Appendix to Robust forecasting of dynamic conditional correlation GARCH models

Kris Boudt* Jón Daníelsson† Sébastien Laurent‡

April 18, 2012

Abstract

In Boudt, Daníelsson, and Laurent (2012) we propose a robust estimator for the cDCC model, originally developed by Engle (2002) and recently modified by Aielli (2009). For the first step estimation of this model, we propose to use a BIP M-estimator with Student t_4 loss function and robust targeting. The choice of that estimator is supported by the simulation study in Section 2 of this webappendix. We further elaborate on the practical relevance of the model, through a case study on Apple in Section 3. Additional figures and tables regarding the application to forecasting the conditional variance and covariance of exchange rates and stock returns are also given in Section 3. But first we describe the estimator for the univariate GARCH model in Section 1.

*KU Leuven/Lessius University College and V.U.University Amsterdam. E-mail: kris.boudt@econ.kuleuven.be

†London School of Economics, United Kingdom. E-mail: j.danielsson@lse.ac.uk

‡Maastricht University, School of Business and Economics, The Netherlands and CORE, Belgium. E-mail: s.laurent@maastrichtuniversity.nl

1 Robust M-estimators of univariate garch models

The central problem is to estimate the parameters of a GARCH(1,1) model in the presence of additive, but once–off, jumps a_t, i.e.

$$
\begin{aligned}
s_t^* &= \mu + s_t && (1.1)\\
s_t &= y_t + a_t && (1.2)\\
y_t &= \sqrt{h_t} z_t \text{ where } z_t \overset{i.i.d.}{\sim} N(0,1) && (1.3)\\
h_t &= \omega + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}. && (1.4)
\end{aligned}
$$

Muler and Yohai (2008) (MY) recommend the estimation of the BIP–GARCH using a M–estimator that minimizes the average value of an objective function $\rho(\cdot)$, called ρ-function, evaluated at the log–transformed squared devolatilized returns, that downweights the extreme observations and hence the jumps, i.e.

$$\hat{\theta}^{\mathrm{M}} = \mathrm{argmin}_{\theta \in \Theta} \frac{1}{T} \sum_{t=1}^{T} \rho \left(\log \frac{s_t^2}{h_t} \right). \tag{1.5}$$

The choice of $\rho(\cdot)$ trades off robustness vs. efficiency. The Gaussian QML estimator is equivalent to M-estimation with ρ-function equal to

$$\rho_0(z) = -z + \exp(z),$$

yielding GARCH estimates that can have a large bias in the presence of outliers.

In the case of Gaussian innovations and no outliers, $\log(s_t^2/h_t)$ has the following density function

$$g_0(z) = \frac{1}{\sqrt{2\pi}} \exp[-(\exp(z) - z)/2].$$

MY recommend $\rho_1(z) = 0.8m(g_0(z)/0.8)$, where the m–function is a smoothed version of $m_1(x) = xI(x \leq 4) + 4I(x > 4)$, i.e.,

$$m(x) = \begin{cases} x & \text{if } \leq 4 \\ P(x) & \text{if } 4 < x \leq 4.3 \\ 4.15 & \text{if } x > 4.3, \end{cases}$$

and

$$\begin{aligned} P(x) &= \frac{2}{(x_1 - x_0)^3}\left(\frac{1}{4}(x^4 - x_0^4) - \frac{1}{3}(2x_0 + x_1)(x^3 - x_0^3) + \frac{1}{2}(x_0^2 + 2x_0x_1)(x^2 - x_0^2)\right) \\ &- \frac{2x_0^2x_1}{(x_1 - x_0)}(x - x_0) - \frac{1}{3(x_1 - x_0)^2}(x - x_0)^3 + x, \end{aligned}$$

where $x_0 = 4$ and $x_1 = 4.3$.

The MY ρ–function ($\rho_1(\cdot)$) cannot readily be extended to the multivariate case, and we hence employ the ρ–function proposed by Boudt and Croux (2010) (denoted $\rho_2(\cdot)$). It is based on the Student t_4 density function and easily extends to the multivariate case:

$$\rho_2(z) = -z + \sigma_{1,4}\rho_{t_{1,4}}(\exp(z)),$$

where

$$\rho_{t_{N,\nu}}(u) = (N + \nu)\log(1 + \frac{u}{\nu - 2}) \tag{1.6}$$

and

$$\sigma_{N,\nu} = N/E[\rho'_{t_{N,\nu}}(u)u], \tag{1.7}$$

with u a chi–squared random variable with N degrees of freedom. Under general conditions, MY also establish the consistency and asymptotic normality of these estimators.

A second modification of the MY procedure is that we integrate reweighted estimates of the mean and variance in the forecasting procedure. The above definitions of the BIP–GARCH model and M–estimators are for $s_t = s_t^* - \mu$. MY assume that $\mu = 0$ and thus only focus on the conditional variance. Unfortunately, this assumption may not hold in practice and a jump robust estimator of μ is therefore needed. Furthermore, MY estimate the intercept ω jointly with the parameters α and β. As noted by Engle and Mezrich (1996), this is especially difficult if α and β sum to a number very close to one, as the intercept will be very small but must remain positive. Engle and Mezrich (1996) propose variance targeting as an estimation procedure where ω is reparameterized as $\hat{h}(1-\alpha_1-\beta_1)$ (with $\hat{h}$ a consistent estimator of h) before estimating the remaining parameters. Francq, Horvath, and Zakoian (2011) show that when the model is misspecified, the variance targeting estimator can be superior to the QMLE for long-term prediction or Value–at–Risk calculations.

In absence of outliers, natural choices for $\hat{\mu}$ and $\hat{h}$ are the sample mean and the sample variance of the returns. However, these estimators are known to be very sensitive to outliers (e.g. outliers causing a large upward bias in the sample variance). We therefore consider in the simulation study of the next section the use of the robust reweighed mean and variance estimators proposed by Boudt, Croux, and Laurent (2011) and described in Appendix A of Boudt, Daníelsson, and Laurent (2012).

2 Simulation

In this section we study the finite sample properties (bias and RMSE) of different estimators for the model in (1.1), more precisely:

- No variance targeting
  1. QML-estimator (M-estimator with loss function ρ_0);
  2. BIP M-estimator with loss function ρ_1;
  3. BIP M-estimator with loss function ρ_2;
- Variance targeting
  4. QML-estimator (M-estimator with loss function ρ_0), with targeting towards the sample mean and variance;
  5. BIP M-estimator with loss function ρ_1, with targeting towards the robust weighted mean and variance;

6. BIP M-estimator with loss function ρ_2, with targeting towards the robust weighted mean and variance.

Simulation setup: We generate bivariate returns S_t^* as the sum of a standard bivariate GARCH(1,1)–cDCC process and a jump process A_t. Let $t_1, \dots, t_l$ be the times when jumps are observed. The simulated returns are given by:

$$\begin{aligned}
S_t^* &= \begin{pmatrix} 0.05 \\ -0.05 \end{pmatrix} + \begin{cases} Y_t + A_t & \text{if } t = t_i, 1 \leq i \leq l = \varepsilon T \\ Y_t & \text{elsewhere} \end{cases}, \\
Y_t &= H_t^{1/2} Z_t, Z_t \sim N(0, I_2) \\
h_{1,t} &= 0.1 + 0.1 y_{1,t-1}^2 + 0.8 h_{1,t-1} \\
h_{2,t} &= 0.1 + 0.2 y_{2,t-1}^2 + 0.7 h_{2,t-1} \\
Q_t &= (1 - 0.1 - 0.8)\overline{Q} + 0.1 P_{t-1} \tilde{Y}_{t-1} \tilde{Y}_{t-1}' P_{t-1} + 0.8 Q_{t-1},
\end{aligned}$$

where $\overline{Q}_{1,2} = \overline{Q}_{2,1} = 0.4$ and $t = 1, \dots, T$, with $T = 2000$. The values $t_1, \dots, t_l$ were chosen equally spaced and $\varepsilon = 0\%$, 1% or 5%. The jump size is the conditional standard deviation of the corresponding elements of Y_t times d for the first series and negative d for the second series, with $d = 3$ or 4. The two assets have the same jump probability and 40% of jumps are cojumps. Consequently, $\varepsilon = 1\%$ (resp. 5%) corresponds to on average 0.7% (resp. 3.5%) of jumps on each series.

2.1 Choice of loss function

Table 1 presents the bias and RMSE for the parameters of the univariate GARCH model for $s_{1,t}^*$. Note that for this series, jumps are positive and ε corresponds to the total percentage of jumps on both series meaning that $\varepsilon = 1\%$ (resp. 5%) corresponds to 0.7% (resp. 3.5%) of jumps on each series. We consider six estimators: the Gaussian QML estimator, the M-estimator with ρ_1 and the more simple estimator ρ_2 based on the Student t_4, with and without variance targeting.

First consider the effect of variance targeting which in practically all cases reduces the RMSE of the robust estimators. A combination of a robust M-estimator with variance targeting is therefore recommended over the classical MY robust M-estimation.

Second, under variance targeting, the M-estimator with ρ_2 (i.e., t_4) has a lower RMSE than the one with ρ_1. Recall that both estimators rely on the BIP-GARCH model.

The estimation of the GARCH parameters using the misspecified BIP-GARCH model does not seem to create any significant bias in the estimated parameter values. Of course, we see that in the absence of additive jumps (i.e., $\varepsilon = 0\%$), we pay the price of a loss of efficiency with respect to the QML estimator. But when $\varepsilon = 1$ or 5%, the QML estimator is largely biased.

Table 1: Bias and RMSE of the GARCH(1,1) estimated by Gaussian QML estimator and BIP-GARCH(1,1) estimated by the M-estimator with ρ_1 (MY) and ρ_2 (Student t_4) with and without variance targeting when the DGP is a normal GARCH(1,1) model with a proportion ε of returns affected by jumps of size d times the conditional standard deviation, $\delta = 0.975$ and $T = 2000$.

			Bias			RMSE			Bias			RMSE		
		True	QMLE	ρ_1	ρ_2	QMLE	ρ_1	ρ_2	QMLE	ρ_1	ρ_2	QMLE	ρ_1	ρ_2
			No variance targeting						*Variance targeting*					
$\varepsilon = 0\%$	μ	0.050	0.000	0.001	0.001	0.023	0.026	0.026	0.001	0.001	0.001	0.023	0.026	0.026
	ω	0.100	0.008	0.009	0.009	0.036	0.045	0.041	0.008	0.009	0.008	0.036	0.045	0.040
	α_1	0.100	0.000	0.006	0.006	0.020	0.026	0.024	0.000	0.006	0.005	0.020	0.025	0.023
	β_1	0.800	-0.009	-0.015	-0.014	0.049	0.062	0.056	-0.009	-0.015	-0.013	0.049	0.062	0.056
	h	1.000	-0.000	0.003	0.012	0.066	0.077	0.075	-0.001	0.003	0.003	0.067	0.074	0.074
$\varepsilon = 1\%$,	μ	0.050	0.020	0.004	0.004	0.030	0.026	0.026	0.020	0.004	0.005	0.031	0.026	0.026
$d = 3$	ω	0.100	0.029	0.012	0.013	0.057	0.049	0.045	0.028	0.012	0.012	0.056	0.050	0.044
	α_1	0.100	-0.003	0.002	0.003	0.024	0.025	0.023	-0.004	0.004	0.001	0.023	0.025	0.022
	β_1	0.800	-0.019	-0.015	-0.013	0.064	0.066	0.059	-0.018	-0.015	-0.013	0.064	0.066	0.058
	h	1.000	0.065	-0.008	0.029	0.097	0.075	0.079	0.058	0.013	0.013	0.091	0.076	0.076
$\varepsilon = 1\%$,	μ	0.050	0.027	0.003	0.003	0.035	0.026	0.026	0.027	0.003	0.003	0.035	0.026	0.026
$d = 4$	ω	0.100	0.062	0.010	0.014	0.099	0.049	0.046	0.060	0.011	0.013	0.098	0.049	0.044
	α_1	0.100	-0.002	0.001	0.003	0.030	0.025	0.024	-0.005	0.004	0.000	0.029	0.025	0.022
	β_1	0.800	-0.043	-0.014	-0.014	0.098	0.066	0.060	-0.041	-0.015	-0.013	0.098	0.066	0.058
	h	1.000	0.122	-0.025	0.034	0.145	0.076	0.081	0.104	0.011	0.011	0.128	0.076	0.076
$\varepsilon = 5\%$,	μ	0.050	0.103	0.022	0.022	0.106	0.034	0.034	0.103	0.022	0.022	0.106	0.034	0.034
$d = 3$	ω	0.100	0.169	0.021	0.042	0.251	0.073	0.087	0.169	0.034	0.031	0.252	0.086	0.073
	α_1	0.100	-0.028	-0.014	-0.014	0.046	0.030	0.029	-0.030	-0.007	-0.018	0.046	0.029	0.030
	β_1	0.800	-0.078	-0.010	-0.014	0.186	0.085	0.085	-0.079	-0.021	-0.006	0.186	0.093	0.079
	h	1.000	0.307	-0.032	0.110	0.319	0.081	0.131	0.300	0.060	0.060	0.311	0.100	0.100
$\varepsilon = 5\%$,	μ	0.050	0.137	0.012	0.013	0.140	0.029	0.029	0.137	0.013	0.013	0.139	0.029	0.029
$d = 4$	ω	0.100	0.306	0.008	0.056	0.468	0.056	0.109	0.397	0.031	0.029	0.539	0.074	0.077
	α_1	0.100	-0.061	-0.014	-0.017	0.079	0.027	0.032	-0.059	-0.001	-0.024	0.075	0.027	0.034
	β_1	0.800	-0.100	-0.005	-0.021	0.281	0.072	0.101	-0.164	-0.025	0.000	0.313	0.082	0.086
	h	1.000	0.547	-0.107	0.135	0.556	0.126	0.154	0.536	0.046	0.046	0.545	0.091	0.091

2.2 Sensitivity to sample size

The bias and RMSE of the BIP–cDCC and the benchmark QML estimator of the parameters of the unobserved GARCH(1,1)–cDCC process Y_t are shown in Table 2. Consider first the bias and RMSE for the parameters of the univariate GARCH model for $s^*_{1,t}$. In line with the results in MY and Carnero, Peña, and Ruiz (2012), we find that the estimation of the GARCH parameters using the misspecified BIP–GARCH model does not seem to create any significant bias in the estimated parameter values. Of course, we see that in the absence of additive jumps (i.e., $\varepsilon = 0\%$), we pay the price of a loss of efficiency with respect to the QML estimator. But when $\varepsilon = 1$ or 5%, the QML estimator is severely biased.

The last three columns of Table 2 present the results for the multivariate case. Like in the univariate case, the estimation of the cDCC parameters using the BIP–cDCC models does not seem to create any significant bias in the estimated parameter values in the absence of jumps (i.e., $\varepsilon = 0\%$). The average of the estimated parameters is very close to the true values. Since the innovations have a conditionally Gaussian distribution, the Gaussian QML estimator based on the correctly specified GARCH model is expected to have (at least asymptotically) the lowest RMSE.

The loss of efficiency of the robust estimator in the absence of additive jumps is moderate compared to the lower bias and gain in efficiency in the presence of these jumps. For $\varepsilon = 1$ or 5% of additive jumps, we find the empirical correlation of the devolatilized returns to be a strongly biased estimate of $\overline{Q}_{1,2}$. Because jumps have the opposite sign and the true correlation is 0.4, we find a negative bias of -5.5% when $\varepsilon = 1\%$ and $d = 3$ and -33.9% when $\varepsilon = 5\%$ and $d = 4$. The persistency parameter β also is largely underestimated. Its bias is -5.7% when $\varepsilon = 1\%$ and $d = 3$ and -21.9% when $\varepsilon = 5\%$ and $d = 4$. When $\varepsilon = 1\%$, the bias in the QML estimate of α is still negligible, but for $\varepsilon = 5\%$ with $d = 4$, we find a bias of -8.4%.

Importantly, in all cases, the bias and RMSE of the estimates of the proposed robust estimator remains small in the presence of additive jumps.

Table 2: Bias and RMSE of the Gaussian QML and robust estimator for the 2-dimensional cDCC model in presence of ε jumps of size d conditional standard deviation, with $\delta = 0.975$ and $T = 2000$.

			μ	ω	α_1	β_1	h	$\overline{Q}_{12}$	α	β
			0.050	0.100	0.100	0.800	1.000	0.400	0.100	0.800
$\varepsilon = 0\%$	QML	bias	0.001	0.008	0.000	-0.009	-0.001	-0.001	-0.001	-0.007
		RMSE	0.023	0.036	0.020	0.049	0.067	0.039	0.019	0.045
	Robust	bias	0.001	0.008	0.005	-0.013	0.003	-0.011	-0.002	-0.006
		RMSE	0.026	0.040	0.023	0.056	0.074	0.040	0.021	0.051
$\varepsilon = 1\%$	QML	bias	0.020	0.028	-0.004	-0.018	0.058	-0.055	-0.010	-0.057
$d = 3$		RMSE	0.031	0.056	0.023	0.064	0.091	0.067	0.033	0.124
	Robust	bias	0.005	0.012	0.001	-0.013	0.013	-0.016	-0.007	-0.006
		RMSE	0.026	0.044	0.022	0.058	0.076	0.041	0.022	0.055
$\varepsilon = 1\%$	QML	bias	0.027	0.060	-0.005	-0.041	0.104	-0.092	-0.008	-0.125
$d = 4$		RMSE	0.035	0.098	0.029	0.098	0.128	0.101	0.043	0.203
	Robust	bias	0.003	0.013	0.000	-0.013	0.011	-0.016	-0.007	-0.006
		RMSE	0.026	0.044	0.022	0.058	0.076	0.041	0.022	0.055
$\varepsilon = 5\%$	QML	bias	0.103	0.169	-0.030	-0.079	0.300	-0.230	-0.070	-0.166
$d = 3$		RMSE	0.106	0.252	0.046	0.186	0.311	0.233	0.080	0.277
	Robust	bias	0.022	0.031	-0.018	-0.006	0.060	-0.037	-0.034	0.015
		RMSE	0.034	0.073	0.030	0.079	0.100	0.053	0.042	0.094
$\varepsilon = 5\%$	QML	bias	0.137	0.397	-0.059	-0.164	-0.536	-0.339	-0.084	-0.219
$d = 4$		RMSE	0.106	0.252	0.046	0.186	0.311	0.341	0.090	0.284
	Robust	bias	0.013	0.029	-0.024	0.000	0.046	-0.033	-0.039	0.026
		RMSE	0.029	0.077	0.034	0.086	0.091	0.051	0.046	0.100

The bias and RMSE of the parameters underlying $h_{1,t}$ and $h_{2,t}$ are similar. To save space, we only report those for $h_{1,t}$.

In Table 3 of this webappendix, we repeat the analysis for $T = 1000$ and obtain similar conclusions.

2.3 Sensitivity to parameter choice

In the main paper we report the bias and RMSE of the Gaussian QML and robust BIP M-estimator for unconditional correlation $\overline{Q}$ and the dependence parameters α and β of the 2-dimensional cDCC model in presence of ε jumps of size d conditional standard deviation, with $\delta = 0.975$ and $T = 2000$ and $\alpha + \beta = 0.95$. In Figures 2-2 we see that similar results are obtained for

Table 3: Bias and RMSE of the QML and BIP M–estimates of the parameters of the cDCC model, $\delta = 0.975$ and $T = 1000$ or $T = 2000$.

		$T = 1000$				$T = 2000$			
		bias		RMSE		bias		RMSE	
		QMLE	BIP–M	QMLE	BIP–M	QMLE	BIP–M	QMLE	BIP–M
$\varepsilon = 0\%$	$\overline{Q}_{1,2}$	-0.002	-0.011	0.055	0.054	-0.001	-0.011	0.039	0.040
	α	-0.001	-0.002	0.027	0.029	-0.001	-0.002	0.019	0.021
	β	-0.016	-0.017	0.073	0.083	-0.007	-0.006	0.045	0.051
$\varepsilon = 1\%$	$\overline{Q}_{1,2}$	-0.052	-0.016	0.075	0.056	-0.055	-0.016	0.067	0.041
$d = 3$	α	-0.007	-0.006	0.043	0.030	-0.010	-0.007	0.033	0.022
	β	-0.068	-0.017	0.161	0.088	-0.057	-0.006	0.124	0.055
$\varepsilon = 1\%$	$\overline{Q}_{1,2}$	-0.087	-0.016	0.104	0.056	-0.092	-0.016	0.101	0.041
$d = 4$	α	-0.006	-0.006	0.055	0.031	-0.008	-0.007	0.043	0.022
	β	-0.119	-0.018	0.214	0.089	-0.125	-0.006	0.203	0.055
$\varepsilon = 5\%$	$\overline{Q}_{1,2}$	-0.227	-0.038	0.233	0.066	-0.230	-0.037	0.233	0.053
$d = 3$	α	-0.067	-0.032	0.082	0.047	-0.070	-0.034	0.080	0.042
	β	-0.142	-0.005	0.248	0.134	-0.166	0.015	0.277	0.094
$\varepsilon = 5\%$	$\overline{Q}_{1,2}$	-0.334	-0.334	0.338	0.338	-0.339	-0.033	0.341	0.051
$d = 4$	α	-0.079	-0.079	0.089	0.089	-0.084	-0.039	0.090	0.046
	β	-0.203	-0.203	0.277	0.277	-0.219	0.026	0.284	0.100

$\alpha + \beta = 0.98$.

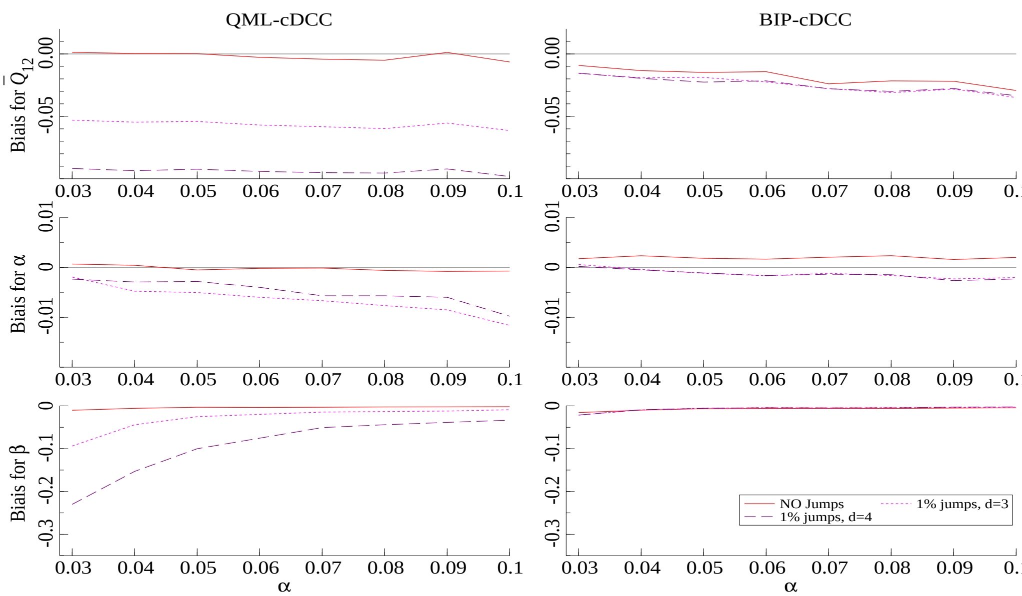

Figure 1: Bias of the Gaussian QML and robust BIP M-estimator for unconditional correlation $\overline{Q}$ and the dependence parameters α and β of the 2-dimensional cDCC model in presence of ε jumps of size d conditional standard deviation, with $\delta = 0.975$ and $T = 2000$. $\alpha + \beta = 0.98$.

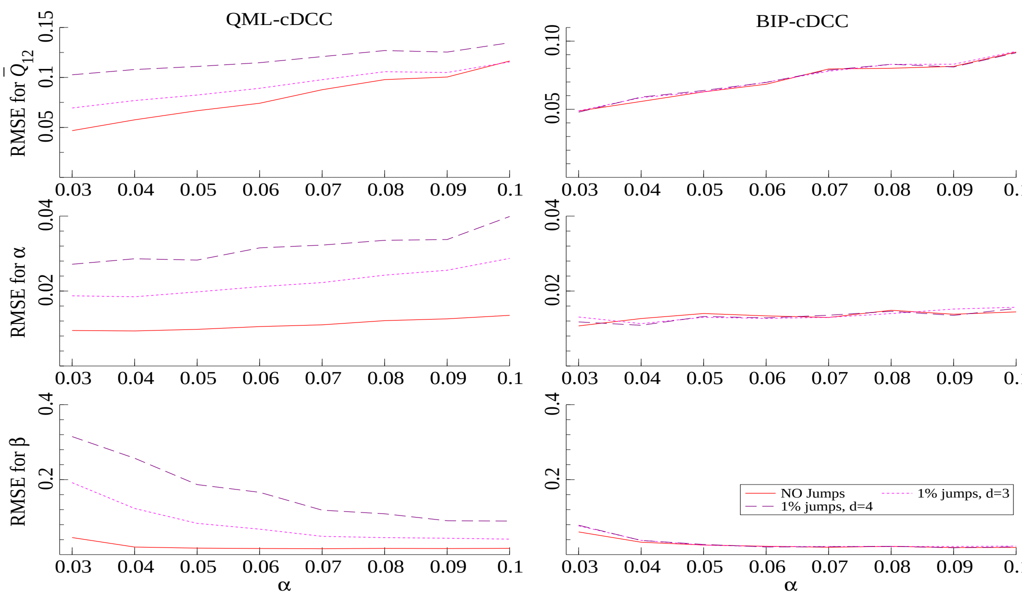

Figure 2: RMSE of the Gaussian QML and robust BIP M-estimator for unconditional correlation $\overline{Q}$ and the dependence parameters α and β of the 2-dimensional cDCC model in presence of ε jumps of size d conditional standard deviation, with $\delta = 0.975$ and $T = 2000$. $\alpha + \beta = 0.98$.

3 Applications

To save space in the main manuscript, the case study on Apple is presented in this webappendix (Subsection 3.1). Additional figures and tables regarding the application to forecasting the conditional variance and covariance of exchange rates and stock returns are given in Subsections 3.2-3.3.

3.1 Case study on Apple

Many volatility models, such as GARCH, are based on the assumption that each return observation has the same relative impact on future volatility, regardless of the magnitude of the return. This assumption is at odds with an increasing body of evidence indicating that the largest return observations have a relatively smaller effect on future volatility than smaller shocks (see for instance Andersen, Bollerslev, and Diebold, 2007).

One reason is extremely large shocks caused by once–off events that cannot be expected to influence future volatility much. One example is the stock price of Apple, which fell 52% on September 29, 2000 after it warned its fourth–quarter profit would fall well short of Wall Street forecasts.

We employed Gaussian quasi–maximum likelihood (QML) to estimate a GARCH$(1,1)$ model on the daily returns on Apple with a sample of one thousand days starting on the first day of 2000 and ending in December 2003. The results as well as the GARCH specification are reported in Table 4.

Table 4: Impact on GARCH of an extreme Apple return

Sample and method	α_1	β_1	$\alpha_1+\beta_1$	$\frac{\omega\times 10^4}{1-\alpha_1-\beta_1}$	$\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t$	VoV
GARCH full sample	0.157	0.824	0.981	29.421	15.140	1.543
GARCH after outlier	0.019	0.976	0.995	8.000	12.038	0.827
BIP–GARCH full sample	0.029	0.969	0.999	11.630	10.997	0.681

Note: The GARCH$(1,1)$ specification for the daily return series of Apple (y_t) is $y_t = \sqrt{h_t}z_t$ where $z_t \overset{i.i.d.}{\sim} N(0,1)$ and $h_t = \omega + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}$. The robust GARCH specification is given in (**??**). VoV (volatility of volatility) is the standard deviation of $\sqrt{\hat{h}_t}$. The estimated volatilities are expressed in percentage points.

We then estimated the model using returns only after September 29, 2000 (we dummied out that day and got the same results). Including the extreme observation increases α_1 from 0.019 to 0.157, decreases β_1 from 0.979 to 0.824 and increases the long–run variance from 8×10^{-4} to about 30×10^{-4}. Including this once–off explainable event in the sample thus strongly blows up the parameter estimates and, as a consequence, the out–of–sample forecasts. We denote this as *extreme bias*.

For comparison, we estimated a model taking this into account, the BIP–GARCH model discussed below, and found that the results are not affected much by the extreme observation. This can be seen in Figure 3 where we plot the daily Apple returns and the volatility forecasts obtained by the GARCH and BIP–GARCH model. We find that in the standard model the volatility shoots up sharply following the event, and conditional volatility forecasts are higher on average throughout the sample than when the BIP procedure is used.

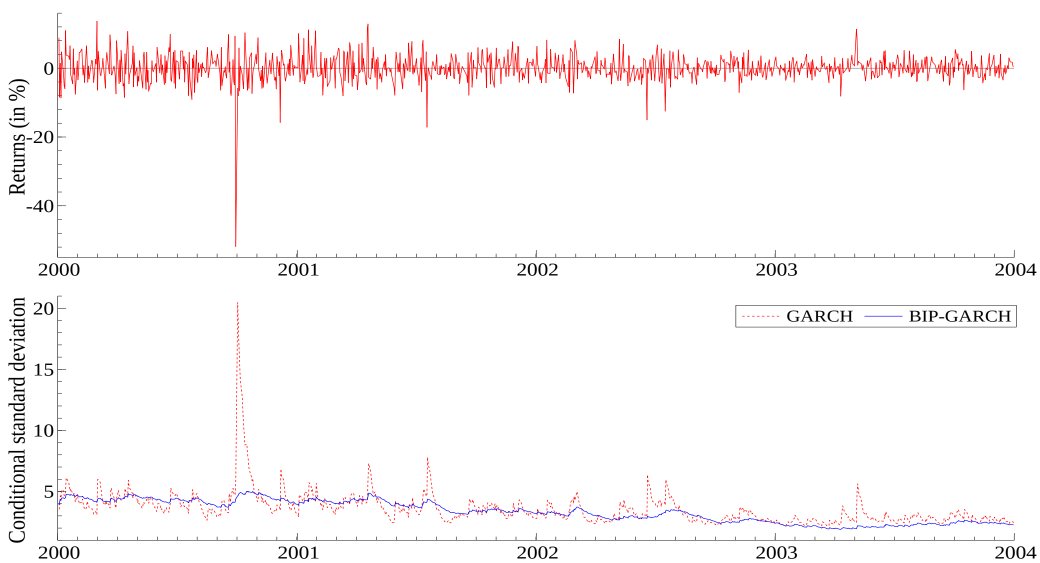

Figure 3: Daily returns in % (upper panel) for Apple and estimated conditional standard deviation for the GARCH and BIP–GARCH (lower panel) on the period 2000-2003.

Table 4 also reports two summary statistics on the estimated conditional volatilities. We see that the mean variance estimate from the GARCH model is only half of its value predicted by the model parameters, while for the BIP GARCH model, these values are very close. A final interesting observation is the difference in the estimated volatility of volatility for the two models. It is 1.543 for the GARCH model and only 0.681 for the BIP–GARCH model.

In conclusion, with this sample, the use of the BIP–GARCH model to forecast volatility lead to much more stable volatility forecasts, as is also clear from the time series plot of volatilities in Figure 3.

The effect of extremes on univariate volatility forecasting can equally be expected to be present in the forecasting of correlations. While little research has demonstrated the impact of extreme observations for correlations, it is readily demonstrated, e.g. by adding one asset to the example with Apple, e.g. Microsoft. Figure 4 plots the daily returns (in %) for these two series. Note the 20% return on the Microsoft stock price, triggered by the once–off event of Microsoft posting first–quarter net income of 46 cents per share, 12 percent above the mean analyst estimate of 41 cents. The same day the stock price of Apple fell by 6%.

Table 5 shows that the effect of this extreme is to cause cDCC conditional correlations to drop in one day from 21.5% to only 3.2%, while historically the average conditional correlation is around 45%. The effect of this extreme is persistent, since it takes more than a month for the estimated conditional correlation to return to its level before the once–off event.

Table 5: Apple/MSFT return and extremes, impact on cDCC–GARCH. Estimation on the full sample (i.e., January 2000 – December 2003).

	cDCC parameters			Correlations t_0 = Oct 19, 2000				
Method	$\overline{Q}_{12}$	α	β	$\hat{R}_{12,t_0}$	$\hat{R}_{12,t_0+1}$	$\hat{R}_{12,t_0+5}$	$\hat{R}_{12,t_0+15}$	$\hat{R}_{12,t_0+20}$
cDCC	0.449	0.026	0.958	0.215	0.032	0.041	0.091	0.167
BIP–cDCC	0.549	0.021	0.956	0.326	0.293	0.288	0.326	0.379

Note: $\hat{R}_{12,t}$ corresponds to the estimated conditional correlation (between Apple and Microsoft) of the classical and robust DCC models on day t.

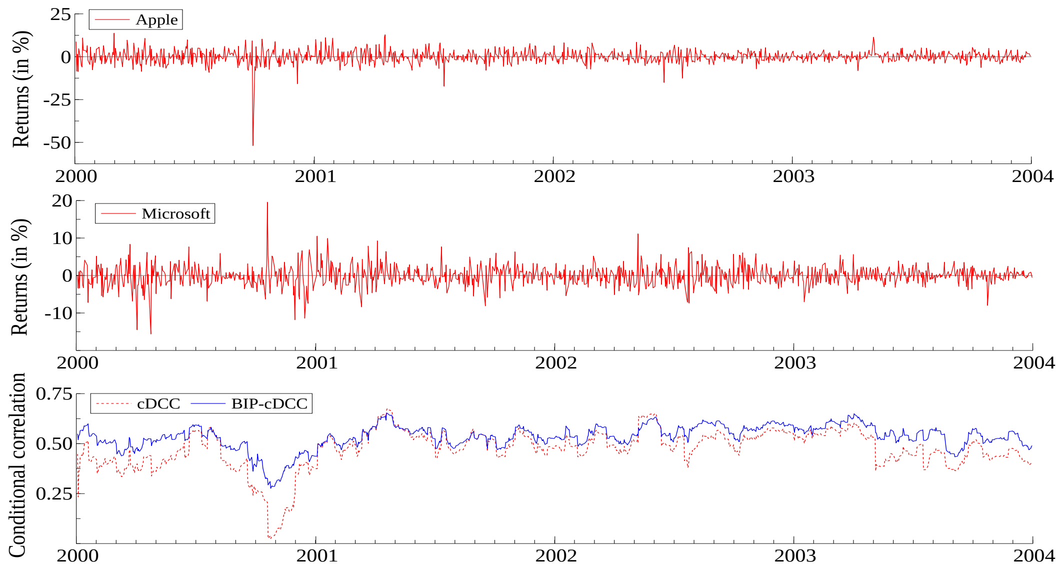

Figure 4: Daily returns in % for Apple and Microsoft (first two panels) and estimated cDCC and BIP-cDCC conditional correlation (lower panel) for the period 2000-2003.

For comparison, we estimated a BIP version of the cDCC model and found the conditional correlation only dropping by 3 percentage points, not 18 points like the baseline model. We further note the strong difference between the two unconditional correlation estimates (about 45% for the cDCC and about 55% for the BIP version) leading the BIP–cDCC correlation to be significantly higher than the cDCC correlation for almost all days in the sample, as can be seen in the lower panel of Figure 4.

3.2 Covariance forecasts exchange rates

The second application is on forecasting the r–step ahead daily conditional covariance matrix of the EUR/USD and Yen/USD exchange rates over the period 2004–2009. From the daily returns, rolling estimation samples of 2303 observations are used to produce the out–of–sample r–step ahead daily covariance forecasts, with $r = 1, \ldots, 10$. In Figures 5-7 we report the time series and scatter plots of the forecasted cDCC and BIP–cDCC variances and covariances.

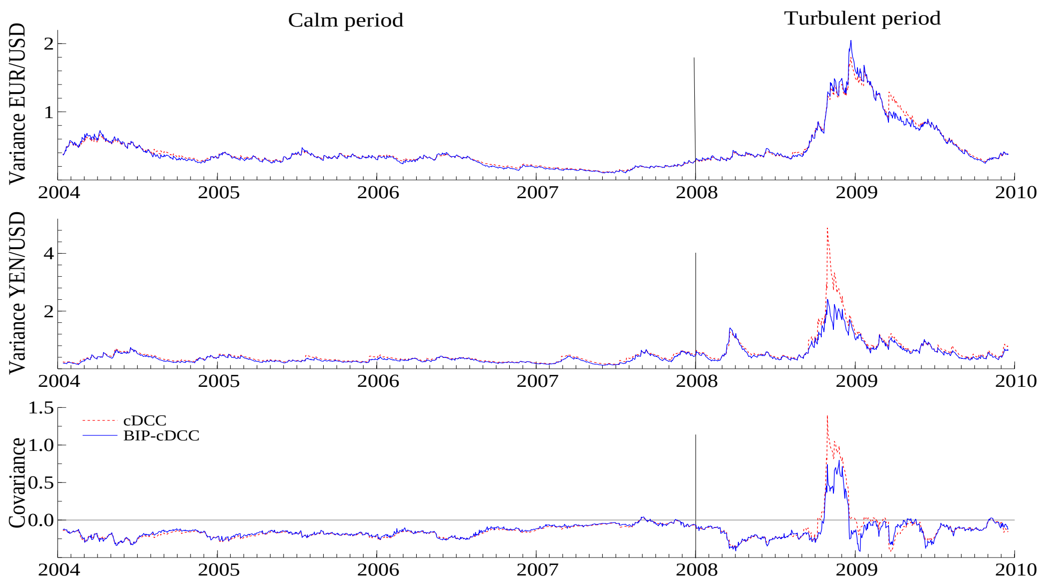

Figure 5: Time series plot of daily variance and covariance forecasts for EUR/USD and Yen/USD over the period 2004–2009.

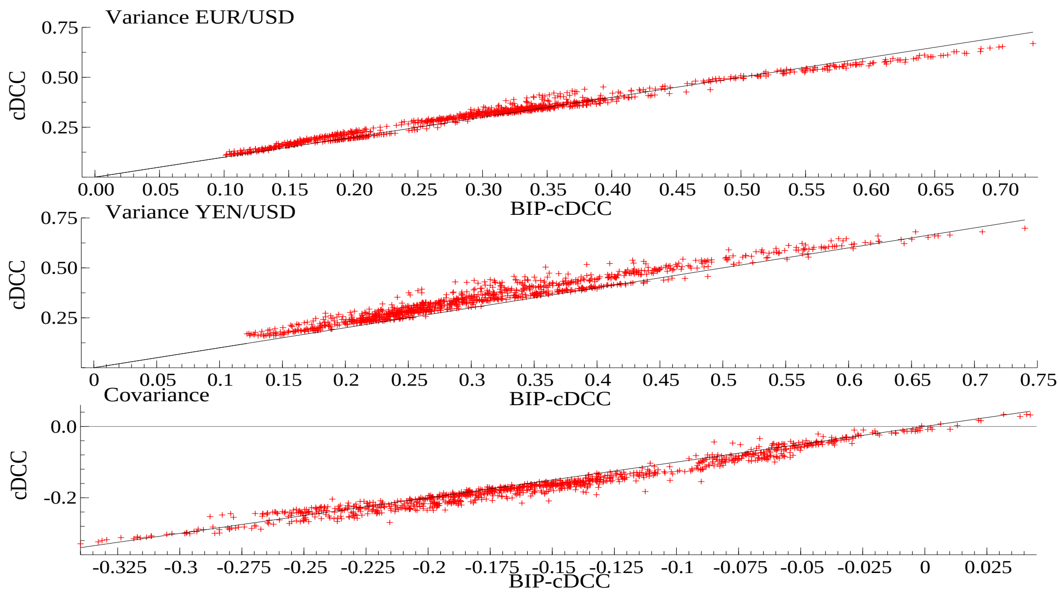

Figure 6: Scatter plot of daily variance and covariance forecasts for EUR/USD and Yen/USD over the calm period 2004–2006.

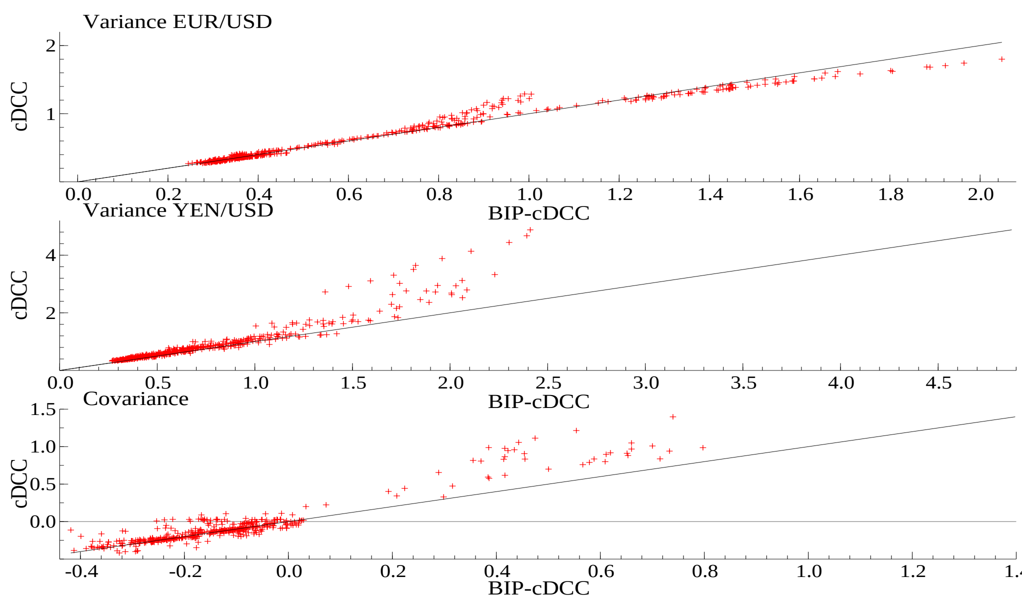

Figure 7: Scatter plot of daily variance and covariance forecasts for EUR/USD and Yen/USD over the turbulent period 2007–2009.

3.3 Industry portfolios

In the application we apply the BIP model to the problem of optimal portfolio allocation under bound constraints on the portfolio weights (long only, and maximum 10%).[1] In Table 6 we report the results for the same portfolio strategy but without bound constraints on the portfolio weights. We see that the use of the BIP method has no significant impact on the average gross portfolio return, but for most sectors it reduces significantly the portfolio standard deviation and turnover.

[1]The list of tickers for each of the sector portfolios is the following: Consumer Discretionary (tickers: AN, AZO, BBBY, BBY, HRB, CCL, CMCSA, DHI, DV, EK, FDO, F, GCI, GPS, GPC, GT, HOG, HAR, HAS, HD, IGT, JCI, KSS, LEG, LEN, LTD, LOW, M, MAR, MAT, MCD, MHP, MDP, NYT, NWL, NKE, JWN, ORLY, ODP, JCP, PHM, RSH, ROST, SHW, SNA, SWK, SPLS, SBUX, HOT, TGT, TIF, TWX, TJX, URBN, VFC, DIS, WPO, WHR), Consumer Staples (ADM, AVP, CAG, CCE, CL, CLX, COST, CPB, CVS, GIS, HNZ, HRL, HSY, K, KMB, KO, KR, MKC, MO, PEP, PG, SLE, STZ, SVU, SWY, SYY, TAP, TSN, WAG, WFM, WMT), Energy (APA, APC, BHI, CHK, COG, COP, CVX, DVN, EOG, EP, HAL, HES, HP, MEE, MRO, MUR, NBL, NBR, NE, NFX, OXY, RDC, RRC, SLB, SUN, SWN, TSO, VLO, WMB, XOM), Financials (ACE, AFL, AIG, ALL, AON, AXP, BAC, BBT, BEN, BK, C, CB, CINF, CMA, EFX, EQR, FHN, FITB, HBAN, HCN, HCP, HST, JPM, KEY, KIM, L, LM, LNC, LUK, MI, MMC, MS, MTB, NTRS, PBCT, PCL, PGR, PNC, PSA, RF, SCHW, SLM, SPG, STI, STT, TMK, TROW, TRV, UNM, USB, VNO, WFC, XL, ZION), Healthcare (ABT, AET, AGN, AMGN, BCR, BAX, BDX, BIIB, BSX, BMY, CAH, CELG, CEPH, CERN, CI, CVH , XRAY, ESRX, FRX, GILD, HUM, JNJ, LH, LLY, MDT, MRK, MYL, PDCO, PKI, PFE, STJ, SYK, THC, TMO, UNH, VAR, WPI), Industrials (APH, AVY, BA, CAT, CMI, CSX, CTAS, DE, DHR, DNB, DOV, EMR, ETN, EXPD, FAST, FDX, FLS, GD, GE, GLW, GR, GWW, HON, IR, ITW, JEC, LMT, LUV, MAS, MMM, NOC, NSC, PBI, PCAR, PCP, PH, PLL, R, RHI, ROK, ROP, RRD, RTN, TXT, TYC, UNP, UTX, WM), IT (AAPL, ADBE, ADI, ADP, ADSK, ALTR, AMAT, AMD, BMC, CA, CPWR, CSC, CSCO, DELL, EMC, ERTS, FISV, FLIR, HPQ, HRS, IBM, INTC, INTU, JBL, JDSU, KLAC, LLTC, LSI, MCHP, MOLX, MSFT, MSI, MU, NSM, NVLS, ORCL, PAYX, QCOM, SYMC, TER, TLAB, TSS, TXN, WDC, XLNX, XRX), and Materials (AA, APD, ARG, BLL, BMS, CLF, DD, DOW, ECL, EMN, FMC, IFF, IP, MWV, NEM, NUE, OI, PPG, PX, SEE, SIAL, VMC, WY, X).

Table 6: Summary statistics on out–of–sample performance of minimum variance portfolios based on the BIP–cDCC vs cDCC model: gross returns (annualized mean, standard deviation), portfolio turnover and difference in annualized Sharpe ratio when the proportional trading cost is κ.

	cDCC			BIP–cDCC			Δ SR BIP–cDCC vs cDCC		
	mean	SD	Turn	mean	SD	Turn	$\kappa = 0$	$1e-4$	$1e-3$
2004-2006									
Cons.Staples	0.072	0.101***	0.482***	0.102	0.088***	0.244***	0.443	0.490	0.933**
Energy	0.518	0.297***	1.953***	0.307	0.186***	0.707***	-0.095	-0.032	0.601
Financials	0.354**	0.158	1.961***	0.076**	0.138	0.665***	-1.692	-1.502	0.221
Healthcare	0.024	0.173***	1.018***	0.102	0.112***	0.329***	0.759	0.838	1.502***
Industrials	0.065	0.150***	1.427***	0.154	0.097***	0.465***	1.138**	1.265**	2.324***
IT	0.075	0.138	0.641***	0.141	0.135	0.496***	0.506	0.522	0.743
Materials	0.221	0.307***	1.262***	0.127	0.126***	0.394***	0.285	0.316	0.538
2007-2009									
Cons.Staples	-0.050	0.178**	0.602***	-0.041	0.167**	0.351***	0.032	0.063	0.348
Energy	-0.042	0.381***	1.494***	-0.044***	0.348	1.002***	-0.016	0.016	0.253
Financials	-0.128	0.333***	1.514***	-0.063	0.236***	0.603***	0.126	0.174	0.617
Healthcare	-0.023	0.202***	0.652***	-0.048	0.174***	0.390***	-0.158	-0.126	0.095
Industrials	-0.107	0.264***	1.355***	-0.139	0.215***	0.667***	-0.237	-0.190	0.269
IT	-0.043	0.272**	0.768	-0.047	0.245**	0.748	-0.032	-0.032	-0.095
Materials	0.097	0.325***	0.944***	-0.041	0.266***	0.529***	-0.459	-0.427	-0.221

***, ** and * indicate significant differences between BIP–cDCC and cDDC at the 1%, 5% and 10% level respectively.

References

Aielli, G. (2009): "Dynamic Conditional Correlations: on Properties and Estimation," *Unpublished paper: Department of Statistics, University of Florence.*

Andersen, T. G., T. Bollerslev, and F. Diebold (2007): "Roughing it up: Including Jump Components in the Measurement, Modelling and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89, 701–720.

Boudt, K., and C. Croux (2010): "Robust M-Estimation of Multivariate GARCH Models," *Computational Statistics and Data Analysis*, 54, 2459–2469.

Boudt, K., C. Croux, and S. Laurent (2011): "Outlyingness Weighted Covariation," *Journal of Financial Econometrics*, 9, 657–684.

Boudt, K., J. Daníelsson, and S. Laurent (2012): "Robust Forecasting of Dynamic Conditional Correlation GARCH Models," *Mimeo.*

Carnero, M., D. Peña, and E. Ruiz (2012): "Estimating GARCH volatility in the presence of outliers," *Economics Letters*, 114, 86–90.

Engle, R. (2002): "Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R., and J. Mezrich (1996): "GARCH for Groups," *RISK*, 9, 36–40.

Francq, C., L. Horvath, and J.-M. Zakoian (2011): "Merits and Drawbacks of Variance Targeting in GARCH Models," *Journal of Financial Econometrics*, 9, 619–656.

Muler, N., and V. Yohai (2008): "Robust Estimates for GARCH Models," *Journal of Statistical Planning and Inference*, 138, 2918–2940.