Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

Dynamic Conditional Correlation: On Properties and Estimation

Gian Piero Aielli

To cite this article: Gian Piero Aielli (2013) Dynamic Conditional Correlation: On Properties and Estimation, Journal of Business & Economic Statistics, 31:3, 282-299, DOI: 10.1080/07350015.2013.771027

To link to this article: https://doi.org/10.1080/07350015.2013.771027

Published online: 22 Jul 2013.

Dynamic Conditional Correlation: On Properties and Estimation

Gian Piero Aielli
(*gian.piero.aielli@gmail.com*)

This article addresses some of the issues that arise with the Dynamic Conditional Correlation (DCC) model. It is proven that the DCC large system estimator can be inconsistent, and that the traditional interpretation of the DCC correlation parameters can result in misleading conclusions. Here, we suggest a more tractable DCC model, called the *c*DCC model. The *c*DCC model allows for a large system estimator that is heuristically proven to be consistent. Sufficient stationarity conditions for *c*DCC processes of interest are established. The empirical performances of the DCC and *c*DCC large system estimators are compared via simulations and applications to real data.

KEY WORDS: Generalized profile likelihood; Integrated correlation; Multivariate GARCH model; Quasi-maximum-likelihood; Two-step estimation.

1. INTRODUCTION

During the last decade, the focus on the variance/correlation decomposition of the asset conditional covariance matrix has become one of the most popular approaches to the modeling of multivariate volatility. Seminal works in this area are the Constant Conditional Correlation (CCC) model by Bollerslev (1990), the Dynamic Conditional Correlation (DCC) model by Engle (2002), and the Varying Correlation (VC) model by Tse and Tsui (2002). Extensions of the DCC model have been proposed by among others, Billio, Caporin, and Gobbo (2006), Cappiello, Engle, and Sheppard (2006), Pesaran and Pesaran (2007), and Franses and Hafner (2009). Related examples are the models by Pelletier (2006), McAleer, Chan, Hoti, and Lieberman (2008), Silvennoinen and Teräsvirta (2009), and Kwan, Li, and Ng (2010).

In the DCC model, the conditional variances follow univariate generalized autoregressive conditional heteroscedasticity (GARCH) models. The conditional correlations are then modeled as peculiar functions of the past GARCH standardized returns. In its original intention, such a modeling approach should have been capable of providing two major advantages. First, due to the modular structure of the DCC conditional covariance matrix, a consistent three-step estimator for large systems (DCC estimator) should have been easily available. Second, due to the explicit parameterization of the conditional correlation process, testing for certain correlation hypotheses, such as whether or not the correlation process is integrated, should have been more immediate than with other data-driven volatility models. Aielli (2006) pointed out that the DCC model is less tractable than expected. For example, the conjecture on the consistency of the second step of the DCC estimator, in which the location correlation parameter is estimated, turns out to be unsubstantiated. As a consequence of this, the third step of the DCC estimator, in which the dynamic correlation parameters are estimated, could be inconsistent in turn. The author then suggested reformulating the DCC correlation driving process as a linear multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) process. The resulting model, called *c*DCC model, allows for an intuitive three-step estimator (*c*DCC estimator) that is feasible with large systems. Compared with the DCC estimator, the *c*DCC estimator requires only a minor amount of additional computational effort.

This article extends Aielli's (2006) results with new theoretical and empirical results. Regarding the DCC model, it is proven that the second step of the DCC estimator can be inconsistent. It is also shown that the traditional interpretation of the dynamic correlation parameters can result in misleading conclusions. For example, in the presence of a unit root, which is commonly tested to check for integrated correlations (see, e.g., Engle and Sheppard 2001; Pesaran and Pesaran 2007), the correlation driving process is weakly stationary.

Regarding the *c*DCC model, sufficient stationarity conditions are established as a modular stationarity principle that is capable of a wide range of applications. In simple terms, if the correlation/standardized return process is stationary, checking for the stationarity of the covariance/return process reduces checking for the stationarity of the univariate GARCH variances. As for the stationarity of the correlation/standardized return process, explicit conditions are derived from a recent result by Boussama, Fuchs, and Stelzer (2011) on the stationarity of linear MGARCH processes. As for the stationarity of the univariate GARCH variances, relatively easy-to-check conditions are available for many specifications adopted in practice (e.g., GARCH, EGARCH, and TARCH; see Francq and Zakoïan 2010).

The behavior of the *c*DCC integrated correlation process is investigated via simulation. It is proven that, in the presence of a unit root in the dynamic correlation parameters, the correlation processes degenerate either to 1 or to -1 in the long run. Otherwise, the correlation processes are mean reverting. Thus, testing for integrated correlation has an unambiguous meaning with the *c*DCC model.

The *c*DCC estimator is shown to be a generalized profile quasi-log-likelihood estimator (Severini 1998). The dynamic

Journal of Business & Economic Statistics
July 2013, Vol. 31, No. 3
DOI: 10.1080/07350015.2013.771027

correlation parameters are the parameters of interest, and the variance parameters and the location correlation parameters are the nuisance parameters. The estimator of the nuisance parameters conditional on the true value of the parameter of interest is proven to be consistent. Relying on this property, a heuristic proof of the consistency of the *c*DCC estimator is provided.

The finite sample performances of the *c*DCC and DCC estimators are compared by means of applications to simulated and real data. It is shown that, under the hypothesis of a correctly specified model, for parameter values that are common in financial applications the bias of the DCC estimator of the location correlation parameter is negligible. For less common parameter values, it can be substantial. In general, the bias is an increasing function of the persistence of the correlation process and of the impact of the innovations. The *c*DCC estimator of the location correlation parameter, instead, appears to be practically unbiased. Some simulation experiments under misspecification are also discussed, in which the DCC and *c*DCC estimators prove to perform very similarly. Regarding the applications to the real data, a small dataset of 10 equity indices and a large dataset of 100 equities are considered. With both datasets, the *c*DCC correlation forecasts are proven to perform as well as, or significantly better than, the corresponding DCC correlation forecasts.

The remainder of the article is organized as follows: Section 2 illustrates a theoretical comparison of the DCC and *c*DCC models; Section 3 discusses the DCC and *c*DCC estimators; Section 4 compares the empirical performances of the two estimators, and Section 5 concludes the article. The proofs of the propositions are compiled in the Appendix.

2. STRUCTURAL PROPERTIES

2.1 The DCC Model

Let $\boldsymbol{y}_t \equiv [y_{1,t}, \ldots, y_{N,t}]'$ denote the vector of excess returns at time $t = 0, \pm 1, \pm 2, \ldots$. It is assumed that $\boldsymbol{y}_t$ is a martingale difference, or $E_{t-1}[\boldsymbol{y}_t] = \boldsymbol{0}$, where $E_{t-1}[\cdot]$ denotes expectations conditional on $\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \ldots$. The asset conditional covariance matrix, $\boldsymbol{H}_t \equiv E_{t-1}[\boldsymbol{y}_t \boldsymbol{y}_t']$, can be written as

$$\boldsymbol{H}_t = \boldsymbol{D}_t^{1/2} \boldsymbol{R}_t \boldsymbol{D}_t^{1/2}, \tag{1}$$

where $\boldsymbol{R}_t \equiv [\rho_{ij,t}]$ is the asset conditional correlation matrix, and $\boldsymbol{D}_t \equiv \mathrm{diag}(h_{1,t}, \ldots, h_{N,t})$ is a diagonal matrix with the asset conditional variances as diagonal elements. By construction, $\boldsymbol{R}_t$ is the conditional covariance matrix of the vector of the standardized returns, $\boldsymbol{\varepsilon}_t \equiv [\varepsilon_{1,t}, \varepsilon_{2,t}, \ldots, \varepsilon_{N,t}]'$, where $\varepsilon_{i,t} \equiv y_{i,t}/\sqrt{h_{i,t}}$. In the DCC model, the diagonal elements of $\boldsymbol{D}_t$ are modeled as univariate GARCH models, that is,

$$h_{i,t} = h_i(\boldsymbol{\theta}_i; y_{i,t-1}, y_{i,t-2}, \ldots), \tag{2}$$

where $h_i(\cdot\,; \cdot, \cdot, \ldots)$ is a known function, and $\boldsymbol{\theta}_i$ is a vector of parameters, $i = 1, 2, \ldots, N$. The conditional correlation matrix is then modeled as a function of the past standardized returns, namely,

$$\boldsymbol{R}_t = \boldsymbol{Q}_t^{*-1/2} \boldsymbol{Q}_t \boldsymbol{Q}_t^{*-1/2}, \tag{3}$$

where

$$\boldsymbol{Q}_t = (1 - \alpha - \beta)\boldsymbol{S} + \alpha \boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}' + \beta \boldsymbol{Q}_{t-1}, \tag{4}$$

where $\boldsymbol{Q}_t \equiv [q_{ij,t}]$, $\boldsymbol{Q}_t^* \equiv \mathrm{diag}(q_{11,t}, \ldots, q_{NN,t})$, $\boldsymbol{S} \equiv [s_{ij}]$, and α and β are scalars. For $\boldsymbol{R}_t$ to be positive definite (pd) it suffices that $\boldsymbol{Q}_t$ is pd, which is the case if $\alpha \geq 0, \beta \geq 0, \alpha + \beta < 1$, and $\boldsymbol{S}$ is pd. Relying on the conjecture that $\boldsymbol{S}$ is the second moment of $\boldsymbol{\varepsilon}_t$ (see below), $\boldsymbol{S}$ is commonly assumed to be unit-diagonal.

Some simulated series of DCC conditional correlations are plotted in Figure 1. The persistence of the correlation process is an increasing function of $\alpha + \beta$. The higher the impact of the innovations, as measured by α, the higher the variance of the correlation process. The parameter s_{ij} is a location parameter.

More general DCC models can be obtained setting

$$\boldsymbol{Q}_t = (\boldsymbol{\iota}\boldsymbol{\iota}' - \boldsymbol{A} - \boldsymbol{B}) \odot \boldsymbol{S} + \boldsymbol{A} \odot \boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}' + \boldsymbol{B} \odot \boldsymbol{Q}_{t-1}, \tag{5}$$

(Engle 2002, eq. (24)), where $\boldsymbol{A} \equiv [\alpha_{ij}]$, $\boldsymbol{B} \equiv [\beta_{ij}]$, $\boldsymbol{\iota}$ denotes the $N \times 1$ vector with unit entries, and $\odot$ denotes the elementwise (Hadamard) matrix product. For $\boldsymbol{Q}_t$ to be pd, it suffices that $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive semidefinite (psd), and $(\boldsymbol{\iota}\boldsymbol{\iota}' - \boldsymbol{A} - \boldsymbol{B}) \odot \boldsymbol{S}$ is pd (Ding and Engle 2001).

The DCC correlation driving process, $\boldsymbol{Q}_t$, is often treated as a linear MGARCH process (see, e.g., Engle 2002, eq. (18)). The role of the parameters α, β, and $\boldsymbol{S}$, is then interpreted accordingly. Indeed, since the conditional covariance matrix of $\boldsymbol{\varepsilon}_t$ is $\boldsymbol{R}_t$ (not $\boldsymbol{Q}_t$), the process $\boldsymbol{Q}_t$ is not a linear MGARCH. A consequence of this is that the traditional interpretation of α, β, and $\boldsymbol{S}$ can lead to misleading conclusions. Consider for example the location correlation parameter, $\boldsymbol{S}$. Applying a standard result on linear MGARCH processes (Engle and Kroner 1995), $\boldsymbol{S}$ is thought of as the second moment of $\boldsymbol{\varepsilon}_t$, or,

$$\boldsymbol{S} = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] \tag{6}$$

(Engle 2002, eqs. (18), (23), and (24)). Relying on this, during the fitting of large systems, $\boldsymbol{S}$ is replaced by the sample second moment of the estimated standardized returns as a feasible estimator (Engle and Sheppard 2001; Billio, Caporin, and Gobbo 2006; Cappiello et al. 2006; Pesaran and Pesaran 2007; Franses and Hafner 2009). Unfortunately, as proven by the following proposition, Equation (6) does not hold in general.

Proposition 2.1. For the model in Equation (5), suppose that $N = 2$, $\alpha_{11} > 0$, $\alpha_{22} = \alpha_{12} = \beta_{11} = \beta_{22} = \beta_{12} = 0$, and $0 < s_{12} < \sqrt{1 - \alpha_{11}}$. Then, $\boldsymbol{R}_t$ is pd, and $E[\varepsilon_{1,t}\, \varepsilon_{2,t}] > s_{12}$.

Proof. All proofs are reported in the Appendix.

The only case in which Equation (6) holds seems to be the case of constant conditional correlations ($\alpha = \beta = 0$). Note in fact that, from Equation (4), under stationarity it follows that $\boldsymbol{S} = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ if and only if $E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] = E[\boldsymbol{Q}_t]$. But $E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] = E[E_{t-1}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']] = E[\boldsymbol{R}_t] = E[\boldsymbol{Q}_t^{*-1/2} \boldsymbol{Q}_t \boldsymbol{Q}_t^{*-1/2}]$, where, if $\boldsymbol{Q}_t$ is dynamic, it holds that a.s. $\boldsymbol{Q}_t^{*-1/2} \boldsymbol{Q}_t \boldsymbol{Q}_t^{*-1/2} \neq \boldsymbol{Q}_t$.

Proposition 2.2. Suppose that $\alpha + \beta < 1$ and that $E[\boldsymbol{Q}_t]$ and $E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ are independent of t. Then,

$$\boldsymbol{S} = \frac{1 - \beta}{1 - \alpha - \beta} E\left[\boldsymbol{Q}_t^{*1/2} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \boldsymbol{Q}_t^{*1/2}\right] - \frac{\alpha}{1 - \alpha - \beta} E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']. \tag{7}$$

Equation (7) shows that the expression of $\boldsymbol{S}$ in terms of process moments is more complicated than Equation (6). Specifically,

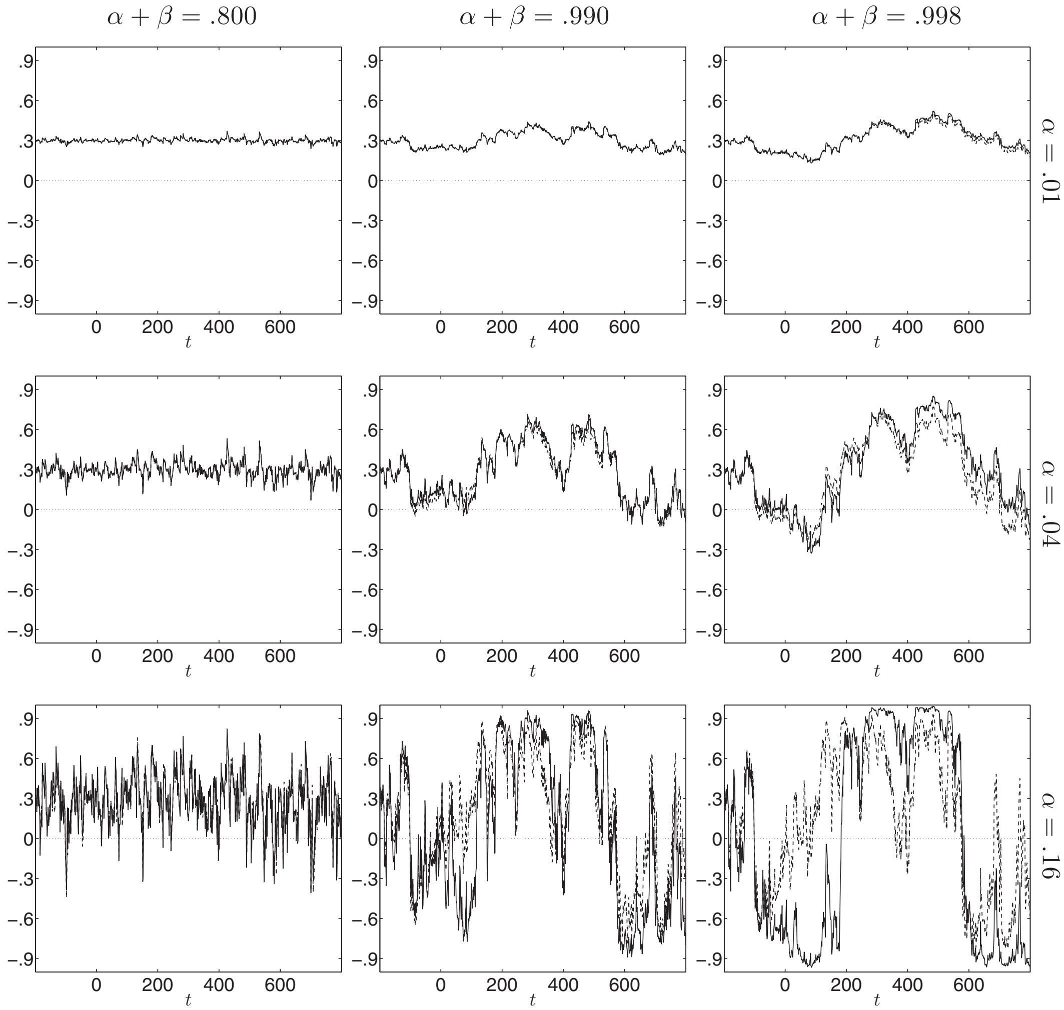

Figure 1. DCC and cDCC conditional correlations. Simulated series of $\rho_{ij,t}$. The DGP parameter values are reported at the top of the panel for $\alpha + \beta$ and on the right-hand side of the panel for α. The location parameter is set as $s_{ij} = 0.3$. DCC in straight line and cDCC in dashed line.

replacing $\boldsymbol{S}$ with the sample second moment of $\boldsymbol{\varepsilon}_t$ proves not to be an obvious estimation device.

Again relying on the conjecture that $\boldsymbol{Q}_t$ follows a linear MGARCH process, if there is a unit root in the dynamic correlation parameters (i.e., for $\alpha + \beta = 1$), the elements of $\boldsymbol{Q}_t$ are thought of as integrated GARCH processes (Engle 2002, eq. (17)). By analogy, the hypothesis $\alpha + \beta = 1$ is often tested as a test of integrated correlation (see, e.g., Engle and Sheppard 2001; Pesaran and Pesaran 2007). Indeed, as proven by the following arguments, for $\alpha + \beta = 1$ the elements of $\boldsymbol{Q}_t$ are weakly stationary. Recalling that $\boldsymbol{S}$ is unit-diagonal, for $\alpha + \beta = 1$ the diagonal elements of $\boldsymbol{Q}_t$ can be written as

$$q_{ii,t} = (1 - \beta) + \beta q_{ii,t-1} + \alpha\left(\varepsilon_{it-1}^2 - 1\right). \tag{8}$$

Since $\alpha(\varepsilon_{it-1}^2 - 1)$ is a martingale difference, $q_{ii,t}$ follows an AR(1) process whose autoregressive parameter is β. For $\alpha + \beta = 1$ and $\alpha > 0$, the process is weakly stationary. For $\alpha + \beta = 1$ and $\alpha = 0$, the error term is zero and $q_{ii,t}$ is constant. Since $\boldsymbol{Q}_t$ is pd, it holds that $|q_{ij,t}| \leq \sqrt{q_{ii,t}\, q_{jj,t}}$. Therefore, for $\alpha + \beta = 1$, the second moment of $q_{ij,t}$ is finite, in which case, if $q_{ij,t}$ is strictly stationary, it is weakly stationary.

After recognizing that $\boldsymbol{Q}_t$ is not a linear MGARCH, finding a tractable representation of $\boldsymbol{Q}_t$ proves to be difficult. A consequence of this is that the dynamic properties of $\boldsymbol{R}_t$ turn out to be essentially unknown. Similar problems also arise with a major competitor of the DCC model, the VC model by Tse and Tsui (2002). In this model, the conditional correlation process is defined as

$$\boldsymbol{R}_t = (1 - \alpha - \beta)\boldsymbol{S} + \alpha \boldsymbol{\mathcal{R}}_{M,t-1} + \beta \boldsymbol{R}_{t-1}, \tag{9}$$

where $M \geq N$ is fixed arbitrary and $\boldsymbol{\mathcal{R}}_{M,t}$ is the sample correlation matrix of $\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-1}, \ldots, \boldsymbol{\varepsilon}_{t-M+1}$. Since $\boldsymbol{R}_t$ is generally not the conditional expectation of $\boldsymbol{\mathcal{R}}_{M,t}$, the process $\boldsymbol{R}_t$ is not standard. Taking the expectations of both members of Equation (9), under stationarity it yields

$$\boldsymbol{S} = \frac{1 - \beta}{1 - \alpha - \beta} E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] - \frac{\alpha}{1 - \alpha - \beta} E[\boldsymbol{\mathcal{R}}_{M,t}]. \tag{10}$$

As with the DCC model, $\boldsymbol{S}$ turns out to be neither easy to interpret nor simple to estimate by means of a moment estimator.

2.2 The *c*DCC Model

The tractability of the DCC model can be substantially improved by reformulating the correlation driving process as

$$\boldsymbol{Q}_t = (1-\alpha-\beta)\,\boldsymbol{S} + \alpha\left\{\boldsymbol{Q}_{t-1}^{*1/2}\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}'\,\boldsymbol{Q}_{t-1}^{*1/2}\right\} + \beta\boldsymbol{Q}_{t-1} \tag{11}$$

(Aielli 2006). The resulting model is called corrected DCC model or *c*DCC model. Some simulated series of *c*DCC conditional correlations were generated, starting with the same seed of the corresponding DCC series, and are plotted in Figure 1. It is shown that, for typical parameter values (i.e., for small α and large $\alpha+\beta$) the DCC and *c*DCC data-generating processes are very similar. Pre- and post-multiplying the right-hand side and the left-hand side of Equation (11) by an arbitrary pd diagonal matrix, $\boldsymbol{Z}$, yields

$$\bar{\boldsymbol{Q}}_t = (1-\alpha-\beta)\,\bar{\boldsymbol{S}} + \alpha\left\{\bar{\boldsymbol{Q}}_{t-1}^{*1/2}\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}'\,\bar{\boldsymbol{Q}}_{t-1}^{*1/2}\right\} + \beta\bar{\boldsymbol{Q}}_{t-1}, \tag{12}$$

where $\bar{\boldsymbol{S}} \equiv \boldsymbol{ZSZ}$, $\bar{\boldsymbol{Q}}_t \equiv \boldsymbol{Z}\boldsymbol{Q}_t\boldsymbol{Z}$, and $\bar{\boldsymbol{Q}}_t^{*1/2} \equiv \boldsymbol{Q}_t^{*1/2}\boldsymbol{Z}$. Noting that $\boldsymbol{Q}_t^{*1/2}\boldsymbol{Z} = \boldsymbol{Z}\boldsymbol{Q}_t^{*1/2}$, it follows that $\bar{\boldsymbol{Q}}_t^{*-1/2}\bar{\boldsymbol{Q}}_t\bar{\boldsymbol{Q}}_t^{*-1/2} = \boldsymbol{Q}_t^{*-1/2}\boldsymbol{Q}_t\boldsymbol{Q}_t^{*-1/2} = \boldsymbol{R}_t$. Thus, $\boldsymbol{Q}_t$ and $\bar{\boldsymbol{Q}}_t$ deliver the same conditional correlation processes or $\boldsymbol{S}$ is not identifiable. As a natural identification condition, it is assumed that $\boldsymbol{S}$ is unit-diagonal. Note that with the DCC model, this assumption is an overidentifying restriction.

One might argue that the *c*DCC model is not really a correlation model in that $\boldsymbol{R}_t$ is modeled only implicitly, as a byproduct of the model of $\boldsymbol{Q}_t$. A more explicit representation of $\rho_{ij,t}$ can be obtained as follows. Let us divide the numerator and denominator of the right-hand side of $\rho_{ij,t} = q_{ij,t}/\sqrt{q_{ii,t}q_{jj,t}}$ by $\sqrt{q_{ii,t-1}q_{jj,t-1}}$. This yields

$$\rho_{ij,t} = \frac{\omega_{ij,t-1} + \alpha\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \beta\rho_{ij,t-1}}{\sqrt{\left\{\omega_{ii,t-1} + \alpha\varepsilon_{i,t-1}^2 + \beta\rho_{ii,t-1}\right\}\left\{\omega_{jj,t-1} + \alpha\varepsilon_{j,t-1}^2 + \beta\rho_{jj,t-1}\right\}}}, \tag{13}$$

where $\omega_{ij,t} \equiv (1-\alpha-\beta)s_{ij}/\sqrt{q_{ii,t}\,q_{jj,t}}$. The above equation shows that with the *c*DCC formula of $\rho_{ij,t}$ the relevant innovations and past correlations are combined into a correlation-like ratio. The denominator of the time-varying parameters, $\omega_{ij,t}$, $\omega_{ii,t}$, and $\omega_{jj,t}$, can be thought of as an ad hoc correction required for purposes of tractability. Let us now consider the DCC analog of Equation (13), which turns out to be

$$\rho_{ij,t} = \frac{\omega_{ij,t-1} + \alpha_{t-1}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \beta\rho_{ij,t-1}}{\sqrt{\left\{\omega_{ii,t-1} + \alpha_{t-1}\varepsilon_{i,t-1}^2 + \beta\rho_{ii,t-1}\right\}\left\{\omega_{jj,t-1} + \alpha_{t-1}\varepsilon_{j,t-1}^2 + \beta\rho_{jj,t-1}\right\}}},$$

where $\omega_{ij,t} \equiv (1-\alpha-\beta)\,s_{ij}/\sqrt{q_{ii,t}\,q_{jj,t}}$ and $\alpha_t \equiv \alpha/\sqrt{q_{ii,t}\,q_{jj,t}}$. Compared with the *c*DCC formula, the DCC formula involves more time-varying parameters, none of them being supported by any apparent motivation. As for the VC model (see Equation (9)), the correlation process is modeled explicitly, but at the cost of an ad hoc innovation term, $\boldsymbol{\mathcal{R}}_{M,t}$, which makes the model difficult to study (see Section 2.1).

With the *c*DCC model, $q_{ii,t}$ follows a GARCH(1,1) process (Bollerslev 1986). For $\alpha+\beta=1$, the intercept of $q_{ii,t}$ is zero, in which case $q_{ii,t} \longrightarrow 0$ for $t \longrightarrow \infty$ (Nelson 1990, prop. 1). This can cause numerical problems if, for $\alpha+\beta=1$ and large t, one computes the conditional correlations as $\rho_{ij,t} = q_{ij,t}/\sqrt{q_{ii,t}q_{jj,t}}$. Such problems disappear by computing $\rho_{ij,t}$ as in Equation (13), which, for $\alpha+\beta=1$, does not depend on Q_t.

Setting $\boldsymbol{\varepsilon}_t^* \equiv \boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t$, it follows that $E_{t-1}[\boldsymbol{\varepsilon}_t^*] = \boldsymbol{0}$ and $E_{t-1}[\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*\prime}] = \boldsymbol{Q}_t$, where $\boldsymbol{Q}_t = (1-\alpha-\beta)\,\boldsymbol{S} + \alpha\,\boldsymbol{\varepsilon}_{t-1}^*\boldsymbol{\varepsilon}_{t-1}^{*\prime} + \beta\,\boldsymbol{Q}_{t-1}$. Therefore, $\boldsymbol{\varepsilon}_t^*$ follows a linear MGARCH process (Engle and Kroner 1995). Specifically, $\boldsymbol{\varepsilon}_t^*$ follows a MARCH model (Ding and Engle 2001). A more general *c*DCC model can be obtained by assuming a more general linear MGARCH model for $\boldsymbol{\varepsilon}_t^*$. Assuming $\boldsymbol{\varepsilon}_t^*$ as a Baba-Engle-Kraft-Kroner (BEKK) model (Engle and Kroner 1995), after replacing $\boldsymbol{\varepsilon}_t^*$ with $\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t$ yields

$$\boldsymbol{Q}_t = \boldsymbol{C} + \sum_{q=1}^{\bar{Q}}\sum_{k=1}^{K}\bar{\boldsymbol{A}}_{q,k}\left\{\boldsymbol{Q}_{t-\mathrm{q}}^{*1/2}\boldsymbol{\varepsilon}_{t-\mathrm{q}}\boldsymbol{\varepsilon}_{t-\mathrm{q}}'\,\boldsymbol{Q}_{t-\mathrm{q}}^{*1/2}\right\}\bar{\boldsymbol{A}}_{q,k}' + \sum_{p=1}^{\bar{P}}\sum_{k=1}^{K}\bar{\boldsymbol{B}}_{p,k}\,\boldsymbol{Q}_{t-p}\,\bar{\boldsymbol{B}}_{p,k}', \tag{14}$$

where

$$\boldsymbol{C} \equiv \boldsymbol{S} - \sum_{q=1}^{\bar{Q}}\sum_{k=1}^{K}\bar{\boldsymbol{A}}_{q,k}\,\boldsymbol{S}\bar{\boldsymbol{A}}_{q,k}' - \sum_{p=1}^{\bar{P}}\sum_{k=1}^{K}\bar{\boldsymbol{B}}_{p,k}\,\boldsymbol{S}\bar{\boldsymbol{B}}_{p,k}'. \tag{15}$$

With this model, $\boldsymbol{Q}_t$ is pd provided that $\boldsymbol{C}$ is pd. A special case of the BEKK model is the Diagonal Vech model (Bollerslev, Engle, and Wooldridge 1988), which, after replacing $\boldsymbol{\varepsilon}_t^*$ with $\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t$, yields

$$\boldsymbol{Q}_t = \left(\boldsymbol{\iota\iota}' - \sum_{q=1}^{Q}\boldsymbol{A}_q - \sum_{p=1}^{P}\boldsymbol{B}_p\right)\odot\boldsymbol{S} + \sum_{q=1}^{Q}\boldsymbol{A}_q\odot\left\{\boldsymbol{Q}_{t-q}^{*1/2}\,\boldsymbol{\varepsilon}_{t-q}\,\boldsymbol{\varepsilon}_{t-q}'\,\boldsymbol{Q}_{t-q}^{*1/2}\right\} + \sum_{p=1}^{P}\boldsymbol{B}_p\odot\boldsymbol{Q}_{t-p}. \tag{16}$$

With this model, $\boldsymbol{Q}_t$ is pd if $\boldsymbol{A}_{\mathrm{q}} \equiv [\alpha_{ij,\mathrm{q}}]$ is psd for $q = 1, 2, \ldots, Q$, $\boldsymbol{B}_{\mathrm{p}} \equiv [\beta_{ij,\mathrm{p}}]$ is psd for $p = 1, 2, \ldots, P$, and the intercept is pd (Ding and Engle 2001). Setting $\boldsymbol{A}_{\mathrm{q}} = \alpha_{\mathrm{q}}\,\boldsymbol{\iota\iota}'$ for $\mathrm{q} = 1, 2, \ldots, Q$, and $\boldsymbol{B}_{\mathrm{p}} = \beta_{\mathrm{p}}\,\boldsymbol{\iota\iota}'$ for $\mathrm{p} = 1, 2, \ldots, P$, provides the Scalar *c*DCC model (Ding and Engle 2001), which, for $\mathrm{P} = \mathrm{Q} = 1$, is the *c*DCC model in Equation (11).

2.3 *c*DCC Stationarity Principle

Relying on the modular structure of the *c*DCC model, the *c*DCC stationarity conditions can be formulated as a flexible stationarity principle (*c*DCC stationarity principle). In simple terms, if the correlation/standardized return process, $(\boldsymbol{R}_t, \boldsymbol{\varepsilon}_t)$, is stationary, checking for the stationarity of the covariance/return process, $(\boldsymbol{H}_t, \boldsymbol{y}_t)$, reduces checking for the stationarity of the univariate GARCH processes, $h_{i,t}$, $i = 1, 2, \ldots, N$. Given that relatively easy-to-check stationarity conditions are available for many univariate GARCH specifications adopted in practice (see Francq and Zakoïan 2010), the *c*DCC stationarity principle is capable of a wide range of applications. A similar principle holds for the DCC and VC models as well, but with the cDCC model it takes on a practical interest thanks to a recent result by Boussama, Fuchs, and Stelzer (2011), relying on which explicit stationarity conditions for the cDCC specification of $(\boldsymbol{R}_t, \boldsymbol{\varepsilon}_t)$ can be derived. Let us now introduce some notations. For the BEKK cDCC model in Equations (14) and (15), set $\boldsymbol{\mathcal{A}}_{\mathrm{q}} \equiv \boldsymbol{U}\{\sum_{\mathrm{k}=1}^{\mathrm{K}} \bar{\boldsymbol{A}}_{\mathrm{q,k}} \otimes \bar{\boldsymbol{A}}_{\mathrm{q,k}}\}\boldsymbol{W}'$ for $q = 1, 2, \ldots, \bar{Q}$, and $\boldsymbol{\mathcal{B}}_{\mathrm{p}} \equiv \boldsymbol{U}\{\sum_{\mathrm{k}=1}^{\mathrm{K}} \bar{\boldsymbol{B}}_{\mathrm{p,k}} \otimes \bar{\boldsymbol{B}}_{\mathrm{p,k}}\}\boldsymbol{W}'$ for $\mathrm{p} = 1, 2, \ldots, \bar{P}$, where $\otimes$ denotes the Kronecker product. The matrices $\boldsymbol{U}$ and $\boldsymbol{W}$ are the unique $N(N+1)/2 \times N^2$ matrices such that $\mathrm{vech}(\boldsymbol{Z}) = \boldsymbol{U}\mathrm{vec}(\boldsymbol{Z})$, $\mathrm{vec}(\boldsymbol{Z}) = \boldsymbol{W}'\mathrm{vech}(\boldsymbol{Z})$, and $\boldsymbol{U}\boldsymbol{W}' = \boldsymbol{I}_{N(N+1)/2}$ for any N-dimensional symmetric matrix $\boldsymbol{Z}$, where $\boldsymbol{I}_M$ is the M-dimensional identity matrix and vec and vech, respectively, denote the operators stacking the columns and the lower triangular part of the matrix argument. The existence and uniqueness of $\boldsymbol{U}$ and $\boldsymbol{W}$ hold by linearity of the vech and vec operators. Let $h_{i,t} = \mathscr{H}_i(\boldsymbol{\theta}_i; \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots)$ denote the representation of $h_{i,t}$ in terms of standardized innovations, obtained by recursively substituting backward for lagged $y_{i,t} = \varepsilon_{i,t}\sqrt{h_{i,t}}$ into $h_{i,t} = h_i(\boldsymbol{\theta}_i; y_{i,t-1}, y_{i,t-2}, \ldots)$. It is assumed that $\boldsymbol{\varepsilon}_t = \boldsymbol{R}_t^{1/2}\boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t$ is iid such that $E[\boldsymbol{\eta}_t] = 0$ and $E[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \boldsymbol{I}_N$, and $\boldsymbol{R}_t^{1/2}$ is the unique psd matrix such that $\boldsymbol{R}_t^{1/2}\boldsymbol{R}_t^{1/2} = \boldsymbol{R}_t$ (the matrix $\boldsymbol{R}_t^{1/2}$ is computed from the spectral decomposition of $\boldsymbol{R}_t$).

Proposition 2.3. (cDCC stationarity principle). For the BEKK cDCC model in Equations (14) and (15), suppose that:

(H1) the density of $\boldsymbol{\eta}_t$ is absolutely continuous with respect to the Lebesgue measure, positive in a neighborhood of the origin;
(H2) $\boldsymbol{C}$ is pd; and
(H3) the in-modulus largest eigenvalue of $\sum_{\mathrm{q}=1}^{\bar{\mathrm{Q}}} \boldsymbol{\mathcal{A}}_{\mathrm{q}} + \sum_{\mathrm{p}=1}^{\bar{\mathrm{P}}} \boldsymbol{\mathcal{B}}_{\mathrm{p}}$ is less than 1.
Then, (i) the process $[\mathrm{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}_t']'$ admits a nonanticipative, strictly and weakly stationary, and ergodic solution. In addition to H1–H3, suppose that:
(H4) $\mathscr{H}_i(\boldsymbol{\theta}_i; \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots)$ is measurable for $i = 1, 2, \ldots, N$.
Then, (ii) the process $[\mathrm{vech}(\boldsymbol{H}_t)', \boldsymbol{y}_t', \mathrm{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}_t']'$ admits a nonanticipative, strictly stationary, and ergodic solution. In addition to H1–H4, suppose that:
(H5) $E[y_{i,t}^2] < \infty$ for $i = 1, 2, \ldots, N$.

Then, (iii) the process $\boldsymbol{y}_t$ admits a weakly stationary solution. Assumption H1 is typically fulfilled in any practical application. As for assumptions H2 and H3, they are relatively easy to check. As for the variance assumptions, H4 and H5, their explicit form (if any) will depend on the model of $h_{i,t}$, $i = 1, 2, \ldots, N$. It is worth noting that H4 requires the measurability of $\mathscr{H}_{i,t}$ under strictly stationary $\varepsilon_{i,t}$, which is a less stringent condition than the usual one of iid $\varepsilon_{i,t}$.

As an illustrative application of Proposition 2.3, consider the Gaussian cDCC model with GARCH(1,1) conditional variances. Since $\boldsymbol{\varepsilon}_t$ is Gaussian, $\boldsymbol{\eta}_t$ is Gaussian, which implies that H1 holds. The intercept of $\boldsymbol{Q}_t$ satisfies $\boldsymbol{C} = (1 - \alpha - \beta)\boldsymbol{S}$; therefore, if $\boldsymbol{S}$ is pd and $\alpha + \beta < 1$, assumption H2 holds. The matrix $\sum_{\mathrm{q}=1}^{\bar{\mathrm{Q}}} \boldsymbol{\mathcal{A}}_{\mathrm{q}} + \sum_{\mathrm{p}=1}^{\bar{\mathrm{P}}} \boldsymbol{\mathcal{B}}_{\mathrm{p}}$ turns out to be scalar and equal to $(\alpha + \beta)\boldsymbol{I}_{N(N+1)/2}$. All the eigenvalues of $(\alpha + \beta)\boldsymbol{I}_{N(N+1)/2}$ are equal to $\alpha + \beta$; therefore, if $\alpha + \beta < 1$, assumption H3 holds. The GARCH(1,1) model is defined as $h_{i,t} = c_i + a_i\, y_{i,t-1}^2 + b_i\, h_{i,t-1}$, where $c_i > 0$, $a_i \geq 0$, and $b_i \geq 0$. If $\varepsilon_{i,t}$ is strictly stationary, assumption H4 holds provided that $E[\log(a_i\, \varepsilon_{i,t}^2 + b_i)] < 0$ and $E[\max\{0, \log(a_i\, \varepsilon_{i,t}^2 + b_i)\}] < \infty$ (Brandt 1986; Bougerol and Picard 1992). If $\varepsilon_{i,t}$ is iid, as in the Gaussian cDCC model, for H4 to hold it suffices that $E[\log(a_i\, \varepsilon_{i,t}^2 + b_i)] < 0$ (Nelson 1990). The latter condition is less stringent than $a_i + b_i < 1$, in which case assumption H5 holds (Bollerslev 1986). In summary, if $\boldsymbol{S}$ is pd and $\alpha + \beta < 1$, the correlation/standardized return process is strictly and weakly stationary; if, in addition, $a_i + b_i < 1$ for $i = 1, 2, \ldots, N$, the covariance/return process is strictly stationary, and the return process is weakly stationary.

According to a result by Engle and Kroner (1995), $\boldsymbol{S}$ is the sample second moment of $\boldsymbol{\varepsilon}_t^*$. Equivalently, $\boldsymbol{S}$ is the sample second moment of $\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t$.

Proposition 2.4. For the BEKK cDCC model in Equations (14) and (15), suppose that assumptions H1–H3 of Proposition 2.3 hold. Then, $\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t$ is covariance stationary, and

$$\boldsymbol{S} = E\left[\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\boldsymbol{Q}_t^{*1/2}\right]. \tag{17}$$

Equation (17) shows that, as with the DCC and VC models (see Equations (7) and (10)), also with the cDCC model $\boldsymbol{S}$ is generally not the second moment of $\boldsymbol{\varepsilon}_t$. However, differing from the DCC and VC models, with the cDCC model the sample counterpart of $\boldsymbol{S}$ is by construction pd. Relying on this property, a pd generalized profile quasi-log-likelihood cDCC estimator for large systems can easily be constructed (see Section 3.2). Note also that $E[\boldsymbol{Q}_t^{*1/2}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\boldsymbol{Q}_t^{*1/2}] = E[\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*\prime}] = E[E_{t-1}[\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*\prime}]] = E[\boldsymbol{Q}_t]$, which, jointly with Equation (17), yields $\boldsymbol{S} = E[\boldsymbol{Q}_t]$; therefore, $\boldsymbol{S}$ is the expectation of the correlation driving process.

2.4 *c*DCC Integrated Correlations

The case of $\alpha + \beta = 1$ (integrated correlations) is not covered by Proposition 2.3. In fact, in this case, $\boldsymbol{C} = (1 - \alpha - \beta)\boldsymbol{S} = 0$, which implies that H2 does not hold. The behavior of the cDCC conditional correlations for $\alpha + \beta = 1$ is illustrated in this section by means of simulations. It will be proven that, for $\alpha + \beta = 1$, the correlation processes degenerate either to 1 or to -1 in the long run. Thus, recalling that, for $\alpha + \beta < 1$, the correlation processes are mean reverting (see Proposition 2.3), it follows that testing for integrated correlations has a clear meaning with the cDCC model. The simulation experiment is planned as follows. For $\alpha + \beta = 1$, where $\alpha = 0.01, 0.04, 0.16, 0.64$, a set of 250 bivariate scalar cDCC series is generated, stopping each series at time $\bar{t} + M$ if either $\rho_{12,\bar{t}+m} > 1 - \epsilon$, or $\rho_{12,\bar{t}+m} < -1 + \epsilon$, for $m = 1, 2, \ldots, M$, where $M = 100{,}000$ and $\epsilon = 10^{-3}$. The occurrence of such a stopping rule is assumed as providing evidence that, in the long run, $\rho_{12,t}$ degenerates either to 1 or to -1. As conditional distributions for $[\varepsilon_{1,t}, \varepsilon_{2,t}]'$ some elliptical distributions are considered (McNeil, Frey, and Embrechts 2005), namely, the multivariate student t distribution with degrees of freedom (d.f.) $\nu = 3, 6, 12$, and the multivariate exponential power (MEP) distribution with kurtosis parameter $\kappa = 1, 1.25, 2, 8$. Such distributions are rescaled to get unit variances and covariance equal to $\rho_{12,t}$. The total number of experiments is 28 (four α-values times seven distributional

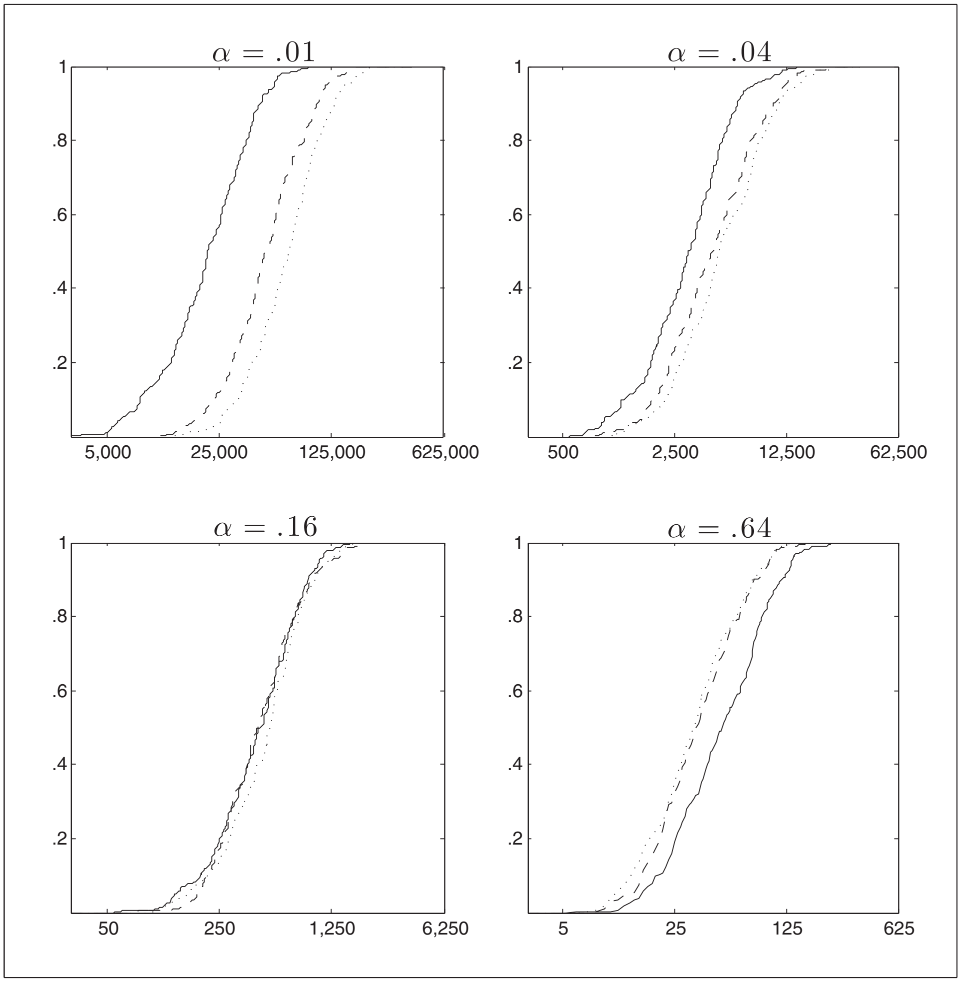

Figure 2. cDCC integrated conditional correlations: student t innovations. Empirical distribution function of the absorbing time, $\bar{t}$ (x-axis in logarithmic scale). The straight, dashed, and dotted lines correspond, respectively, to d.f. equal to 3, 6, and 12.

assumptions). The marginals of the MEP distribution are generalized error distributions (Nelson 1991). Their kurtosis is a decreasing function of κ. For $\kappa = 1$ they are double exponential and for $\kappa = 2$ they are Gaussian. By the properties of the elliptical distributions and the structure of the cDCC model, it can be easily shown that the dynamic properties of $\rho_{12,t}$ in a bivariate model are the same as the dynamic properties of $\rho_{ij,t}$, $i \neq j$, in a model with $N > 2$ assets, provided that the conditional distribution of the two models is elliptical of the same kind.

For each α-value, Figure 2 reports the empirical distribution of the 250 absorbing times, $\bar{t}$, under student t innovations. It is shown that all correlation series degenerate. The smaller the α, the longer the absorbing time, which is coherent with the fact that the variance of the correlation process is an increasing function of α (see Figure 1). For typical parameter values (e.g., $\alpha \leq 0.04$), the absorbing time is a decreasing function of the kurtosis. Analogous results arise under MEP innovations (the corresponding plots are not reported here for reasons of space).

The amount of series such that the absorbing state, $\rho_{12,\bar{t}}$, is positive (resp. negative) is an estimate of the probability that the assets become perfectly positively (resp. negatively) correlated in the long run. None of the 28 experiments provides an estimated probability that is significantly different from 1/2. This is due to having set $\rho_{12,0} = 0$ as a starting point of the correlation process. From another experiment run with Gaussian innovations, where $\alpha = 0.04$ and $\rho_{12,0} = -0.95$, -0.90, -0.60, -0.30, $+0.30$, $+0.60$, $+0.90$, $+0.95$, the estimated probabilities turned out to be equal, respectively, to 0.10, 0.15, 0.32, 0.40, 0.59, 0.70, 0.84, 0.91. As expected, the probability of perfectly positive correlation in the long run is an increasing function of the initial correlation.

3. LARGE SYSTEM ESTIMATION

3.1 The DCC Estimator

Let $L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi}) = \sum_{t=1}^{T} l_t(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ denote the (log) factorization of the DCC quasi-log-likelihood (QLL), where $\boldsymbol{\theta} \equiv [\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \ldots, \boldsymbol{\theta}_N']'$, $\boldsymbol{\phi} \equiv [\alpha, \beta]'$, $l_t(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi}) = -(1/2)\{N \log(2\pi) + \log|\tilde{\boldsymbol{H}}_t| + y_t' \tilde{\boldsymbol{H}}_t^{-1} y_t\}$, and $\tilde{\boldsymbol{H}}_t$ denotes the DCC conditional covariance matrix (see Equations (1)–(4)) evaluated at $(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ (hereafter, the symbol ~ will denote series evaluated at given parameter values). The matrix $\boldsymbol{S}$ includes $N(N-1)/2$ distinct parameters to estimate; therefore, even for known $\boldsymbol{\theta}$, the joint quasi-maximum-likelihood (QML) estimation of the DCC model is infeasible for large N. As a feasible estimator, Engle (2002) suggested a three-step procedure called DCC estimator. Before introducing it, let us define the following estimator of $\boldsymbol{S}$ conditional on $\boldsymbol{\theta}$.

Definition 3.1. (DCC conditional estimator of $\boldsymbol{S}$). For fixed $\boldsymbol{\theta}$, set $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}} \equiv T^{-1}\sum_{t=1}^{T} \tilde{\boldsymbol{\varepsilon}}_t \tilde{\boldsymbol{\varepsilon}}_t'$, where $\tilde{\boldsymbol{\varepsilon}}_t \equiv [\tilde{\varepsilon}_{1,t}, \ldots, \tilde{\varepsilon}_{N,t}]'$, $\tilde{\varepsilon}_{i,t} \equiv y_{i,t}/\sqrt{\tilde{h}_{i,t}}$, and $\tilde{h}_{i,t} \equiv h_{i,t}(\boldsymbol{\theta}_i; y_{i,t-1}, y_{i,t-2}, \ldots)$.

The estimator $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}}$ is the sample second moment of the standardized returns evaluated at $\boldsymbol{\theta}$. Alternatively, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}}$ can be defined as the sample covariance matrix of $\tilde{\boldsymbol{\varepsilon}}_t$, or as the sample (centered or uncentered) correlation matrix of $\tilde{\boldsymbol{\varepsilon}}_t$ (Engle and Sheppard 2001). In the latter case, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}}$ is unit-diagonal, like $\boldsymbol{S}$.

Definition 3.2. (DCC estimator)

Step 1: Set $\hat{\boldsymbol{\theta}} \equiv [\hat{\boldsymbol{\theta}}_1', \hat{\boldsymbol{\theta}}_2', \ldots, \hat{\boldsymbol{\theta}}_N']'$, where $\hat{\boldsymbol{\theta}}_i$ is the univariate QML estimator of $\boldsymbol{\theta}_i$, $i = 1, 2, \ldots, N$.
Step 2: Set $\hat{\boldsymbol{S}} \equiv \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}}}$.
Step 3: Set $\hat{\boldsymbol{\phi}} \equiv \text{argmax}_{\boldsymbol{\phi}} L_T(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}, \boldsymbol{\phi})$ subject to $\alpha \geq 0, \beta \geq 0$, and $\alpha + \beta < 1$.

Step 1 consists of a set of N univariate fittings; Step 2 is easy to compute after Step 1; finally, Step 3 is a maximization in two variables. Therefore, the DCC estimator is feasible for large N. A further reduction of computing time can be obtained by setting $L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi}) \equiv \sum_{i=2}^{N} L_{T,i,i-1}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1}, s_{i,i-1}, \boldsymbol{\phi})$, where $L_{T,i,j}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j, s_{i,j}, \boldsymbol{\phi})$ is the QLL of the bivariate DCC submodel of $(y_{i,t}, y_{j,t})$. The function $L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ in this case is called bivariate composite QLL. Engle, Shephard, and Sheppard (2008) proved that using a bivariate composite QLL in place of the full QLL can result in a dramatic reduction of the bias of $\hat{\boldsymbol{\phi}}$.

Let us denote the true parameter values with a superscript zero. Relying on the conjecture that $\boldsymbol{S}^0 = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ (see Section 2.1), $\hat{\boldsymbol{S}}$ is thought of as a consistent estimator (Engle and Sheppard 2001; Cappiello et al. 2006; Billio, Caporin, and Gobbo 2006; Pesaran and Pesaran 2007; Franses and Hafner 2009). As proven in Section 2.1, the equality $\boldsymbol{S}^0 = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ does not hold in general. Under consistency of $\hat{\boldsymbol{\theta}}$, if $\boldsymbol{S}^0 \neq E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ and plim $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}}$ is finite in a neighborhood of $\boldsymbol{\theta}^0$, it follows that

$$\text{plim}\, \hat{\boldsymbol{S}} = \text{plim}\, \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}}} \neq \boldsymbol{S}^0$$

(Wooldridge 1994, lemma A.1). If $\hat{\boldsymbol{S}}$ is inconsistent, $\hat{\boldsymbol{\phi}}$ is inconsistent in turn unless $\boldsymbol{S}$ and $\boldsymbol{\phi}$ are orthogonal (Newey and McFadden 1986, sec. 6.2, p. 2179). The inconsistency of $\hat{\boldsymbol{S}}$ could be a source of inconsistency also for the adjusted standard errors of $\hat{\boldsymbol{\phi}}$ computed relying on the formulas by Newey and McFadden (1986, theorem 6.1; see Engle 2002, p. 342, eq. 33) or by White (1996; see Engle and Sheppard 2001). In fact, the proof of the consistency of the adjusted standard errors requires a consistent estimator of $(\boldsymbol{\theta}, \boldsymbol{S})$.

3.2 The *c*DCC Estimator

As a large system estimator for the *c*DCC model, Aielli (2006) suggested a three-step estimator called *c*DCC estimator. In the following, $L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi}) = \sum_{t=1}^{T} l_t(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ will denote the (log) factorization of the QLL of the *c*DCC model. The parameters $(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ enter the QLL through the *c*DCC defining Equations (1)–(3) and (11). Before introducing the *c*DCC estimator, let us define the following estimator of $\boldsymbol{S}$ conditional on $(\boldsymbol{\theta}, \boldsymbol{\phi})$.

Definition 3.3. (*c*DCC conditional estimator of $\boldsymbol{S}$). For fixed $(\boldsymbol{\theta}, \boldsymbol{\phi})$, set $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}} \equiv T^{-1}\sum_{t=1}^{T} \tilde{\boldsymbol{Q}}_t^{*1/2} \tilde{\boldsymbol{\varepsilon}}_t \tilde{\boldsymbol{\varepsilon}}_t' \tilde{\boldsymbol{Q}}_t^{*1/2}$, where $\tilde{\boldsymbol{Q}}_t^* \equiv \text{diag}(\tilde{q}_{11,t}, \ldots, \tilde{q}_{NN,t})$, where

$$\tilde{q}_{ii,t} = (1 - \alpha - \beta) + \alpha\, \tilde{\varepsilon}_{i,t-1}^2 \tilde{q}_{ii,t-1} + \beta\, \tilde{q}_{ii,t-1}. \tag{18}$$

The estimator $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is the sample counterpart of $\boldsymbol{S}$ evaluated at $(\boldsymbol{\theta}, \boldsymbol{\phi})$ (see Equation (17)). By construction, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is pd. Alternatively, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ can be defined as the sample covariance matrix, or as the sample correlation (centered or uncentered) matrix, of $\tilde{\boldsymbol{Q}}_t^{*1/2} \tilde{\boldsymbol{\varepsilon}}_t$. In the latter case, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is unit-diagonal, like $\boldsymbol{S}$.

Definition 3.4. (*c*DCC estimator)

Step 1: Set $\hat{\boldsymbol{\theta}}$ as in Definition 3.2.
Step 2: Set $\hat{\boldsymbol{\phi}} \equiv \text{argmax}_{\boldsymbol{\phi}}\, L_T(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}, \boldsymbol{\phi})$ subject to $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta < 1$.
Step 3: Set $\hat{\boldsymbol{S}} \equiv \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}}}$.

Step 1 is the same as Step 1 of the DCC estimator; Step 2 is a maximization in two variables; finally, Step 3 does not actually need to be calculated, in that $\hat{\boldsymbol{S}}$ is nothing but the value of $\hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}$ at the end of Step 2. Note that with the DCC estimator the estimator of $\boldsymbol{S}$ is computed only once, before estimating $\boldsymbol{\phi}$, whereas, with the *c*DCC estimator, it is recomputed at each evaluation of the objective fuction of $\boldsymbol{\phi}$. This, however, does not require a sensible additional computational effort, as the calculations required by $\hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}$ basically reduce running the N univariate GARCH recursions providing $\tilde{q}_{ii,t}$ for fixed $\boldsymbol{\phi}$ and $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, $i = 1, 2, \ldots, N$ (see Definition 3.3). Under appropriate conditions (Newey and McFadden 1986, theorem 3.5), a Newton-Raphson one-step iteration from the *c*DCC estimation output will deliver an estimator with the same asymptotic efficiency as the joint QML estimator. As with the DCC estimator, a bivariate composite version of the *c*DCC estimator can be obtained replacing the full *c*DCC QLL with the bivariate composite *c*DCC QLL.

Recalling that $\boldsymbol{S}^0$ is the second moment of $\boldsymbol{Q}_t^{*1/2} \boldsymbol{\varepsilon}_t$, it follows that $\hat{\boldsymbol{S}}_{(\boldsymbol{\theta},\boldsymbol{\phi})}$ is consistent if evaluated at the true values of $(\boldsymbol{\theta}, \boldsymbol{\phi})$.

Proposition 3.1. For the *c*DCC model in Equation (11), suppose that assumptions H1–H3 of Proposition 2.3 hold. Then, plim $\hat{\boldsymbol{S}}_{\boldsymbol{\theta}^0,\boldsymbol{\phi}^0} = \boldsymbol{S}^0$.

Since $\boldsymbol{S}^0$ is unit-diagonal and $\boldsymbol{Q}_t^{*1/2} \boldsymbol{\varepsilon}_t$ is zero-mean, Proposition 3.1 holds whether $\hat{\boldsymbol{S}}_{(\boldsymbol{\theta},\boldsymbol{\phi})}$ is computed as a sample second moment, or as a sample covariance matrix, or as a sample (centered or uncentered) correlation matrix, of $\tilde{\boldsymbol{Q}}_t^{*1/2} \tilde{\boldsymbol{\varepsilon}}_t$. Relying on Proposition 3.1, a heuristic proof of the consistency of the *c*DCC estimator can be provided as follows (see Figure 3 for a graphical illustration of it). Denote as $\Phi \equiv \{(\alpha, \beta) : \alpha \geq 0, \beta \geq 0, \alpha + \beta < 1\}$ the constraint on $\boldsymbol{\phi}$. The *c*DCC estimator, $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}, \hat{\boldsymbol{\phi}})$, can be defined as the maximizer of the *c*DCC QLL subject to $\{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{S} = \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}, \boldsymbol{\phi} \in \Phi\}$. If $\hat{\boldsymbol{\theta}}$ is consistent and plim $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is finite for all $(\boldsymbol{\theta}, \boldsymbol{\phi})$,

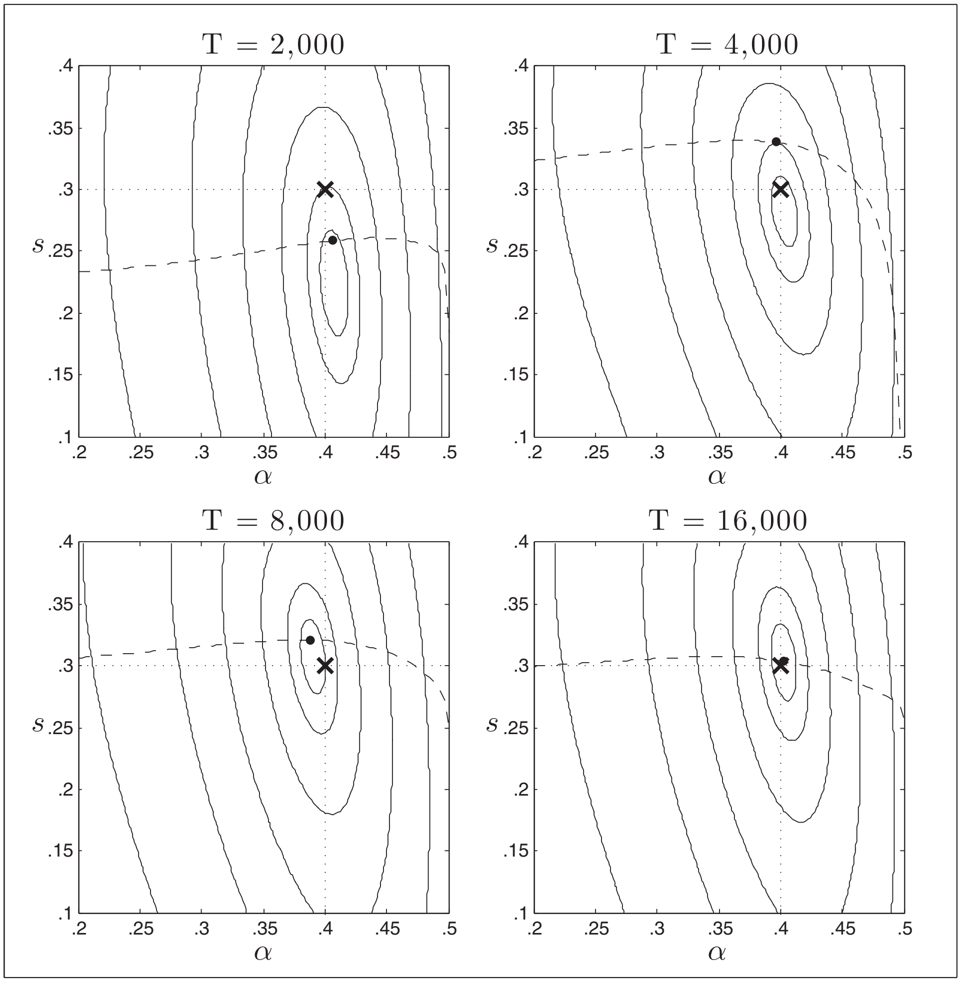

Figure 3. Performance of the cDCC estimator for increasing T. The figure provides an example of the behavior of the cDCC estimator for increasing sample sizes. It is assumed that $\boldsymbol{\theta}$ and β are known. The DGP is bivariate Gaussian with parameters set as $(s, \alpha) = (0.3, 0.4)$, where $s \equiv s_{12}$. The contour plots refer to the scaled cDCC QLL, which in this case can be written as $T^{-1}L_T(s, \alpha)$. The dashed line refers to the constraint under which the scaled cDCC QLL is maximized by the cDCC estimator (see Section 3.2). Within the considered model, the constraint is a curve of the plane, namely, $\{s = \hat{s}_\alpha, \alpha \in [0, 1)\}$. For varying α, the scaled objective function of α, that is, $T^{-1}L_T(\hat{s}_\alpha, \alpha)$, describes the value of the scaled cDCC QLL along the constraint (the curve of the plane). The cDCC estimator, $(\hat{\alpha}, \hat{s})$, is denoted with a bullet and the true value with a cross. The more the observations, the more the scaled cDCC QLL centers on the true value of the parameters, the more the cDCC constraint approaches a correctly specified constraint.

it follows that

$$\begin{aligned}\operatorname{plim}\{&\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{S} = \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}, \boldsymbol{\phi} \in \Phi\}\\ &= \{\boldsymbol{\theta} = \boldsymbol{\theta}^0, \boldsymbol{S} = \operatorname{plim}\hat{\boldsymbol{S}}_{\boldsymbol{\theta}^0,\boldsymbol{\phi}}, \boldsymbol{\phi} \in \Phi\}\end{aligned}$$

(Wooldridge 1994, lemma A.1). Recalling that $\operatorname{plim}\hat{\boldsymbol{S}}_{\boldsymbol{\theta}^0,\boldsymbol{\phi}^0} = \boldsymbol{S}^0$ (see Proposition 3.1), from the above equation it follows that the limit in probability of the cDCC constraint is a correctly specified constraint. Therefore, if $\operatorname{plim} T^{-1}L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ is finite for all $(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$, and uniquely maximized at $(\boldsymbol{\theta}^0, \boldsymbol{S}^0, \boldsymbol{\phi}^0)$, which is a common assumption in QML settings (see Bollerslev and Wooldridge 1992), $\operatorname{plim} T^{-1}L_T(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}, \boldsymbol{\phi})$ is uniquely maximized at $\boldsymbol{\phi}^0$ (Wooldridge 1994, lemma A.1). If, in addition, $\operatorname{plim} T^{-1}L_T(\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\phi})$ converges uniformly to its limit, $\hat{\boldsymbol{\phi}}$ is consistent (Newey and McFadden 1986, theorem 2.1). Regarding the consistency of $\hat{\boldsymbol{S}}$, if $\operatorname{plim}\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is finite for all $(\boldsymbol{\theta}, \boldsymbol{\phi})$, and $\operatorname{plim}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}) = (\boldsymbol{\theta}^0, \boldsymbol{\phi}^0)$, then $\operatorname{plim}\hat{\boldsymbol{S}} = \operatorname{plim}\hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}}} = \boldsymbol{S}^0$ (Wooldridge 1994, lemma A.1). Finally, regarding the consistency of $\hat{\boldsymbol{\theta}}$, appropriate conditions can be found for several univariate GARCH specifications adopted in practice (see Francq and Zakoïan 2010). Note that the above arguments on the consistency of the cDCC estimator of $(\boldsymbol{S}, \boldsymbol{\phi})$ do not apply to the DCC estimator of $(\boldsymbol{S}, \boldsymbol{\phi})$ because the DCC conditional estimator of $\boldsymbol{S}$ is not generally consistent if evaluated at the true value of $\boldsymbol{\theta}$.

The cDCC estimator of $\boldsymbol{\phi}$ can be thought of as a generalized profile QLL estimator (Severini 1998), in which $\boldsymbol{\phi}$ is the parameter of interest and $(\boldsymbol{\theta}, \boldsymbol{S})$ is the nuisance parameter. The estimator of the nuisance parameter conditional on $\boldsymbol{\phi}$ is $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}})$. The function $L_T(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{S}}_{\hat{\boldsymbol{\theta}},\boldsymbol{\phi}}, \boldsymbol{\phi})$ in Step 2 of Definition 3.4 is the generalized profile QLL. A generalized profile QLL estimator of the DCC model, totally analogous to the cDCC estimator, can be obtained defining $\hat{\boldsymbol{S}}_{(\boldsymbol{\theta},\boldsymbol{\phi})}$ as the sample counterpart of Equation (7) for fixed $(\boldsymbol{\theta}, \boldsymbol{\phi})$. Unfortunately, such an estimator is not guaranteed to be pd. Similar problems of positive definiteness would also arise with the VC model (see Equation 10).

A large system estimator for the diagonal vech cDCC model, Equation (16), can be obtained as a direct extension of the cDCC estimator. The parameter $\boldsymbol{\phi}$ is defined as the vector collecting the distinct elements of the dynamic parameter matrices, $\boldsymbol{A}_{\mathrm{q}} \equiv [\alpha_{ij,\mathrm{q}}]$ for $q = 1, 2, \ldots, \mathrm{Q}$, and $\boldsymbol{B}_{\mathrm{p}} \equiv [\beta_{ij,\mathrm{p}}]$ for $p = 1, 2, \ldots, \mathrm{P}$. The conditional estimator of $\boldsymbol{S}$ is computed replacing Equation (18) with

$$\tilde{q}_{ii,t} = \left(1 - \sum_{q=1}^{Q} \alpha_{ii,q} - \sum_{p=1}^{P} \beta_{ii,p}\right) s_{ii} + \sum_{q=1}^{Q} \alpha_{ii,q}\, \tilde{\varepsilon}_{i,t-q}^{2}\, \tilde{q}_{ii,t-q} + \sum_{p=1}^{P} \beta_{ii,p}\, \tilde{q}_{ii,t-p},$$

where $s_{ii} = 1$. The proof of the consistency of $\hat{\boldsymbol{S}}(\boldsymbol{\theta}^0, \boldsymbol{\phi}^0)$ is totally analogous to the proof of Proposition 3.1. For suitably restricted $\boldsymbol{\phi}$, the estimator is feasible (see, e.g., the block partition by Billio, Caporin, and Gobbo 2006). The generalized profile QLL must be maximized over the set of the vectors $\boldsymbol{\phi}$ such that the dynamic parameter matrices are psd, and the conditional estimator of the intercept, namely,

$$\left(\boldsymbol{\iota}\boldsymbol{\iota}' - \sum_{q=1}^{Q} \boldsymbol{A}_q - \sum_{p=1}^{P} \boldsymbol{B}_p\right) \odot \hat{\boldsymbol{S}}_{(\hat{\boldsymbol{\theta}},\boldsymbol{\phi})},$$

is pd. Ensuring that the latter constraint holds can be a difficult task.

3.3 Inferences From *c*DCC Estimations

The approach by Newey and McFadden (1986), to the estimation of the asymptotic covariance matrix of two-step estimators, can be adopted to compute inferences based on the (possibly diagonal vech) *c*DCC estimator. Let $\boldsymbol{s}$, $\hat{\boldsymbol{s}}(\boldsymbol{\theta}, \boldsymbol{\phi})$, and $\boldsymbol{\lambda}_t(\boldsymbol{\theta}, \boldsymbol{\phi})$ denote the vectors stacking the lower off-diagonal elements of $\boldsymbol{S}$, $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$, and $\tilde{\boldsymbol{Q}}_t^{*1/2} \tilde{\boldsymbol{\varepsilon}}_t\, \tilde{\boldsymbol{\varepsilon}}_t'\, \tilde{\boldsymbol{Q}}_t^{*1/2}$, respectively. Let $\boldsymbol{\gamma} \equiv [\boldsymbol{\theta}', \boldsymbol{\phi}', \boldsymbol{s}']'$, $l_t(\boldsymbol{\theta}, \boldsymbol{s}, \boldsymbol{\phi})$, and $l_{i,t}(\boldsymbol{\theta}_i) \equiv -(1/2)\{\log(2\pi) + \log \tilde{h}_{i,t} + y_{i,t}^2/\tilde{h}_{i,t}\}$, where $i = 1, 2, \ldots, N$, denote the *c*DCC parameter, the individual *c*DCC QLL at time t, and the ith individual GARCH QLL at time t. The *c*DCC estimator, denoted as $\hat{\boldsymbol{\gamma}} \equiv [\hat{\boldsymbol{\theta}}', \hat{\boldsymbol{\phi}}', \hat{\boldsymbol{s}}']'$, is a solution of the estimating equations $T^{-1}\sum_{t=1}^{T} \boldsymbol{g}_t(\boldsymbol{\gamma}) = 0$, where the individual score is defined as $\boldsymbol{g}_t(\gamma) \equiv [\{\boldsymbol{u}_t(\boldsymbol{\theta})\}', \{\boldsymbol{v}_t(\boldsymbol{\theta}, \boldsymbol{\phi})\}', \{\boldsymbol{p}_t(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{s})\}']'$, where $\boldsymbol{u}_t(\boldsymbol{\theta}) = [\{\boldsymbol{u}_{1,t}(\boldsymbol{\theta}_1)\}', \ \ldots, \{\boldsymbol{u}_{N,t}(\boldsymbol{\theta}_N)\}']'$, where $\boldsymbol{u}_{i,t} = (\partial/\partial\boldsymbol{\theta}_i) l_{i,t}(\boldsymbol{\theta}_i)$, $\boldsymbol{v}_t(\boldsymbol{\theta}, \boldsymbol{\phi}) \equiv (\partial/\partial\boldsymbol{\phi}) l_t(\boldsymbol{\theta}, \hat{\boldsymbol{s}}(\boldsymbol{\theta}, \boldsymbol{\phi}), \boldsymbol{\phi})$ and $\boldsymbol{p}_t(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{s}) \equiv \boldsymbol{\lambda}_t(\boldsymbol{\theta}, \boldsymbol{\phi}) - \boldsymbol{s}$. Under the conditions by Newey and McFadden (1986, theorem 3.2, with $\widehat{\boldsymbol{W}}$ replaced by the identity), it follows that $\sqrt{T}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}^0) \overset{A}{\approx} N(0, \{\boldsymbol{\mathcal{C}}^0\}^{-1}\boldsymbol{\mathcal{D}}^0\{\boldsymbol{\mathcal{C}}^{0\prime}\}^{-1})$, where

$$\boldsymbol{\mathcal{C}}^0 \equiv \operatorname{plim}\left\{ T^{-1} \sum_{t=1}^{T} \frac{\partial}{\partial \boldsymbol{\gamma}'} \boldsymbol{g}_t(\boldsymbol{\gamma}) \bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0} \right\} \quad \text{and}$$

$$\boldsymbol{\mathcal{D}}^0 \equiv \lim_{T\longrightarrow\infty} \mathrm{VAR}\left[T^{-1/2} \sum_{t=1}^{T} \boldsymbol{g}_t(\boldsymbol{\gamma}^0) \right].$$

The matrix $\boldsymbol{\mathcal{C}}^0$ is block lower triangular with block partition determined by the partition of $\boldsymbol{\gamma}$ into $\boldsymbol{\theta}$, $\boldsymbol{\phi}$, and $\boldsymbol{s}$. It can be proven that the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}_i$ coincides with its univariate QML asymptotic covariance matrix. The asymptotic covariance matrix of $\hat{\boldsymbol{\phi}}$ does not depend on derivatives with respect to the $\mathcal{O}(N^2)$ term $\boldsymbol{s}$. Estimates of the asymptotic covariance matrices of $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\phi}}$, and, if needed, $\hat{\boldsymbol{s}}$, can be computed by replacing the relevant blocks of $\boldsymbol{\mathcal{C}}^0$ and $\boldsymbol{\mathcal{D}}^0$ with the sample counterparts evaluated at $\hat{\boldsymbol{\gamma}}$. For $\boldsymbol{\mathcal{D}}^0$, a heteroscedastic and autocorrelation (HAC)-robust estimator is required (see, e.g., Newey and West 1987). If $\hat{\boldsymbol{S}}_{\boldsymbol{\theta},\boldsymbol{\phi}}$ is computed as a sample (centered or uncentered) correlation matrix, the individual score of $\boldsymbol{s}$, namely, $\boldsymbol{p}_t(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{s})$, must be designed accordingly.

4. EMPIRICAL APPLICATIONS

In this section, the empirical performances of the DCC and *c*DCC estimators are compared. The calculations are run using a MATLAB code based on a sequential quadratic programming optimizer. In the simulations under the hypothesis of correctly specified model, the true value of the parameters is used as a starting point for the maximization algorithms. In the simulations under misspecification and in the applications to the real data, the starting points are computed as the maximizers over a grid of the related objective functions. The constraints $\epsilon \le \alpha \le 1-\epsilon$, $\epsilon \le \beta \le 1-\epsilon$, and $\alpha+\beta \le 1-\epsilon$, where $\epsilon = 10^{-5}$, are imposed to achieve more stable maximizations. The conditional estimators of $\boldsymbol{S}$ are computed as centered correlations.

4.1 Simulations Under Correctly Specified Model

In this section, the simulation performances of the DCC and *c*DCC estimators under correctly specified model are compared. For $s \equiv s_{12} = 0, \pm0.3, \pm0.6, \pm0.9$, $\alpha = 0.01, 0.04, 0.16$, and $\alpha+\beta = 0.800, 0.990, 0.998$, a set of $M = 500$ independent DCC bivariate Gaussian series of length $T = 1750$ are generated. A burning period of 500 observations for each series is discarded to alleviate the effect of the initial values. Since the focus of the simulation study is on the correlation fitting performances, the DCC estimator is computed assuming that the standardized returns are known. An analogous experiment is run in which the *c*DCC estimator is computed from *c*DCC bivariate Gaussian generated series. The results of the simulations are reported in the panels of Figures 4–7. Regarding the DCC estimator, the estimator of s exhibits a positive bias for $s < 0$ and a negative bias for $s > 0$ (see Figure 4). Such a shrinkage effect increases with the persistence of the correlation process as measured by $\alpha+\beta$ and with the impact of the innovations as measured by α. For varying s, the bias moves according to a sinusoidal pattern. This is particularly evident for $\alpha+\beta = 0.998$ and $\alpha = 0.16$. For typical values of the dynamic parameters (i.e., for $\alpha+\beta \ge 0.990$ and $\alpha \le 0.04$), the bias is negligible. For $s = 0$, the estimator appears to be unbiased irrespective of the value of the dynamic parameters. The panel of Figure 5 reports the boxplots of the relative estimation error of $\hat{\alpha}$. For $\alpha+\beta = 0.998$ and $\alpha = 0.16$, the DCC boxplots are symmetric and well centered around zero. The same holds for the DCC boxplots of the relative estimation error of β (see Figure 6). This seems to suggest that the bias of the DCC estimator of s does not seriously affect the performances of the DCC estimator of (α, β). Figure 7 reports the boxplots of the mean absolute error of the conditional correlation estimator,

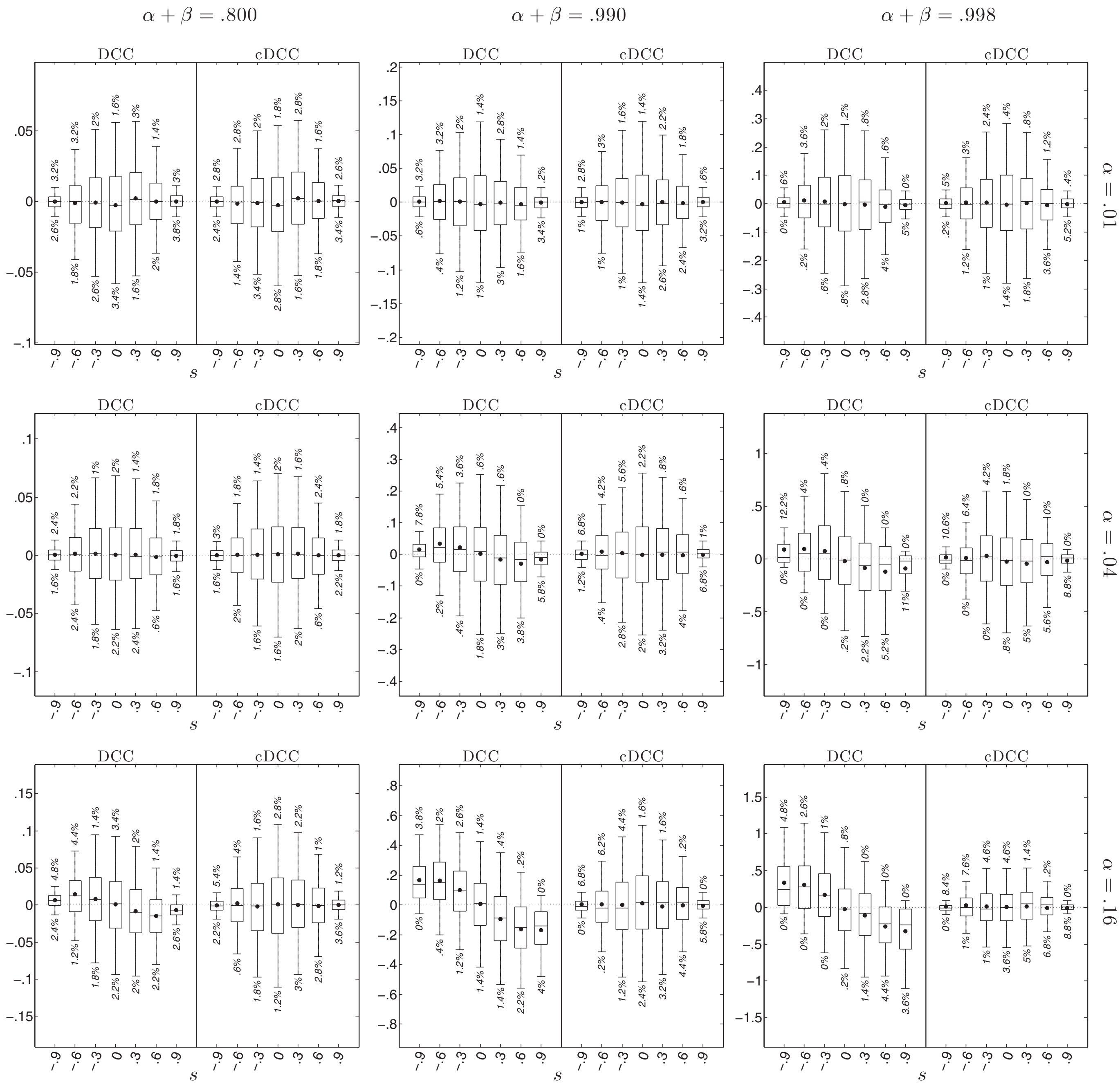

Figure 4. Estimation error of $\hat{s}$. Boxplots of $\hat{s} - s$. The DGP parameter values are reported at the top of the panel for $\alpha + \beta$, on the right hand side of the panel for α, and along the x-axis of each plot for s. For each boxplot, the line inside the boxplot and the bullet denote, respectively, the median and the average of the estimates. The maximum length of the whiskers is set to the interquartile range. The percentages at the end of the whiskers refer to the outliers.

computed as $M^{-1}T^{-1}\sum_{m=1}^{M}\sum_{t=1}^{T}|\hat{\rho}_t - \rho_t|$, where $\hat{\rho}_t \equiv \hat{\rho}_{12,t}$ and $\rho_t \equiv \rho_{12,t}$. Again, in spite of the large bias affecting the DCC estimator of s for $\alpha + \beta = 0.998$ and $\alpha = 0.16$, the performances of the DCC estimator turns out to be relatively good. This can be explained as follows. By backward substitutions, for large t the DCC equation of $q_{12,t}$ can be approximated as

$$q_{12,t} \approx \frac{1-\alpha-\beta}{1-\beta}s + \frac{\alpha}{1-\beta}\left\{(1-\beta)\sum_{n=1}^{t-1}\beta^{n-1}\varepsilon_{1,t-n}\,\varepsilon_{2,t-n}\right\}.$$

The right-hand side of the above equation is a weighted mean of s and of the term in brackets. If both $\alpha + \beta$ and α are large (such as for $\alpha + \beta = 0.998$ and $\alpha = 0.16$), the weight of s is small. In this case, the effect on the estimates of α, β, and ρ_t due to replacing s with a biased estimator will be small in turn.

Regarding the cDCC estimator, the estimator of s does not exhibit any apparent bias (see Figure 4). All the related boxplots are symmetric and well centered around zero. As for α, β, and ρ_t (see Figures 5–7), apart for $\alpha + \beta \geq 0.990$ and $\alpha = 0.16$, the performances of the cDCC estimator are practically identical to

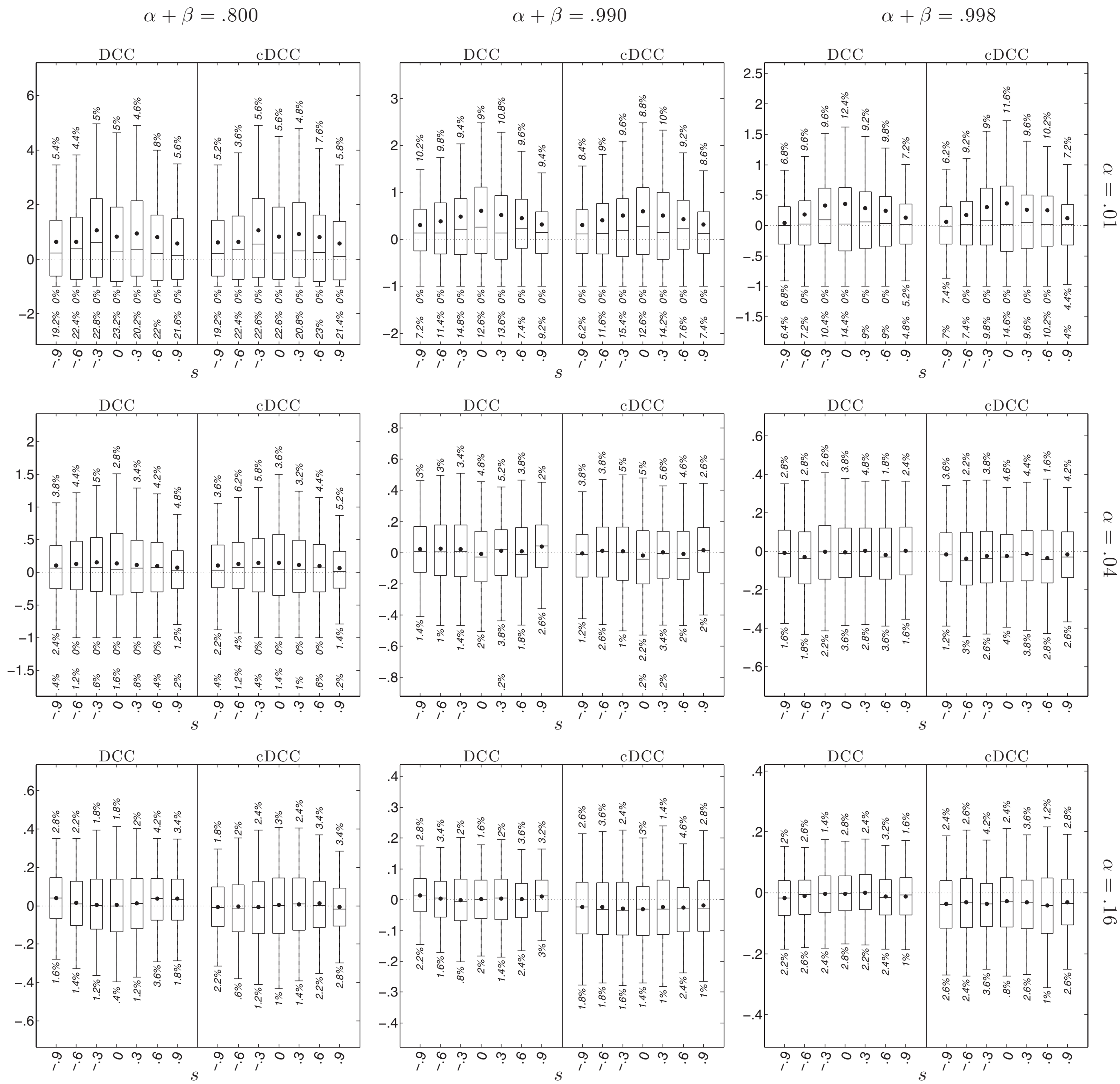

Figure 5. Estimation error of $\hat{\alpha}$. Boxplots of $(\hat{\alpha} - \alpha)/\alpha$ (for the layout of the panel and the construction of the box plots, see the caption of Figure 4). If there are estimates on the upper boundary, the related percentage is reported at the top of the plot. In the presence of estimates on the upper boundary the end of the upper whisker can coincide with the boundary, in which case the percentage of outliers is zero. Symmetric definitions hold for the percentages reported below the lower whisker.

those of the DCC estimator. For $\alpha + \beta \geq 0.990$ and $\alpha = 0.16$, instead, the cDCC estimator underperforms the DCC estimator, especially in terms of correlation mean absolute error (see Figure 7). For such parameter values, however, a comparison of the two estimators under a correctly specified model is not appropriate because the data-generating processes of the two models are rather different. An illustration of this is reported in Figure 8, in which the autocorrelation function (ACF) of ρ_t for the two models are depicted. For $\alpha + \beta \geq 0.99$ and $\alpha = 0.16$, the cDCC conditional correlation process exhibits less memory than the DCC conditional correlation process.

4.2 Simulations Under Misspecification

To compare the performances of the DCC and cDCC estimators under misspecification, a set of 500 bivariate Gaussian return series of length $T = 1000$ are generated in which the conditional variances follow GARCH processes and the correlation process is fixed arbitrary. For each series the DCC and cDCC estimators are computed, and a test of correctly specified model from the two estimation outputs is carried out. The percentage of rejections at a given level is then considered as a measure of the estimator performances under misspecification. The

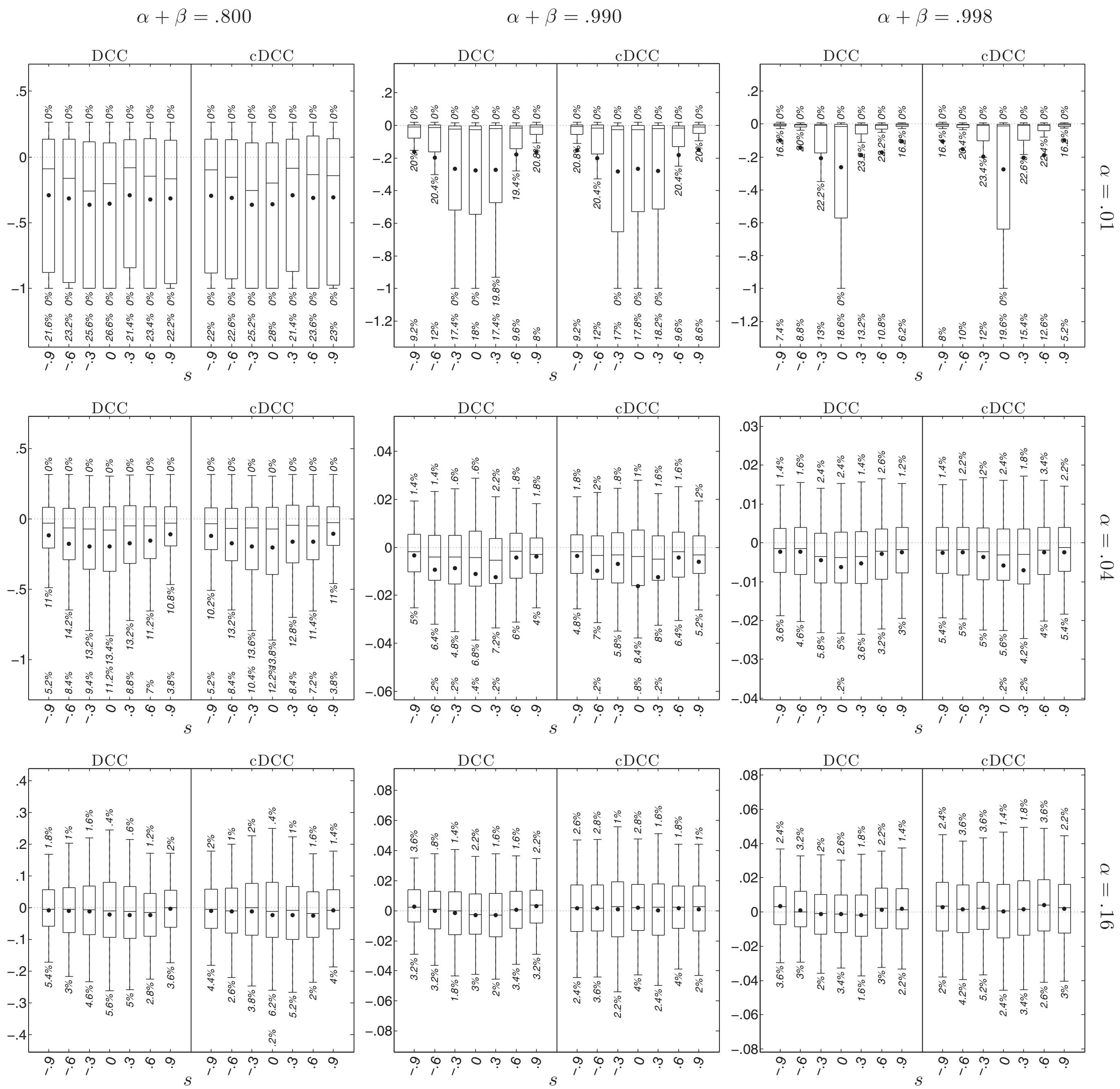

Figure 6. Estimation error of $\hat{\beta}$. Boxplots of $(\hat{\beta} - \beta)/\beta$ (for the layout of the panel, the construction of the boxplots, and the percentages appearing in each plot, see the caption of Figure 5).

best estimator is the one providing the smallest percentage of rejections.

The conditional variances are GARCH(1,1), set as $h_{1,t} = 0.01 + 0.05\, y_{1,t-1}^2 + 0.94\, h_{1,t-1}$ and $h_{2,t} = 0.30 + 0.20\, y_{2,t-1}^2 + 0.50\, h_{2,t-1}$. One variance process is highly persistent and the other is not. The orders of the GARCH variances (not the parameters) are assumed as known. Following Engle (2002), the conditional correlation processes are set as CONSTANT $\equiv \{\rho_t = 0.9\}$, STEP $\equiv \{\rho_t = 0.9 - 0.5(t \geq 500)\}$, SINE $\equiv \{\rho_t = 0.5 + 0.4.\cos(2\pi t/200)\}$, FASTSINE $\equiv \{\rho_t = 0.5 + 0.4.\cos(2\pi t/20)\}$, and RAMP $\equiv \{\rho_t = \mathrm{mod}(t/200)/200\}$. The choice of such processes ensures a variety of correlation dynamics, such as rapid changes, gradual changes, and periods of constancy (see Engle 2002). To test for a correctly specified model, the following three regression-based tests computed from the return series of a portfolio $\boldsymbol{w}'\boldsymbol{y}_t$, where $\boldsymbol{w}$ is the vector of the portfolio weights, are considered.

- *Engle-Colacito regression*. The Engle-Colacito (E&C) regression (Engle and Colacito 2006) is defined as

$$\{(\boldsymbol{w}'\boldsymbol{y}_t)^2/(\boldsymbol{w}'\boldsymbol{H}_t\boldsymbol{w})\} - 1 = \lambda + \xi_t,$$

where ξ_t is an innovation term. Recalling that the conditional variance of $\boldsymbol{w}'\boldsymbol{y}_t$ is $\boldsymbol{w}'\boldsymbol{H}_t\boldsymbol{w}$, if the model is correctly

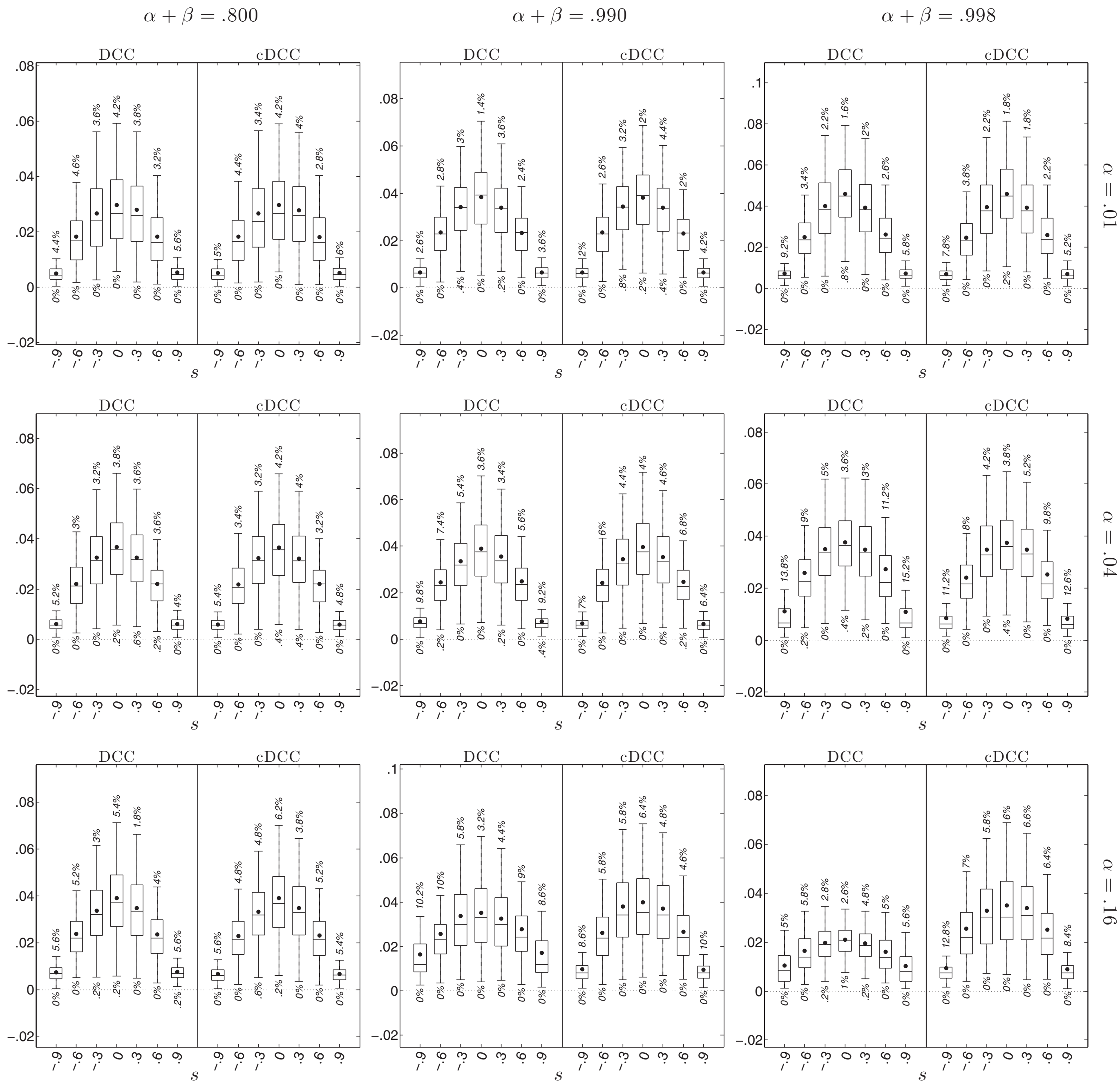

Figure 7. Mean absolute error of $\hat{\rho}_t$ (for the layout of the panel, the construction of the boxplots, and the percentages appearing in each plot, see the caption of Figure 5).

specified, it follows that $\lambda = 0$. The E&C test is a test of the null hypothesis that $\lambda = 0$. An HAC-robust estimator of the standard error of ξ_t is required.

- *Dynamic quantile test*. Denote the $\tau \times 100\%$-quantile of the conditional distribution of $\boldsymbol{w}'\boldsymbol{y}_t$ as $\mathrm{VaR}_t(\tau)$ (where VaR stands for *value at risk*). For fixed τ, set $\mathrm{HIT}_t \equiv 1$ if $\boldsymbol{w}'\boldsymbol{y}_t < \mathrm{VaR}_t(\tau)$, and $\mathrm{HIT}_t \equiv 0$ otherwise. If the model is correctly specified, $\{\mathrm{HIT}_t - \tau\}$ is zero-mean iid. The dynamic quantile (DQ) test by Engle and Manganelli (2004) is an F-test of the null hypothesis that all coefficients, as well as the intercept, are zero in a regression of $\{\mathrm{HIT}_t - \tau\}$ on past values, $\mathrm{VaR}_t(\tau)$, and any other variables. In this article, it is set $\mathrm{VaR}_t(\tau) = -1.96\sqrt{\boldsymbol{w}'\boldsymbol{H}_t\boldsymbol{w}}$, which corresponds to the 2.5% estimated quantile under Gaussianity. Five lags and the current estimated VaR are used as regressors.
- *LM test of ARCH effects*. If the model is correctly specified, the series $(\boldsymbol{w}'\boldsymbol{y}_t)^2/(\boldsymbol{w}'\boldsymbol{H}_t\boldsymbol{w})$ does not exhibit serial correlation. The LM test of ARCH effects (Engle 1982) is a test of the null hypothesis that $(\boldsymbol{w}'\boldsymbol{y}_t)^2/(\boldsymbol{w}'\boldsymbol{H}_t\boldsymbol{w})$ is serially uncorrelated. In this article, five lags are used.

Under the null hypothesis, the considered tests are asymptotically normal. In practice, because of the replacement of true quantities with estimated quantities, the asymptotic size of the tests can be different from the nominal size. Each

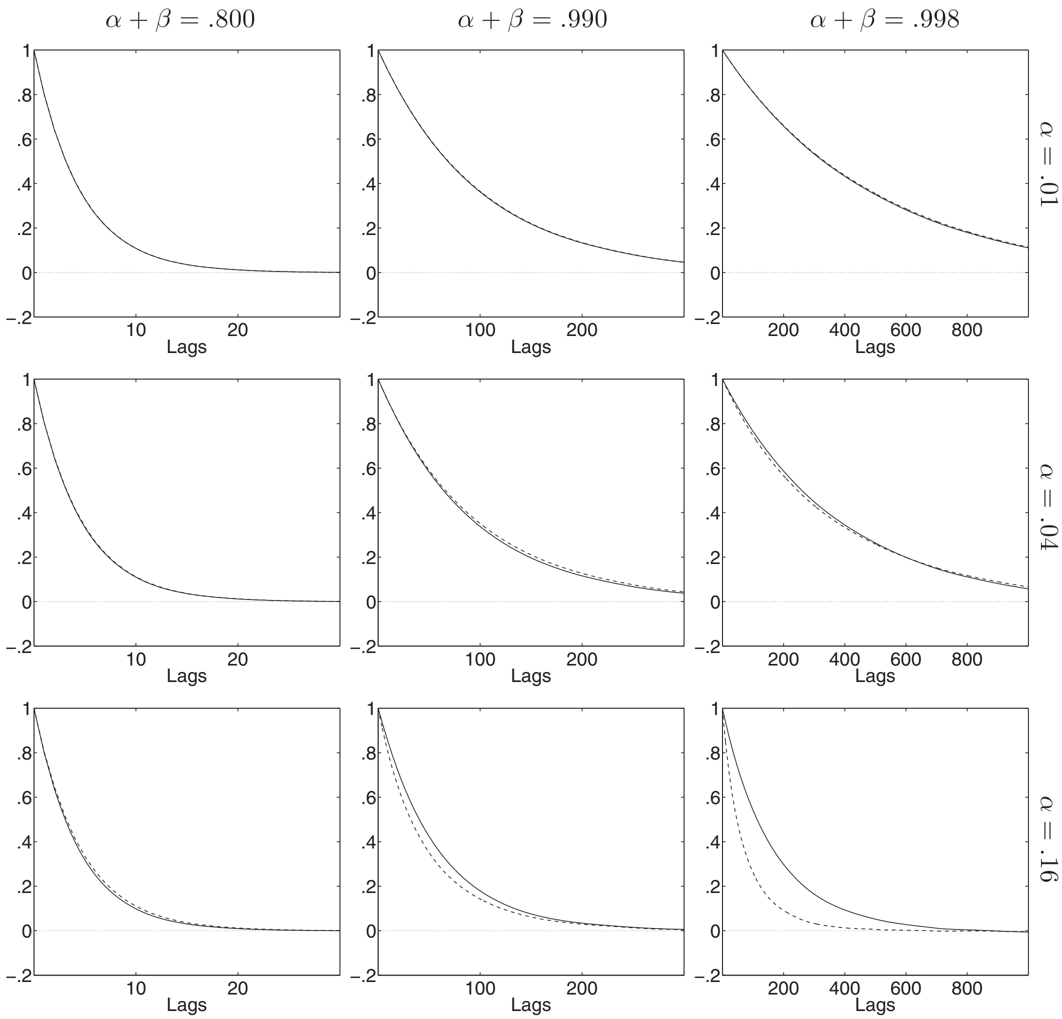

Figure 8. Autocorrelation function of $\rho_{ij,t}$. ACF computed as the mean of the sample ACF's of 100 series of length $T = 100{,}000$. The DGP parameter values are reported at the top of the panel for $\alpha + \beta$, and on the right-hand side of the panel for α. The location parameter is set as $s_{ij} = 0.3$. DCC in straight line and *c*DCC in dashed line.

test is computed from three portfolios, namely, the equally weighted portfolio (EW), the minimum variance portfolio with short selling (MV), and the minimum variance portfolio without short selling (*MV). The vector of EW weights is $\boldsymbol{\iota}/N$. The MV and *MV weights, which are functions of the true conditional covariance matrix, $\boldsymbol{H}_t$, are computed from the estimated conditional covariance matrix, $\hat{\boldsymbol{H}}_t$. For each estimator, $5 \times 3 \times 3 = 45$ percentages of rejections at a 5% level are computed, where 5 is the number of correlation specifications, 3 is the number of regression-based tests, and 3 is the number of portfolios. The null hypothesis of equal percentage of rejections for the two estimators is not rejected at any standard levels for any of the 45 cases. Thus, at least under the considered correlation misspecifications, the DCC and *c*DCC estimators provide very similar performances.

4.3 Applications to Real Data

To compare the performances of the DCC and *c*DCC estimators on real data, two datasets are considered. One dataset includes the S&P 500 composite index and the related nine S&P Depositary Receipts (SPDR) sector indices for a total of $N = 10$ assets. The sample period is May 2003–February 2010, which results in $T = 1750$ daily returns. Another dataset includes $N = 100$ randomly selected equities from the S&P 1500 index industrial and consumer goods components. The sample period is June 2003–March 2010, again for a total of $T = 1750$ daily returns. Both datasets are extracted from Datastream. The two estimators are computed in their bivariate composite version (see Sections 3.1 and 3.2) assuming GARCH(1,1) variances. As performance measures, three sets of out-of-sample forecast criteria are considered: (i) equal predictive ability (EPA) tests of one-step-ahead forecasts, (ii) regression-based tests computed from one-step-ahead forecasts, and (iii) EPA tests of multi-step-ahead correlation forecasts.

- *EPA tests of one-step-ahead forecasts*. Let d_t and d_t^c denote the loss due to replacing the true one-step-ahead asset conditional covariance matrix, $\boldsymbol{H}_t$, with the related DCC and *c*DCC estimates, denoted as $\hat{\boldsymbol{H}}_{t|t-1}$ and $\hat{\boldsymbol{H}}^c_{t|t-1}$, respectively, computed from a rolling window of $\bar{T} < T$ estimated excess returns, $\hat{\boldsymbol{y}}_{t-j}$, $j = 1, 2, \ldots, \bar{T}$.

Table 1. EPA tests of one-step-ahead forecasts

	Small dataset				Large dataset			
Loss type	CORR	EW	MV	*MV	CORR	EW	MV	*MV
MSE	−1.44	−0.17	−0.65	0.26	**−2.41**	**−2.50**	−1.82	−1.11
SCORE	−1.88	−1.31	0.28	0.12	**−2.68**	−1.43	−0.93	**−2.46**

NOTE: Negative (resp. positive) values are in favor of the cDCC (resp. DCC) estimator. Numbers in boldface denote significance at 5% level.

The number of predictions is $T - \bar{T}$. The null hypothesis $\mathcal{H}_0 : E[\bar{d}^c - \bar{d}\ ] = 0$, where $\bar{d}$ and $\bar{d}^c$ are the average losses, denotes EPA for the two estimators. Under appropriate conditions (Diebold and Mariano 1995), it holds that

$$\text{EPA} \equiv \frac{\sqrt{T-\bar{T}}(\bar{d}^c - \bar{d}\ )}{\sqrt{\widehat{VAR}[\sqrt{T-\bar{T}}(\bar{d}^c - \bar{d}\ )]}} \overset{A}{\sim} N(0,1), \quad (19)$$

where $\widehat{VAR}[\sqrt{T-\bar{T}}(\bar{d}^c - \bar{d}\ )]$ is an HAC-robust estimate of the variance of $\sqrt{T-\bar{T}}(\bar{d}^c - \bar{d}\ )$. Negative (resp. positive) values of EPA provide evidence in favor of cDCC (resp. DCC) forecasts. The length of the rolling window is set to $\bar{T} = 1250$, for a total of $T - \bar{T} = 500$ forecasts. The estimated excess returns are computed as demeaned observed returns using the mean of the data in the rolling window. As loss functions, two mean square error (MSE)-based losses (Diebold and Mariano 1995), and two score-based losses (Amisano and Giacomini 2007), are considered. Regarding the DCC estimator, the MSE-based losses are defined as

$$d_t \equiv \text{EWMSE}_t \equiv ((\boldsymbol{w}'\hat{\boldsymbol{y}}_{t|t-1})^2 - \boldsymbol{w}'\hat{\boldsymbol{H}}_{t|t-1}\boldsymbol{w}\ )^2, \quad (20)$$

where $\boldsymbol{w}$ is the vector of the EW weights and $\hat{y}_{i,t|t-1}$ is the one-step-ahead estimated excess return, and

$$d_t \equiv \text{CORRMSE}_t \equiv \frac{1}{N(N-1)/2} \times \sum_{i<j=2,\ldots,N} \left(\hat{\varepsilon}_{i,t|t-1}\,\hat{\varepsilon}_{j,t|t-1} - \hat{\rho}_{ij,t|t-1}\right)^2, \quad (21)$$

where $\hat{\rho}_{ij,t|t-1} \equiv \hat{h}_{ij,t|t-1}/\sqrt{\hat{h}_{ii,t|t-1}\hat{h}_{jj,t|t-1}}$ and $\hat{\varepsilon}_{i,t|t-1} \equiv \hat{y}_{i,t|t-1}/\sqrt{\hat{h}_{i,t|t-1}}$, where $\hat{h}_{ij,t|t-1}$ is the ijth element of $\hat{\boldsymbol{H}}_{t|t-1}$. The score-based losses are defined as

$$d_t \equiv \text{EWSCORE}_t \equiv \log(\boldsymbol{w}'\hat{\boldsymbol{H}}_{t|t-1}\boldsymbol{w}) + (\boldsymbol{w}'\hat{\boldsymbol{y}}_{t|t-1})^2/(\boldsymbol{w}'\hat{\boldsymbol{H}}_{t|t-1}w), \quad (22)$$

and

$$d_t \equiv \text{CORRSCORE}_t \equiv \log|\hat{\boldsymbol{R}}_{t|t-1}| + \hat{\boldsymbol{\varepsilon}}'_{t|t-1}\{\hat{\boldsymbol{R}}_{t|t-1}\}^{-1}\hat{\boldsymbol{\varepsilon}}_{t|t-1}, \quad (23)$$

where $\hat{\boldsymbol{\varepsilon}}_{t|t-1} \equiv [\hat{\varepsilon}_{1,t|t-1}, \ldots, \hat{\varepsilon}_{N,t|t-1}]'$ and $\hat{\boldsymbol{R}}_{t|t-1}$ is the correlation matrix associated to $\hat{\boldsymbol{H}}_{t|t-1}$. The losses (20) and (22) focus on the performances of the EW conditional variance forecast, whereas the losses (21) and (23) focus on the performances of the asset correlation forecast. The losses (20) and (22) are also computed for the MV and *MV portfolios, in which case the related portfolio weights are computed from $\hat{\boldsymbol{H}}_{t|t-1}$. The formulas of the corresponding losses for the cDCC estimator are obtained replacing $\hat{\boldsymbol{H}}_{t|t-1}$ with $\hat{\boldsymbol{H}}^c_{t|t-1}$ in Equations (20–23). It is worth noting that the only model-dependent estimation error entering the considered losses is that due to the correlation estimator. Therefore, the related EPA tests are essentially tests of EPA of correlation estimators. The resulting test statistics are reported in Table 1.

With the small dataset, all tests are insignificant at a 5% level. With the large dataset the message is in favor of the cDCC estimator, as shown by the fact that the sign of the EPA tests is always negative, and that four tests of eight are significant at a 5% level (one among them is significant at a 1% level).

- *Regression-based tests computed from one-step-ahead forecasts*. The regression-based tests of Section 4.2 can be used to assess the prediction performances of the DCC (resp. cDCC) estimator after replacing $\boldsymbol{y}_t$ and $\boldsymbol{H}_t$ with $\hat{\boldsymbol{y}}_{t|t-1}$ and $\hat{\boldsymbol{H}}_{t|t-1}$ (resp. $\hat{\boldsymbol{H}}^c_{t|t-1}$). If the estimated forecasts coincide with the true one-step-ahead quantities, the resulting test statistics are asymptotically normal. Table 2 reports the test statistics. The performances of the DCC and cDCC estimators are similar. With the small dataset the cDCC estimator turns out to be slightly better than the DCC estimator, in that six cDCC test statistics of nine are smaller than the corresponding DCC test statistics.

Table 2. Regression-based tests computed from one-step-ahead forecasts

		E&C Test			DQ Test			ARCH Test		
Dataset	Estimator	EW	MV	*MV	EW	MV	*MV	EW	MV	*MV
Small	cDCC	**3.77**	**19.55**	**6.19**	−0.67	−0.47	−0.50	0.59	**3.72**	**3.53**
	DCC	**3.90**	**25.39**	**6.31**	−0.67	−0.48	−0.50	0.68	**4.63**	**3.53**
Large	cDCC	**34.63**	**83.23**	**72.11**	−0.25	−0.87	−0.77	**4.28**	**2.84**	1.32
	DCC	**33.83**	**85.66**	**73.94**	−0.23	−0.91	−0.75	**3.92**	**2.24**	1.37

NOTE: Numbers in boldface denote significance at 5% level.

Table 3. EPA tests of multi-step-ahead correlation forecasts

Dataset	EPA test	Forecast horizon 2	3	5	8	13	21	34	55	89	144	233
Small	$\bar{\rho}^c$ vs. $\bar{\varrho}^c$	−1.31	−1.39	−1.45	−1.30	−0.88	−0.77	−0.69	**−2.47**	**−2.72**	**−2.61**	**−2.15**
	$\bar{\rho}$ vs. $\bar{\varrho}$	−1.41	−1.61	−1.71	−1.51	−0.51	0.48	**2.41**	0.46	−1.43	0.02	−1.92
	$\bar{\rho}^c$ vs. $\bar{\rho}$	−1.49	−1.51	−1.50	−1.39	−1.33	−1.36	−1.76	**−2.69**	**−2.64**	**−2.60**	**−2.15**
	$\bar{\varrho}^c$ vs. $\bar{\varrho}$	−1.51	−1.54	−1.56	−1.49	−1.52	−1.58	**−2.11**	**−2.72**	**−2.02**	**−2.11**	−1.56
Large	$\bar{\rho}^c$ vs. $\bar{\varrho}^c$	**−2.19**	**−2.84**	**−2.88**	**−2.84**	**−2.45**	**−2.18**	−1.51	−1.17	0.42	1.14	1.44
	$\bar{\rho}$ vs. $\bar{\varrho}$	−1.86	**−2.69**	**−2.93**	**−2.99**	**−2.93**	**−2.91**	**−2.39**	**−2.29**	−0.60	−0.77	1.58
	$\bar{\rho}^c$ vs. $\bar{\rho}$	**−2.31**	**−2.21**	**−2.16**	−1.68	−1.71	−0.90	0.19	−0.07	0.79	0.98	1.63
	$\bar{\varrho}^c$ vs. $\bar{\varrho}$	**−2.25**	**−2.07**	−1.93	−1.31	−1.43	−0.47	0.62	0.58	1.09	0.19	**2.10**

NOTE: Negative (resp. positive) values of the EPA test, "X vs. Y," are in favor of X (resp. Y). Numbers in boldface denote significance at 5% level. A superscript c denotes cDCC forecasts.

- *EPA tests of multi-step-ahead correlation forecasts*. The forecast of $\rho_{ij,t+m}$ at time t is defined as $E_t[\rho_{ij,t+m}]$, $m \geq 1$. Apart from $m = 1$, in which case $E_t[\rho_{ij,t+1}] = \rho_{ij,t+1}$, with both the DCC model and the cDCC model $E_t[\rho_{ij,t+m}]$ is infeasible. For the DCC model, Engle and Sheppard (2001) suggested to approximate $E_t[\rho_{ij,t+m}]$ either with

$$\bar{\rho}_{ij,t+m|t} \equiv \frac{q_{ij,t+m|t}}{\sqrt{q_{ii,t+m|t}\, q_{jj,t+m|t}}}, \tag{24}$$

where

$$q_{ij,t+m|t} \equiv s_{ij}\,(1-\alpha-\beta)\sum_{n=0}^{m-2}(\alpha+\beta)^n + q_{ij,t+1}(\alpha+\beta)^{m-1}, \tag{25}$$

or with

$$\bar{\varrho}_{ij,t+m|t} \equiv \bar{\varrho}_{ij}\,(1-\alpha-\beta)\sum_{n=0}^{m-2}(\alpha+\beta)^n + \rho_{ij,t+1}(\alpha+\beta)^{m-1}, \tag{26}$$

where $\bar{\varrho}_{ij} \equiv E[\varepsilon_{i,t}\,\varepsilon_{j,t}]$. In practice, the unknown quantities appearing in Equations (25) and (26) are replaced by their estimates based on the rolling window $\hat{\boldsymbol{y}}_{t-1}, \hat{\boldsymbol{y}}_{t-2}, \ldots, \hat{\boldsymbol{y}}_{t-\bar{T}}$. Specifically, $\bar{\varrho}_{ij}$ is replaced by the sample correlation of the estimated standardized returns. Analogous correlation forecast approximations can be defined for the cDCC model. A MSE-based loss for $\bar{\rho}_{ij,t+m|t}$ can be defined as

$$d_t \equiv \frac{1}{N(N-1)/2}\sum_{i<j=2,\ldots,N}\{\hat{\varepsilon}_{i,t+m|t+m-1}\,\hat{\varepsilon}_{i,t+m|t+m-1} - \bar{\rho}_{ij,t+m|t}\}^2, \tag{27}$$

where $t = \bar{T}+1, \bar{T}+2, \ldots, T-m$, and $m \geq 2$. The rolling window one-step-ahead estimated standardized returns appearing in the formula are preferred to the estimates of the standardized returns based on the whole sample to alleviate the bias due to possible structural breaks in the univariate variance processes. Replacing $\bar{\rho}_{ij,t+m|t}$ with $\bar{\varrho}_{ij,t+m|t}$ in Equation (27) will provide a MSE-based loss for $\bar{\varrho}_{ij,t+m|t}$. The considered loss functions can be used to compare via EPA tests the performances of the correlation forecasts $\bar{\rho}_{ij,t+m|t}$ and $\bar{\varrho}_{ij,t+m|t}$ for a given estimator, and the performances of the DCC and cDCC estimators for a given correlation forecast. The related EPA test statistics, computed for the forecast horizons $m = 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233$, are reported in Table 3. The selected forecast horizons are the Fibonacci numbers greater than 1 and less than 250, where 250 is about the number of daily returns per year. Of course, for varying the forecast horizons and the correlation forecasts the EPA tests are not independent.

  Let us consider the comparison of $\bar{\rho}$ and $\bar{\varrho}$ for a fixed estimator.

  Starting from the small dataset, with the cDCC estimator the preference is for $\bar{\rho}$ (see the first row of the table, in which all test statistics are negative). Specifically, for $m \geq 55$ the EPA test is significant at a 5% level. With the DCC estimator, the message is less conclusive (see the second row of the table). Moving to the large dataset, the preference is for $\bar{\rho}$ with both estimators (see the fifth and sixth rows of the table). Let us now consider the comparison of the DCC and cDCC estimators for a fixed correlation forecast. With the small dataset the test statistic is always negative, which is in favor of cDCC forecasts (see the third and fourth rows of the table). Specifically, for long forecast horizons ($m \geq 34$) most tests are significant at a 5% level. With the large dataset, apart for $m = 233$ when $\bar{\varrho}$ is used, in all cases the test is either not significant or significant and in favor of cDCC forecasts (see the seventh and eighth rows of the table). In summary, at least for the considered datasets, there is some evidence that $\bar{\rho}$ outperforms $\bar{\varrho}$, and that the cDCC correlation forecasts outperform the DCC correlation forecasts.

5. CONCLUSIONS

This article addressed some problems that arise with the DCC model. It has been proven that the second step of the DCC large system estimator can be inconsistent, and that the traditional GARCH-like interpretation of the DCC dynamic correlation parameters can lead to misleading conclusions. A more tractable alternative to the DCC model, called cDCC model, has been described. A large system estimator for the cDCC model has been discussed and heuristically proven to be consistent. Sufficient conditions for the stationarity of cDCC processes of interest have also been derived. The empirical performances of the DCC and cDCC large system estimators have been compared by means of applications to simulated and real data. It has been shown that, under correctly specified model, if the persistence of the correlation process and the impact of the innovations are high, the DCC estimator of the location correlation parameter can be seriously biased. The corresponding cDCC estimator resulted practically uniformly unbiased. Regarding the applications to the real data, the cDCC correlation forecasts have been proven to perform equally or significantly better than the DCC correlation forecasts.

APPENDIX: PROOFS

Proof of Proposition 2.1. It can be easily verified that $\boldsymbol{A}$ is psd, $\boldsymbol{B}$ is psd, and $(\boldsymbol{\iota\iota}' - \boldsymbol{A} - \boldsymbol{B}) \odot \boldsymbol{S}$ is pd, which ensures that $\boldsymbol{R}_t$ is pd. Recalling that $s_{11} = s_{22} = 1$, the elements of Q_t satisfy $q_{11,t} = (1 - \alpha_{11}) + \alpha_{11}\varepsilon^2_{1,t-1}$, $q_{22,t} = 1$, and $q_{12,t} = s_{12}$. By Jensen's inequality and $E[\varepsilon^2_{1,t}] = 1$, it follows that

$$
\begin{aligned}
E[\varepsilon_{1,t}\,\varepsilon_{2,t}] = E[\rho_{12,t}] &= E[q_{12,t}/\sqrt{q_{11,t}q_{22,t}}] \\
&= E\left[s_{12}/\sqrt{(1-\alpha_{11}) + \alpha_{11}\varepsilon^2_{1,t-1}}\right] \\
&= s_{12}\,E\left[1/\sqrt{(1-\alpha_{11}) + \alpha_{11}\varepsilon^2_{1,t-1}}\right] \\
&> s_{12}/\sqrt{(1-\alpha_{11}) + \alpha_{11}E[\varepsilon^2_{1,t-1}]} = s_{12}.
\end{aligned}
$$

□

Proof of Proposition 2.2. Taking the expectations of both members of Equation (4) and rearranging, yields $\boldsymbol{S} = \{(1-\beta)E[\boldsymbol{Q}_t] - \alpha E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t]\}/(1-\alpha-\beta)$. Noting that $E[\boldsymbol{Q}_t] = E[\boldsymbol{Q}^{*1/2}_t\boldsymbol{R}_t\,\boldsymbol{Q}^{*1/2}_t] = E[\boldsymbol{Q}^{*1/2}_t E_{t-1}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t]\boldsymbol{Q}^{*1/2}_t] = E[E_{t-1}[\boldsymbol{Q}^{*1/2}_t\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t\,\boldsymbol{Q}^{*1/2}_t]] = E[\boldsymbol{Q}^{*1/2}_t\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t\,\boldsymbol{Q}^{*1/2}_t]$, proves the proposition. □

Proof of Proposition 2.3. From $\boldsymbol{\varepsilon}^*_t = \boldsymbol{Q}^{*1/2}_t\boldsymbol{\varepsilon}_t$, where $\boldsymbol{\varepsilon}_t = \boldsymbol{R}^{1/2}_t\boldsymbol{\eta}_t$, it follows that $\boldsymbol{\varepsilon}^*_t = \boldsymbol{Q}^{1/2}_t\boldsymbol{\eta}_t$, where $\boldsymbol{Q}^{1/2}_t \equiv \boldsymbol{Q}^{*1/2}_t\boldsymbol{R}^{1/2}_t$ is the unique psd matrix such that $\boldsymbol{Q}^{1/2}_t\boldsymbol{Q}^{1/2}_t = \boldsymbol{Q}_t$. This fact, together with H1–H3, ensures that the BEKK process, $[\text{vech}(\boldsymbol{Q}_t)', \boldsymbol{\varepsilon}^{*\prime}_t]'$, admits a nonanticipative, strictly stationary, and ergodic solution (Boussama, Fuchs, and Stelzer 2011, theorem 2.3). Therefore, any time-invariant function of $[\text{vech}(\boldsymbol{Q}_t)', \boldsymbol{\varepsilon}^{*\prime}_t]'$, such as $[\text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$, admits a nonanticipative, strictly stationary, and ergodic solution (Billingsley 1995, theorem 36.4). Since the elements of $[\text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$ have finite variance, by Cauchy-Schwartz inequality the second moment of $[\text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$ exists finite, which completes the proof of point (i). To prove point (ii), it suffices to note that $[\text{vech}(\boldsymbol{H}_t)', \boldsymbol{y}'_t, \text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$ is a time-invariant function of $[h_{1,t}, \ldots, h_{N,t}, \text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$, which, under H1–H4, is a measurable function of the nonanticipative, strictly stationary, and ergodic process $[\text{vech}(\boldsymbol{R}_t)', \boldsymbol{\varepsilon}'_t]'$. Under H1–H5, point (iii) follows by strict stationarity of $\boldsymbol{y}_t$ and Cauchy-Schwartz inequality. □

Proof of Proposition 2.4. Under H1–H3 of Proposition 2.3, $\boldsymbol{\varepsilon}^*_t$ is weakly stationary (Boussama, Fuchs, and Stelzer 2011, theorem 2.3). If $\boldsymbol{\varepsilon}^*_t$ is weakly stationary, the second moment of $\boldsymbol{\varepsilon}^*_t$ is the matrix $\boldsymbol{S}$ in Equation (15) (Engle and Kroner 1995). Since $\boldsymbol{\varepsilon}^*_t = \boldsymbol{Q}^{*1/2}_t\boldsymbol{\varepsilon}_t$, the proposition is proven. □

Proof of Proposition 3.1. Under H1–H3 of Proposition 2.3, $\boldsymbol{\varepsilon}^*_t$ is weakly stationary and ergodic with stationary second moment $\boldsymbol{S}^0$ (see Proposition 2.4). Hence, under H1–H3 of Proposition 2.3, the matrix process $\boldsymbol{\varepsilon}^*_t\boldsymbol{\varepsilon}^{*\prime}_t$, which is a time-invariant function of $\boldsymbol{\varepsilon}^*_t$, is ergodic with finite first moment $\boldsymbol{S}^0$. This yields $\text{plim}\, T^{-1}\sum_{t=1}^T \boldsymbol{\varepsilon}^*_t\boldsymbol{\varepsilon}^{*\prime}_t = S^0$, which proves the proposition in that, for $(\boldsymbol{\theta}, \boldsymbol{\phi}) = (\boldsymbol{\theta}^0, \boldsymbol{\phi}^0)$, we have $\tilde{\boldsymbol{Q}}^{*1/2}_t\,\tilde{\boldsymbol{\varepsilon}}_t = \boldsymbol{\varepsilon}^*_t$. □

ACKNOWLEDGMENTS

I thank Farid Boussama, Christian T. Brownlees, Giorgio Calzolari, Massimiliano Caporin, Gabriele Fiorentini, Giampiero M. Gallo, Matteo Guastini, Davide Perilli, Pietro Rigo, Paolo Santucci De Magistris, Nicola Sartori, Enrique Sentana, and two anonymous referees for their careful and useful suggestions. I also thank my brother Gian Luca and Katerina Zamit for their precious support.

[Received November 2009. Revised December 2012.]

REFERENCES

Aielli, G. P. (2006), "Consistent Estimation of Large Scale Dynamic Conditional Correlations," unpublished paper, University of Florence. [282,285,288]

Amisano, G., and Giacomini, R. (2007), "Comparing Density Forecasts via Weighted Likelihood Ratio Tests," *Journal of Business & Economic Statistics*, 25, 177–190. [296]

Billingsley, P. (1995), *Probability and Measure*, New York: John Wiley. [298]

Billio, M., Caporin, M., and Gobbo, M. (2006), "Flexible Dynamic Conditional Correlation Multivariate GARCH Models for Asset Allocation," *Applied Financial Economics Letters*, 2, 123–130. [282,283,288,290]

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327. [285,286]

——— (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics*, 72, 498–505. [282]

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988), "A Capital Asset Pricing Model With Time-Varying Covariances," *Journal of Political Economy*, 96, 116–131. [285]

Bollerslev, T., and Wooldridge, J. M. (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models With Time-Varying Covariances," *Econometric Reviews*, 11, 143–172. [289]

Bougerol, P., and Picard, N. (1992), "Strict Stationarity of Generalized Autoregressive Processes," *Annals of Probability*, 20, 1714–1730. [286]

Boussama, F., Fuchs, F., and Stelzer, R. (2011), "Stationarity and Geometric Ergodicity of BEKK Multivariate GARCH Models," *Stochastic Processes and Their Applications*, 121, 2331–2360. [282,286,298]

Brandt, A. (1986), "The Stochastic Equation $Y_{n+1} = A_nY_n + B_n$ With Stationary Coefficients," *Advances in Applied Probability*, 18, 211–220. [286]

Cappiello, L., Engle, R. F., and Sheppard, K. (2006), "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4, 537–572. [282,283,288]

Diebold, F. X., and Mariano, R. S. (1995), "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13, 253–263. [296]

Ding, Z., and Engle, R. F. (2001), "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing," *Academia Economic Papers*, 29, 157–184. [283,285]

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007. [294]

——— (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models," *Journal of Business & Economic Statistics*, 20, 339–350. [282,283,284,287,293]

Engle, R. F., and Colacito, R. (2006), "Testing and Valuing Dynamic Correlations for Asset Allocation," *Journal of Business & Economic Statistics*, 24, 238–253. [293]

Engle, R. F., and Kroner, K. F. (1995), "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, 122–150. [283,285,286,298]

Engle, R. F., and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business & Economic Statistics*, 22, 367–381. [294]

Engle, R. F., and Sheppard, K. (2001), *Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH*, NBER Working Papers 8554, Cambridge, MA: National Bureau of Economic Research. [282,283,284,288,297]

Engle, R. F., Shephard, N., and Sheppard, K. (2008), "Fitting Vast Dimensional Time-Varying Covariance Models," Economics Series Working Papers 403, Oxford: Department of Economics, University of Oxford. [288]

Francq, C., and Zakoïan, J. M. (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications*, Chichester: Wiley. [282,285,289]

Franses, P. H., and Hafner, C. (2009), "A Generalized Dynamic Conditional Correlation Model for Many Asset Returns," *Econometric Reviews*, 28, 612–631. [282,283,288]

Kwan, W., Li, W. K., and Ng, K. W. (2010), "A Multivariate Threshold Varying Conditional Correlations Model," *Econometric Reviews*, 29, 20–38. [282]

McAleer, M., Chan, F., Hoti, S., and Lieberman, O. (2008), "Generalized Autoregressive Conditional Correlation," *Econometric Theory*, 24, 1554–1583. [282]

McNeil, A., Frey, R., and Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton Series in Finance, Princeton: Princeton University Press. [286]

Nelson, D. B. (1990), "Stationarity and Persistence in the GARCH(1, 1) Model," *Econometric Theory*, 6, 318–334. [285,286]

——— (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370. [287]

Newey, W. K., and McFadden, D. (1986), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics* (vol. 4), eds. R. F. Engle and D. McFadden, New York: Elsevier Science, chap. 36, pp. 2111–2245 [288,289,290]

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [290]

Pelletier, D. (2006), "Regime Switching for Dynamic Correlations," *Journal of Econometrics*, 131, 445–473. [282]

Pesaran, B., and Pesaran, M. (2007), "Modelling Volatilities and Conditional Correlations in Futures Markets With a Multivariate *t* Distribution," Cambridge Working Papers in Economics 0734, Cambridge: Faculty of Economics, University of Cambridge. [282,283,284,288]

Severini, T. A. (1998), "Likelihood Functions for Inference in the Presence of a Nuisance Parameter," *Biometrika*, 85, 507–522. [282,289]

Silvennoinen, A., and Teräsvirta, T. (2009), "Modeling Multivariate Autoregressive Conditional Heteroskedasticity With the Double Smooth Transition Conditional Correlation GARCH Model," *Journal of Financial Econometrics*, 7, 373–411. [282]

Tse, Y. K., and Tsui, A. K. C. (2002), "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model With Time-Varying Correlations," *Journal of Business & Economic Statistics*, 20, 351–362. [282,284]

White, H. (1996), *Estimation, Inference and Specification Analysis*, New York: Cambridge University Press. [288]

Wooldridge, J. M. (1994), "Estimation and Inference for Dependent Processes," in *Handbook of Econometrics* (vol. 4), eds. R. Engle and D. McFadden, New York: Elsevier Science, pp. 2639–2738. [288,289]