedasticity in Multiple Time Series Models," unpublished manuscript, University of California, San Diego.

Kroner, K. F., and S. Claessens, 1991, "Optimal Dynamic Hedging Portfolios and the Currency Composition of External Debt," *Journal of International Money and Finance*, 10, 131–148.

Kroner, K. F., and J. Sultan, 1993, "Time Varying Distributions and Dynamic Hedging with Foreign Currency Futures," *Journal of Financial and Quantitative Analysis*, 28, 35–551.

Nelson, D., 1990, "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Ng, V., R. Engle, and M. Rothschild, 1992, "A Multi Dynamic Factor Model for Stock Returns," *Journal of Econometrics*, 52, 245–266.

Pagan, A., and G. W. Schwert, 1990, "Alternative Models for Conditional Stock Volatility," *Journal of Econometrics*, 45, 267–290.

Wooldridge, J., 1990, "A Unified Approach to Robust, Regression-Based Specification Tests," *Econometric Theory*, 6, 17–43.