VOLATILITY SPILLOVERS: A SPARSE MULTIVARIATE GARCH APPROACH WITH AN APPLICATION TO COMMODITY MARKETS

Geert Dhaene

KU Leuven
geert.dhaene@kuleuven.be

Piet Sercu

KU Leuven
piet.sercu@kuleuven.be

Jianbin Wu[†]

Nanjing University
wujianbin@nju.edu.cn

January 18, 2022

We propose sparse DCC-GARCH and BEKK-GARCH models based on L_1 regularization. We use the models to study daily return volatility and correlation spillovers for the 24 constituents of the Bloomberg commodity index in the period 2000–2018. The sparse models outperform the diagonal models out-of-sample in terms of model fit and other criteria. We also test whether the higher visibility of metals and energy markets compared to agricultural commodities affects the speed of information processing. We find correlation spillovers from metals and energy to agricultural commodities even though the latter tend to settle somewhat later.

Data availability: The data that support the findings of this study are available from Datastream. Restrictions apply to the availability of these data, which were used under license for this study. Data are available on Datastream with the permission of Datastream.

Keywords: *multivariate GARCH, regularization, penalized estimation, volatility spillovers, correlation spillovers, commodity markets, price discovery*

[†] Corresponding author. Address: School of Economics, Nanjing University, Anzhong Building 2401, Nanjing University, Hankou Rd 22, Nanjing, Jiangsu 210000, China. Wu acknowledges the financial support from the National Natural Science Foundation of China (project code 71803080).

1. INTRODUCTION

The estimation of conditional covariances between asset returns is central to many areas of empirical finance, including portfolio selection, asset pricing, and hedging. A large literature on multivariate GARCH models has developed; see Bauwens, Laurent, and Rombouts (2006) for a survey. Two prominent and widely used models are the Baba, Engle, Kraft, and Kroner (BEKK) model, proposed by Engle and Kroner (1995), and the scalar dynamic conditional correlation (DCC) model, proposed by Engle (2002) and Tse and Tsui (2002). These models, however, were are not set up to easily handle full-scale volatility or correlation spillover effects across many assets. In the scalar DCC model, for example, the asset return correlations are assumed to evolve identically for all assets. This entails no restriction for bivariate systems, but when the number of assets is large this assumption becomes harder to defend. To address this problem, "diagonal" and "full" versions of the DCC model have been proposed; see, e.g., Engle (2002), Cappiello, Engle, and Sheppard (2006), Hafner and Franses (2009), and Billio and Caporin (2009). But these do not fully solve the problem. While the diagonal DCC model allows for idiosyncratic correlation dynamics, it still ignores correlation spillover effects. The full DCC model does allow for correlation spillovers, but here the number of parameters is of order n^2, with n the number of assets considered, so estimation of the full DCC model is feasible only when n is small or, at most, moderately large. Essentially the same holds for BEKK models, where the diagonal model version ignores volatility spillovers and the general model version runs into estimation problems unless n is relatively small. In short, in multivariate GARCH modeling there is a conflict between flexibility and feasibility of estimation.

A recent account from the perspective of the DCC model is given in Bauwens, Grigoryeva, and Ortega (2016), along with a proposal of several new non-scalar DCC models and improved estimation of the full DCC model when n is moderately large. In addition, these authors identify a second issue, the risk of overfitting: in an empirical study of the thirty DJIA assets, the full DCC model is often dominated by less richly parameterized models. In an application to pricing DJIA options, Rombouts, Stentoft, and Violante (2014) compared the predictive accuracy of 444 DCC model variants with different specifications of the conditional variances, conditional correlations, and innovation distribution. They reached the complementary conclusion that allowing for correlation spillovers in a parsimonous way yields substantial gains in pricing performance.

In this paper we develop sparse versions of the DCC and BEKK models, in an effort to mitigate the conflict between flexibility and feasibility as well as the risk of overfitting.[1] The proposed sparse models are more flexible than the diagonal models, yet more parsimonious than the full models. They are intended to capture correlation and volatility spillover effects while still being estimable when the dimension, n, is moderately large ($n = 24$ in our application). Prior to estimation, the sparse models can be viewed as full models, with parameter restrictions being imposed along

[1] These difficulties are not unique to the DCC and BEKK models. They occur generally in multivariate GARCH models and could, in principle, be addressed using sparse model versions, similar to those presented here for the DCC and BEKK models. Alternative ways to estimate large-dimensional GARCH models—but without regularization—were considered by many authors, including Ledoit, Santa-Clara, and Wolf (2003) for the diagonal VEC model (Bollerslev, Engle, and Wooldridge 1988), Engle (2009) for the DCC and factor DCC models, Francq and Zakoïan (2016) for the BEKK and DCC models, and Pakel, Shephard, Sheppard, and Engle (2021) for the BEKK, DCC, and cDCC (Aielli 2013) models.

the estimation in a data-compatible way rather than a priori. The sparse parameter structure is obtained by regularization. Specifically, we add an L_1 (or lasso) penalty (Tibshirani 1996) to the log-likelihood function to penalize the off-diagonal elements of the coefficient matrices, and we select the degree of regularization by cross-validation. This drives many of the off-diagonal elements to zero, so that a sparse structure of correlation or volatility spillover effects obtains. Thus, one of the key aims of our proposal is exploratory analysis: out of the many possible directed correlation or volatility spillovers in a large-dimensional system of asset returns, how to detect the true ones or, realistically, how to detect the discernable ones? Our approach is similar in spirit to work by Sun and Lin (2012) and Cattivelli and Gallo (2018), who regularize, respectively, the loadings of the full-factor multivariate GARCH model of Vrontos, Dellaportas, and Politis (2003) and the coefficient matrices of the vector multiplicative error model of Cipollini, Engle, and Gallo (2006, 2013). It is also related to Callot, Kock, and Medeiros (2017) and Croux, Rombouts, and Wilms (2018), who proposed equation-by-equation regularization to forecast realized (co)variances. Although our focus here is on GARCH models, we note that our approach can also be applied to models involving multivariate realized measures, for example the multivariate HEAVY model of Noureldin, Shephard, and Sheppard (2012) or multivariate extensions of the realized GARCH model of Hansen, Huang, and Shek (2012).

To explore the merits of the sparse GARCH approach we use both simulations and empirical data. In simulations we examine how well a sparse GARCH parameter structure, with unknown degree of sparsity, can be detected by L_1 regularization coupled with cross-validation. To give an idea of the performance, in a sparse 24-dimensional BEKK model with 6000 observations, a high degree of persistence (GARCH coefficients between 0.985 and 0.990), and 96% of the spillover coefficients (the off-diagonal ARCH coefficients) being zero or nearly zero and 4% being non-negligible (around one-fourth of the value of the diagonal ARCH coefficients), 87% of the former coefficients are estimated as zero and 83% of the latter as nonzero.

Our empirical analysis studies the daily returns on the 24 individual commodities that are the constituents of the daily Bloomberg Commodity Index. Like Bloomberg, we select the nearest futures price, not the spot price. One reason is that not all commodity spot markets are equally well organized; spot soy and wheat prices, for instance, are averages of prices bilaterally arranged throughout the morning at the silos in Illinois and Indiana. Second, futures markets often react faster (see, e.g., Tse 1999; Yang, Bessler, and Leatham 2001; and Zhong, Darrat, and Otero 2004).

In selecting Bloomberg's commodity prices as our data set, an important consideration was synchronicity. Our data are settlement prices fixed around noon EST, and while some timing issues remain, they are minor compared to what one would face in international stock-index returns; see, e.g., Jung and Maderitsch (2014). Equally important, the commodity data allow us to judge whether the reported spillover patterns are just the result of in-sample overfitting; below, we articulate our priors as to which market(s) should be a leader or a follower, absent timing issues.[2]

[2] Prior work on multivariate GARCH patterns related to commodity markets often focuses on links between commodities as a group and other asset classes like bonds, stocks, and currencies, or on links among closely related commodities (e.g., energy). Our interest, instead, is in links between broad subclasses of commodities, namely energy, metals, and agricultural goods. Within that narrower literature, prior work has focused on explaining the changing

Being the source versus the receiver of spillovers reflects the relative speed of information processing. In perfect markets, all markets should react instantly to any relevant news, leaving no room for spillovers. In reality, we hypothesize that oil and gold are comparatively fast, with other metals coming up second and agricultural goods bringing up the rear. This pattern is roughly in line with the empirical results cited below, but it also fits an economic logic. One prime argument in support of the above ranking is that, in reality, the acquisition and processing of information is costly. With partially informed players, crowds become wiser the more opinions they reflect. For that reason we expect more alertness in markets that enjoy more visibility and attract more analyst attention. Energy and precious metals, especially oil and gold, do enjoy a very high visibility and are even covered in general news media. Metals and energy-related commodities also have a wider analyst following, being relevant to more industrial firms than are agricultural goods, relevant to just food & beverages and textile sectors. It is not a coincidence, in that light, that energy and metals are heavily traded in London too.[3] For New York and Chicago traders, one benefit is that, every morning, they observe London's consensus digest of the overnight news, which reduces the tâtonnement noise that normally arises when a daily market re-opens. In addition, energy and metals markets tend to have higher turnovers, and this is associated with better liquidity and lower trading costs. Lastly, metals resemble financial assets better, compared to agricultural commodities. Their value per unit of volume is higher, for instance, which means that shipment costs are lower relative to the price; and there is no seasonality in supply and therefore in convenience yield, which simplifies pricing.[4]

The empirical literature on this subject tends to agree that energy is a main source of spillovers, followed by metals, with agricultural commodities being the net receivers. The statistical approaches behind these findings are quite varied: bivariate E-GARCH with time-varying correlation (Ji and Fan 2012), Diebold-Yilmaz generalized VAR (Chevallier and Ielpo 2013), causality-in-variance tests (Nazlioglu, Erdem, and Soytas 2013), quantile-regression-based conditional value-at-risk (CoVaR) (Algieri and Leccadito 2017), multivariate DECO-GARCH (Kang, McIver, and Yoon 2017), and regime-switching CoVaR copulas (Ji, Bouri, Roubaud, and Shahzad 2018). The consensus is not complete, though. Kaltalioglu and Soytas (2011) and Cabrera and Schulz (2016), using causality-in-variance tests and DCC-GARCH, respectively, find essentially no spillover links. Lu, Yang, and Liu (2019), lastly, who work with a HAR model, detect bidirectional links before the 2008 crisis, and spillover links towards oil thereafter.

levels of correlations, associating these with the 2008 financial crisis, financialization, the introduction of bio-fuels, or OPEC announcements. See, e.g., Du, Yu, and Hayes (2011), Creti, Jots, and Mignon (2013), Liu, Ji, and Fan (2013), Mensi, Hammoudeh, Nguyen, and Yoon (2014), Adams and Glck (2015), Antonakakis, Floros, and Kizys (2016), Roy and Roy (2017), Bollerslev, Hood, Huss, and Pedersen (2018), Green, Larsson, Lunina, and Nilsson (2018), Shahzad, Hernandez, Yahyaee, and Jammazi (2018), Uddin, Hernandez, Shahzad, and Hedstrm (2018), and Pal and Mitra (2019).

[3] London does trade some of New York's agricultural commodities too, but rather as part of the round-the-clock trading arrangement set up by U.S. exchanges; the key trading sessions remain American-based. In metals and Brent oil, London is a leading venue.

[4] Buying 'cash' (i.e., holding physical inventory) is similar to buying forward, to an industrial user of the good. Unlike a futures position, a cash position also offers the option of actually using the goods if there is a surge in demand or a problem with deliveries. This extra benefit, the convenience yield, is valuable and leads to a premium in cash prices if and when market-wide inventories are low, which happens more often for seasonally-produced and weather-sensitive agricultural goods.

Our empirical findings in terms of model evaluation and spillover patterns are as follows. For the purpose of out-of-sample prediction of commodity price risks and portfolio choice, the sparse DCC model has the overall best performance, followed by the diagonal and scalar DCC variants. The sparse BEKK model is a distant fourth, but it still beats the diagonal BEKK version and static portfolio choices. Regarding the spillover patterns, our empirical results are in line with our priors, with most of the literature, and with the out-of-sample results. Using the sparse DCC model, we find that among the ten commodities that rank highest in terms of being a net source of spillovers, only one is agricultural, while among the ten commodities that rank lowest, eight are agricultural. This suggests that our findings are based on real effects, not overfitting. The pattern is still present if we employ the sparse BEKK model, but it is much weaker.

Section 2 introduces the sparse DCC and BEKK models. Section 3 presents the results of a simulation experiment for the sparse BEKK model. In Section 4 we estimate and evaluate sparse DCC and BEKK models for daily excess returns on 24 commodities in the period 2000–2018. We compare the performance of the sparse models with the diagonal BEKK model and with the scalar and diagonal DCC models using the Giacomini and White (2006) test, compute the model confidence set of Hansen, Lunde, and Nason (2011), and report on the estimated covariance and correlation spillover effects. Section 5 concludes.

2. SPARSE MULTIVARIATE GARCH MODELS

2.1. Specification

Let r_t be the vector of returns on n assets in period t. Write $r_t = E_{t-1}r_t + e_t$, where E_{t-1} is the conditional expectation given past information, so that $E_{t-1}e_t = Ee_t = 0$. Define the conditional and unconditional variance and correlation matrices

$$H_t = E_{t-1}(e_t e_t'), \qquad \overline{H} = E(e_t e_t'),$$
$$R_t = E_{t-1}(\varepsilon_t \varepsilon_t') = D_t^{-1} H_t D_t^{-1}, \qquad \overline{R} = E(\varepsilon_t \varepsilon_t'),$$

where ε_t is the vector of standardized returns and D_t is the diagonal matrix with the conditional standard deviations of the returns on the diagonal, i.e.,

$$\varepsilon_t = D_t^{-1} e_t, \qquad D_t = (I_n \odot H_t)^{1/2},$$

where $\odot$ is the Hadamard product. Multivariate GARCH models specify how H_t and R_t evolve over time, often through a first-order ARMA-type structure.

One challenge in multivariate GARCH modeling is to keep the model sufficiently flexible while preventing the number of parameters from growing too rapidly with n. See, for example, the discussion in Bauwens, Laurent, and Rombouts (2006). Leaving other differences aside, multivariate GARCH models typically come, in increasing order of generality, in "scalar", "diagonal", and "general" versions, with $O(1)$, $O(n)$, and $O(n^2)$ parameters, respectively. It is generally acknowledged that the richly parameterized models, with $O(n^2)$ parameters, can only be estimated sensibly when n is small enough. For greater n, researchers tend to resort to scalar or diagonal model versions, with $O(1)$ or $O(n)$ free parameters. These more tightly parameterized models result from imposing

prior restrictions on the coefficient matrices. Our aim is to avoid such a priori restrictions. Starting from a rich model specification with $O(n^2)$ parameters, we impose parameter sparsity through L_1 regularization. In this way, the sparsity structure is the result of a data-driven procedure instead of being imposed ex ante.

Consider the general DCC model with the correlation part specified as

$$R_t = (I_n \odot Q_t)^{-1/2} Q_t (I_n \odot Q_t)^{-1/2}, \tag{2.1}$$

$$Q_t = \overline{R} - A\overline{R}A' - B\overline{R}B' + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + BQ_{t-1}B', \tag{2.2}$$

where A and B are coefficient matrices. This is the model of Cappiello, Engle, and Sheppard (2006) without the asymmetry term. For given $\overline{R}$ (which can be pre-estimated by correlation targeting), this model has $2n^2$ correlation parameters. The scalar version of the model is the standard DCC model of Engle (2002), with $A = aI_n$, $B = bI_n$, and scalar parameters a and b. The diagonal version restricts A and B to be diagonal matrices and has $2n$ correlation parameters (Cappiello, Engle, and Sheppard 2006; Hafner and Franses 2009). Other variants include models with regime switching correlations (Pelletier 2006) or a block structure on A and B, possibly obtained via clustering (Billio, Caporin, and Gobbo 2006; Billio and Caporin 2009; Otranto 2010). A common motivation in these papers is to specify the asset return correlation dynamics flexibly, yet tractably for estimation. Our approach is to obtain sparsely parameterized correlation dynamics via an L_1 penalized log-likelihood with penalty function

$$\text{pen}_{\lambda_A,\lambda_B}(A,B) = \lambda_A \sum_{i\neq j} |A_{ij}| + \lambda_B \sum_{i\neq j} |B_{ij}|$$

for chosen tuning parameters $\lambda_A > 0$ and $\lambda_B > 0$. Note that only the off-diagonal elements of A and B enter the penalty function. The effect of L_1 penalization is that estimates of the off-diagonal elements of A and B are being shrunk towards zero, typically resulting in many estimates being identically zero. Therefore, the estimated model lies between the diagonal and the general model versions.

The penalization approach can be applied in the same way to multivariate GARCH models that specify H_t directly. For example, the first-order BEKK model of Engle and Kroner (1995), subject to the variance targeting constraint, specifies

$$H_t = \overline{H} - A\overline{H}A' - B\overline{H}B' + Ae_{t-1}e_{t-1}'A' + BH_{t-1}B', \tag{2.3}$$

which is analogous to (2.2) and has analogous scalar and diagonal versions. Hence, penalization of A and B in the BEKK model can proceed in exactly the same way as in the DCC model. Note, however, that despite the technical similarity the role of penalization in the DCC and BEKK models is different. In the BEKK model, penalization affects the variance and covariance spillovers whereas in the DCC model only correlation spillovers are affected, that is, the DCC variance estimates are unaffected by the penalization.

The distinction between different types of tuning parameters (here, λ_A and λ_B) in the penalty function allows additional modeling and penalization flexibility. For example, setting $0 < \lambda_A < \infty$ and $\lambda_B = \infty$ imposes diagonality on B and sparsity on the off-diagonal elements of A. Furthermore, the model can easily be extended to incorporate slowly changing unconditional variances

or correlations (Engle and Rangel 2008; Hafner and Linton 2010; Bauwens, Hafner, and Pierret 2013) or additional effects such as asymmetries (Cappiello, Engle, and Sheppard 2006). Additional effects typically entail additional parameter matrices, which may be penalized as above to the desired degree. Note, furthermore, that penalization of the diagonal model version also fits into our framework, by writing $A = aI_n + \text{diag}(\alpha)$ and $B = bI_n + \text{diag}(\beta)$, where a and b are scalars and α and β are vectors, and using $\text{pen}_{\lambda_\alpha,\lambda_\beta}(\alpha,\beta) = \lambda_\alpha \sum_i |\alpha_i| + \lambda_\beta \sum_i |\beta_i|$ as penalty function.

2.2. Estimation

To estimate the sparse DCC model we use the two-step procedure of Engle (2002), augmented with penalization in the second step.[5] The volatility part of the DCC model consists of n univariate GARCH(1,1) models, one for each asset, with parameters jointly denoted as θ. The correlation part consists of (2.1)–(2.2), with parameters $\phi = (A, B)$. We use a Gaussian quasi-likelihood function. Although the distribution of financial returns has fatter tails than the normal distribution, maximizing a Gaussian (quasi-)likelihood gives consistent estimates of GARCH models under non-normality, provided that the first two moments are correctly specified (Bollerslev and Wooldridge 1992). The penalized Gaussian quasi log-likelihood, for given $\overline{R}$ and tuning parameters $\lambda = (\lambda_A, \lambda_B)$, is

$$\begin{aligned} l_{\text{pen}}(\theta,\phi) &= -\frac{1}{2}\sum_t (n\log(2\pi) + \log|H_t| + e_t' H_t^{-1} e_t) - \text{pen}_\lambda(\phi) \\ &= l_{\text{v}}(\theta) + l_{\text{c}}(\theta,\phi) - \text{pen}_\lambda(\phi), \end{aligned}$$

where l_{v} and l_{c} correspond to the volatility and correlation parts,

$$\begin{aligned} l_{\text{v}}(\theta) &= -\frac{1}{2}\sum_t (n\log(2\pi) + 2\log|D_t| + r_t' D_t^{-2} r_t), \\ l_{\text{c}}(\theta,\phi) &= -\frac{1}{2}\sum_t (\log|R_t| + \varepsilon_t' R_t^{-1}\varepsilon_t - \varepsilon_t'\varepsilon_t), \end{aligned}$$

as in Engle (2002). In step one, $l_{\text{v}}(\theta)$ is maximized by fitting a GARCH(1,1) model for each asset separately. This gives $\widehat{\theta}$, $\widehat{D}_t$, $\widehat{\varepsilon}_t = \widehat{D}_t^{-1} r_t$, and $l_{\text{c}}(\widehat{\theta},\phi)$, with $T^{-1}\sum_t \widehat{\varepsilon}_t\widehat{\varepsilon}_t'$ as the correlation-targeting estimate of $\overline{R}$. The second step is to solve

$$\max_\phi \left\{ l_{\text{c}}(\widehat{\theta},\phi) - \text{pen}_\lambda(\phi) \right\},$$

for which we use the block-coordinate update method. Dividing $\phi = (\phi_{\text{pen}}, \phi_{\text{unp}})$ into a block of penalized parameters ϕ_{pen} and a block of unpenalized parameters ϕ_{unp}, we update one block at the time, cycling over the two blocks until convergence. In each cycle, we update ϕ_{unp} with one step of the Newton-Raphson method and ϕ_{pen} with a single pass of the coordinate ascent optimization algorithm (given that $\text{pen}_\lambda(\phi)$ is not differentiable in ϕ_{pen} at the origin). The latter algorithm updates one parameter at the time, $\phi_j \in \phi_{\text{pen}}$, with all others held fixed, as follows:

[5] It should be noted that Aielli (2013) pointed out an inconsistency problem with the two-step estimation method of Engle (2002) and suggested a solution, the cDDC model. In practice, however, the differences between cDCC and two-step DCC estimates tend to be small; see Aielli (2013) and Bauwens, Grigoryeva, and Ortega (2016). Therefore, we employ the (simpler) two-step estimation method.

if $|\nabla_{\phi_j} l_c(\widehat{\theta}, \phi)|_{\phi_j=0}$ is less than the tuning parameter, ϕ_j is set to zero; else ϕ_j is set such that $\nabla_{\phi_j} l_c(\widehat{\theta}, \phi)$ equals the tuning parameter.

The sparse BEKK model with volatility specification (2.3) and parameters $\phi = (A, B)$ can be estimated along the same lines in one step by maximizing $l_{\text{pen}}(\phi) = l_v(\phi) - \text{pen}_\lambda(\phi)$, where $l_v(\phi) = -\frac{1}{2}\sum_t (n \log(2\pi) + \log|H_t| + r_t' H_t^{-1} r_t)$ and with $T^{-1}\sum_t e_t e_t'$ as the variance-targeting estimate of $\overline{H}$.

At each iteration of the optimization, we check positive definiteness of Q_t or H_t (in the DCC or BEKK model, respectively) for all t in the estimation data. If positive definiteness does not hold at some iteration, the step size of the parameter update is reduced by a factor $1/2$ until positive definiteness is restored. This guarantees positive definiteness at the converged estimates in the estimation data, but not necessarily outside the estimation data, although we did not encounter this problem. Should it occur, one may impose a positive lower bound on the eigenvalues of Q_t or H_t as in, e.g., Callot, Kock, and Medeiros (2017). Without further restriction, the sparse GARCH models do not guarantee positive definiteness.

Multivariate GARCH models with high-dimensional parameters are numerically challenging to estimate. The penalization step is numerically slow, adding to the challenge. Furthermore, the degree of regularization is controlled by the tuning parameters λ_A and λ_B, which have to be chosen. At the present stage, we set $\lambda_B = \infty$ (enforcing B to be diagonal) and, in the DCC model, we impose the further restriction $B = bI_n$, where b is a scalar. Hafner and Franses (2009) noted that, in the diagonal DCC model, the parameters associated with the autoregressive part Q_{t-1} are less varying than those associated with the innovations $\varepsilon_t \varepsilon_t'$. So, broadly speaking, B may be more tightly parameterized than A. With $B = bI_n$ in the DCC model, we have $\phi = (A, b)$; in the BEKK model, with $B = \text{diag}(\beta)$, we have $\phi = (A, \beta)$. In both models, ϕ_{pen} consists of the off-diagonal elements of A only. We choose λ_A by cross-validation, using the first 90% of the data as training data and the remaining 10% as validation data, involving the following steps:

(i) based on the training data, we estimate ϕ_{unp} with ϕ_{pen} set to zero;
(ii) at these values of ϕ_{unp} and ϕ_{pen}, we compute the log-likelihood gradient vector for ϕ_{pen}, that is, $G_{\phi_{\text{pen}}} = \nabla_{\phi_{\text{pen}}} l_c(\widehat{\theta}, \phi)$ (in the DCC model) or $G_{\phi_{\text{pen}}} = \nabla_{\phi_{\text{pen}}} l_v(\phi)$ (in the BEKK model);
(iii) we compute the 60–98th percentiles, in steps of 2%, of the elements of $|G_{\phi_{\text{pen}}}|$;
(iv) for λ_A equal to each of these percentiles, we compute the penalized estimate of ϕ based on the training data and evaluate the unpenalized log-likelihood on the validation data at this value of ϕ;
(v) we choose the percentile that maximizes this log-likelihood, and set λ_A equal to that percentile of the new $|G_{\phi_{\text{pen}}}|$ that is computed based on both the training and the validation data.

3. SIMULATIONS

This section reports on simulations for the sparse BEKK model. The simulation setup broadly mimics the dimension and properties of the daily excess return data of the 24 Bloomberg commodity subindexes for 2000–2018 that we use in the empirical application discussed in the next section.

Table 1. Estimated versus true sparsity (BEKK model, first 5 replications)

		percentile of $\lvert G_{\phi_{\text{pen}}}\rvert$								
	diag. BEKK	95	90	85	80	75	70	65	60	55
loglik.	-40.3342	-40.3051	-40.2927	-40.2811	-40.2920	-40.2699	**-40.2695**	-40.2699	–	–
zero	552	516	509	501	470	476	459	453	–	–
"large"	0	6	10	12	15	16	16	16	–	–
"small"	0	0	0	0	1	1	3	3	–	–
loglik.	-40.7406	-40.4661	-40.4606	-40.4571	-40.4564	-40.4562	-40.4561	**-40.4558**	-40.4567	–
zero	552	509	494	470	445	425	405	393	365	–
"large"	0	11	12	15	16	16	16	16	17	–
"small"	0	0	1	1	2	2	3	4	5	–
loglik.	-40.2537	-40.0338	-40.0276	-40.0256	-40.0235	-40.0229	-40.0223	-40.0219	**-40.0215**	-40.0218
zero	552	513	504	497	490	472	442	432	415	381
"large"	0	8	12	15	16	16	17	17	17	17
"small"	0	0	0	0	0	0	0	2	2	2
loglik.	-40.7528	-40.3729	-40.3673	**-40.3646**	-40.3657	–	–	–	–	–
zero	552	491	462	434	389	–	–	–	–	–
"large"	0	14	18	19	19	–	–	–	–	–
"small"	0	1	1	1	1	–	–	–	–	–
loglik.	-40.5466	-40.3318	-40.3253	-40.3220	-40.3208	**-40.3207**	-40.3210	–	–	–
zero	552	507	490	468	449	435	410	–	–	–
"large"	0	12	16	16	17	17	17	–	–	–
"small"	0	0	0	0	0	2	2	–	–	–

Model: $r_t \sim N(0, H_t), H_t = \overline{H} - A\overline{H}A' - \text{diag}(\beta)\overline{H}\text{diag}(\beta) + Ae_{t-1}e'_{t-1}A' + \text{diag}(\beta)H_{t-1}\text{diag}(\beta)$, with sparsity imposed on the off-diagonal elements of A. Loglik: log-likelihood, normalized by the number of observations and evaluated at validation data and at penalized estimates obtained from the estimation data with λ_A set to a percentile of $|G_{\phi_{\text{pen}}}|$; the maximum log-likelihood is indicated in boldface. "zero" is the number of zero estimated parameters whose true value is 0. "large" and "small" are the number of nonzero estimated parameters whose true values are ± 0.02 and ± 0.002, respectively.

Our aim here is to explore how well the estimator can detect the sparse parameter structure in a moderately high-dimensional, highly parameterized BEKK model. We generated data

$$r_t = e_t \sim N(0, H_t), \qquad t = 1, \ldots, 6000,$$

(in addition to a burn-in sample of 1000 periods) with $n = 24$, H_t as in (2.3) with $B = \text{diag}(\beta)$ and $\beta_1, \ldots, \beta_n$ independently drawn from the uniform distribution $U[0.985, 0.990]$, $\overline{H}$ equal to the empirical daily return covariance matrix, and A chosen as follows. We drew the diagonal elements of A independently from $U[0.08, 0.11]$. For the off-diagonal elements of A, we randomly set 10 of them equal to 0.02, 10 equal to -0.02, 5 equal to 0.002, 5 equal to -0.002, and the other 522 equal to zero. For each simulated data set, we estimated $\phi = (A, \beta)$, with the covariance matrix of the simulated data as an estimate of $\overline{H}$.

The estimation, including the selection of the tuning parameter λ_A as described above, is quite time-consuming. We only ran 20 replications, but this already gives an idea of the performance of the cross-validation to select λ_A. In addition, to speed up the algorithm, we slightly changed steps (iii)–(v). We started with λ_A equal to the 95th and 90th percentiles of $|G_{\phi_{\text{pen}}}|$, and compared the corresponding log-likelihoods, evaluated at the validation data. If the log-likelihood associated with the 95th percentile dominates the other, we selected λ_A as the 95th percentile of $|G_{\phi_{\text{pen}}}|$ and stopped. Otherwise, we continued and compared the log-likelihoods (again evaluated at the validation data) with λ_A equal to the 90th and 85th percentiles of $|G_{\phi_{\text{pen}}}|$, and so forth, until

Table 2. Estimated versus true sparsity (BEKK model, 20 replications)

	estimated zero	estimated nonzero	total
true zero	456.95	65.05	522
true "small"	8.30	1.70	10
true "large"	3.70	16.70	20
total	468.95	83.45	552

Model: $r_t \sim N(0, H_t), H_t = \overline{H} - A\overline{H}A' - \text{diag}(\beta)\overline{H}\text{diag}(\beta) + Ae_{t-1}e'_{t-1}A' + \text{diag}(\beta)H_{t-1}\text{diag}(\beta)$, with sparsity imposed on the off-diagonal elements of A. "large" refers to values ± 0.02; "small" refers to values ± 0.002.

the log-likelihood started to decrease. At that point, λ_A was set equal to the percentile of $|G_{\phi_{\text{pen}}}|$ yielding the greatest log-likelihood. These modifications shorten (iii)–(iv) because the step size is larger, the algorithm stops as soon as the log-likelihood decreases, and (v) is not executed at all.

Table 1 reports, for the first 5 replications, the sequence of log-likelihood values ("loglik."), normalized by the number of observations and evaluated at the validation data and at penalized parameter estimates obtained from the training data with λ_A equal to different percentiles of $|G_{\phi_{\text{pen}}}|$. In all 20 replications, the log-likelihood value first increases as λ_A decreases (and more nonzero parameter estimates appear), but at some point it starts to decrease, a sign of overfitting as too many nonzero estimates appear. When λ_A is set to the 95th percentile of $|G_{\phi_{\text{pen}}}|$ (a heavy penalty for nonzero estimates), on average 506 out of the 522 true zero parameters are estimated to be zero, 0.5 of the 10 parameters with true value ± 0.002 are estimated to be nonzero, and 12.3 out of the of 20 parameters with true value ± 0.02 are estimated to be nonzero. As λ_A decreases to the 90th percentile of $|G_{\phi_{\text{pen}}}|$, on average 486.8 of the 522 true zeros are estimated as zero; and 1.05 of the 10 true ± 0.002's are estimated as nonzero, and 14.95 out of the 20 true ± 0.02's. Naturally, the lesser the penalization, the more frequently true nonzeros are estimated as nonzero, at the cost of more frequently estimating true zeros as nonzero. On average, λ_A is selected as the 79.25th percentile of $|G_{\phi_{\text{pen}}}|$.

Table 2 is a contingency table of the true and estimated off-diagonal elements of A, averaged across the simulations. As the table shows, the underlying sparsity structure is uncovered reasonably well, with 83.5% of the "large" nonzero parameter values ± 0.02 being detected and 87.54% of the zeros being estimated as zero. As expected, most "small" nonzero parameter values, ± 0.002, are shrunk to zero.

Table 3 gives further details for each simulation separately, showing that the estimation results are reasonably stable across the replications. The first column gives the percentile of $|G_{\phi_{\text{pen}}}|$ that was selected by cross-validation. The next three columns pertain to the off-diagonal elements of A: the number of true zeros estimated as zero and the number of "small" and "large" values, respectively, estimated as nonzero. The last four columns report the true and estimated values of the average of the diagonal elements of A and $\text{diag}(\beta)$. These estimates are very close to the true values.

Table 3. Estimation results for each replication (BEKK model, 20 replications)

	percentile of $\lvert G_{\phi_{\text{pen}}} \rvert$	zero	"small"	"large"	$\sum_i A_{ii}/n$	$\sum_i \widehat{A}_{ii}/n$	$\sum_i \beta_i/n$	$\sum_i \widehat{\beta}_i/n$
	70	464	2	16	0.0954	0.0961	0.9878	0.9876
	65	410	2	16	0.0956	0.0945	0.9878	0.9880
	55	404	1	18	0.0950	0.0955	0.9876	0.9873
	85	441	1	19	0.0951	0.0906	0.9872	0.9878
	75	363	3	19	0.0954	0.0969	0.9875	0.9860
	70	424	1	19	0.0946	0.0936	0.9875	0.9875
	85	487	2	16	0.0980	0.0969	0.9881	0.9880
	80	459	1	17	0.0952	0.0940	0.9876	0.9879
	70	412	6	19	0.0942	0.0935	0.9874	0.9874
	90	478	1	18	0.0945	0.0934	0.9873	0.9871
	85	487	1	15	0.0955	0.0964	0.9875	0.9876
	75	462	2	16	0.0938	0.0920	0.9875	0.9871
	80	474	0	18	0.0967	0.0986	0.9870	0.9863
	95	508	2	16	0.0933	0.0946	0.9872	0.9869
	85	472	1	17	0.0943	0.0920	0.9870	0.9876
	90	508	1	16	0.0934	0.0928	0.9875	0.9873
	95	512	0	12	0.0917	0.0916	0.9873	0.9869
	65	437	3	16	0.0950	0.0929	0.9877	0.9878
	90	502	2	13	0.0962	0.0971	0.9876	0.9869
	80	435	2	18	0.0934	0.0906	0.9878	0.9881
mean	79.25	456.95	1.70	16.70	0.0948	0.0942	0.9875	0.9873

Model: $r_t \sim N(0, H_t), H_t = \overline{H} - A\overline{H}A' - \text{diag}(\beta)\overline{H}\text{diag}(\beta) + Ae_{t-1}e'_{t-1}A' + \text{diag}(\beta)H_{t-1}\text{diag}(\beta)$, with sparsity imposed on the off-diagonal elements of A. The first column gives the cross-validated percentile of $\lvert G_{\phi_{\text{pen}}} \rvert$ to select λ_A. The next three columns refer to the off-diagonal elements of A (522 zeros; 10 "small" values ± 0.002; 20 "large" values ± 0.02) and report the numbers of zeros estimated as zero, "small" values estimated as nonzero, and "large" values estimated as nonzero, respectively.

The appendix reports on simulations for the DCC model (under a slightly different setup), which yield qualitatively similar results.

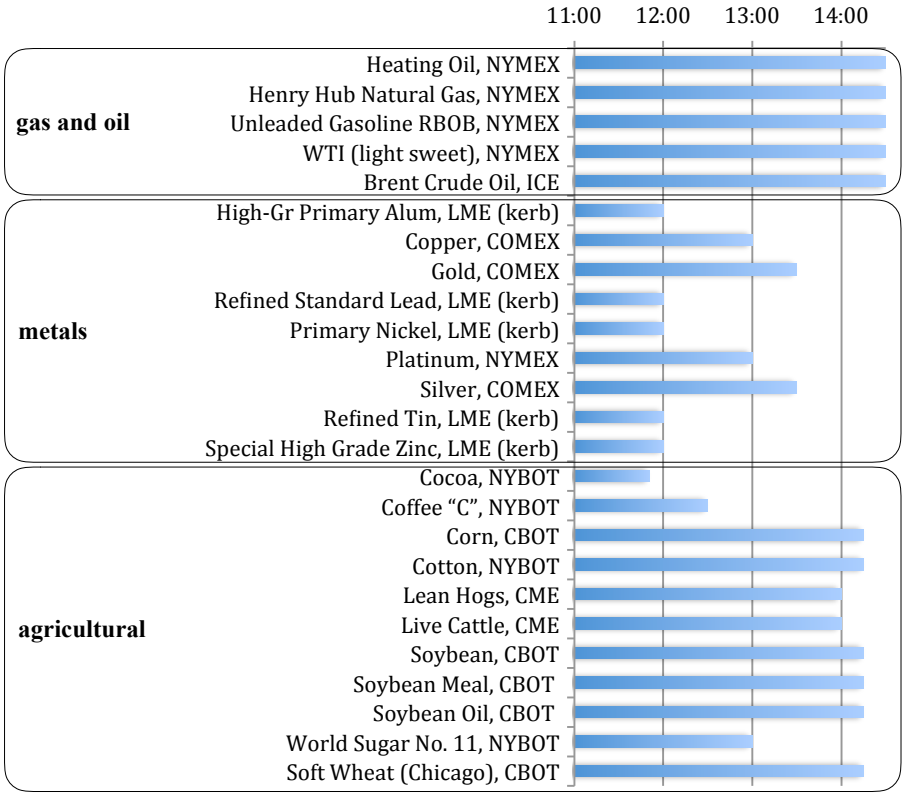
4. APPLICATION TO COMMODITY MARKETS, 2000–2018

In this section we estimate and compare scalar, diagonal, and sparse multivariate GARCH models for (close-to-close) daily excess returns on the 24 commodities eligible for inclusion in the Bloomberg Commodity Index over the period from January 4, 2000, to December 31, 2018. The commodities are aluminum, cocoa, coffee, copper, corn, cotton, composite crude oil (with two designated contracts: Brent and WTI), gold, heating oil, lead, lean hogs, live cattle, natural gas, nickel, platinum, silver, soybean meal, soybean oil, soybean, sugar, tin, unleaded gas, wheat,[6] and zinc.

The excess returns were obtained from Datastream using codes '.BCOM**', e.g., 'BCOMCR' for composite crude oil. Figure 1 provides more detail about the commodities, the trading venue, and the times when prices settle. The Bloomberg Commodity Index is intended to reflect New York prices at noon or shortly thereafter. Eighteen of the prices settle between 13:00 and 14:30 EST. Among the earliest fixings are aluminum, lead, nickel, tin, and zinc, where Bloomberg goes

[6] The wheat subindex included in the Bloomberg Commodity Index has two designated contracts: soft (Chicago) and hard red winter (KC HRW), but the latter is only available as of 2012. Therefore, we use the Chicago wheat contract.

Figure 1. Commodity price settlement times (EST)



The figure shows settlement times for each of the 25 goods (of which Bloomberg merges two, Brent and WTI, into a single 'crude' series). The earliest price (and therefore the most stale price, by the time the last markets settle) is cocoa at NYBOT, which settles at 11:50 EST, followed by five LME kerbstone prices settled at 12:00 EST. The abbreviated names mean the following. ICE: InterContinental Exchange; NYMEX: New York Mercantile Exchange; LME: London Metal Exchange; COMEX: Commodities Exchange, NY; NYBOT: New York Board of Trade; CME: Chicago Mercantile Exchange; CBOT: Chicago Board of Trade. Some of these are legacy names: COMEX, NYMEX, and CBOT are now part of CME, and NYBOT is part of ICE (the owner of NYSE).

for 12:00 EST 'kerbstone' prices from the London Metal Exchange.[7] Also cocoa and coffee settle early, as the graph shows.

4.1. Estimation

For each commodity excess return subindex, $r_{i,t}$, we assume the mean equation is an AR(1) process, $r_{i,t} = \beta_{0,i} + \beta_{1,i} r_{i,t-1} + e_{i,t}$, and apply the multivariate GARCH models to the least-squares residuals, $\widehat{e}_{i,t}$. We divide the data into three parts, with the first 80% as the in-sample training data, the next 10% as the in-sample validation data, and the remaining 10% as the out-of-sample testing data. The cross-validation algorithm (i)–(v) selects the tuning parameter λ_A as the 76th percentile of $|G_{\phi_{\text{pen}}}|$ for the sparse DCC model and as the 74th percentile for the sparse BEKK model; see Table 4. The selections, indicated in boldface, result in 117 and 78 nonzero estimated

[7] The term kerbstone originally refers to informal trading in the street, outside the exchange, after the official closing. Nowadays it refers to electronic trading after the traditional session, which used to be open outcry.

Table 4. Sparse DCC and BEKK estimates for different levels of regularization

sparse DCC				sparse BEKK			
percentile of $\lvert G_{\phi_{\text{pen}}}\rvert$	λ_A	loglik.	no. of param.	percentile of $\lvert G_{\phi_{\text{pen}}}\rvert$	λ_A	loglik.	param.
60	0.1042	−38.8486	158.	60	0.2047	−39.2133	122
62	0.1094	−38.8472	149	62	0.2179	−39.2127	116
64	0.1138	−38.8463	145	64	0.2316	−39.2124	112
66	0.1177	−38.8456	143	66	0.2511	−39.2121	101
68	0.1217	−38.8451	138	68	0.2638	−39.2122	96
70	0.1258	−38.8448	136	70	0.2758	−39.2123	87
72	0.1306	−38.8447	126	72	0.2981	−39.2120	80
74	0.1348	−38.8446	121	**74**	**0.3155**	**−39.2118**	**78**
76	**0.1390**	**−38.8446**	**117**	76	0.3311	−39.2119	73
78	0.1438	−38.8447	111	78	0.3654	−39.2120	69
80	0.1547	−38.8460	92	80	0.3826	−39.2124	68
82	0.1616	−38.8470	86	82	0.4265	−39.2125	66
84	0.1688	−38.8489	81	84	0.4487	−39.2132	61
86	0.1759	−38.8509	75	86	0.4925	−39.2138	59
88	0.1825	−38.8527	71	88	0.5321	−39.2142	58
90	0.1965	−38.8569	60	90	0.6053	−39.2153	56
92	0.2088	−38.8592	52	92	0.7326	−39.2158	51
94	0.2339	−38.8639	41	94	0.8231	−39.2158	50
96	0.2864	−38.8677	32	96	0.9984	−39.2160	49
98	0.3656	−38.8692	28	98	1.2097	−39.2158	48

DCC specifications: $Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2Q_{t-1}$ with $A = aI_n$ in the scalar DCC, A diagonal in the diagonal DCC, and A unrestricted in the sparse DCC. BEKK specifications: $H_t = \overline{H} - A\overline{H}A' - B\overline{H}B' + Ar_{t-1}r_{t-1}'A' + BH_{t-1}B'$ with A and B diagonal in the diagonal BEKK, and A unrestricted and B diagonal in the sparse BEKK. For the sparse DCC model, the number of parameters refers to the correlation part only. Loglik: log-likelihood, normalized by the number of observations and evaluated at validation data. The cross-validated percentile of $\lvert G_{\phi_{\text{pen}}}\rvert$ to select the regularization parameter λ_A is indicated in boldface.

Table 5. Summary of model estimates and log-likelihood values

	log-likelihood in sample	log-likelihood out-of-sample	number of parameters	a or $\sum_i A_{ii}/n$	b or $\sum_i B_{ii}/n$
sparse DCC	−39.1938	−35.9882	107+72	0.0668	0.9922
diagonal DCC	−39.2497	−36.0046	25+72	0.0727	0.9902
scalar DCC	−39.3213	−36.0388	2+72	0.0640	0.9955
sparse BEKK	−39.7197	−36.4939	84	0.1060	0.9898
diagonal BEKK	−39.7909	−36.5467	48	0.1070	0.9897

DCC specifications: $Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2Q_{t-1}$ with $A = aI_n$ in the scalar DCC, A diagonal in the diagonal DCC, and A unrestricted in the sparse DCC. BEKK specifications: $H_t = \overline{H} - A\overline{H}A' - B\overline{H}B' + Ar_{t-1}r_{t-1}'A' + BH_{t-1}B'$ with A and B diagonal in the diagonal BEKK, and A unrestricted and B diagonal in the sparse BEKK. For the DCC models the number of parameters is split between those in the correlation part and the $3 \times 24 = 72$ parameters in the volatility part. The log-likelihood values are normalized by the number of observations.

parameters in the sparse DCC (correlation part) and the sparse BEKK model with the training data, respectively.

Using both the in-sample training and validation data, we estimated the sparse DCC and BEKK models (with the corresponding λ_A obtained), the scalar and diagonal DCC models, and the di-

agonal BEKK model. Table 5 reports a summary of the parameter estimates and the average log-likelihood in and out-of-sample (computed with the parameters fixed at the in-sample estimates) for each model. The sparse DCC model has the greatest in-sample and out-of-sample average log-likelihood values and the diagonal BEKK model has the least. There are 107 and 84 nonzero estimated parameters in the sparse DCC (correlation part) and the sparse BEKK model, respectively. Given the number of unpenalized parameters (DCC: 24 diagonal elements in A and one scalar parameter in B; BEKK: 24 diagonal elements in A and 24 in B), there are 82 nonzero estimated off-diagonal elements of A in the DCC model, while there are only 36 in the BEKK model. The average estimated a (or $\sum_i A_{ii}/n$) is 0.0668, 0.0727, and 0.0640 in the sparse DCC, the diagonal DCC, and the scalar DCC model, respectively, much smaller than those in the sparse BEKK and diagonal BEKK models (0.1060 and 0.1070). In contrast, the estimated b (or $\sum_i B_{ii}/n$) of the DCC models are larger than those in the BEKK models. Finally, the estimated diagonal elements in the sparse models are very close to those in the diagonal (or scalar) models, and the estimated b (or $\sum_i B_{ii}/n$) is still very close to one, especially in the DCC model.

Our sample period includes the financial crisis of 2007–2008 and the Great Recession. One may wonder if this negatively affects the predictive performance of the models due to a possible structural break. Formal statistical inference regarding the presence of a structural break in models estimated by regularization is challenging and beyond the scope of this paper. Our approach, instead, is to compare the estimated models (and the out-of-sample fit in particular) with the same set of models estimated with post-crisis data only. In view of the NBER declaration that the U.S. recession ended in June 2009, we consider the period from July 1, 2009 to December 31, 2018 as our post-crisis subsample. We divide this subsample into three parts such that the in-sample validation period and the out-of-sample testing period are exactly the same as in our earlier division of the full sample period. Thus, the only difference between our full-sample analysis and our subsample analysis is the in-sample training data set, which spans a much shorter period in the subsample (less than 6 years) than in the full sample (more than 15 years). The estimation of the models based on the subsample data now proceeds as before. The cross-validation algorithm (i)-(v) selects the tuning parameter λ_A as the 72nd percentile of $|G_{\phi_{\text{pen}}}|$ for the sparse DCC model and as the 88th percentile for the sparse BEKK model. The selected values are, respectively, $\lambda_A = 0.0969$ and $\lambda_A = 0.1621$, resulting in considerably more nonzero estimated parameters using the training data than earlier: 154 in the sparse DCC (correlation part) and 126 in the sparse BEKK model. Maximizing the penalized likelihood over the joint training and validation data with fixed λ_A ultimately gives 146 nonzero parameters in the sparse DCC (correlation part) and 126 in the sparse BEKK model. Table 6 reports summary results, in the same format as Table 5, and yields some interesting results. First, in all models the in-sample fit is better in the subsample than in the full sample. This is hardly surprising since the model fit worsens during the crisis period. Second, the estimated diagonals of the coefficient matrices in Tables 5–6 are broadly similar, although the estimates in Table 6 show less persistence, in particular for the BEKK models. Third, and of interest for our purpose, the out-of-sample fit is better (for all models) if we use the full sample rather than the post-crisis subsample only. Perhaps this may come as a surprise, but recall that the training

Table 6. Summary of model estimates and log-likelihood values (based on post-crisis data only)

	log-likelihood in sample	log-likelihood out-of-sample	number of parameters	a or $\sum_i A_{ii}/n$	b or $\sum_i B_{ii}/n$
sparse DCC	−36.7622	−36.9612	146+72	0.0650	0.9880
diagonal DCC	−36.8509	−36.9966	25+72	0.0705	0.9845
scalar DCC	−36.9675	−37.1661	2+72	0.0719	0.9905
sparse BEKK	−37.2976	−37.7943	131	0.1080	0.9570
diagonal BEKK	−37.5654	−38.6035	48	0.1519	0.8226

DCC specifications: $Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2 Q_{t-1}$ with $A = aI_n$ in the scalar DCC, A diagonal in the diagonal DCC, and A unrestricted in the sparse DCC. BEKK specifications: $H_t = \overline{H} - A\overline{H}A' - B\overline{H}B' + Ar_{t-1}r_{t-1}'A' + BH_{t-1}B'$ with A and B diagonal in the diagonal BEKK, and A unrestricted and B diagonal in the sparse BEKK. For the DCC models the number of parameters is split between those in the correlation part and the $3 \times 24 = 72$ parameters in the volatility part. The log-likelihood values are normalized by the number of observations.

data period is much shorter in the subsample analysis than in the full-sample analysis. Because of the better out-of-sample fit, we will continue to use the full-sample models in the remainder of the paper.

4.2. Further out-of-sample evaluation

We further evaluate the estimated models using the 10% out-of-sample data with two methods: the pairwise model comparison test of Giacomini and White (2006) and the model confidence set of Hansen, Lunde, and Nason (2011).

Our comparison of the five estimated GARCH models using the Giacomini-White test adopts two alternative loss functions. The first is a standard statistical loss function: minus the out-of-sample (Gaussian quasi-)log-likelihood. This loss function is evaluated with $r_t r_t'$ as a noisy but unbiased proxy for the latent conditional variance H_t, against which the forecasts of H_t are evaluated to determine the loss. The use of a proxy for H_t potentially affects the ranking of the various forecasts, a point that received considerable attention in the literature. The Gaussian quasi-log-likelihood loss function, however, enjoys robustness, in the sense of Patton (2011), to using an unbiased proxy for H_t instead of H_t. That is, the asymptotic ranking of the forecasts is unaffected (see also Patton and Sheppard 2009 and Laurent, Rombouts, and Violante 2013). Table 7 reports the p values of the tests. The sparse DCC model has the lowest loss and outperforms the diagonal DCC and scalar DCC models, though not significantly. The sparse BEKK model significantly improves on the diagonal BEKK model. All three DCC models significantly outperform the BEKK models.

Our second loss function in the Giacomini-White test adopts a more economic criterion, based on the asset-allocation methodology proposed by Engle and Colacito (2006). Consider an asset allocation problem for n assets with return vector r_t whose conditional variance matrix is H_t. The variance minimization problem is

$$\min_{w_t} w_t' H_t w_t \qquad \text{subject to } w_t' 1_n = 1,$$

Table 7. Giacomini-White tests based on the out-of-sample log-likelihood value

	sparse DCC	diag. DCC	scalar DCC	sparse BEKK	diag. BEKK
sparse DCC	–	0.5396	0.2801	0.0000	0.0000
diagonal DCC	0.5396	–	0.2970	0.0000	0.0000
scalar DCC	0.2801	0.2970	–	0.0000	0.0000
sparse BEKK	0.0000	0.0000	0.0000	–	0.0000
diagonal BEKK	0.0000	0.0000	0.0000	0.0000	–
average loss	35.9882	36.0046	36.0388	36.4939	36.5467

Entries: p values of the Giacomini-White test of the null that the corresponding row and column models have equal expected loss, with minus the out-of-sample log-likelihood as the loss function.

Table 8. Giacomini-White tests based on out-of-sample asset allocation

	sparse DCC	diagonal DCC	scalar DCC	sparse BEKK	diagonal BEKK	constant weight	equal weight
sparse DCC	–	0.0184	0.3724	0.1537	0.0549	0.0000	0.0000
diagonal DCC	0.0184	–	0.2587	0.1306	0.0453	0.0000	0.0000
scalar DCC	0.3724	0.2587	–	0.2266	0.0773	0.0000	0.0000
sparse BEKK	0.1537	0.1306	0.2266	–	0.0003	0.0000	0.0000
diagonal BEKK	0.0549	0.0453	0.0773	0.0003	–	0.0000	0.0000
constant weight	0.0000	0.0000	0.0000	0.0000	–	0.0000	0.0060
equal weight	0.0000	0.0000	0.0000	0.0000	0.0000	0.0060	–
average loss	0.1602	0.1593	0.1619	0.1699	0.1743	0.2426	0.3351

Entries: p values of the Giacomini-White test of the null that the corresponding row and column models have equal expected loss, with the out-of-sample squared portfolio return multiplied by 10000 as the loss function.

where 1_n is an $n \times 1$ vector of ones. The solution is $w_t = (1_n' H_t^{-1} 1_n)^{-1} H_t^{-1} 1_n$ and the minimum-variance portfolio has return $w_t'(r_t - E_{t-1} r_t)$ in excess of the expected return. With the out-of-sample squared return $(w_t'(r_t - E_{t-1} r_t))^2$ as the loss function, we compare each pair of GARCH models using the Giacomini-White test. In addition to the portfolios constructed from the GARCH models, we also consider the equally-weighted portfolio, with weights $w_t = n^{-1} 1_n$, and the constantly-weighted portfolio with weights $w_t = (1_n' \overline{H}^{-1} 1_n)^{-1} \overline{H}^{-1} 1_n$ based on the unconditional variance of $r_t - E_{t-1} r_t$. Table 8 shows that in terms of asset allocation the sparse BEKK model significantly outperforms the diagonal BEKK model, while the sparse DCC model does not outperform the diagonal DCC model. The constantly-weighted portfolio is significantly dominated by all BEKK and DCC models. The equally-weighted portfolio is dominated by all other portfolios and the dominations are statistically significant (in the absence of transaction costs). Again, the DCC models dominate the BEKK models in terms of average loss, although this time the dominations are, at best, only marginally significant.

To complement the Giacomini-White tests we also report the out-of-sample model confidence set of Hansen, Lunde, and Nason (2011). Table 9 provides the model confidence sets calculated with the same loss functions as above, a significance level of 5%, 1000 bootstrap replications, and an average block length of 10. The model confidence set is indicated in boldface. With the log-likelihood-based loss function, only the DCC models are in the model confidence set. With the

Table 9. Out-of-sample model confidence sets

	log-likelihood value		asset allocation	
	loss	rank	loss	rank
sparse DCC	**35.9882**	**1**	**0.1602**	**2**
diagonal DCC	**36.0046**	**2**	**0.1593**	**1**
scalar DCC	**36.0388**	**3**	**0.1619**	**3**
sparse BEKK	36.4939	4	**0.1699**	**4**
diagonal BEKK	36.5467	5	0.1743	5
constant weight	38.0810	6	0.2426	6
equal weight	48.1407	7	0.3351	7

Boldface numbers indicate the models in the model confidence set with significance level 5%. Computed with 1000 bootstrap replications and an average block length of 10.

asset allocation loss function, the DCC models and the sparse BEKK model are in the confidence set.

4.3. Covariance and correlation spillovers

The main advantage of the sparse BEKK and DCC models, relative to their scalar and diagonal versions, is that they allow covariance and correlation spillovers through the off-diagonal elements of A. Consider the sparse BEKK model. If A_{ij} is nonzero, then a shock to market j's return at time $t-1$ will affect market i's (co)volatility at time t. Figure 2 depicts the estimated covariance spillover effects based on the in-sample estimates. Each directed arrow corresponds to a nonzero estimated off-diagonal element of A in the sparse BEKK model, with darker lines representing stronger effects. In the same way, Figure 3 depicts the estimated correlation spillover effects associated with the sparse DCC model. When interpreting the patterns of covariance and correlation spillovers, one should keep in mind the very different role of the coefficients in the BEKK and DCC models and, correspondingly, the different role of penalization in the two models. Simply stated, given that covariance and correlation spillovers are rather different objects, one should not necessarily expect their patterns to be roughly similar.

We note that the graph for the sparse DCC model, the approach ranked as the best-performing among all models (sparse or not), contains far more links than that of the sparse BEKK model. Finding many links is helpful only if those links pick up useful features, though. Below, we argue that the DCC results again rank best here.

As we saw, the likelihood and portfolio-management results suggest that DCC's larger catch, in terms of spillover relations deemed relevant, improves the out-of-sample predictive power. The same holds for the economic interpretation. In the introduction we argued that, everything else being the same, energy and metals should be the leaders in the spillover network since they have more analyst following, media attention, volume, and liquidity. Imperfections in the synchronicity could possibly interfere with this. The five LME non-ferrous metal markets settle early, which makes them miss the early-afternoon information and potentially creates the impression that they are slow. The same holds for cocoa and coffee.

Figure 2. Covariance spillover estimates in the sparse BEKK model

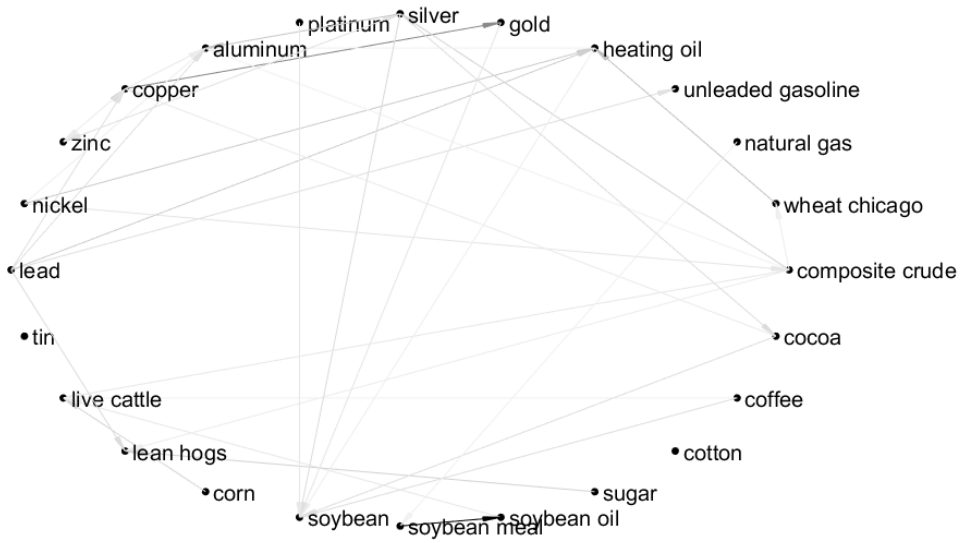
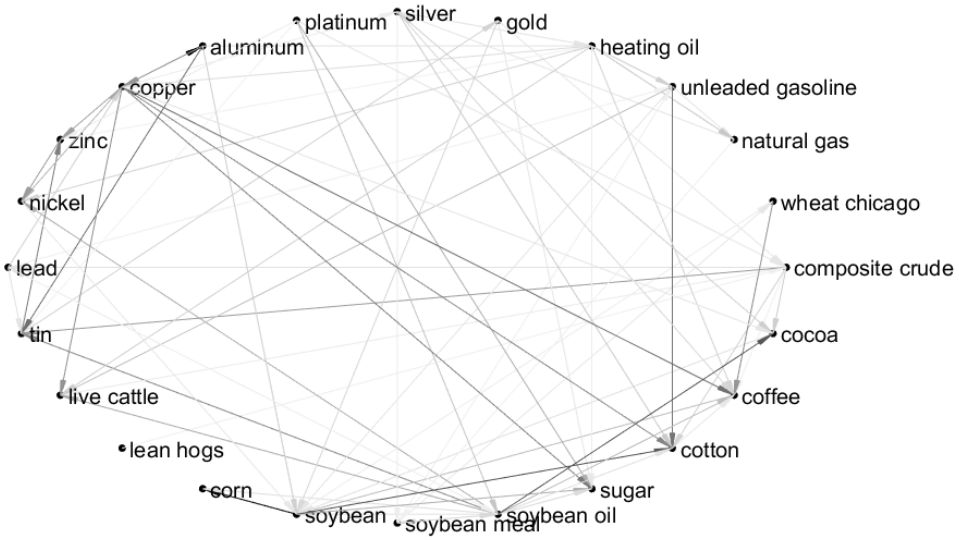


Figure 3. Correlation spillover estimates in the sparse DCC model



To check whether the estimates do detect these patterns, we rank the commodities by their 'net' volatility spillover or correlation spillover count, i.e., the number of outgoing links minus the number of incoming links. Table 10 shows the results. The sparse DCC model fits the predictions remarkably well. Of the ten goods with the strongest net 'out' profile, nine turn out to be energy and metals. Two of these, aluminum and lead, manage to stay quite up to date even though their price is 2–2.5 hours old by the time most markets settle. The group of net recipients of spillovers, to the right of the table, also fits the predictions. Eight of the ten in this list are agricultural goods, and the two metals that strayed into this group come with comparatively stale LME price tags, which helps explain why they do not act as discoverers.

The results for the BEKK estimates are less clear. The classification of energy and metals versus

Table 10. Ranking of commodities in terms of net lead-minus-lag spillover links

DCC correlation spillovers							
top 10 sources of spillovers				top 10 destinations of spillovers			
commodity	#outs	#ins	outs − ins	commodity	#outs	#ins	outs − ins
soybean oil	12	5	7	*wheat Chicago*	2	3	−1
copper	10	5	5	**nickel**	3	5	−2
platinum	6	1	5	*cocoa*	3	5	−2
heating oil	8	4	4	*corn*	1	3	−2
aluminum	5	2	3	**tin**	3	6	−3
composite crude	7	5	2	*soybean*	5	9	−4
silver	6	4	2	*cotton*	3	7	−4
lead	5	3	2	*live cattle*	2	6	−4
gold	4	2	2	*sugar*	2	7	−5
gasoline	5	4	1	*coffee*	2	8	−6
BEKK covariance spillovers							
top 10 sources of spillovers				top 10 destinations of spillovers			
commodity	#outs	#ins	outs − ins	commodity	#outs	#ins	outs − ins
lead	6	1	5	*soybean meal*	2	2	0
composite crude	6	2	4	*soybean oil*	2	2	0
nickel	4	1	3	**tin**	1	1	0
silver	5	3	2	*cotton*	1	1	0
coffee	3	1	2	**heating oil**	3	4	−1
cocoa	3	2	1	**gasoline**	1	2	−1
natural gas	2	1	1	*lean hogs*	1	4	−3
platinum	2	1	1	*live cattle*	1	5	−4
corn	2	1	1	**aluminum**	1	7	−6
sugar	2	1	1	*soybean*	1	7	−6

We rank the commodities on the basis of the number of outgoing spillover relations minus that of incoming relations in the sparse DCC or BEKK model. Energy and metals are printed in bold, agricultural commodities in italics. Markets that settle early (and therefore should seem to lag the others) are printed in grey font.

agricultural goods fits in six cases out of ten in each column, which is hardly better than a random outcome. Of the seven markets in this list that close early, five end in the net 'out' group, against the expectations. In defence of the sparse-BEKK results, we note that the expected pattern is actually still detectable in the tails: 4 out of 4 of the top-'out' goods are energy and metals while 3 out of 4 of the strongest 'in' goods are agricultural goods, with the stray 'follower' metal, tin, possibly explained by its 12:00 settlement time. In short, the sparse BEKK model does detect useful links, but in this application it seems less good at this than the DCC model. This is in line with what we found in terms of likelihoods and out-of-sample performance at portfolio risk management.

5. CONCLUDING REMARKS

In this paper, we propose a sparse modeling approach towards multivariate GARCH. The focus is on GARCH$(1,1)$ structures, the generalization to higher orders being obvious, as well as other extensions. Our approach allows to explore the dynamics of moderately large-dimensional financial time series, with particular attention to uncovering volatility or correlation spillover effects. As the number of potential spillover effects increases quadratically with the dimension of the system,

some form of regularization is needed, resulting in a sparse structure of identified spillover effects. We use a lasso penalty to impose selectivity on the detection of spillover effects. One limitation of this approach is that it becomes difficult to construct parameter confidence sets and to carry out statistical tests on the parameters. But the merits are evident from our work on both simulated and real-world data.

Specifically, in our application to daily returns for 24 commodities over the period 2000–2018, we find that the sparse DCC model dominates the standard DCC models that exclude correlation spillover effects. The dominance is evident both in-sample (i.e., fit) and out-of-sample in terms of a likelihood-based loss function. The sparse BEKK model, likewise, dominates the standard diagonal BEKK variants that exclude volatility spillovers. DCC models, whether sparse or not, do better than BEKK ones. Looking for an economic interpretation of the links brought forward, we find empirical support for the pattern we had advanced on a priori grounds: the sparse DCC model almost unambiguously deems energy and metals to be the sources of spillovers, and agricultural goods the receivers. Two metals are classified as followers, but this may reflect the fact that their prices are at least two hours older than, e.g., those of the energy products. The BEKK model, in contrast, seems to be overly selective: the predicted pattern is discernible only in the four most extreme cases at each side.

REFERENCES

Adams, Z. and T. Glck (2015). Financialization in commodity markets: A passing trend or the new normal? *Journal of Banking & Finance 60*, 93–111.

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics 31*, 282–299.

Algieri, B. and A. Leccadito (2017). Assessing contagion risk from energy and non-energy commodity markets. *Energy Economics 62*, 312–322.

Antonakakis, N., C. Floros, and R. Kizys (2016). Dynamic spillover effects in futures markets: UK and US evidence. *International Review of Financial Analysis 48*, 406–418.

Bauwens, L., L. Grigoryeva, and J.-P. Ortega (2016). Estimation and empirical performance of non-scalar dynamic conditional correlation models. *Computational Statistics and Data Analysis 100*, 17–36.

Bauwens, L., C. M. Hafner, and D. Pierret (2013). Multivariate volatility modeling of electricity futures. *Journal of Applied Econometrics 28*, 743–761.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics 21*, 79–109.

Billio, M. and M. Caporin (2009). A generalized dynamic conditional correlation model for portfolio risk evaluation. *Mathematics and Computers in Simulation 79*, 2566–2578.

Billio, M., M. Caporin, and M. Gobbo (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters 2*, 123–130.

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

Bollerslev, T., B. Hood, J. Huss, and L. H. Pedersen (2018). Risk everywhere: Modeling and managing volatility. *The Review of Financial Studies 31*, 2729–2773.

Bollerslev, T. and J. M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews 11*, 143–172.

Cabrera, B. L. and F. Schulz (2016). Volatility linkages between energy and agricultural commodity prices. *Energy Economics 54*, 190–203.

Callot, L. A., A. B. Kock, and M. C. Medeiros (2017). Modeling and forecasting large realized covariance matrices and portfolio choice. *Journal of Applied Econometrics 32*, 140–158.

Cappiello, L., R. F. Engle, and K. Sheppard (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics 4*, 537–572.

Cattivelli, L. and G. M. Gallo (2018). Adaptive lasso for vector multiplicative error models. Working paper (available at SSRN).

Chevallier, J. and F. Ielpo (2013). Volatility spillovers in commodity markets. *Applied Economics Letters 20*, 1211–1227.

Cipollini, F., R. F. Engle, and G. M. Gallo (2006). Vector multiplicative error models: Representation and inference. NBER, Technical Working Paper 331.

Cipollini, F., R. F. Engle, and G. M. Gallo (2013). Semiparametric vector MEM. *Journal of Applied Econometrics 28*, 1067–1086.

Creti, A., M. Jots, and V. Mignon (2013). On the links between stock and commodity markets' volatility. *Energy Economics 37*, 16–28.

Croux, C., J. Rombouts, and I. Wilms (2018). Multivariate lasso-based forecast combinations for stock market volatility. EDHEC, Working paper 18-003.

Du, X., C. L. Yu, and D. J. Hayes (2011). Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis. *Energy Economics 33*, 497–503.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*, 339–350.

Engle, R. and R. Colacito (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics 24*, 238–253.

Engle, R. F. (2009). High dimension dynamic correlations. In J. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics: Festschrift for David Hendry*, pp. 122–148. Oxford University Press.

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

Engle, R. F. and J. G. Rangel (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies 21*, 1187–1222.

Francq, C. and J.-M. Zakoïan (2016). Estimating multivariate volatility models equation by equation. *Journal of the Royal Statistical Society, Series B 78*, 613–635.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*, 1545–1578.

Green, R., K. Larsson, V. Lunina, and B. Nilsson (2018). Cross-commodity news transmission and volatility spillovers in the German energy markets. *Journal of Banking & Finance 95*, 231–243.

Hafner, C. M. and P. H. Franses (2009). A generalized dynamic conditional correlation model: Simulation and application to many assets. *Econometric Reviews 28*, 612–631.

Hafner, C. M. and O. Linton (2010). Efficient estimation of a multivariate multiplicative volatility model. *Journal of Econometrics 159*, 55–73.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*, 877–906.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*, 453–497.

Ji, Q., E. Bouri, D. Roubaud, and S. J. H. Shahzad (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching CoVaR-copula model. *Energy Economics 75*, 14–27.

Ji, Q. and Y. Fan (2012). How does oil price volatility affect non-energy commodity markets? *Applied Energy 89*, 273–280.

Jung, R. and R. Maderitsch (2014). Structural breaks in volatility spillovers between international financial markets: Contagion or mere interdependence? *Journal of Banking & Finance 47*, 331–342.

Kaltalioglu, M. and U. Soytas (2011). Volatility spillover from oil to food and agricultural raw material markets. *Modern Economy 2*, 71–76.

Kang, S. H., R. McIver, and S.-M. Yoon (2017). Dynamic spillover effects among crude oil, precious metal, and agricultural commodity futures markets. *Energy Economics 62*, 19–32.

Laurent, S., J. V. Rombouts, and F. Violante (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics 173*, 1–10.

Ledoit, O., P. Santa-Clara, and M. Wolf (2003). Flexible multivariate GARCH modeling with an application to international stock markets. *Review of Economics and Statistics 85*, 735–747.

Liu, M.-L., Q. Ji, and Y. Fan (2013). How does oil market uncertainty interact with other markets? An empirical analysis of implied volatility index. *Energy 55*, 860–868.

Lu, Y., L. Yang, and L. Liu (2019). Volatility spillovers between crude oil and agricultural commodity markets since the financial crisis. *Sustainability 11*, in press.

Mensi, W., S. Hammoudeh, D. K. Nguyen, and S.-M. Yoon (2014). Dynamic spillovers among major energy and cereal commodity prices. *Energy Economics 43*, 225–243.

Nazlioglu, S., C. Erdem, and U. Soytas (2013). Volatility spillover between oil and agricultural commodity markets. *Energy Economics 36*, 658–665.

Noureldin, D., N. Shephard, and K. Sheppard (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics 27*, 907–933.

Otranto, E. (2010). Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics and Data analysis 54*, 1–15.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics 39*, 652–668.

Pal, D. and S. K. Mitra (2019). Correlation dynamics of crude oil with agricultural commodities: A comparison between energy and food crops. *Economic Modelling 82*, 453–466.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*, 246–256.

Patton, A. J. and K. Sheppard (2009). Evaluating volatility and correlation forecasts. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 801–838. Springer.

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics 131*, 445–473.

Rombouts, J., L. Stentoft, and F. Violante (2014). The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options. *International Journal of Forecasting 30*, 78–98.

Roy, R. P. and S. S. Roy (2017). Financial contagion and volatility spillover: An exploration into Indian commodity derivative market. *Economic Modelling 67*, 368–380.

Shahzad, S. J. H., J. A. Hernandez, K. H. A. Yahyaee, and R. Jammazi (2018). Asymmetric risk spillovers between oil and agricultural commodities. *Energy Policy 118*, 182–198.

Sun, Y. and X. Lin (2012). Regularization for stationary multivariate time series. *Quantitative Finance 12*, 573–586.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B 58*, 267–288.

Tse, Y. (1999). Price discovery and volatility spillovers in the DJIA index and futures markets. *Journal of Futures Markets 19*, 911–930.

Tse, Y. K. and A. K. C. Tsui (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics 20*, 351–362.

Uddin, G. S., J. A. Hernandez, S. J. H. Shahzad, and A. Hedstrm (2018). Multivariate dependence and spillover effects across energy commodities and diversification potentials of carbon assets. *Energy Economics 71*, 35–46.

Vrontos, I. D., P. Dellaportas, and D. N. Politis (2003). A full-factor multivariate GARCH model. *The Econometrics Journal 6*, 312–334.

Yang, J., D. A. Bessler, and D. J. Leatham (2001). Asset storability and price discovery in commodity futures markets: A new look. *Journal of Futures Markets 21*, 279–300.

Zhong, M., A. F. Darrat, and R. Otero (2004). Price discovery and volatility spillovers in index futures markets: Some evidence from mexico. *Journal of Banking & Finance 28*, 3037–3054.

[dataset] Datastream, 2019. Daily settlement prices for aluminum, cocoa, coffee, copper, corn, cotton, composite crude oil (Brent and WTI), gold, heating oil, lead, lean hogs, live cattle, natural gas, nickel, platinum, silver, soybean meal, soybean oil, soybean, sugar, tin, unleaded gas, wheat(Chicago) and zinc. The excess returns were obtained from Datastream using codes '.BCOM**', e.g., 'BCOMCR' for composite crude oil.

Appendix: Simulations for the DCC model

Here we present simulation results for the sparse DCC model. The setup is qualitatively similar to that for the BEKK model, but differs in some details. The setup here is tailored to a different application: the daily market index return data of 24 developed countries for 1994–2014. We have not re-run the sparse DCC simulations for the commodity data setup because the estimation is very slow, but we believe that the results would be very similar to those presented below.

As in Hafner and Franses (2009), we focus on the DCC model's correlation part only, ignoring the volatility part. So we set $D_t = I_n$ and only carried out step 2 of the DCC estimation. We generated data $r_t = \varepsilon_t$ for $t = 1, \ldots, 5000$ (and a burn-in sample of 1000 periods) according to

$$\varepsilon_t \sim N(0, R_t), \qquad R_t = (I_n \odot Q_t)^{-1/2} Q_t (I_n \odot Q_t)^{-1/2},$$
$$Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2 Q_{t-1},$$

with $n = 24$, $b^2 = 0.995$, $\overline{R}$ equal to the empirical daily return correlation matrix, and A chosen as follows. We drew the diagonal elements of A from the uniform distribution $U[.8c, 1.2c]$ with mean $c = .07$, set 20 randomly chosen off-diagonal elements of A equal to the values in the set $\pm c \cdot \{.01, .02, .1, .15, .2\}$ (each value being repeated twice), and set the other 532 off-diagonal elements of A equal to zero. We generated 20 simulated data sets in this way.

For each simulated data set, we estimated $\phi = (A, b)$ as outlined above, with the correlation matrix of the simulated data as an estimate of $\overline{R}$. To reduce the computation time, we fixed the tuning parameter λ_A at the 88th percentile of $|G_{\phi_{\text{pen}}}|$ (computed from the full simulated data set) instead of determining λ_A by cross-validation. Table 11 is a contingency table of the true and estimated off-diagonal elements of A, averaged across the simulations. As the table shows, the underlying sparsity structure is uncovered reasonably well, with two thirds of the "large" nonzero parameter values (those in $\pm c \cdot \{.1, .15, .2\}$) being detected and 95% of the zeros being estimated as zero. As expected, "small" nonzero parameter values (those in $\pm c \cdot \{.01, .02\}$) are much harder to detect: only 8% are estimated to be nonzero.

Table 11. Estimated versus true sparsity (DCC model, 20 replications)

	estimated zero	estimated nonzero	total
true zero	503.85	28.15	532
true "small"	7.35	0.65	8
true "large"	4.00	8.00	12
total	515.20	36.80	552

Model: $r_t = \varepsilon_t \sim N(0, R_t)$, $R_t = (I_n \odot Q_t)^{-1/2} Q_t (I_n \odot Q_t)^{-1/2}$, $Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2 Q_{t-1}$, with sparsity imposed on the off-diagonal elements of A. "large" refers to values in $\pm c \cdot \{.1, .15, .2\}$; "small" refers to values in $\pm c \cdot \{.01, .02\}$.

Table 12 gives details for each simulation separately. The first three columns pertain to the off-diagonal elements of A, giving the number of true zeros estimated as zero, and the number of

"small" and "large" values, respectively, estimated to be nonzero. The last four columns report the true and estimated values of the average of the diagonal elements of A and b^2. These estimates are very close to the true values, although $\widehat{b}^2$ tends to slightly underestimate b^2.

Table 12. Estimation results for each replication (DCC model, 20 replications)

zero	"small"	"large"	$\sum_i A_{ii}/n$	$\sum_i \widehat{A}_{ii}/n$	b^2	$\widehat{b}^2$
503	1	9	0.0697	0.0688	0.995	0.9945
498	1	7	0.0689	0.0666	0.995	0.9946
506	2	10	0.0714	0.0709	0.995	0.9948
507	0	5	0.0700	0.0718	0.995	0.9943
500	0	7	0.0719	0.0726	0.995	0.9943
508	0	9	0.0687	0.0679	0.995	0.9946
502	0	9	0.0660	0.0650	0.995	0.9946
502	2	9	0.0666	0.0641	0.995	0.9946
511	1	9	0.0683	0.0683	0.995	0.9946
508	1	7	0.0670	0.0674	0.995	0.9947
505	1	9	0.0688	0.0697	0.995	0.9944
502	1	7	0.0710	0.0691	0.995	0.9947
505	0	10	0.0687	0.0690	0.995	0.9946
517	0	9	0.0690	0.0695	0.995	0.9944
503	0	7	0.0685	0.0690	0.995	0.9944
484	0	5	0.0666	0.0648	0.995	0.9946
513	0	9	0.0697	0.0703	0.995	0.9943
504	1	8	0.0668	0.0665	0.995	0.9948
502	0	8	0.0674	0.0664	0.995	0.9946
497	2	7	0.0705	0.0688	0.995	0.9949

Model: $r_t = \varepsilon_t \sim N(0, R_t), R_t = (I_n \odot Q_t)^{-1/2} Q_t (I_n \odot Q_t)^{-1/2}, Q_t = \overline{R} - A\overline{R}A' - b^2\overline{R} + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + b^2 Q_{t-1}$, with sparsity imposed on the off-diagonal elements of A. The first three columns refer to the off-diagonal elements of A (532 zeros, 8 "small" values, and 12 "large" values) and report the numbers of zeros estimated as zero, "small" values estimated as nonzero, and "large" values estimated as nonzero, respectively.