Econometric Theory, **21**, 2005, 1058–1086. Printed in the United States of America.
DOI: 10.1017/S026646660505053X

BIVARIATE ARCH MODELS: FINITE-SAMPLE PROPERTIES OF QML ESTIMATORS AND AN APPLICATION TO AN LM-TYPE TEST

Emma M. Iglesias
Michigan State University

Garry D.A. Phillips
Cardiff University

This paper provides two main new results: the first shows theoretically that large biases and variances can arise when the quasi-maximum likelihood (QML) estimation method is employed in a simple bivariate structure under the assumption of conditional heteroskedasticity; and the second demonstrates how these analytical theoretical results can be used to improve the finite-sample performance of a test for multivariate autoregressive conditional heteroskedastic (ARCH) effects, suggesting an alternative to a traditional Bartlett-type correction. We analyze two models: one proposed in Wong and Li (1997, *Biometrika* 84, 111–123) and another proposed by Engle and Kroner (1995, *Econometric Theory* 11, 122–150) and Liu and Polasek (1999, *Modelling and Decisions in Economics*; 2000, working paper, University of Basel). We prove theoretically that a relatively large difference between the intercepts in the two conditional variance equations, which leads to the two series having correspondingly different volatilities in the restricted case, may produce very large variances in some QML estimators in the first model and very severe biases in some QML estimators in the second. Later we use our bias expressions to propose an LM-type test of multivariate ARCH effects and show through simulations that small-sample improvements are possible, especially in relation to the size, when we bias correct the estimators and use the expected hessian version of the test.

1. INTRODUCTION

The multivariate-ARCH (autoregressive conditional heteroskedastic) model was first introduced by Kraft and Engle (1983) and Bollerslev, Engle, and Wool-

Both authors thank H. Wong for providing us with the Gauss program to simulate the Wong and Li (1997) model. We also thank three anonymous referees for extremely helpful comments, and we are grateful for the comments received at seminars given at Cardiff University, Michigan State University, Queen Mary London, University of Exeter, and University of Montreal. We acknowledge gratefully also the financial support from an ESRC grant (award number T026 27 1238). A previous version of this paper appeared as IVIE Working paper 2004-09. Address correspondence to Emma M. Iglesias, Department of Economics, Michigan State University, 101 Marshall-Adams Hall, East Lansing, MI 48824-1038, USA; e-mail: iglesia5@msu.edu.

  0266-4666/05 $12.00

dridge (1988). Since then, new combinations of this specification of the variance equation with different structures in the mean equation have been proposed; see, e.g., Baba, Engle, Kraft, and Kroner (1991); Harmon (1988); Engle and Kroner (1995); Calzolari and Fiorentini (1994); Polasek and Kozumi (1996); and Bauwens, Laurent, and Rombouts (2005) for a review of recent developments.

The multivariate model implies that the conditional variance-covariance matrix (H_t) of the disturbances (ε_t) depends on the information set (I_{t-1}). The main problem to be faced in this specification is the relatively large number of parameters that are involved. There are, however, many possible parameterizations for H_t that reduce the number of parameters to estimate. One possibility is to consider the "vech" (vec-half) representation. However, even for the estimation of this model, it is necessary to restrict the number of parameters still further. Another possible specification is the diagonal representation, where each element of the covariance matrix $h_{jk,t}$ is a function only of past values of itself and past values of $\varepsilon_{j,t}\varepsilon_{k,t}$. The drawback in this case is that we must still ensure that H_t is a positive definite matrix for all values of the ε_t, and it can be a difficult task to check this in the previous specifications. This is why Engle and Kroner (1995) proposed a new parameterization: the BEKK (Baba, Engle, Kraft, and Kroner, 1991) representation. A recent discussion of all these models (and how they can be nested) can be found in Bauwens et al. (2005).

Nowadays there exists an extensive literature about multivariate-ARCH models that have been applied to different varieties of data. Most of them use (quasi)-maximum likelihood (QML-ML) as the estimation procedure. However, there are relatively few theoretical papers that examine the consequences of this.

If we define y_t as an M-dimensional finite-order vector of time series variables, the relevant part of the (conditional) log-likelihood function in these models is denoted by

$$L(y,\theta) = \sum_{t=1}^{T} L_t(y_t,\theta) \propto -\frac{1}{2}\sum_{t=1}^{T}\log|H_t| - \frac{1}{2}\sum_{t=1}^{T}(y_t-\mu_t)'H_t^{-1}(y_t-\mu_t). \tag{1.1}$$

Liu and Polasek (1999) gave the following representation of the conditional information matrix of the ML estimator ($I(\theta)$) in a general multivariate heteroskedastic model:

$$I(\theta) = \frac{1}{2}\sum_{t=1}^{T}\left(\frac{\partial \mathrm{vech} H_t}{\partial\theta'}\right)' D'(H_t^{-1}\otimes H_t^{-1})D\,\frac{\partial \mathrm{vech} H_t}{\partial\theta'} + \sum_{t=1}^{T}\left(\frac{\partial\mu_t}{\partial\theta'}\right)' H_t^{-1}\frac{\partial\mu_t}{\partial\theta'} \tag{1.2}$$

(see Liu and Polasek, 1999, p. 103), where $\mu_t = E(y_t/I_{t-1})$ is an $M \times 1$ conditional mean vector, $H_t = \text{var}(y_t/I_{t-1})$ is an $M \times M$ conditional variance matrix, D is the $M^2 \times M(M+1)/2$ duplication matrix, and $\otimes$ indicates Kronecker product.

Regarding asymptotic theory, Tuncer (1994, 2000), Bauwens and Vandeuren (1995), Jeantheau (1998), and Comte and Lieberman (2003) have established the strong consistency of the quasi-maximum likelihood estimator (QMLE) in a simple multivariate-ARCH model. Asymptotic normality is proved provided that the initial state is either stationary or fixed. More recently, Ling and McAleer (2003) have shown the asymptotic normality in a vector autoregressive moving average–generalized autoregressive conditional heteroskedasticity (VARMA-GARCH) model requiring only the existence of the second-order unconditional moment and a finite fourth-order conditional moment of the errors, which represents an important advance. However these papers have nothing to say about the finite-sample properties of QMLE, and in this paper we provide results that go some way toward addressing this.

In relation to finite samples, in a more recent paper, Liu and Polasek (2000) have compared through Monte Carlo simulation the biases that are generated using the Splus + GARCH program package of MathSoft (1996), the BASEL package of Polasek et al. (1999), and the application of the method of scoring for MLE using the exact information matrix (given previously). The generated biases are seen to be striking, and the Bayesian method seems to be the best alternative; see Polasek et al. (1999) for a discussion of this method. For a sample of 200 observations, their results show the existence of severe biases. There are, in fact, other recent papers that analyze different types of Bayesian bivariate-ARCH models applied to economic data, such as Osiewalski and Pipien (2004). On the other hand, Wong and Li (1997) reported through Monte Carlo simulation that in their model the biases in the parameters were very small (see Wong and Li, 1997, pp. 119–122). It is precisely this apparent conflict over the nature and the size of the bias in bivariate-ARCH models that has motivated the work in our present paper.

It is interesting to note too that in a recent paper, Jensen and Rahbek (2004) prove how in univariate ARCH processes the QMLE is always asymptotically normal provided that the fourth moment of the innovation process exists, whether or not the process is stationary. This gives support to the estimation of ARCH processes without being subject to constraints, and in this paper we carry out the estimation through unrestricted QML.

The plan of the paper is as follows. In the next section we will begin analyzing a bivariate model under two important specifications that have been proposed in the literature so far: the one given in Wong and Li (1997), where they allow the two disturbances to be dependent but not correlated, and the one proposed in Engle and Kroner (1995) and Liu and Polasek (1999, 2000), where linear dependence between the disturbances is introduced. We provide theoretical results on the $O(T^{-1})$ biases for the QML estimators in each specification under the assumption of conditional heteroskedasticity. We impose the restriction that the variance parameters are zero (overspecification of ARCH effects), hence following an approach that can be found in a number of other studies (see Engle, Hendry, and Trumble, 1985; Linton, 1997). For ease of manipulation, we assume also that the intercept in the mean equation is known, although more complicated structures could, in principle, be analyzed following the same methodology. In fact, the results given in Iglesias and Phillips (2003) that showed that it is the number of exogenous variables in the mean equation (and not their individual characteristics) that determines the bias also apply in this setting. Although our theoretical results are obtained in a very restricted model, we are able to prove how, in the Wong and Li (1997) model, the variances for some estimators can be large when there is a relatively large difference between the intercepts in the variance equations (they only showed results for cases when the intercept parameters had very similar numerical values), and in the second model that is examined in Liu and Polasek (1999, 2000) we show that a large difference in the intercepts under the null of no ARCH effects (when the two series can have very different volatilities) can produce very large biases in some of the QMLEs. We demonstrate also theoretically how, in the Wong and Li and Liu and Polasek models, some assumptions should be imposed for the QML estimator to be well defined. We provide evidence that the biases can be very different depending on both the structure we impose on the model and the combinations of the parameters we study. We also analyze some invariance properties, extending the Lumsdaine (1995) work in a univariate framework. Later, in Section 3, we consider an LM (Lagrange multiplier) test for multivariate ARCH effects. We find that an LM test based upon the expected hessian is available that completely dominates the outer product and hessian versions. We also show how the bias approximations obtained in the null case can be used to improve the finite-sample performance of the test. There are many papers that propose improving the finite-sample performance of likelihood ratio (LR) tests by using Bartlett-type corrections (see, e.g., the recent paper by Johansen, 2002); however, such a correction is not available for the LM test. We show that, in the context of an LM test, the novel approach of bias correcting the QML estimators may be a suitable alternative. Finally, Section 4 concludes.

2. SOME FINITE-SAMPLE RESULTS FOR BIVARIATE-ARCH MODELS

2.1. Case 1: Allowing for Dependent but Uncorrelated Disturbances

We begin by analyzing the framework proposed in Wong and Li (1997) for the variance equation, where the model is specified as

$$y_t = \beta + \varepsilon_t \tag{2.1}$$

and where $y_t = (y_{1t}, y_{2t})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, and $E(\varepsilon_t) = 0$. The intercept vector $\beta = (\beta_{10}\beta_{20})'$ is assumed known. We could, in principle, allow for the estimation of the intercept and the introduction of any number of exogenous variables in the mean equation along the lines of the results of Iglesias and Phillips (2003).

This would entail a considerable increase in the complexity of the analysis. However, our main interest here is to examine some well-known models in the literature, and our specification in (2.1) adequately allows for this. The conditional variance equation follows the structure

$$H_t = \begin{pmatrix} h_{11t} & 0 \\ 0 & h_{22t} \end{pmatrix},$$

where

$$h_{11t} = E(\varepsilon_{1t}^2/I_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{1t-1}^2 + \alpha_2 \varepsilon_{2t-1}^2, \tag{2.2}$$

$$h_{22t} = E(\varepsilon_{2t}^2/I_{t-1}) = \gamma_0 + \gamma_1 \varepsilon_{1t-1}^2 + \gamma_2 \varepsilon_{2t-1}^2. \tag{2.3}$$

Expressions (2.2) and (2.3) can be rewritten as

$$\varepsilon_{1t}^2 = \alpha_0 + \alpha_1 \varepsilon_{1t-1}^2 + \alpha_2 \varepsilon_{2t-1}^2 + \eta_{1t},$$

$$\varepsilon_{2t}^2 = \gamma_0 + \gamma_1 \varepsilon_{1t-1}^2 + \gamma_2 \varepsilon_{2t-1}^2 + \eta_{2t},$$

where, because of the uncorrelatedness of the epsilons,

$$E(\eta_{1t}) = E(\eta_{2t}) = 0; \qquad E(\eta_{1t}\eta_{2t}) = 0,$$

$$E(\eta_{1t}^2) = E(2h_{11t}^2); \qquad E(\eta_{2t}^2) = E(2h_{22t}^2).$$

We assume that the process is at least second-order stationary (see Wong and Li, 1997). After some algebra, we find

$$E(\varepsilon_{1t}^2) = \frac{\alpha_0(1-\gamma_2) + \alpha_2\gamma_0}{(1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2}; \qquad E(\varepsilon_{2t}^2) = \frac{\gamma_0(1-\alpha_1) + \gamma_1\alpha_0}{(1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2}.$$

From the preceding discussion we may deduce the following restrictions on the variance equation parameters:

$$\gamma_2 < 1; \qquad \alpha_1 < 1; \qquad (1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2 > 0.$$

Besides, in the analysis that follows, we will study the case of overspecification of ARCH effects:

$$\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.$$

Our objective is to analyze the QML biases of $O(T^{-1})$ in this simple model. The methodology we will use has been proposed by Cox and Snell (1968) for the MLE and extended by McCullagh (1987) to the QMLE. McCullagh (1987) showed that for independent but not necessarily identically distributed observations, the bias (b) of the QMLE of β ($\hat{\beta}$) reduces to

$$b_s = E(\hat{\beta}_s - \beta_s) = \sum_{i,j,l=1}^{p} k^{si}k^{jl}\left\{\left(\frac{2+\kappa_4}{4}\right)k_{ijl} + k_{ij,l}\right\} + O(T^{-2}) \tag{2.4}$$

for $s = 1,\ldots,p$, where $k_{ij} = E(\partial^2 L/\partial\beta_i\partial\beta_j)$, $k_{ijl} = E(\partial^3 L/\partial\beta_i\partial\beta_j\partial\beta_l)$, $k_{ij,l} = E((\partial^2 L/\partial\beta_i\partial\beta_j)\partial L/\beta_l)$, for $i,j,l = 1,\ldots,p$ (L denotes the log-likelihood function). Here κ_4 is the fourth cumulant of the true distribution. The total Fisher information matrix and its inverse are defined by $K = \{-k_{ij}\}$ and $K^{-1} = \{-k^{ij}\}$, respectively. The formula is valid, even for nonindependent observations, provided that all k's are of $O(T)$ (see Cordeiro and McCullagh, 1991), and this justifies the application of the methodology in our case. When $\kappa_4 = 0$, QML equals ML, and then the formula of McCullagh (1987) equals the one of Cox and Snell (1968). In practice, when we want to use our expressions for the case where our time series vector would present conditional heteroskedasticity, κ_4 can be estimated by using the methodology to estimate cumulants developed in Cox and Hall (2002).

To proceed to obtain the expectations of the second- and third-order derivatives, we can follow the matrix differential calculus techniques of Magnus and Neudecker (1991). Liu and Polasek (1999) provided the expression of the conditional information matrix ($I(\theta)$) of a general VAR(k) − VARCH(q) model for $y_t = (y_{1t}, y_{2t},\ldots, y_{Mt})$, by specializing (1.2):

$$I(\theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$$

with

$$I_{11} = \sum_{t=1}^{T} W_t' H_t^{-1} W_t, \qquad I_{21} = I_{12}' = 0$$

and

$$I_{22} = \frac{1}{2}\sum_{t=1}^{T} V_t' D'(H_t^{-1} \otimes H_t^{-1}) D V_t,$$

where

$$W_t = (I_M, X_{t-1},\ldots, X_{t-k}), \qquad V_t = (I_N, Z_{t-1},\ldots, Z_{t-q}),$$

$$X_{t-i} = \operatorname{diag}(y_{1t-i}, y_{2t-i},\ldots, y_{Mt-i})', \quad \text{for } i = 1,\ldots,k,$$

$$Z_{t-j} = \operatorname{diag}(\varepsilon_{1t-j}^2, \varepsilon_{1t-j}\varepsilon_{2t-j},\ldots, \varepsilon_{Mt-j}^2), \quad \text{for } j = 1,\ldots,q.$$

Note that I_M and I_N are $M \times M$ and $N \times N$ identity matrices, respectively, and D is the duplication matrix defined in (1.2).

This formula is valid only in the situation where there are no parameters to estimate in the mean equation, which is precisely our case. We extend the work by Liu and Polasek (1999) to include all the cumulants we need for our analysis, and Appendix A provides the expressions for the second- and third-order derivatives of (1.1) in our model on applying the differential matrix calculus.

Tables A1 and A2 in Appendix A show the expressions for all the k components that are needed to apply expression (2.4) and obtain the bias results and the variances (given by the information matrix) for the general QML estimator in this model. Unfortunately, although Iglesias and Phillips (2003) were able to find a bias approximation in closed form for the variance parameter estimators in a univariate ARCH(1) model without imposing restrictions, the additional complexity of the multivariate model prevents similar easy to interpret derivations unless restrictions are imposed. To make progress under the assumption that we specify the conditional variance structure given in (2.2) and (2.3), we impose the restrictions that $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$. This type of restriction has been imposed in many of the theoretical analyses that have been carried out in univariate ARCH models so far (e.g., Engle et al., 1985; Linton, 1997), and it facilitates, especially here, the analysis and the interpretation of the results. Table A3 shows the results of the k components when the restrictions are imposed. It is to be noted that if, for example, it is demonstrated through the bias approximations that severe biases and/or large variances are possible in the restricted model, then these characteristics will surely be found in unrestricted models. Of course, if such problems do not arise in the restricted case the same may be true in the unrestricted model, but it need not be so. Hence, care is needed in drawing conclusions from the approximations.

THEOREM 2.1. *If* $y_t = \varepsilon_t$ *where* $y_t = (y_{1t}, y_{2t})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ *is a vector of random variables that has the structure given in (2.2) and (2.3), with* $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$, *then the biases and the variances of the QML estimators to order* T^{-1} *are given by*

$$E(\hat{\alpha}_0 - \alpha_0) = \frac{\alpha_0}{T} + o(T^{-1}) \qquad E(\hat{\gamma}_0 - \gamma_0) = \frac{\gamma_0}{T} + o(T^{-1}),$$

$$E(\hat{\alpha}_1 - \alpha_1) = -\frac{1}{T} + o(T^{-1}) \qquad E(\hat{\alpha}_2 - \alpha_2) = o(T^{-1}),$$

$$E(\hat{\gamma}_1 - \gamma_1) = o(T^{-1}) \qquad E(\hat{\gamma}_2 - \gamma_2) = -\frac{1}{T} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\alpha}_0) = \frac{4\alpha_0^2}{T} + o(T^{-1}) \qquad \operatorname{var}(\hat{\gamma}_0) = \frac{4\gamma_0^2}{T} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\alpha}_1) = \frac{1}{T} + o(T^{-1}) \qquad \operatorname{var}(\hat{\alpha}_2) = \frac{\alpha_0^2}{T\gamma_0^2} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\gamma}_1) = \frac{\gamma_0^2}{T\alpha_0^2} + o(T^{-1}) \qquad \operatorname{var}(\hat{\gamma}_2) = \frac{1}{T} + o(T^{-1}).$$

Proof. Given in Appendix A.

Notice that the biases in the restricted model are relatively small, suggesting that estimation bias may not be a particular problem in this model. However, it is interesting to note how, when the intercept parameters α_0 and γ_0 differ substantially, the preceding model can generate severe and large variances in the QML estimators of the α_2 and γ_1 parameters (at least in one of them). In prac-

Table 2.1. Approximate standard errors when we over-specify the multivariate ARCH effects, $T = 400$

	$\alpha_0 = 0.81$		$\alpha_0 = 0.04$	
	$\gamma_0 = 0.04$		$\gamma_0 = 0.04$	
α_0	0.081	(0.082)	0.004	(0.004)
α_1	0.050	(0.051)	0.050	(0.050)
α_2	1.012	(1.035)	0.050	(0.051)
γ_0	0.004	(0.004)	0.004	(0.004)
γ_1	0.002	(0.002)	0.050	(0.051)
γ_2	0.050	(0.050)	0.050	(0.050)

Note: Simulated values are given in parentheses for 20,000 replications.

tical applications that fit a model with this specification to real data, one should be mindful of this fact when interpreting the estimation results. It is very easy to find an interpretation in this situation: under the null of no ARCH effects, the two intercepts reflect the two unconditional volatilities of the two series. So our results show that severe variances result in this case when the two series have very different volatilities.

Table 2.1 shows the standard errors of $O(T^{-1})$ and a comparison with the simulated errors for different combinations of the intercepts of the conditional variance equation, confirming the results shown previously.

On the other hand, the bias and variances of the QML estimators to $O(T^{-1})$ in a univariate ARCH(1) model, $E(\varepsilon_t^2/I_{t-1}) = \alpha_1 + \alpha_2\varepsilon_{t-1}^2$, when nothing is estimated in the mean equation whereas $\alpha_2 = 0$, are given by (see Engle et al., 1985; Iglesias and Phillips, 2003)

$$E(\hat{\alpha}_1 - \alpha_1) = \frac{\alpha_1}{T} + o(T^{-1}) \qquad E(\hat{\alpha}_2 - \alpha_2) = -\frac{1}{T} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\alpha}_1) = \frac{3\alpha_1^2}{T} + o(T^{-1}) \qquad \operatorname{var}(\hat{\alpha}_2) = \frac{1}{T} + o(T^{-1}).$$

Comparing these biases with those of Theorem 2.1, it is seen that, in the new bivariate specification, the biases in the parameters that are common have the same structure whereas, on the other hand, there is a loss of estimation efficiency to $O(T^{-1})$ in the intercept parameter estimator, and no gain or loss in efficiency for the estimator of the ARCH parameter.

Extending the work in Lumsdaine (1995), the representation of the relevant part of the log-likelihood involves

$$L_t = -\frac{1}{2}\left(\log h_{11t} + \log h_{22t} + \frac{\varepsilon_{1t}^2}{h_{11t}} + \frac{\varepsilon_{2t}^2}{h_{22t}}\right).$$

Using the same argument as the one given in Lumsdaine (1995, p. 10), we can prove that if α_0 and γ_0 change in the same proportion, the biases and t-statistics in $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$ will remain invariant. This result matches with the bias and variance results obtained in Theorem 2.1. However, if α_0 and γ_0 vary in different proportions, the invariance property does not hold.

2.2. Case 2: Allowing for Dependent and Correlated Disturbances

We analyze now the variance specification proposed by Engle and Kroner (1995) and Liu and Polasek (1999, 2000), given by the bivariate model

$$y_t = \beta + \varepsilon_t, \tag{2.5}$$

where $y_t = (y_{1t}, y_{2t})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, $E(\varepsilon_t) = 0$, and we assume again the intercept vector $\beta = (\beta_{10}, \beta_{20})'$ to be known. We allow for possible misspecification of the marginal distribution of the errors; thus the QML estimator is used. The variance representation implies a diagonal structure for the disturbances following an ARCH(1) process:

where $\text{var}(\varepsilon_t/I_{t-1}) = H_t = \begin{pmatrix} h_{11t} & h_{12t} \\ h_{21t} & h_{22t} \end{pmatrix}$ and

$$\begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \\ \alpha_{30} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^2 \\ \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}^2 \end{pmatrix}. \tag{2.6}$$

Then, it follows that

$$E(\varepsilon_{1t}\varepsilon_{2s}/I_{t-1}) = \alpha_{20} + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2s-1}, \qquad t = s, \tag{2.7}$$

$$\qquad 0 \text{ otherwise}$$

$$E(\varepsilon_{1t}^2/I_{t-1}) = \alpha_{10} + \alpha_{11}\varepsilon_{1t-1}^2, \tag{2.8}$$

$$E(\varepsilon_{2t}^2/I_{t-1}) = \alpha_{30} + \alpha_{33}\varepsilon_{2t-1}^2.$$

Following Engle and Kroner (1995) and Liu and Polasek (1999, 2000), we assume that the process is at least second-order stationary. The unconditional expectations become $E(\varepsilon_{1t}^2) = \alpha_{10}/(1-\alpha_{11})$, $E(\varepsilon_{2t}^2) = \alpha_{30}/(1-\alpha_{33})$, $E(\varepsilon_{1t}\varepsilon_{2t}) = \alpha_{20}/(1-\alpha_{22})$, and the unconditional correlation coefficient between both disturbances is $\alpha_{20}\sqrt{(1-\alpha_{11})(1-\alpha_{33})}/(1-\alpha_{22})\sqrt{\alpha_{10}\alpha_{30}}$. This implies that in this model, to guarantee that the correlation coefficient is absolutely smaller than 1, the following restriction is required:

$$\frac{\alpha_{20}^2}{\alpha_{10}\alpha_{30}} < \frac{(1-\alpha_{22})^2}{(1-\alpha_{11})(1-\alpha_{33})}.$$

When $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$, then $\alpha_{20}^2/\alpha_{10}\alpha_{30} < 1$. In addition, $\alpha_{10}, \alpha_{30} > 0$, whereas $0 < \alpha_{11}, \alpha_{33} < 1$.

Our objective is to analyze the biases of $O(T^{-1})$ in this simple model when we use the QML estimation procedure. Tables B1 and B2 in Appendix B show all the k components that are needed to apply (2.4) and to get the bias expressions for the general QML estimator. Again, to get closed-form solutions and for ease of interpretation, our analysis assumes that we specify a diagonal structure in the conditional variance, when, in fact, the true model is the one for which we have $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$. Table B3 in Appendix B shows the k components when they are evaluated under that restriction. The bias results are given in Theorem 2.2.

THEOREM 2.2. *If $y_t = \varepsilon_t$ where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a vector of random variables that has the structure given in (2.6) with $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$, then the biases and the variances of the QML estimators to order T^{-1} are given by*

$$E(\hat{\alpha}_{10} - \alpha_{10}) = \frac{\alpha_{10}^2\alpha_{30}(\alpha_{10}^2\alpha_{30}^2 + \alpha_{10}^2\alpha_{20}^2 + \alpha_{20}^2\alpha_{30}^2 + 2\alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{20}^4)}{T(\alpha_{10}^2\alpha_{30}^2 + 4\alpha_{10}\alpha_{20}^2\alpha_{30} + \alpha_{20}^4)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{10}) = \frac{\alpha_{10}^2(3\alpha_{10}^3\alpha_{30}^3 + 13\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 10\alpha_{10}\alpha_{20}^4\alpha_{30} + 2\alpha_{20}^6)}{T(\alpha_{10}^3\alpha_{30}^3 + 5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 5\alpha_{10}\alpha_{20}^4\alpha_{30} + \alpha_{20}^6)} + o(T^{-1}),$$

$$E(\hat{\alpha}_{20} - \alpha_{20}) = \frac{\alpha_{20}(\alpha_{10}^4\alpha_{30}^2 + \alpha_{10}^2\alpha_{30}^4 + 6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + \alpha_{20}^4\alpha_{10}^2 + \alpha_{20}^4\alpha_{30}^2 - 2\alpha_{20}^6)}{2T(\alpha_{10}^2\alpha_{30}^2 + 4\alpha_{10}\alpha_{20}^2\alpha_{30} + \alpha_{20}^4)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{20}) = \frac{(\alpha_{10}^3\alpha_{30}^3 + 6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 5\alpha_{10}\alpha_{20}^4\alpha_{30} + 2\alpha_{20}^6)}{T(\alpha_{10}^2\alpha_{30}^2 + 4\alpha_{10}\alpha_{20}^4\alpha_{30} + \alpha_{20}^4)} + o(T^{-1}),$$

$$E(\hat{\alpha}_{30} - \alpha_{30}) = \frac{\alpha_{10}\alpha_{30}^2(\alpha_{10}^2\alpha_{30}^2 + \alpha_{10}^2\alpha_{20}^2 + \alpha_{20}^2\alpha_{30}^2 + 2\alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{20}^4)}{T(\alpha_{10}^2\alpha_{30}^2 + 4\alpha_{10}\alpha_{20}^2\alpha_{30} + \alpha_{20}^4)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{30}) = \frac{\alpha_{30}^2(3\alpha_{10}^3\alpha_{30}^3 + 13\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 10\alpha_{10}\alpha_{20}^4\alpha_{30} + 2\alpha_{20}^6)}{T(\alpha_{10}^3\alpha_{30}^3 + 5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 5\alpha_{10}\alpha_{20}^4\alpha_{30} + \alpha_{20}^6)} + o(T^{-1}),$$

$$E(\hat{\alpha}_{11} - \alpha_{11}) = -\frac{\alpha_{10}\alpha_{30}(\alpha_{10}^2\alpha_{30}^2 + \alpha_{10}^2\alpha_{20}^2 + \alpha_{20}^2\alpha_{30}^2 + 2\alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{20}^4)}{T(\alpha_{10}^2\alpha_{30}^2 + 4\alpha_{10}\alpha_{20}^2\alpha_{30} + \alpha_{20}^4)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{11}) = \frac{\alpha_{10}^2\alpha_{30}^2(\alpha_{10}\alpha_{30} + 3\alpha_{20}^2)}{T(\alpha_{10}^3\alpha_{30}^3 + 5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2 + 5\alpha_{10}\alpha_{20}^4\alpha_{30} + \alpha_{20}^6)} + o(T^{-1}),$$

$$E(\hat{\alpha}_{22}-\alpha_{22}) = -\frac{(\alpha_{10}^4\alpha_{30}^2+\alpha_{10}^2\alpha_{30}^4+6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^4+\alpha_{30}^2\alpha_{20}^4-2\alpha_{20}^6)}{2T(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\alpha}_{22}) = \frac{(\alpha_{10}^2\alpha_{30}^2+\alpha_{20}^4)}{T(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^4)} + o(T^{-1}),$$

$$E(\hat{\alpha}_{33}-\alpha_{33}) = -\frac{\alpha_{10}\alpha_{30}(\alpha_{10}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^2+\alpha_{20}^2\alpha_{30}^2+2\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{20}^4)}{T(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)} + o(T^{-1}),$$

$$\operatorname{var}(\hat{\alpha}_{33}) = \frac{\alpha_{10}^2\alpha_{30}^2(\alpha_{10}\alpha_{30}+3\alpha_{20}^2)}{T(\alpha_{10}^3\alpha_{30}^3+5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^6)} + o(T^{-1}).$$

Proof. Given in Appendix B.

In spite of the large and tedious expressions we get, it is important to highlight the utility we can get from them, because they enable us to find approximations to the biases for any combination of parameters and to discover their evolution. Notice that all the coefficient bias approximations contain in the denominator the term $(\alpha_{10}\alpha_{30}-\alpha_{20}^2)$, which is the determinant of the unconditional covariance matrix of the disturbances. Hence one obvious situation in which large estimator biases are to be expected is when there is high correlation between the disturbances. However, the biases can still be large even when this correlation is relatively modest, as will be shown subsequently. Thus we can provide theoretical support for the large biases found by Liu and Polasek (2000) even though our analytical results are obtained under strong restrictions. An additional use for the approximations is for bias correction under the assumption of overspecification of the conditional process; we can use the expressions for bias correction, substituting estimates for the true values of the expressions. The direct applicability of the results for testing will be shown in the next section of the paper.

We have noted that our theoretical analysis supports the results in Liu and Polasek (2000), in the sense that the biases can be very large in these models—even though our setting is different—but our findings provide evidence that when the disturbances are not highly correlated, the biases are only so large for some combinations of parameters. Table 2.2 shows how the larger biases are those for the parameters of the ARCH components, especially when there is a large difference between the intercepts of the two conditional variance equations (the simulated results support the same outcome). For example, the approximate bias of the estimator of α_{22} increases from around -0.005 to -0.249 when the constant terms α_{10} and α_{30} change from being the same and equal at 0.15 to α_{10} being kept constant at 0.15 and α_{30} increasing to 15.

Once we have found the bias expressions of $O(T^{-1})$, we can again extend the work by Lumsdaine (1995) to our model. In this case we need to change

TABLE 2.2. Biases and variances of $O(T^{-1})$ for some different parameter configurations, $\alpha_{10} = 0.15$, $\alpha_{20} = 0.05$, and $T = 200$

	$\alpha_{30} = 0.15$	$\alpha_{30} = 15$
$E(\hat{\alpha}_{10} - \alpha_{10})$	0.00083 (0.0071)	0.00083 (0.0201)
$\mathrm{var}(\hat{\alpha}_{10})$	0.00032 (0.0024)	0.00034 (0.0035)
$E(\hat{\alpha}_{20} - \alpha_{20})$	0.00026 (0.0025)	0.01246 (0.0145)
$\mathrm{var}(\hat{\alpha}_{20})$	0.00013 (0.0097)	0.01127 (0.0092)
$E(\hat{\alpha}_{30} - \alpha_{30})$	0.00083 (0.0081)	0.08322 (0.1032)
$\mathrm{var}(\hat{\alpha}_{30})$	0.00032 (0.0025)	3.37250 (4.2710)
$E(\hat{\alpha}_{11} - \alpha_{11})$	−0.00553 (−0.0071)	−0.00554 (−0.0077)
$\mathrm{var}(\hat{\alpha}_{11})$	0.00041 (0.0050)	0.00498 (0.0051)
$E(\hat{\alpha}_{22} - \alpha_{22})$	−0.00519 (−0.0075)	−0.24921 (−0.2161)
$\mathrm{var}(\hat{\alpha}_{22})$	0.00347 (0.0095)	0.00497 (0.0056)
$E(\hat{\alpha}_{33} - \alpha_{33})$	−0.00553 (−0.0081)	−0.00554 (−0.0081)
$\mathrm{var}(\hat{\alpha}_{33})$	0.00041 (0.0060)	0.00498 (0.0062)

Note: Simulated values are given in parentheses for 20,000 replications.

α_{10}, α_{20}, and α_{30} in the same proportion to get invariance in the bias and t-statistics of $\hat{\alpha}_{11}$, $\hat{\alpha}_{22}$, and $\hat{\alpha}_{33}$. Otherwise, the invariance property becomes invalid (again, this is consistent with the results in Theorem 2.2).

2.2.1. Special Case When the Correlation of the Disturbances is Overspecified. In this case, if we set $\alpha_{20} = 0$, Theorem 2.2 now becomes Corollary 2.1.

COROLLARY 2.1. *If $y_t = \varepsilon_t$ where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a vector of random variables that has the structure given in (2.6) under overspecification of the conditional correlation* $(\alpha_{20} = 0)$*, then the biases and the variances of the QML estimators to order T^{-1} are given by*

$$E(\hat{\alpha}_{10} - \alpha_{10}) = \frac{\alpha_{10}}{T} + o(T^{-1}) \qquad E(\hat{\alpha}_{11} - \alpha_{11}) = -\frac{1}{T} + o(T^{-1}),$$

$$E(\hat{\alpha}_{20} - \alpha_{20}) = o(T^{-1}) \qquad E(\hat{\alpha}_{22} - \alpha_{22}) = -\frac{\alpha_{10}^2 + \alpha_{30}^2}{2T\alpha_{10}\alpha_{30}} + o(T^{-1}),$$

$$E(\hat{\alpha}_{30} - \alpha_{30}) = \frac{\alpha_{30}}{T} + o(T^{-1}) \qquad E(\hat{\alpha}_{33} - \alpha_{33}) = -\frac{1}{T} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{10}) = \frac{3\alpha_{10}^2}{T} + o(T^{-1}) \qquad \mathrm{var}(\hat{\alpha}_{11}) = \frac{1}{T} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{20}) = \frac{\alpha_{10}\alpha_{30}}{T} + o(T^{-1}) \qquad \mathrm{var}(\hat{\alpha}_{22}) = \frac{1}{T} + o(T^{-1}),$$

$$\mathrm{var}(\hat{\alpha}_{30}) = \frac{3\alpha_{30}^2}{T} + o(T^{-1}) \qquad \mathrm{var}(\hat{\alpha}_{33}) = \frac{1}{T} + o(T^{-1}).$$

Proof. In the results given in Theorem 2.2, we set $\alpha_{20} = 0$. ■

The expression for the bias of $\hat{\alpha}_{22}$ is now especially easy to interpret, and it is easy also to analyze the effect of a large distance between the two intercepts. On the other hand, the bias and variances of the QML estimators in a univariate ARCH(1) model, when nothing is estimated in the mean equation, were given at the end of Section 2.1. So we see that the effect of imposing a correlation between the disturbances, when in fact it does not exist, again does not affect the bias structure, although on the other hand, this time there is neither gain nor loss in efficiency to the order of the approximation.

3. AN LM-TYPE TEST ALLOWING FOR BIAS CORRECTION IN THE ESTIMATORS

In this section, we examine how the biases of $O(T^{-1})$ can be used to improve the finite-sample performance of a test for multivariate ARCH effects. We propose that instead of improving the finite-sample behavior of the test by applying a Bartlett-type correction (Bartlett, 1937), which in any case is not available, we proceed by bias correcting the estimates themselves. The justification for this is the following. Let us consider the LM test that takes the form (see Harvey, 1989, p. 169)

$$LM = (D\log L(\tilde{\Psi}_0))' I_{\tilde{\Psi}_0}^{-1} D\log L(\tilde{\Psi}_0), \tag{3.1}$$

where $D\log L(\tilde{\Psi}_0)$ is the vector of first-order derivatives of the log-likelihood function evaluated under the null hypothesis, $\tilde{\Psi}_0$ is the vector of restricted estimates, and $I_{\tilde{\Psi}_0}$ is the estimated information matrix.

Harvey (1989) notes that $D\log L(\tilde{\Psi}_0)' \approx -(\Psi^* - \tilde{\Psi}_0)' D^2 \log L(\Psi^*)$ where Ψ^* is the vector of unrestricted estimates; using this approximation we may write that

$$\begin{aligned} LM &= (D\log L(\tilde{\Psi}_0))' I_{\tilde{\Psi}_0}^{-1} D\log L(\tilde{\Psi}_0) \\ &\approx (\Psi^* - \tilde{\Psi}_0)' D^2 \log L(\Psi^*) I_{\tilde{\Psi}_0}^{-1} D^2 \log L(\Psi^*)(\Psi^* - \tilde{\Psi}_0), \end{aligned} \tag{3.2}$$

which has an asymptotically equivalent form given by

$$(\Psi^* - \tilde{\Psi}_0)' I_{\tilde{\Psi}_0}^{-1} (\Psi^* - \tilde{\Psi}_0). \tag{3.3}$$

In the case where there are no nuisance parameters and the null hypothesis is that $H_0 : \Psi_0 = 0$, we see that $\Psi^* - \tilde{\Psi}_0 = \Psi^*$, in which case the preceding statistic reduces to

$$(\Psi^*)' I_{\tilde{\Psi}_0}^{-1} (\Psi^*). \tag{3.4}$$

Our proposal is to use a bias-corrected estimate of Ψ^*, denoted by Ψ^*_{BC}, in place of Ψ^* in the LM statistic. Because Ψ^*_{BC} is second-order efficient it is anticipated that the statistics will converge to its limiting distribution faster, so the size of the test in small samples will be closer to its nominal level. If the bias correction is nonstochastic, then the information matrix will be unchanged; this is the case for the situation we shall consider subsequently. Thus, if in (3.3) $\tilde{\Psi}_0$ is replaced by $\tilde{\tilde{\Psi}}_0$, the bias approximation for the unrestricted estimate obtained when assuming the null is true, then $\Psi^* - \tilde{\tilde{\Psi}}_0 = \Psi^*_{BC}$, so that the preceding statistic is modified to

$$(\Psi^*_{BC})' I^{-1}_{\tilde{\Psi}_0} (\Psi^*_{BC}). \tag{3.5}$$

Under the alternative, the bias correction that is used is incorrect, so that in addition to improving the size some improvement in power seems likely. The statistic we actually use is

$$(D \log L(\tilde{\tilde{\Psi}}_0))' I^{-1}_{\tilde{\Psi}_0} D \log L(\tilde{\tilde{\Psi}}_0), \tag{3.6}$$

which is asymptotically equivalent to (3.5). To see this note that

$$(D \log L(\tilde{\tilde{\Psi}}_0))' \approx -(\Psi^* - \tilde{\tilde{\Psi}}_0)' D^2 \log L(\Psi^*) = -(\Psi^*_{BC})' D^2 \log L(\Psi^*).$$

On substituting from this approximation into (3.6) we may deduce the required result.

So far we have assumed the absence of parameters not subject to test; however the basic argument is unchanged when such parameters are present. We can choose to ignore them and use the form of the test that tests only a subset of the complete parameter vector, or we can include them, in which case they too can be evaluated at their bias-corrected values. The argument for including them turns mainly on the possibility that the power may be increased because the bias correction is invalid under the alternative.

We show now in more detail through simulation how the bias-correction procedure works. For ease of application we use the Wong and Li model (Case 1 in the previous section) as an example. In particular, because in the LM procedure estimation is conducted only under the null, the bias approximations that, for the conditional variance parameters, are found only in the null case can be employed directly because they are nonstochastic and known. As was seen in Theorem 2.1 bias approximations were found for the constant terms in the variance equations (2.2) and (2.3); these are nuisance parameters for the LM test on the variance parameters, because they are not subject to the test. Bias-corrected estimates for them are easily found. These bias-corrected estimates will be employed in the LM test. However, as has been noted previously, an additional use of the bias approximations for the conditional variance parameters in the null case can also be found. Rather than evaluate these parameters as zero under the null, we may set them at the $O(T^{-1})$ biases because the expected values of the QML estimators are not zero but are close to the bias approximation. This yields the test statistic given in (3.6). To analyze the effect of this use of the bias corrections, we shall first conduct simulations with the bias-corrected constant terms in the LM while setting the parameters under test to zero. Then in further simulations we both use bias-corrected estimates for the constant terms and set the parameters under test to their asymptotic bias values. Hence in this case we are effectively testing a null under which the conditional variance parameters are equal to the expected value of the QML estimator rather than zero.

In this case $\Psi = (\alpha_0, \alpha_1, \alpha_2, \gamma_0, \gamma_1, \gamma_2)'$ is the 6×1 vector of unknown parameters, whereas the null hypothesis we wish to test is

$$H_0 : \alpha_1, \alpha_2, \gamma_1, \gamma_2 = 0.$$

In what follows we shall consider three versions of the LM test statistic.

Model 1. The nuisance parameters are replaced with uncorrected QML estimates, and the parameters under test are set to zero (M1). This is the standard test statistic given in (3.1).

Model 2. The nuisance parameters are replaced with bias-corrected QML estimates, and the parameters under test are set to zero (M2). This case is considered for comparison purposes.

Model 3. The nuisance parameters are replaced with bias-corrected QML estimates, and the parameters under test are set to their asymptotic bias values (M3). This is the statistic given in (3.6).

There are several variants of the LM test, and generally they differ only in the estimator of the information matrix; see, for example, Amemiya (1985) and Dagenais and Dufour (1991) for some related literature. We may distinguish three types of such estimators; the outer product (OP) matrix of the score vector, the hessian (HES) matrix, and the expectation of the hessian (ExpHES) matrix. A nonoperational procedure that we shall examine for comparative purposes uses the true hessian (TrueHES), where the actual values of unknown parameters are employed rather than estimates. Each of these four variants of the LM test will be examined in the simulations in the contexts of Models 1–3.

The LM test based upon the expected hessian is not always available because finding the closed-form solution for the expected hessian may not be possible. In this case, however, it is straightforward. Besides, finding the expected hessian for any higher order specification of the Wong and Li (1997) model would also be straightforward. From Wong and Li (1997) we find on using (2.1)–(2.3) that we may write

$$D \log L(\Psi) = \sum_{t=1}^{T} \left(-\frac{1}{2h_{11t}} \left(1 - \frac{\varepsilon_{1t}^2}{h_{11t}} \right) dh, -\frac{1}{2h_{22t}} \left(1 - \frac{\varepsilon_{2t}^2}{h_{22t}} \right) dh \right)', \tag{3.7}$$

$$Hessian(\Psi) = \begin{pmatrix} hessian_1 & 0 \\ 0 & hessian_2 \end{pmatrix},$$

where

$$hessian_i = -\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2}\left(\frac{2\varepsilon_{it}^2}{h_{iit}}-1\right)\frac{1}{h_{iit}^2}\,dhdh'\right], \qquad i=1,2,$$

$$dh = (1, \varepsilon_{1t-1}^2, \varepsilon_{2t-1}^2)'.$$

On taking expectations through $Hessian(\Psi)$ we have

$$ExpHES(\Psi) = \begin{pmatrix} -\frac{T}{2\alpha_0^2} & -\frac{T}{2\alpha_0} & -\frac{T\gamma_0}{2\alpha_0^2} & 0 & 0 & 0 \\ -\frac{T}{2\alpha_0} & -\frac{3T}{2} & -\frac{T\gamma_0}{2\alpha_0} & 0 & 0 & 0 \\ -\frac{T\gamma_0}{2\alpha_0^2} & -\frac{T\gamma_0}{2\alpha_0} & -\frac{3T\gamma_0^2}{2\alpha_0^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{T}{2\gamma_0^2} & -\frac{T\alpha_0}{2\gamma_0^2} & -\frac{T}{2\gamma_0} \\ 0 & 0 & 0 & -\frac{T\alpha_0}{2\gamma_0^2} & -\frac{3T\alpha_0^2}{2\gamma_0^2} & -\frac{T\alpha_0}{2\gamma_0} \\ 0 & 0 & 0 & -\frac{T}{2\gamma_0} & -\frac{T\alpha_0}{2\gamma_0} & -\frac{3T}{2} \end{pmatrix}$$

with inverse

$$(ExpHES(\Psi))^{-1} = \begin{pmatrix} -\frac{4\alpha_0^2}{T} & \frac{\alpha_0}{T} & \frac{\alpha_0^2}{T\gamma_0} & 0 & 0 & 0 \\ \frac{\alpha_0}{T} & -\frac{1}{T} & 0 & 0 & 0 & 0 \\ \frac{\alpha_0^2}{T\gamma_0} & 0 & -\frac{\alpha_0^2}{T\gamma_0^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{4\gamma_0^2}{T} & \frac{\gamma_0^2}{T\alpha_0} & \frac{\gamma_0}{T} \\ 0 & 0 & 0 & \frac{\gamma_0^2}{T\alpha_0} & -\frac{\gamma_0^2}{T\alpha_0^2} & 0 \\ 0 & 0 & 0 & \frac{\gamma_0}{T} & 0 & -\frac{1}{T} \end{pmatrix}.$$

Notice that all the test statistics that we shall consider can be placed in explicit form using some evaluation of (3.7) together with either the appropriate estimate of the outer product, the hessian, or the expected hessian or with the known expected hessian. We thus have four variants of the LM test. Their size and power are examined in a set of 60,000 simulation experiments. First the test sizes are examined for sample sizes $T = 50$, 100, 200, and 500 where the nuisance parameters are set to $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$. This choice of parameter values and sample sizes was made to ensure that the small-sample biases and variances were not trivial. In the simulations, to examine the power of the tests we considered two sets of values for the variance parameters: (i) $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.16$ and (ii) $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.49$. The first of these represents a moderate departure from the null whereas the second lies close to the stationarity bound and so is a relatively extreme departure.

The results on the test size are given in Table 3.1 and for size-adjusted power in Table 3.2. The first clear result we find is that of the bad size properties in small samples for the HES version of the LM test (see Table 3.1), because it is clearly oversized, even at $T = 500$, in marked contrast to the other tests. The OP, ExpHES, and TrueHes have much better size properties. However, when we check the size-adjusted power of the tests (Table 3.2) the lack of power of the OP test for finite samples is clear whereas the test based on ExpHES is much more powerful than either the OP or HES test. Thus, importantly, we find that among the operational tests the ExpHES test completely dominates the OP and HES tests. At the more extreme alternative the ExpHES and the TrueHES tests have power close to unity at all sample sizes. From the results, the first recommendation in practical applications is to use the ExpHES to test for multivariate ARCH effects. Once we have selected the ExpHES, we can concentrate on the selection among Model 1, Model 2, or Model 3. Model 3 seems to have much better size properties than Model 1 or Model 2. Comparing Models 2 and 3 it is interesting to see the marginal effect of introducing the QML biases in place of zeros in specifying the null hypothesis. As was suggested by our earlier theoretical analysis, the size of the test is improved. Analyzing the TrueHES, the test having the best size properties is again clearly Model 3. We feel this is important evidence because, given that the expected Hessian is known and not estimated, we can more directly attribute the improved size to the bias correction. Thus, the use of bias correction to improve the size of the test, as an alternative to the traditional Bartlett-type correction, is supported in our study. If we consider the size-adjusted power, we observe how the test power in Models 2 and 3 improves on that of Model 1, with Model 3 being slightly superior. So the overall conclusion from the simulations is that, of the operational tests, only ExpHES performs well. Its size is approximately correct even at $T = 50$ whereas it has high power against both the moderate and extreme alternatives at all sample sizes considered. It even dominates the nonoperational TrueHES test for the moderate alternative and has comparable but slightly less power for the extreme alternative.

TABLE 3.1. Size results based on 5% critical values

	OP			HES			ExpHES			TrueHES		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
$T = 500$	0.038	0.038	0.039	0.086	0.081	0.086	0.058	0.056	0.052	0.055	0.059	0.052
$T = 200$	0.047	0.048	0.051	0.146	0.141	0.145	0.057	0.057	0.048	0.062	0.065	0.052
$T = 100$	0.054	0.048	0.054	0.148	0.134	0.146	0.058	0.061	0.044	0.063	0.072	0.053
$T = 50$	0.043	0.040	0.048	0.101	0.095	0.098	0.063	0.062	0.042	0.066	0.080	0.054

Note: The results are based on 60,000 Monte Carlo replications under the null of no ARCH effects; $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$.

TABLE 3.2. Power results based on 5% critical values size-adjusted

	OP			HES			ExpHES			TrueHES		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
When the alternative hypothesis is $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.16$												
$T = 500$	0.940	0.942	0.932	0.996	0.996	0.996	1.000	1.000	1.000	1.000	1.000	1.000
$T = 200$	0.244	0.250	0.218	0.229	0.232	0.229	1.000	1.000	1.000	0.961	0.966	0.962
$T = 100$	0.062	0.079	0.063	0.018	0.020	0.020	0.979	0.979	0.979	0.852	0.857	0.853
$T = 50$	0.040	0.047	0.042	0.025	0.027	0.027	0.815	0.820	0.832	0.708	0.708	0.709
When the alternative hypothesis is $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.49$												
$T = 500$	0.837	0.844	0.846	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$T = 200$	0.537	0.541	0.526	0.719	0.754	0.689	1.000	1.000	1.000	1.000	1.000	1.000
$T = 100$	0.101	0.143	0.087	0.030	0.075	0.015	0.999	1.000	0.999	0.999	0.999	0.999
$T = 50$	0.013	0.018	0.009	0.009	0.009	0.007	0.966	0.969	0.966	0.981	0.982	0.982

Note: The results are based on 60,000 Monte Carlo replications; $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$.

Hence, our simulations support the use of the ExpHES test while bias correcting all the QML estimates.

4. CONCLUSIONS

In this paper we have provided theoretical evidence of the severe biases and large variances that result from unconstrained QML estimation of a simple bivariate-ARCH model under overspecification of the conditional heteroskedasticity processes. When we analyze the model in Wong and Li (1997), we find that some of the estimators can have large variances if the difference between the intercepts in the model is relatively large. In the case of the Engle and Kroner (1995) and Liu and Polasek (1999, 2000) specification, we find that strongly contemporaneously correlated disturbances and/or a large difference between the intercepts can produce large biases in the estimators of the ARCH terms for some combinations of parameters. Under the null of no ARCH effects, the intercepts capture the volatility of the series, and then the results of this paper warn about the testing of multivariate ARCH effects among series that may have very different degrees of volatilities. One rule of thumb in practical applications would be to always standardize the volatilities of the series before they are used in a multivariate model, although the best recommendation is to use the bias expressions that are provided in this paper. We believe that the possibility of extreme biases and variances should be taken into account in practical applications when QML is used as the estimation procedure, and this paper provides an analysis of what happens in a simple bivariate process. In the last section of the paper we show, through Monte Carlo simulations, that the expected hessian form of the LM test for multivariate ARCH effects is much superior to the OP and HES versions; also our bias approximations can be used to improve its finite-sample performance by bias correcting the estimators of the parameters. Our results suggest that this can be considered as an alternative way to improve the finite-sample behavior in testing instead of applying a Bartlett-type correction. The general recommendation from this paper is that when testing for multivariate ARCH effects by performing the LM test, the expected hessian form should be used and all QML estimators should be bias corrected.

REFERENCES

Amemiya, T. (1985) *Advanced Econometrics*. Harvard University Press.

Baba, Y., R.F. Engle, D.F. Kraft, & K.F. Kroner (1991) Multivariate Simultaneous Generalised ARCH. Discussion paper 89-57, University of California, San Diego, Department of Economics.

Bartlett, M.S. (1937) Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London, Series A* 160, 268–282.

Bauwens, L., S. Laurent, and J.V.K. Rombouts (2005) Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, forthcoming.

Bauwens, L. and J.-P. Vandeuren (1995) On the Weak Consistency of the Quasi-Maximum Likelihood Estimator in VAR Models with BEKK-GARCH(1,1) Errors. CORE Discussion paper 9538.

Bollerslev, T., R.F. Engle, & J.M. Wooldridge (1988) A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96, 116–131.

Calzolari, G. & G. Fiorentini (1994) Conditional Heteroscedasticity in Non-linear Simultaneous Equations. EUI Working paper ECO 94/44.

Comte, F. & O. Lieberman (2003) Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84, 61–84.

Cordeiro, G.M. & P. McCullagh (1991) Bias correction in generalised linear models. *Journal of the Royal Statistical Society, Series B* 629–643.

Cox, D.R. & P. Hall (2002) Estimation in a simple random effects model with nonnormal distributions. *Biometrika* 89, 831–840.

Cox, D.R. & E.J. Snell (1968) A general definition of residuals. *Journal of the Royal Statistical Society, Series B* 30, 248–275.

Dagenais, M. & J.M. Dufour (1991) Invariance, nonlinear models, and asymptotic tests. *Econometrica* 50, 1601–1615.

Engle, R.F., D.F. Hendry, & D. Trumble (1985) Small-sample properties of ARCH estimators and tests. *Canadian Journal of Economics* 18, 66–93.

Engle, R.F. & K.F. Kroner (1995) Multivariate simultaneous generalised ARCH. *Econometric Theory* 11, 122–150.

Harmon, R. (1988) The Simultaneous Equations Model with Generalised Autoregressive Conditional Heteroscedasticity: The SEM-GARCH Model. International Finance Discussion paper 322, Board of Governors of the Federal Reserve System, Washington, DC.

Harvey, A.C. (1989) *The Econometric Analysis of Time Series*, 2nd ed. MIT Press.

Iglesias, E.M. & G.D.A. Phillips (2003) Small Sample Estimation Bias in GARCH Models with Any Number of Exogenous Variables in the Mean Equation. Working paper, Michigan State University.

Jeantheau, T. (1998) Strong consistency of estimators of multivariate ARCH models. *Econometric Theory* 14, 70–86.

Jensen, S.T. & A. Rahbek (2004) Asymptotic normality of the QMLE of ARCH in the nonstationary case. *Econometrica* 72, 641–646.

Johansen, S. (2002) A small sample correction of the test for cointegrating rank in the vector autoregressive model. *Econometrica* 70, 1929–1961.

Kraft, D.F. & R.F. Engle (1983) Autoregressive Conditional Heteroscedasticity in Multiple Time Series. Manuscript, University of California, San Diego, Department of Economics.

Ling, S. & M. McAleer (2003) Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19, 278–308.

Linton, O. (1997) An asymptotic expansion in the GARCH(1,1) model. *Econometric Theory* 13, 558–581.

Liu, S. & W. Polasek (1999) Maximum likelihood estimation for the VAR-VARCH model: A new approach. In U. Leopold-Wildburger, G. Feichtinger, & K.-P. Kistner (eds.), *Modelling and Decisions in Economics, Essays in Honour of Franz Ferschl*, pp. 99–113. Physica-Verlag.

Liu, S. & W. Polasek (2000) On Comparing Estimation Methods for VAR-ARCH Models. Working paper, University of Basel.

Lumsdaine, R.L. (1995) Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: A Monte Carlo investigation. *Journal of Business & Economic Statistics* 13(1), 1–10.

McCullagh, P. (1987) *Tensor Methods in Statistics*. Chapman and Hall.

Magnus, J.R. & H. Neudecker (1991) *Matrix Differential Calculus with Applications in Statistics and Econometrics*, rev. ed. Wiley.

MathSoft (1996) *S + GARCH User's Manual, Version 1.0*. Data Analysis Products Division, Mathsoft.

Osiewalski, J. & M. Pipien (2004) Bayesian comparison of bivariate ARCH type models for the main exchange rates in Poland. *Journal of Econometrics* 123, 371–391.

Polasek, W., S. Jin, L. Ren, & R. Vonthein (1999) The BASEL Package. ISO-WWZ, University of Basel.

Polasek, W. and H. Kozumi (1996) The VAR-VARCH Model: A Bayesian Approach. In J.C. Lee, W.O. Johnson, & A. Zellner (eds.), *Modelling and Prediction Honoring Seymour Geisser*, pp. 402–422. Springer-Verlag.

Tuncer, R. (1994) Convergence in Probability of the Maximum Likelihood Estimator of a Multivariate ARMA Model with GARCH(1,1) Errors. CREST Working paper 9441.

Tuncer, R. (2000) Asymptotic Normality of the Maximum Likelihood Estimators of a Multivariate Random Walk with Drift Model having GARCH(1,1) Errors. Paper presented at the ERC-Metu International Conference in Economics IV, Ankara, September 13–16, 2000.

Wong, H. & W.K. Li (1997) On a multivariate conditional heteroscedastic model. *Biometrika* 84, 111–123.

APPENDIX A: Proof of Theorem 2.1

The proof of Theorem 2.1 implies the use of expression (2.4) to find the k_{ij}, the k_{ijl}, and the $k_{ij,l}$ components. Using differential matrix calculus, defining $H_t^{-1} = \begin{pmatrix} h^{11t} & h^{12t} \\ h^{21t} & h^{22t} \end{pmatrix}$, $\varepsilon_t^2 = (\varepsilon_{1t}^2, \varepsilon_{2t}^2)'$, and assuming the parameter vector to be $(\alpha_0, \gamma_0, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$, we obtain the matrix of second-order derivatives shown in Table A1.

Under the assumption of overspecification of multivariate ARCH effects, we get the $K = \{-k_{ij}\}$ matrix and its inverse, respectively (from where the approximations of the variances are obtained):

$$\frac{T}{2}\begin{pmatrix} \frac{1}{\alpha_0^2} & 0 & \frac{1}{\alpha_0} & \frac{\gamma_0}{\alpha_0^2} & 0 & 0 \\ 0 & \frac{1}{\gamma_0^2} & 0 & 0 & \frac{\alpha_0}{\gamma_0^2} & \frac{1}{\gamma_0} \\ \frac{1}{\alpha_0} & 0 & 3 & \frac{\gamma_0}{\alpha_0} & 0 & 0 \\ \frac{\gamma_0}{\alpha_0^2} & 0 & \frac{\gamma_0}{\alpha_0} & \frac{3\gamma_0^2}{\alpha_0^2} & 0 & 0 \\ 0 & \frac{\alpha_0}{c^2} & 0 & 0 & \frac{3\alpha_0^2}{\gamma_0^2} & \frac{\alpha_0}{\gamma_0} \\ 0 & \frac{1}{\gamma_0} & 0 & 0 & \frac{\alpha_0}{\gamma_0} & 3 \end{pmatrix};$$

$$\frac{1}{T}\begin{pmatrix} 4\alpha_0^2 & 0 & -\alpha_0 & -\frac{\alpha_0^2}{\gamma_0} & 0 & 0 \\ 0 & 4\gamma_0^2 & 0 & 0 & -\frac{\gamma_0^2}{\alpha_0} & -\gamma_0 \\ -\alpha_0 & 0 & 1 & 0 & 0 & 0 \\ -\frac{\alpha_0^2}{\gamma_0} & 0 & 0 & \frac{\alpha_0^2}{\gamma_0^2} & 0 & 0 \\ 0 & -\frac{\gamma_0^2}{\alpha_0} & 0 & 0 & \frac{\gamma_0^2}{\alpha_0^2} & 0 \\ 0 & -\gamma_0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Table A1. Second-order derivatives in matrix notation

$$-\frac{T}{2}\begin{pmatrix} (h^{11t})^2 & 0 & (h^{11t})^2\varepsilon_{t-1}^{2\prime} & 0 \\ 0 & (h^{22t})^2 & 0 & (h^{22t})^2\varepsilon_{t-1}^{2\prime} \\ (h^{11t})^2\varepsilon_{t-1}^{2} & 0 & (h^{11t})^2\varepsilon_{t-1}^{2\prime}\varepsilon_{t-1}^{2} & 0 \\ 0 & (h^{22t})^2\varepsilon_{t-1}^{2} & 0 & (h^{22t})^2\varepsilon_{t-1}^{2\prime}\varepsilon_{t-1}^{2} \end{pmatrix}$$

The third-order derivatives k_{ijl} that are different from zero are shown in Table A2.

With these expressions of the second- and third-order derivatives, it is possible to apply the results of McCullagh (1987) to find the bias general expression of the QML estimator under conditional heteroskedasticity. For ease of interpretation, in this paper we offer the closed-form solutions under the case of $\kappa_4 = 0$ and overspecification of multivariate ARCH effects. In this situation, we get the intermediate results to introduce in expression (2.4) that we give next.

The Cox and Snell (1968) expressions that are required (apart from the second-order derivatives, and the third-order derivatives previously given), once we evaluate them when $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$, are (we only give those that are different from zero) shown in Table A3.

Table A2. Third-order derivatives

	Evaluation		Evaluation		Evaluation
k_{111}	$2T(h^{11t})^3$	k_{113}	$2T(h^{11t})^3\varepsilon_{1t-1}^2$	k_{114}	$2T(h^{11t})^3\varepsilon_{2t-1}^2$
k_{222}	$2T(h^{22t})^3$	k_{225}	$2T(h^{22t})^3\varepsilon_{1t-1}^2$	k_{226}	$2T(h^{22t})^3\varepsilon_{2t-1}^2$
k_{133}	$2T(h^{11t})^3\varepsilon_{1t-1}^4$	k_{134}	$2T(h^{11t})^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$	k_{144}	$2T(h^{11t})^3\varepsilon_{2t-1}^4$
k_{255}	$2T(h^{22t})^3\varepsilon_{1t-1}^4$	k_{256}	$2T(h^{22t})^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$	k_{266}	$2T(h^{22t})^3\varepsilon_{2t-1}^4$
k_{333}	$2T(h^{11t})^3\varepsilon_{1t-1}^6$	k_{334}	$2T(h^{11t})^3\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2$	k_{434}	$2T(h^{11t})^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4$
k_{444}	$2T(h^{11t})^3\varepsilon_{2t-1}^6$	k_{555}	$2T(h^{22t})^3\varepsilon_{1t-1}^6$	k_{556}	$2T(h^{22t})^3\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2$
k_{656}	$2T(h^{22t})^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4$	k_{666}	$2T(h^{22t})^3\varepsilon_{2t-1}^6$		

TABLE A3. Evaluation under the case of overspecification of multivariate ARCH effects

	Eval.		Eval.		Eval.		Eval.		Eval.
$\frac{1}{2}k_{131} + k_{13,1}$	$-\frac{T}{2\alpha_0^2}$	$\frac{1}{2}k_{132} + k_{13,2}$	$-\frac{T}{2\alpha_0\gamma_0}$	$\frac{1}{2}k_{133} + k_{13,3}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{134} + k_{13,4}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{135} + k_{13,5}$	$-\frac{T}{2\gamma_0}$
$\frac{1}{2}k_{136} + k_{13,6}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{141} + k_{14,1}$	$-\frac{T\gamma_0}{2\alpha_0^3}$	$\frac{1}{2}k_{142} + k_{14,2}$	$-\frac{T}{2\alpha_0^2}$	$\frac{1}{2}k_{143} + k_{14,3}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{144} + k_{14,4}$	$-\frac{T\gamma_0^2}{2\alpha_0^3}$
$\frac{1}{2}k_{145} + k_{14,5}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{146} + k_{14,6}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{251} + k_{25,1}$	$-\frac{T}{2\gamma_0^2}$	$\frac{1}{2}k_{252} + k_{25,2}$	$-\frac{T\alpha_0}{2\gamma_0^3}$	$\frac{1}{2}k_{253} + k_{25,3}$	$-\frac{T\alpha_0}{2\gamma_0^2}$
$\frac{1}{2}k_{254} + k_{25,4}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{255} + k_{25,5}$	$-\frac{T\alpha_0^2}{2\gamma_0^3}$	$\frac{1}{2}k_{256} + k_{25,6}$	$-\frac{T\alpha_0}{2\gamma_0^2}$	$\frac{1}{2}k_{261} + k_{26,1}$	$-\frac{T}{2\alpha_0\gamma_0}$	$\frac{1}{2}k_{262} + k_{26,2}$	$-\frac{T}{2\gamma_0^2}$
$\frac{1}{2}k_{263} + k_{26,3}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{264} + k_{26,4}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{265} + k_{26,5}$	$-\frac{T\alpha_0}{2\gamma_0^2}$	$\frac{1}{2}k_{266} + k_{26,6}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{331} + k_{33,1}$	$-\frac{3T}{\alpha_0}$
$\frac{1}{2}k_{332} + k_{33,2}$	$-\frac{3T}{\gamma_0}$	$\frac{1}{2}k_{333} + k_{33,3}$	$-3T$	$\frac{1}{2}k_{334} + k_{33,4}$	$-\frac{3T\gamma_0}{\alpha_0}$	$\frac{1}{2}k_{335} + k_{33,5}$	$-\frac{3T\alpha_0}{\gamma_0}$	$\frac{1}{2}k_{336} + k_{33,6}$	$-3T$
$\frac{1}{2}k_{431} + k_{43,1}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{432} + k_{43,2}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{433} + k_{43,3}$	$-\frac{T\gamma_0}{2\alpha_0}$	$\frac{1}{2}k_{434} + k_{43,4}$	$-\frac{T\gamma_0^2}{2\alpha_0^2}$	$\frac{1}{2}k_{435} + k_{43,5}$	$-\frac{T}{2}$
$\frac{1}{2}k_{436} + k_{43,6}$	$-\frac{T\gamma_0}{2\alpha_0}$	$\frac{1}{2}k_{441} + k_{44,1}$	$-\frac{3T\gamma_0^2}{\alpha_0^3}$	$\frac{1}{2}k_{442} + k_{44,2}$	$-\frac{3T\gamma_0}{\alpha_0^2}$	$\frac{1}{2}k_{443} + k_{44,3}$	$-\frac{3T\gamma_0^2}{\alpha_0^2}$	$\frac{1}{2}k_{444} + k_{44,4}$	$-\frac{3T\gamma_0^3}{\alpha_0^3}$
$\frac{1}{2}k_{445} + k_{44,5}$	$-\frac{3T\gamma_0}{\alpha_0}$	$\frac{1}{2}k_{446} + k_{44,6}$	$-\frac{3T\gamma_0^2}{\alpha_0^2}$	$\frac{1}{2}k_{551} + k_{55,1}$	$-\frac{3T\alpha_0}{\gamma_0^2}$	$\frac{1}{2}k_{552} + k_{55,2}$	$-\frac{3T\alpha_0^2}{\gamma_0^3}$	$\frac{1}{2}k_{553} + k_{55,3}$	$-\frac{3T\alpha_0^2}{\gamma_0^2}$
$\frac{1}{2}k_{554} + k_{55,4}$	$-\frac{3T\alpha_0}{\gamma_0^2}$	$\frac{1}{2}k_{555} + k_{55,5}$	$-\frac{3T\alpha_0^3}{\gamma_0^3}$	$\frac{1}{2}k_{556} + k_{55,6}$	$-\frac{3T\alpha_0^2}{\gamma_0^2}$	$\frac{1}{2}k_{651} + k_{65,1}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{652} + k_{65,2}$	$-\frac{T\alpha_0}{2\gamma_0^2}$
$\frac{1}{2}k_{653} + k_{65,3}$	$-\frac{T\alpha_0}{2\gamma_0}$	$\frac{1}{2}k_{654} + k_{65,4}$	$-\frac{T}{2}$	$\frac{1}{2}k_{655} + k_{65,5}$	$-\frac{T\alpha_0^2}{2\gamma_0^2}$	$\frac{1}{2}k_{656} + k_{65,6}$	$-\frac{T\alpha_0}{2\gamma_0}$	$\frac{1}{2}k_{661} + k_{66,1}$	$-\frac{3T}{\alpha_0}$
$\frac{1}{2}k_{662} + k_{66,2}$	$-\frac{3T}{\gamma_0}$	$\frac{1}{2}k_{663} + k_{66,3}$	$-3T$	$\frac{1}{2}k_{664} + k_{66,4}$	$-\frac{3T\gamma_0}{\alpha_0}$	$\frac{1}{2}k_{665} + k_{66,5}$	$-\frac{3T\alpha_0}{\gamma_0}$	$\frac{1}{2}k_{666} + k_{66,6}$	$-3T$

APPENDIX B: Proof of Theorem 2.2

The proof of Theorem 2.2 implies the use of expression (2.4) to find the $k_{ij}, k_{ij,l}$ and the k_{ijl} components. Using differential matrix calculus, defining $H_t^{-1} = \begin{pmatrix} h^{11t} & h^{12t} \\ h^{21t} & h^{22t} \end{pmatrix}$, det $=$ $(h^{11t}h^{22t} + (h^{12t})^2)$, and ordering the parameters as $\alpha_{10}, \alpha_{20}, \alpha_{30}, \alpha_{11}, \alpha_{22}, \alpha_{33}$, we obtain the matrix of second-order derivatives, shown in Table B1.

The third-order derivatives are given in Table B2.

The Cox and Snell (1968) expressions that are required (apart from the second-order derivatives, and the third-order derivatives previously given), once we evaluate them when $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$, are (we only give those that are different from zero) shown in Table B3.

Table B1. Second-order derivatives in matrix notation

$$-\frac{T}{2}\begin{pmatrix}
(h^{11t})^2 & 2h^{11t}h^{12t} & (h^{12t})^2 & (h^{11t})^2\varepsilon^2_{1t-1} & 2h^{11t}h^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1} & (h^{12t})^2\varepsilon^2_{2t-1} \\
2h^{11t}h^{12t} & 2\det & 2h^{12t}h^{22t} & 2h^{11t}h^{12t}\varepsilon^2_{1t-1} & 2\det\varepsilon_{1t-1}\varepsilon_{2t-1} & 2h^{12t}h^{22t}\varepsilon^2_{1t-1} \\
(h^{12t})^2 & 2h^{12t}h^{22t} & (h^{22t})^2 & (h^{12t})^2\varepsilon^2_{1t-1} & 2h^{12t}h^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1} & (h^{22t})^2\varepsilon^2_{2t-1} \\
(h^{11t})^2\varepsilon^2_{1t-1} & 2h^{11t}h^{12t}\varepsilon^2_{1t-1} & (h^{12t})^2\varepsilon^2_{1t-1} & (h^{11t})^2\varepsilon^4_{1t-1} & 2h^{11t}h^{12t}\varepsilon^3_{1t-1}\varepsilon_{2t-1} & (h^{12})^2\varepsilon^2_{1t-1}\varepsilon^2_{2t-1} \\
2h^{11t}h^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1} & 2\det\varepsilon_{1t-1}\varepsilon_{2t-1} & 2h^{12t}h^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1} & 2h^{11t}h^{12t}\varepsilon^3_{1t-1}\varepsilon_{2t-1} & 2\det\varepsilon^2_{1t-1}\varepsilon^2_{2t-1} & 2h^{12t}h^{22t}\varepsilon_{1t-1}\varepsilon^3_{2t-1} \\
(h^{1t})^2\varepsilon^2_{2t-1} & 2h^{12t}h^{22t}\varepsilon^2_{1t-1} & (h^{22t})^2\varepsilon^2_{2t-1} & (h^{12})^2\varepsilon^2_{1t-1}\varepsilon^2_{2t-1} & 2h^{12t}h^{22t}\varepsilon_{1t-1}\varepsilon^3_{2t-1} & (h^{22t})^2\varepsilon^4_{2t-1}
\end{pmatrix}$$

TABLE B2. Third-order derivatives

	Evaluation		Evaluation		Evaluation
k_{111}	$2T(h^{11t})^3$	k_{112}	$4T(h^{11t})^2h^{12t}$	k_{113}	$2T(h^{12t})^2h^{11t}$
k_{114}	$2T(h^{11t})^3\varepsilon^2_{1t-1}$	k_{115}	$4T(h^{11t})^2h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}$	k_{116}	$2Th^{11t}(h^{12t})^2\varepsilon^2_{2t-1}$
k_{212}	$2Th^{11t}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{213}	$2Th^{12t}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{214}	$4T(h^{11t})^2h^{12t}\varepsilon^2_{1t-1}$
k_{215}	$2Th^{11t}\varepsilon_{1t-1}\varepsilon_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{216}	$2Th^{12t}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{222}	$4Th^{12t}(3h^{11t}h^{22t} + (h^{12t})^2)$
k_{223}	$2Th^{22t}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{224}	$2Th^{11t}\varepsilon^2_{1t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{225}	$4Th^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}(3h^{11t}h^{22t} + (h^{12t})^2)$
k_{226}	$2Th^{22t}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{133}	$2T(h^{12t})^2h^{22t}$	k_{134}	$2Th^{11t}(h^{12t})^2\varepsilon^2_{1t-1}$
k_{135}	$2Th^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{136}	$2Th^{22t}(h^{12t})^2\varepsilon^2_{2t-1}$	k_{233}	$4T(h^{22t})^2h^{12t}$
k_{234}	$2Th^{12t}\varepsilon^2_{1t-1}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{235}	$2Th^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{236}	$4Th^{12t}(h^{22t})^2\varepsilon^2_{2t-1}$
k_{144}	$2T(h^{11t})^3\varepsilon^4_{1t-1}$	k_{145}	$4T(h^{11t})^2h^{12t}\varepsilon^3_{1t-1}\varepsilon_{2t-1}$	k_{146}	$2T(h^{12t})^2h^{11t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}$
k_{244}	$4T(h^{11t})^2h^{12t}\varepsilon^4_{1t-1}$	k_{245}	$2Th^{11t}\varepsilon^2_{1t-1}\varepsilon_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{246}	$2Th^{12t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$
k_{155}	$2Th^{11t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{156}	$2Th^{12t}\varepsilon_{1t-1}\varepsilon^3_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{166}	$2T(h^{12t})^2h^{22t}\varepsilon^4_{2t-1}$
k_{255}	$4Th^{12t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}(3h^{11t}h^{22t} + (h^{12t})^2)$	k_{256}	$2Th^{22t}\varepsilon_{1t-1}\varepsilon^3_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{266}	$4T(h^{22t})^2h^{12t}\varepsilon^4_{2t-1}$
k_{333}	$2T(h^{22t})^3$	k_{334}	$2Th^{22t}(h^{12t})^2\varepsilon^2_{1t-1}$	k_{335}	$4T(h^{22t})^2h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}$
k_{336}	$2T(h^{22t})^3\varepsilon^2_{2t-1}$	k_{434}	$4T(h^{12t})^2h^{11t}\varepsilon^4_{1t-1}$	k_{435}	$2Th^{12t}\varepsilon^3_{1t-1}\varepsilon_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$
k_{436}	$2T(h^{12t})^2h^{22t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}$	k_{444}	$2T(h^{11t})^3\varepsilon^6_{1t-1}$	k_{445}	$4T(h^{11t})^2h^{12t}\varepsilon^5_{1t-1}\varepsilon_{2t-1}$
k_{446}	$2T(h^{12t})^2h^{11t}\varepsilon^4_{1t-1}\varepsilon^2_{2t-1}$	k_{335}	$2Th^{22t}\varepsilon^2_{1t-1}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{356}	$4T(h^{22t})^2h^{12t}\varepsilon_{1t-1}\varepsilon^3_{2t-1}$
k_{455}	$2Th^{11t}\varepsilon^4_{1t-1}\varepsilon^2_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$	k_{456}	$2Th^{12t}\varepsilon^3_{1t-1}\varepsilon^3_{2t-1}(h^{11t}h^{22t} + (h^{12t})^2)$	k_{366}	$2T(h^{22t})^3\varepsilon^4_{2t-1}$
k_{466}	$2T(h^{12t})^2h^{22t}\varepsilon^2_{1t-1}\varepsilon^4_{2t-1}$	k_{555}	$4Th^{12t}\varepsilon^3_{1t-1}\varepsilon^3_{2t-1}(3h^{11t}h^{22t} + (h^{12t})^2)$	k_{556}	$2Th^{22t}\varepsilon^2_{1t-1}\varepsilon^4_{2t-1}(h^{11t}h^{22t} + 3(h^{12t})^2)$
k_{656}	$4T(h^{22t})^2h^{12t}\varepsilon_{1t-1}\varepsilon^5_{2t-1}$	k_{666}	$2T(h^{22t})^3\varepsilon^6_{2t-1}$		

TABLE B3. Evaluation under the case of overspecification of multivariate ARCH effects

	Evaluation		Evaluation		Evaluation		Evaluation
$\frac{1}{2}k_{141} + k_{14,1}$	$\frac{-T\alpha_{10}\alpha_{30}^3}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{142} + k_{14,2}$	$\frac{T\alpha_{20}\alpha_{30}^2(\alpha_{10} - \alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{143} + k_{14,3}$	$\frac{T\alpha_{20}^2\alpha_{30}^2}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{144} + k_{14,4}$	$\frac{-T\alpha_{10}^2\alpha_{30}^3}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{145} + k_{14,5}$	$\frac{T\alpha_{20}^2\alpha_{30}^2(\alpha_{10} - \alpha_{30})}{2(\alpha_{10} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{146} + k_{14,6}$	$\frac{T\alpha_{20}^2\alpha_{30}^3}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{241} + k_{24,1}$	$\frac{T\alpha_{10}\alpha_{20}\alpha_{30}^2}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{242} + k_{24,2}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{30} - \alpha_{10})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{243} + k_{24,3}$	$\frac{-T\alpha_{20}^3\alpha_{30}}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{244} + k_{24,4}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}^2}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{245} + k_{24,5}$	$\frac{T\alpha_{20}^3\alpha_{30}(\alpha_{30} - \alpha_{10})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{246} + k_{24,6}$	$\frac{T\alpha_{20}^3\alpha_{30}^2}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{151} + k_{15,1}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{30} - \alpha_{10})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{152} + k_{15,2}$	$\frac{T\alpha_{20}\alpha_{30}(\alpha_{30}^2 + \alpha_{10}^2 - 2\alpha_{20}^2)}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{153} + k_{15,3}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{10} - \alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{154} + k_{15,4}$	$\frac{T\alpha_{10}\alpha_{20}^2\alpha_{30}(\alpha_{30} - \alpha_{10})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{155} + k_{15,5}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{30}^2 + \alpha_{10}^2 - 2\alpha_{20}^2)}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{156} + k_{15,6}$	$\frac{T\alpha_{20}^2\alpha_{30}^2(\alpha_{10} - \alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{161} + k_{16,1}$	$\frac{T\alpha_{20}^4}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{162} + k_{16,2}$	$\frac{T\alpha_{20}^3(\alpha_{30} - \alpha_{10})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{253} + k_{25,3}$	$\frac{T\alpha_{20}(\alpha_{30} - \alpha_{10})(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{254} + k_{25,4}$	$\frac{T\alpha_{10}\alpha_{20}(\alpha_{10} - \alpha_{30})(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{163} + k_{16,3}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{164} + k_{16,4}$	$\frac{T\alpha_{10}\alpha_{20}^4}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{255} + k_{25,5}$	$\frac{T\alpha_{20}(2\alpha_{20}^2 - \alpha_{10}^2 - \alpha_{30}^2)(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{256} + k_{25,6}$	$\frac{T\alpha_{20}\alpha_{30}(\alpha_{30} - \alpha_{10})(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{165} + k_{16,5}$	$\frac{T\alpha_{20}^4(\alpha_{30} - \alpha_{10})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{166} + k_{16,6}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}^2}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{251} + k_{25,1}$	$\frac{T\alpha_{20}(\alpha_{10} - \alpha_{30})(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{252} + k_{25,2}$	$\frac{T(2\alpha_{20}^2 - \alpha_{10}^2 - \alpha_{30}^2)(\alpha_{20}^2 + \alpha_{10}\alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{261} + k_{26,1}$	$\frac{T\alpha_{10}\alpha_{20}^3}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{262} + k_{26,2}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{10} - \alpha_{30})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{263} + k_{26,3}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{264} + k_{26,4}$	$\frac{-T\alpha_{10}^2\alpha_{20}^3}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{265} + k_{26,5}$	$\frac{T\alpha_{10}\alpha_{20}^3(\alpha_{10} - \alpha_{30})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{266} + k_{26,6}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}^2}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{431} + k_{43,1}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{432} + k_{43,2}$	$\frac{T\alpha_{20}^3(\alpha_{10} - \alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{433} + k_{43,3}$	$\frac{T\alpha_{20}^4}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{434} + k_{43,4}$	$\frac{-T\alpha_{10}^2\alpha_{20}^2\alpha_{30}}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{445} + k_{44,5}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2(\alpha_{10} - \alpha_{30})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{442} + k_{44,2}$	$\frac{3T\alpha_{10}\alpha_{20}\alpha_{30}^2(\alpha_{10} - \alpha_{30})}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{446} + k_{44,6}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^3}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{435} + k_{43,5}$	$\frac{T\alpha_{20}^4(\alpha_{10} - \alpha_{30})}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$
$\frac{1}{2}k_{352} + k_{35,2}$	$\frac{T\alpha_{10}\alpha_{20}(\alpha_{30}^2 + \alpha_{10}^2 - 2\alpha_{20}^2)}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{355} + k_{35,5}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{30}^2 + \alpha_{10}^2 - 2\alpha_{20}^2)}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{436} + k_{43,6}$	$\frac{T\alpha_{20}^4\alpha_{30}}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$	$\frac{1}{2}k_{441} + k_{44,1}$	$\frac{-3T\alpha_{10}^2\alpha_{30}^3}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^3}$

$\frac{1}{2}k_{451}+k_{45,1}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}(3\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{356}+k_{35,6}$	$\frac{T\alpha_{10}\alpha_{20}^2\alpha_{30}(\alpha_{10}-\alpha_{30})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{443}+k_{44,3}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{444}+k_{44,4}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^3}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{351}+k_{35,1}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{353}+k_{35,3}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{10}-\alpha_{30})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{354}+k_{35,4}$	$\frac{T\alpha_{10}^2\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{361}+k_{36,1}$	$\frac{T\alpha_{10}^2\alpha_{20}^2}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{363}+k_{36,3}$	$\frac{-T\alpha_{10}^3\alpha_{30}}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{364}+k_{36,4}$	$\frac{T\alpha_{10}^3\alpha_{20}^2}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{365}+k_{36,5}$	$\frac{T\alpha_{10}^2\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{366}+k_{36,6}$	$\frac{-T\alpha_{10}^3\alpha_{30}^2}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$

	Evaluation		Evaluation
$\frac{1}{2}k_{452}+k_{45,2}$	$\frac{3T\alpha_{20}\alpha_{30}[\alpha_{10}((\alpha_{10}^2-4\alpha_{20}^2)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)+\alpha_{30}^2(\alpha_{10}\alpha_{30}+\alpha_{20}^2))-2\alpha_{20}^4\alpha_{30}]}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$	$\frac{1}{2}k_{656}+k_{65,6}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2(3\alpha_{10}-\alpha_{30})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{453}+k_{45,3}$	$\frac{3T\alpha_{20}^2\alpha_{30}[\alpha_{10}^2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)-\alpha_{10}\alpha_{30}(\alpha_{10}\alpha_{30}+\alpha_{20}^2)+2\alpha_{20}^4]}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$	$\frac{1}{2}k_{661}+k_{66,1}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{454}+k_{45,4}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}(3\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{662}+k_{66,2}$	$\frac{3T\alpha_{10}^2\alpha_{20}\alpha_{30}(\alpha_{30}-\alpha_{10})}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{455}+k_{45,5}$	$\frac{3T\alpha_{20}^2\alpha_{30}[\alpha_{10}((\alpha_{10}^2-4\alpha_{20}^2)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)+\alpha_{30}^2(\alpha_{10}\alpha_{30}+\alpha_{20}^2))-2\alpha_{20}^4\alpha_{30}]}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$	$\frac{1}{2}k_{663}+k_{66,3}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^2}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{456}+k_{45,6}$	$\frac{3T\alpha_{20}^2\alpha_{30}^2[\alpha_{10}^2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)-\alpha_{10}\alpha_{30}(\alpha_{10}\alpha_{30}+\alpha_{20}^2)+2\alpha_{20}^4]}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$	$\frac{1}{2}k_{664}+k_{66,4}$	$\frac{3T\alpha_{10}^3\alpha_{20}^2\alpha_{30}}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{362}+k_{36,2}$	$\frac{T\alpha_{10}^2\alpha_{20}(\alpha_{30}-\alpha_{10})}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$	$\frac{1}{2}k_{665}+k_{66,5}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}(\alpha_{30}-\alpha_{10})}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{461}+k_{46,1}$	$\frac{T\alpha_{20}^2[(2\alpha_{20}^2\alpha_{30}+\alpha_{10}\alpha_{30}^2+6\alpha_{10}\alpha_{20}^2)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)+6\alpha_{20}^4\alpha_{30}-3\alpha_{10}\alpha_{30}^2(\alpha_{10}\alpha_{30}+\alpha_{20}^2)]}{4(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$	$\frac{1}{2}k_{666}+k_{66,6}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^3}{(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^3}$
$\frac{1}{2}k_{462}+k_{46,2}$	$\frac{T\alpha_{20}^3[-(2\alpha_{20}^2+\alpha_{10}\alpha_{30}+3\alpha_{10}^2+3\alpha_{30}^2)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)+3\alpha_{10}\alpha_{20}^2\alpha_{30}-6\alpha_{20}^4+3\alpha_{10}^2\alpha_{30}^2]}{2(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$		
$\frac{1}{2}k_{463}+k_{46,3}$	$\frac{T\alpha_{20}^2[(2\alpha_{20}^2\alpha_{10}+\alpha_{10}^2\alpha_{30}+6\alpha_{30}\alpha_{20}^2)(\alpha_{10}\alpha_{30}-\alpha_{20}^2)+6\alpha_{20}^4\alpha_{10}-3\alpha_{30}\alpha_{10}^2(\alpha_{10}\alpha_{30}+\alpha_{20}^2)]}{4(\alpha_{10}\alpha_{30}-\alpha_{20}^2)^4}$		

(*continued*)

TABLE B3. *Continued*

	Evaluation		Evaluation
$\frac{1}{2}k_{464} + k_{46,4}$	$\frac{T\alpha_{10}\alpha_{20}^2[(2\alpha_{20}^2\alpha_{30} + \alpha_{10}\alpha_{30}^2 + 6\alpha_{10}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 6\alpha_{20}^4\alpha_{30} - 3\alpha_{10}\alpha_{30}^2(\alpha_{10}\alpha_{30} + \alpha_{20}^2)]}{4(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$	$\frac{1}{2}k_{651} + k_{65,1}$	$\frac{3T\alpha_{20}^2\alpha_{10}[\alpha_{30}^2(\alpha_{10}\alpha_{30} - \alpha_{20}^2) - \alpha_{10}\alpha_{30}(\alpha_{10}\alpha_{30} + \alpha_{20}^2) + 2\alpha_{20}^4]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$
$\frac{1}{2}k_{465} + k_{46,5}$	$\frac{T\alpha_{20}^4[-(2\alpha_{20}^2 + \alpha_{10}\alpha_{30} + 3\alpha_{10}^2 + 3\alpha_{30}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 3\alpha_{10}\alpha_{20}^2\alpha_{30} - 6\alpha_{20}^4 + 3\alpha_{10}^2\alpha_{30}^2]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$	$\frac{1}{2}k_{652} + k_{65,2}$	$\frac{3T\alpha_{20}\alpha_{10}[\alpha_{30}((\alpha_{30}^2 - 4\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + \alpha_{10}^2(\alpha_{10}\alpha_{30} + \alpha_{20}^2)) - 2\alpha_{20}^4\alpha_{10}]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$
$\frac{1}{2}k_{466} + k_{46,6}$	$\frac{T\alpha_{20}^2\alpha_{30}[(2\alpha_{20}^2\alpha_{10} + \alpha_{10}^2\alpha_{30} + 6\alpha_{30}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 6\alpha_{20}^4\alpha_{10} - 3\alpha_{30}\alpha_{10}^2(\alpha_{10}\alpha_{30} + \alpha_{20}^2)]}{4(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$	$\frac{1}{2}k_{654} + k_{65,4}$	$\frac{3T\alpha_{20}^2\alpha_{10}^2[\alpha_{30}^2(\alpha_{10}\alpha_{30} - \alpha_{20}^2) - \alpha_{10}\alpha_{30}(\alpha_{10}\alpha_{30} + \alpha_{20}^2) + 2\alpha_{20}^4]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$
$\frac{1}{2}k_{551} + k_{55,1}$	$\frac{T(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[(2\alpha_{20}^2\alpha_{30} + \alpha_{30}^2\alpha_{10} + 6\alpha_{10}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 3\alpha_{30}(2\alpha_{20}^4 - \alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{10}^2\alpha_{30}^2)]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$	$\frac{1}{2}k_{655} + k_{65,5}$	$\frac{3T\alpha_{20}^2\alpha_{10}[\alpha_{30}((\alpha_{30}^2 - 4\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + \alpha_{10}^2(\alpha_{10}\alpha_{30} + \alpha_{20}^2)) - 2\alpha_{20}^4\alpha_{10}]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$
$\frac{1}{2}k_{552} + k_{55,2}$	$\frac{T(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[3(\alpha_{10}\alpha_{20}^3\alpha_{30} + \alpha_{10}^2\alpha_{20}\alpha_{30}^2 - \alpha_{20}^5) - \alpha_{20}(2\alpha_{20}^2 + \alpha_{10}\alpha_{30} + 3\alpha_{10}^2 + 3\alpha_{30}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)]}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$		
$\frac{1}{2}k_{553} + k_{55,3}$	$\frac{T(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[(2\alpha_{20}^2\alpha_{10} + \alpha_{10}^2\alpha_{30} + 6\alpha_{30}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 3\alpha_{10}(2\alpha_{20}^4 - \alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{10}^2\alpha_{30}^2)]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$		
$\frac{1}{2}k_{554} + k_{55,4}$	$\frac{T\alpha_{10}(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[(2\alpha_{20}^2\alpha_{30} + \alpha_{30}^2\alpha_{10} + 6\alpha_{10}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 3\alpha_{30}(2\alpha_{20}^4 - \alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{10}^2\alpha_{30}^2)]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$		
$\frac{1}{2}k_{555} + k_{55,5}$	$\frac{T\alpha_{20}(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[3(\alpha_{10}\alpha_{20}^3\alpha_{30} + \alpha_{10}^2\alpha_{20}\alpha_{30}^2 - \alpha_{20}^5) - \alpha_{20}(2\alpha_{20}^2 + \alpha_{10}\alpha_{30} + 3\alpha_{10}^2 + 3\alpha_{30}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2)]}{(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$		
$\frac{1}{2}k_{556} + k_{55,6}$	$\frac{T\alpha_{30}(\alpha_{20}^2 + \alpha_{10}\alpha_{30})[(2\alpha_{20}^2\alpha_{10} + \alpha_{10}^2\alpha_{30} + 6\alpha_{30}\alpha_{20}^2)(\alpha_{10}\alpha_{30} - \alpha_{20}^2) + 3\alpha_{10}(2\alpha_{20}^4 - \alpha_{10}\alpha_{20}^2\alpha_{30} - \alpha_{10}^2\alpha_{30}^2)]}{2(\alpha_{10}\alpha_{30} - \alpha_{20}^2)^4}$		