Energy Economics 85 (2020) 104555

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

Volatility spillovers in commodity markets: A large *t*-vector autoregressive approach

Luca Barbaglia [a,b,*], Christophe Croux [c], Ines Wilms [d]

[a] European Commission, Joint Research Centre (JRC), Ispra, Italy
[b] Faculty of Economics and Business, KU Leuven, Belgium
[c] EDHEC Business School, Lille, France
[d] Department of Quantitative Economics, Maastricht University, The Netherlands

ARTICLE INFO

Article history:
Received 11 February 2019
Received in revised form 26 August 2019
Accepted 18 October 2019
Available online 2 November 2019

JEL classification:
C58
C32
Q02

Keywords:
Commodities
Forecasting
Lasso
Multivariate *t*-distribution
Vector autoregressive model
Volatility spillover

ABSTRACT

Prices of commodities have shown large fluctuations. A high volatility of one commodity today may impact the volatility of another commodity tomorrow. As such, agricultural and energy commodities are closely dependent due to the expansion of the biofuel industry. We study volatility spillovers among a large number of energy, agriculture and biofuel commodities using the vector auto regressive (VAR) model. To account for the possible fat-tailed distribution of the model errors, we propose the *t*-lasso method for obtaining a large VAR. The *t*-lasso is shown to have excellent properties, and a forecast analysis shows that the *t*-lasso attains better forecast accuracy than standard estimators. Our empirical analysis shows the existence of volatility spillovers between energy and biofuel, and between energy and agricultural commodities.

1. Introduction

Commodity markets are of great interest to financial market analysts and investors who uses them as alternative investment areas or for risk hedging purposes. In the last decade, commodity prices have experienced periods of high volatility and unexpected price swings have occurred increasingly more frequently. These events combined with the growing role of commodities in global asset markets (e.g., Kang and Yoon, 2009; Chkili et al., 2014; Hasanov et al., 2018) have led to a focus shift in the commodity literature from price levels to *volatility*. As volatility is a risk measure, financial analysts are interested in better modeling, understanding and forecasting the volatility of commodities. Portfolio optimization, option pricing, value at risk modeling, and dynamic hedging all require accurate volatility forecasts as key inputs (Hasanov et al., 2018).

In recent years, the high volume of financial transactions in commodity markets has led to a detachment from simple supply-demand dynamics. This makes it harder for analysts to describe swings in volatilities simply recurring to the economic fundamentals. In addition, commodity markets are volatile by their very nature: a policy change, a natural disaster or a technological breakthrough might cause an unexpected period of high volatility. The recent 2007/2008 crisis has further stressed the importance of understanding risk transmission mechanisms between different commodity markets or commodity types (see Serra and Zilberman, 2013, and references therein). Knowing whether the high volatility of one commodity today leads to high volatility of another commodity tomorrow can be key information to reduce risk exposure to future potential losses. We focus on these lagged effects among volatilities and refer to them as *volatility spillovers*.

Traditionally, the commodity volatility literature has primarily focused on non-renewable energy and agricultural commodities. High volatility in the energy markets unidirectionally affects the volatility of agricultural products via energy intensive agricultural inputs (e.g., fuel and fertilizers). We refer to Hasanov et al. (2018) and references therein. The emergence of biofuel production,

* Corresponding author at: European Commission, Joint Research Centre (JRC), Ispra, Italy.
E-mail address: luca.barbaglia@ec.europa.eu (L. Barbaglia).

https://doi.org/10.1016/j.eneco.2019.104555

however, has changed the link between energy and agricultural commodities, as certain crops represent simultaneously food, feedstock and fuel sources. In 2012, for instance, 42% of US total corn harvest was destined for ethanol production (Beckman et al., 2013). Bioethanol and biodiesel have become technological substitutes for traditional fuels such as diesel and gasoline (Chang and Su, 2010). At the same time, their production heavily depends on the supply of agriculture commodities (i.e. bioethanol and biodiesel have corn and soybeans as respective primary source). As a consequence, energy and agricultural commodities have shown a tighter and tighter relation since 2006, resulting in stronger volatility spillovers among the two commodity types (e.g., Nazlioglu et al., 2013). The financialization of commodities motivated the inclusion of biofuel feedstocks in the commodity volatility literature, but research on it remains scarce (Hasanov et al., 2018). In this paper, we take a joint modeling approach for volatilities among energy, biofuel and agricultural commodities, allowing to capture volatility spillovers among these different types of commodities.

We analyze volatility spillovers using the vector autoregressive (VAR) model for multiple time series. The logarithmic transformed volatilities of the different commodities, measured by realized variances or ranges, are the time series. This VAR model has been successfully applied in many empirical applications, but two main concerns remain: (i) the number of time series in the VAR model is limited since the number of parameters to be estimated increases quadratically with the number of time series included (Diebold and Yilmaz, 2015, p. 183) and, (ii) the standard estimation procedure of the VAR does not account for fat-tailed errors and thus for the frequent occurrence of extreme observations in the volatility series (e.g., Callot et al., 2017).

We introduce the *t-lasso* for estimating a large VAR. This approach allows for (i) a large number of time series relative to the time series length. To ensure that the estimation of this large VAR model is feasible, we use a penalized estimator of the VAR model in the spirit of the lasso (Tibshirani, 1996). The standard lasso estimator has become popular in energy economics and energy finance (e.g., Ziel and Weron (2018) for short-term electricity price forecasting, Miao et al. (2017) for oil price forecasting). (ii) Multivariate t-distributed errors where the degrees of freedom are estimated in contrast to, for instance, Finegold and Drton (2011) who take them to be fixed. Note that Schmidt and Makalic (2017) also use a t-lasso kind of approach but only for the univariate regression model. We employ the proposed t-lasso estimator of the large VAR to address the following research questions.

Q1: Are there volatility spillovers among energy, biofuel and agricultural commodities?
Q2: Are these spillovers varying over time?
Q3: Does accounting for fat tails improve forecast accuracy?

In our empirical analysis, we do find volatility spillovers between energy and biofuels. Among energy commodities, most of the volatility spillovers involve gasoline, which is often blended with ethanol for consumption. Furthermore, we observe volatility spillovers between energy and agriculture, regardless of the fact whether those crops can be used for biofuel production or not. We also find bidirectional volatility spillovers between biofuels and those agricultural commodities that can be used as inputs in biofuel production. These volatility spillovers are not observed in times of low energy prices. Finally, the results from our commodity analysis reveal fat-tailedness of the VAR errors with the estimated degrees of freedom of the t-distribution quite low. A forecast study shows that the t-lasso does produce more accurate volatility forecasts than other estimators.

The remainder of this article is structured as follows. Section 2 reviews the VAR model and introduces the t-lasso estimator. Section 3 presents the data, the definition of volatility spillovers and the network analysis tool to visualize them. Section 4 presents the results. A simulation experiment showing the good performance of the proposed t-lasso estimator is discussed in Section 5. Finally, Section 6 concludes.

2. Model and estimators

We model volatility spillovers between 10 commodities using the VAR model, following the approach of Diebold and Yilmaz (2009, 2012). Alternatively, multivariate GARCH or stochastic volatility models may be used, see for instance Serra (2011), Du et al. (2011), Arouri et al. (2011) and Karali and Ramirez (2014). These models are, however, difficult to estimate in higher dimensions.

Let $\boldsymbol{y}_t$ be a J-dimensional vector of log volatilities for $1 \leq t \leq T$, with T the time series length. We take the logarithmic transformation of the volatilities to ensure positivity of the volatility forecasts (e.g., Callot et al., 2017). We consider a stationary VAR of order P for the log volatilities

$$\boldsymbol{y}_t = \sum_{p=1}^{P} \boldsymbol{B}_p \boldsymbol{y}_{t-p} + \boldsymbol{e}_t, \tag{1}$$

where $\boldsymbol{B}_p$ is the $J \times J$ matrix of autoregressive coefficients for lags $1 \leq p \leq P$, and $\boldsymbol{e}_t$ is a J-dimensional vector of innovations (or error terms) with zero mean and covariance matrix $\boldsymbol{\Sigma}$. Without loss of generality, we assume that all log volatility time series are mean centered such that no intercept is included in (1).

Several recent papers propose methods to estimate large VAR models, that is models containing a large number of time series relative to the time series length (e.g., Davis et al., 2016; Gelper et al., 2016; Barbaglia et al., 2016; Derimer et al., 2018). The extension to large VAR models with fat-tailed errors has, however, not been addressed yet. It has been argued that log volatilities are asymptotically normal (Barndorff-Nielsen and Shephard, 2002), which is supported by some empirical evidence (e.g., Andersen et al., 2001). Nevertheless, subsequent studies emphasize that this might not hold for all volatility measures. For realized variances, Corsi et al. (2008) and Hassler et al. (2016) highlight that the log volatility of the S&P 500 index based on 5min intra-day returns exhibits deviations from the Gaussian distribution. For realized ranges, Caporin and Velo (2015) show that the log volatility has heavier tails than the normal. Qu et al. (2018) reveal the non-Gaussianity of the residuals in heterogeneous autoregressive models for the Australian New South Wales electricity prices.

For ease of notation, we rewrite model (1) in matrix form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}, \tag{2}$$

where $\boldsymbol{Y} = [\boldsymbol{y}'_{P+1}, \ldots, \boldsymbol{y}'_T]'$ is a $N \times J$ matrix and with $N = T - P$. Let $\boldsymbol{X} = [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_P]$ be the $N \times JP$ matrix of lagged time series, where $\boldsymbol{X}_p = [\boldsymbol{y}'_{P+1-p}, \ldots, \boldsymbol{y}_{T-p}c]'$ is a $N \times J$ matrix for $1 \leq p \leq P$. Finally, the $JP \times J$ matrix of autoregressive coefficients is $\boldsymbol{B} = [\boldsymbol{B}'_1, \ldots, \boldsymbol{B}'_P]'$ and $\boldsymbol{E}$ is the $N \times J$ innovation matrix.

In the next subsections we first review the lasso for the VAR model, then we introduce the t-lasso. To this end, we build on the t-lasso estimator for sparse graphical models proposed by Finegold and Drton (2011), but we extend their approach by (i) turning their static framework into a time series model and (ii) allowing for an unknown degrees of freedom of the t-distribution.

2.1. Gaussian innovations

Assume that the innovations $\boldsymbol{e}_t$ follow a multivariate normal distribution $N(\boldsymbol{0}, \boldsymbol{\Sigma})$. The lasso equals the following penalized least squares estimator of $\boldsymbol{B}$

$$\hat{\boldsymbol{B}} = \underset{\boldsymbol{B}}{arg\,min} \frac{1}{2N} \text{tr}\left[(\boldsymbol{Y} - \boldsymbol{XB})(\boldsymbol{Y} - \boldsymbol{XB})'\right] + \lambda \sum_{i,j=1}^{J} \sum_{p=1}^{P} |B_{p,ij}|,$$

where $\text{tr}(\cdot)$ is the trace operator, and $\lambda > 0$ is a regularization parameter (Hastie et al., 2015, p. 8). The ijth entry of the matrix $\boldsymbol{B}_p$ is denoted by $[\boldsymbol{B}_p]_{ij} = B_{p,ij}$. This penalty ensures that the estimation of $\boldsymbol{B}$ is feasible even if the number of parameters exceeds the time series length. Moreover, it sets some elements of $\hat{\boldsymbol{B}}$ exactly equal to zero. The larger the regularization parameter λ, the sparser $\hat{\boldsymbol{B}}$, that is the more of its elements are exactly zero.

In line with Rothman et al. (2010), we explicitly account for correlated innovations by including the estimation of the inverse innovation covariance matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$. To this end, we turn to the maximum likelihood framework and jointly estimate $\boldsymbol{B}$ and $\boldsymbol{\Omega}$ by minimizing the negative log likelihood:

$$(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\Omega}}) = \underset{\boldsymbol{B},\boldsymbol{\Omega}}{arg\,min} \frac{1}{2N} \text{tr}\left[(\boldsymbol{Y} - \boldsymbol{XB})\boldsymbol{\Omega}(\boldsymbol{Y} - \boldsymbol{XB})'\right] - \frac{1}{2}\log|\boldsymbol{\Omega}| + \lambda \sum_{i,j=1}^{J} \sum_{p=1}^{P} |B_{p,ij}| + \gamma \sum_{i \neq j}^{J} |\omega_{ij}|, \tag{3}$$

where ω_{ij} is the ijth element of $\boldsymbol{\Omega}$, $|\boldsymbol{\Omega}|$ is the determinant of $\boldsymbol{\Omega}$, and $\gamma > 0$ is another regularization parameter. This penalty ensures that the estimation of $\boldsymbol{\Omega}$ is feasible even if the number of parameters exceeds the time series length. Furthermore, it sets some elements of $\hat{\boldsymbol{\Omega}}$ to zero. The larger the regularization parameter γ, the sparser $\hat{\boldsymbol{\Omega}}$. We refer to the estimator in (3) as the *Gaussian lasso*.

2.2. Student t-innovations

We depart from the normality assumption and assume that the innovations are distributed according to a multivariate t-distribution $t_\nu(\boldsymbol{0}, \boldsymbol{\Psi})$, where $\nu > 0$ are the degrees of freedom and $\boldsymbol{\Psi}$ is the scale matrix. If the covariance matrix of the innovations exists, for $\nu > 2$, $\boldsymbol{\Sigma} = \boldsymbol{\Psi}\nu/(\nu - 2)$. The associated density function is given by

$$\frac{\Gamma((\nu + J)/2)|\boldsymbol{\Omega}|^{1/2}}{(\pi\nu)^{J/2}\Gamma(\nu/2)[1 + \boldsymbol{e}_t'\boldsymbol{\Omega}\boldsymbol{e}_t/\nu]^{(\nu+J)/2}}, \tag{4}$$

with $\boldsymbol{\Omega} = \boldsymbol{\Psi}^{-1}$ (Kotz and Nadarajah, 2004, p. 1). Recall that J is the dimension of the time series.

The joint estimator of $\boldsymbol{B}$ and $\boldsymbol{\Omega}$ is defined as in (3), replacing the Gaussian by the t-density given in (4), and keeping the penalty terms. We call this estimator the *t-lasso*. An outline of the algorithm for computing the t-lasso can be found in Appendix A.

3. Data and volatility spillover analysis

In this section, we first present the data. Then, we define volatility spillovers based on the forecast error variance decomposition. Finally, we present a network analysis tool visualizing these volatility spillovers.

3.1. Data

We study agricultural (corn, wheat, soybean, sugar, cotton, coffee), energy (crude oil, gasoline, natural gas) and biofuel (ethanol) commodities, so $J = 10$. We use daily information about the opening, highest and lowest prices of the nearest to maturity contracts traded in the corresponding future markets. Data are available on Thomson Reuters Eikon. The time span ranges from January 3rd, 2012 to October 28th, 2016, thus $T = 1070$ daily observations. Table 1 reports the label and the description of each variable.

Table 1
Variable description as in Thomson Reuters Eikon for Future Continuation 1 commodities, nearest to maturity contracts.

Label	Commodity	Description
CRUO	Crude oil	NYMEX Light Sweet Crude Oil (WTI) Composite Energy
GASO	Gasoline	RBc1 NYMEX RBOB Gasoline Composite Energy
NATG	Natural gas	NGc1 NYMEX Henry Hub Natural Gas Composite Energy
ETHA	Ethanol	CBoT Denatured Fuel Ethanol Electronic Energy
CORN	Corn	Cc1 CBoT Corn Composite Commodity
WHEA	Wheat	Wc1 CBoT Wheat Composite Commodity
SOYB	Soybeans	Sc1 CBoT Soybeans Composite Commodity
SUGA	Sugar	ICE-US Sugar No. 11 Futures Electronic Commodity
COTT	Cotton	ICE-US Cotton No. 2 Futures Electronic Commodity
COFF	Coffee	KCc1 ICE-US Coffee C Futures Electronic Commodity

We obtain volatility measures using the realized daily range proposed by Parkinson (1980). Consider the highest price $H_{t,j}$ and the lowest price $L_{t,j}$ attained on date $1 \leq t \leq T$ for commodity $1 \leq j \leq J$. Then the realized range, as an estimator of the daily realized variance, is given by

$$v_{t,j} = \frac{1}{4\log(2)}(\log(H_{t,j}) - \log(L_{t,j}))^2.$$

The time series entering the VAR of Eq. (1) are the log transformations of the realized daily ranges, that is $\mathbf{y}_t = [\log(v_{t,j}), \ldots, \log(v_{t,J})]'$. For more on realized range measures of volatility, see Shu and Zhang (2006) or Martens and van Dick (2007). Other measures of daily volatility, like the realized variance (e.g., McAleer and Medeiros, 2008), can be taken as well, but they require more detailed intra-day data. For commodity futures, however, data on the opening, lowest and highest daily prices are easier to obtain than intra-day high-frequency returns, motivating our choice for daily realized ranges. We check for stationarity of the log volatilities with univariate unit root tests and the pooled unit root test of Levin et al. (2002) and find strong evidence in favor of stationarity (p-values < 0.01).

3.2. Spillover indices

We compute volatility spillovers by means of the generalized forecast error variance decomposition. We take the definitions from Diebold and Yilmaz (2015). Consider the vector moving average (VMA) representation of the vector autoregressive model:

$$\boldsymbol{y}_t = \sum_{p=0}^{\infty} \boldsymbol{\theta}_p \boldsymbol{e}_{t-p},$$

where $\boldsymbol{\theta}_p$ is the moving average coefficient matrix at lag p (cfr. Wold's representation theorem, Lütkepohl, 2005, p. 25). The h-step-ahead forecast error for the jth component of $\boldsymbol{y}_t$ is

$$\hat{y}_{t+h,j} - y_{t+h,j} = \sum_{p=0}^{h-1} \theta_{p,j1} e_{t+h-p,1} + \cdots + \theta_{p,jJ} e_{t+h-p,J}, \tag{5}$$

where $e_{t,j}$ is the jth component of $\boldsymbol{e}_t$, and $\theta_{p,jk}$ is the jkth entry of $\boldsymbol{\theta}_p$.

If an impulse to $e_{t,k}$ of size of one standard deviation is given, then the expected value of the vector of innovations equals

$$E(\boldsymbol{e}_t | e_{t,k} = \sqrt{\sigma_{kk}}) = \frac{\boldsymbol{\Sigma}\boldsymbol{\delta}_k}{\sqrt{\sigma_{kk}}}, \tag{6}$$

where σ_{kk} is the kkth entry of $\boldsymbol{\Sigma}$ and $\boldsymbol{\delta}_k$ is the selection vector of length J with unity entry as its kth element and zeros elsewhere. Eq. (6) defines a generalized impulse, as in Pesaran and Shin (1998), with response vector $\boldsymbol{\theta}_p \boldsymbol{\Sigma} \boldsymbol{\delta}_k / \sqrt{\sigma_{kk}}$, for $p = 1, \ldots, P$. Note that generalized impulses, in contrast to the orthogonalized impulses, are invariant with respect to the ordering of the time series. It is important to note that (6) does not only hold for a normal distribution, but also for a t-distribution (e.g., Ding, 2016). If $\nu \leq 2$ and the covariance matrix of the t-distribution is not existing, the $\boldsymbol{\Sigma}$ in (6) should be replaced by the $\boldsymbol{\Psi}$ scale matrix. The corresponding generalized forecasts error variance decomposition is then obtained as in Lanne and Nyberg (2016). The proportion of the variance of the h-step-ahead forecast error (5) accounted for by the innovation in component k of the multivariate time series, given by

$$w_{h,jk} = \frac{\sigma_{kk}^{-1} \sum_{p=0}^{h-1} (\boldsymbol{\delta}_j' \boldsymbol{\theta}_p \boldsymbol{\Sigma} \boldsymbol{\delta}_k)^2}{\sum_{p=0}^{h-1} \boldsymbol{\delta}_j' \boldsymbol{\theta}_p \boldsymbol{\Sigma} \boldsymbol{\theta}_p' \boldsymbol{\delta}_j}.$$

The *volatility spillover* from commodity k to commodity j at horizon h is defined as

$$s_{h,k \to j} = 100 * \frac{w_{h,jk}}{\sum_{k=1}^{J} w_{h,jk}}. \tag{7}$$

Following Diebold and Yilmaz (2015), the *volatility spillover index*

$$s_h = \sum_{\substack{j \neq k \\ j,k=1}}^{J} s_{h,k \to j}, \tag{8}$$

is a single measure of the overall volatility spillovers at horizon h. In the remainder, we consider several horizons, namely $h = 1$ (one day), $h = 5$ (one week), $h = 20$ (one month) and $h = 60$ (three months). These choices aim to capture the short term ($h = 1, 5$) and long term ($h = 20, 60$) behavior of different types of financial investors: for instance, dealers would trade at higher frequencies, and are hence expected to be more responsive to short term dynamics and forecasts, while institutional investors trade less frequently and are therefore expected to be more interested in long term dynamics and forecasts (Corsi, 2009).

3.3. Network analysis

We visualize the volatility spillovers using a network analysis tool (e.g., Hautsch et al., 2015). The nodes in the network are the different commodities. An edge from commodity k to commodity j is drawn if $s_{h,k \to j}$ from equation (7) is non-zero; the edge width represents the magnitude of the volatility spillover. Hence, our network is directed and weighted, however, not necessarily sparse. Indeed, the sparsity of $\hat{\boldsymbol{B}}$ is not necessarily preserved in the estimated VMA coefficient matrices $\hat{\boldsymbol{\theta}}$. Consequently, the network will picture many non-zero volatility spillovers and hence contain many edges, although many of them may be quite small. Therefore, for better visualization, we give the network representing only the 15% largest volatility spillovers.

4. Results

We employ a rolling window over the period 2012-2016 with window width $W = 220$ days. At each time point $t = W, \ldots, T$, we estimate a VAR(P) model using the t-lasso from the log-volatilities in the time window $t - W + 1$ until t. The order P of the VAR is selected using the Bayesian Information Criterion. We first present the time evolution of the estimated degrees-of-freedom and of the volatility spillovers (Section 4.1). Then, we picture the networks representing the volatility spillovers (Section 4.2). Finally, we show the good performance of the t-lasso in terms of forecast accuracy (Section 4.3).

The results discussed in Sections 4.1 and 4.2 are based on a volatility spillover analysis with horizon $h = 5$, a week, thereby following Bubák et al. (2011). Our main findings are stable across the different forecast horizons and, hence, omitted but available from the authors upon request. Furthermore, we redid the analysis using impulses resulting from Cholesky's decomposition (Diebold and Yilmaz, 2015, p. 14) or the spectral decomposition (Hafner and Herwartz, 2006) and obtained only minor changes in the magnitudes of the volatility spillovers.

4.1. Rolling window analysis

Estimated degrees of freedom. Fig. 1 reports the estimated degrees of freedom of the multivariate t-distribution of the VAR innovations at each time point t, with t the end point of a 220-day rolling time window. The average value of $\hat{\nu}$ is 1.57, with maximum 1.80 and minimum 1.43. Overall, we observe that the estimated degrees of freedom are very low. This confirms the existence of heavy tails and justifies the use of the t-lasso rather than the Gaussian lasso.

If we look at the evolution of $\hat{\nu}$ over time, we detect lower values of the estimated degrees of freedom in the time windows ending between the second half of 2014 and the first months of 2015. Volatile commodity markets characterized the second half of 2014, mainly driven by the drop of almost 50% in crude oil price (Knittel and Pindyck, 2016). Smaller values of $\hat{\nu}$ indicate more pronounced extreme realizations in the VAR innovations.

Volatility spillover index. Fig. 2 reports the evolution of the volatility spillover index (cfr. Eq. (8)) for the t-lasso as a function of time t, being t the end point of each time window. We observe that the volatility spillover index experiences large swings over time. In particular, in the second half of 2014 we detect a large drop in the overall level of volatility spillovers. This drop is not permanent and during 2015 the volatility spillover index returns to the level prior to 2014.

The downturn in volatility spillover index can also be linked to the fall of crude oil price that occurred in the second half of 2014. Crude oil price experienced large downturns driven by (i) a larger supply from OPEC and non-OPEC countries and (ii) a weak global demand due to the slowdown of the world economy, notably the Chinese one. The drop in energy prices reduced the probability of observing a substitution effect between standard fuels and biofuels at price level (Chang and Su, 2010; Hassler and Sinn, 2016), and, at the same moment, we observe lesser and weaker volatility connectedness among energy, biofuel and agriculture commodities.

4.2. Network analysis

We visualize the volatility spillovers in a network of commodities. We take three networks to represent spillovers (i) at the end of the sample, and (ii) during periods of low connectedness, and (iii) during periods of high connectedness. Fig. 3 top presents the network for the time window ending on October 28th, 2016, the last time point of the sample. For instance, a directed edge is drawn from gasoline (GASO) to crude oil (CRUO) since the volatility spillover from gasoline towards crude oil is different from zero and belongs to the largest 15% volatility spillovers. Recall that the edge width represents the magnitude of the volatility spillover. Fig. 3 also reports the networks for the time windows ending on February 5th, 2015 (bottom left), where the value of the volatility spillover index was lowest, and ending on June 17th, 2014 (bottom right), with highest value of the volatility spillover index. For simplicity, from now on we refer to the three networks only by the year of the ending date of their windows.

Link energy-biofuel. In all networks we find volatility spillovers among energy and biofuel commodities. In the 2015 and 2014 networks, gasoline is the only energy commodity connected by an edge

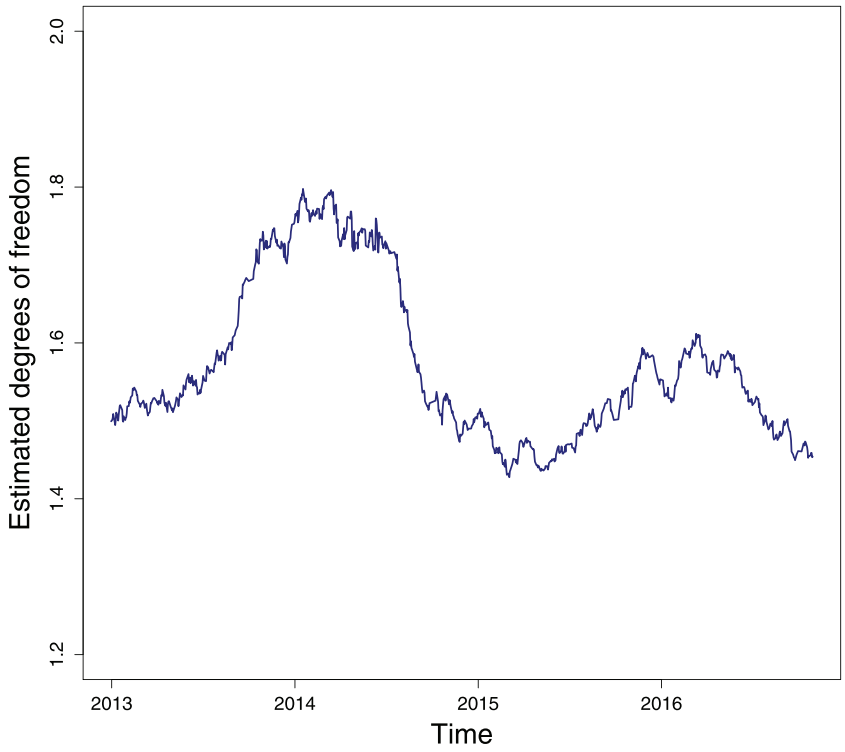

Fig. 1. Estimated degrees of freedom of the *t*-distribution. The horizontal axis represents the ending date of a 220-day time window.

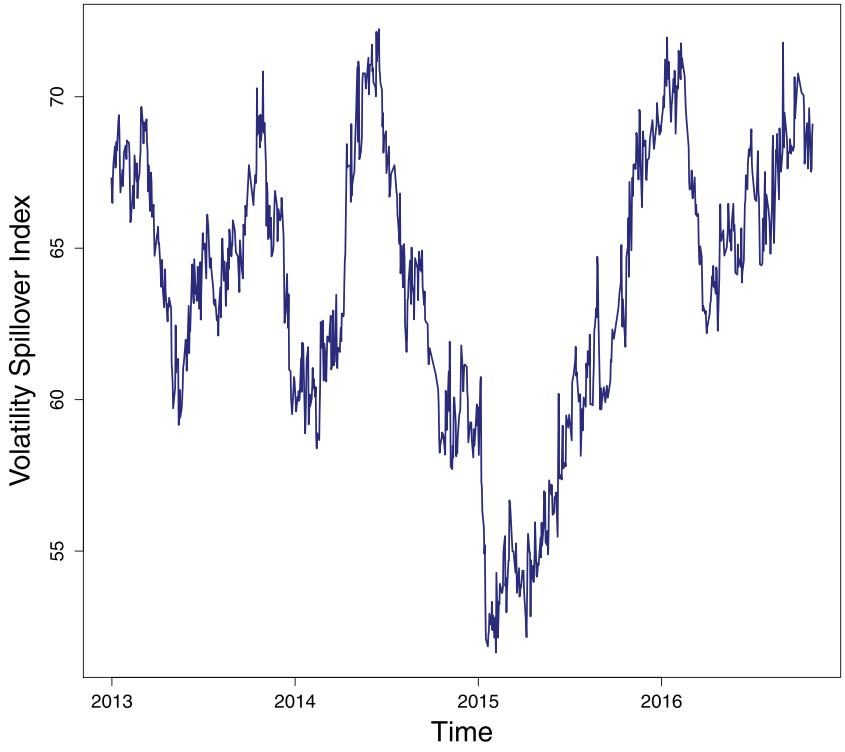

Fig. 2. Volatility spillover index. The horizontal axis represents the ending date of a 220-day time window.

with ethanol. As ethanol is often blended in gasoline for consumption (Serra and Zilberman, 2013), it is not surprising to observe volatility spillovers between the two commodities in times of high energy prices (i.e. network 2014) and in times of large energy price changes (i.e. network 2015).

As pointed out by a referee, there is a mandatory percentage blending of ethanol in gasoline. For instance, the US Energy Information Administration categorizes gasoline as E10, E15 and E85, when blended with 10%, 15% and 85% ethanol, respectively, and similar blending decisions have been taken by other countries. So the percentage of blending is not fully free. At price level, the presence of a substitution effect between standard fuels and biofuels has been investigated and detected (e.g., Serra et al., 2010; Hassler and Sinn, 2016; Chang and Su, 2010). This substitution effect is not

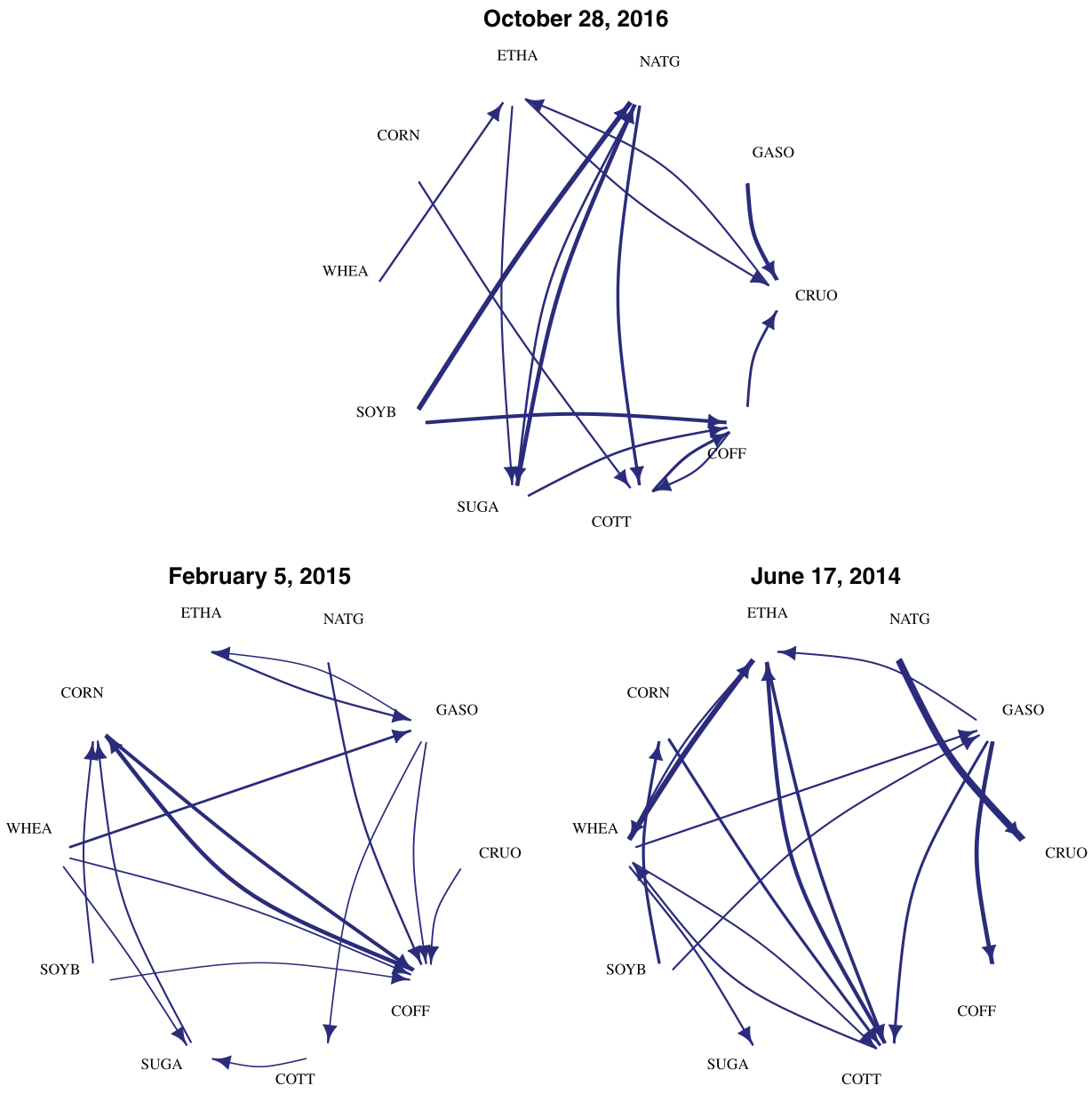

Fig. 3. Commodity networks of volatility spillovers for different ending times of the 220-day rolling windows: October 28th, 2016 (top; end of sample), February 5th, 2015 (bottom left; low connectedness) and June 17*th* 2014 (bottom right; high connectedness).

so much due to a blending decision of the final customer, but rather due to the adoption by a government of higher blending standards, the introduction of new engines running on higher biofuel percentage blends and the higher usage of biofuels for industrial purposes. Note that our paper in on the volatility of the prices, not the prices as such.

Link energy-agriculture. Bidirectional volatility spillovers between energy and agriculture are detected in all networks, confirming the findings of Rezitis (2015). In the 2016 network, natural gas shows numerous and important spillovers from and towards agricultural commodities, whereas in the 2015 and 2014 networks, gasoline is the most connected energy commodity with agriculture. Both natural gas and gasoline are key commodities in the agricultural sector: the former is the major production cost for fertilizers, whereas the latter is the most important fuel (together with diesel). If energy prices are high (e.g., network 2014), fuel has a large impact on agricultural commodities, which could explain the numerous volatility spillovers from/to gasoline. Conversely, as energy prices are low and fuel is cheap (e.g., network 2016), variations in fertilizer prices largely affect agricultural commodities volatilities, which could explain the many volatility spillovers from/to natural gas.

In all networks, various agricultural commodities are involved in volatility spillovers to and from energy. In general, volatility spillovers from and to energy involve both agricultural crops that are primary inputs for biofuel production – like wheat, soybean and sugar, and crops that do not have a direct link with biofuels – like coffee, as in Rezitis (2015).

Link biofuel-agriculture. Volatility spillovers between biofuel and agriculture are found only in the 2016 and 2014 networks. These volatility spillovers are bidirectional and involve wheat, sugar and cotton: this is consistent with our expectation since these crops can be used for ethanol production. In the 2015 network, we observe no volatility spillovers between biofuel and agriculture. In that period energy prices substantially dropped (Knittel and Pindyck, 2016), making biofuel less attractive compared to other standard fuels and

this could have resulted in weaker volatility spillovers from and to agriculture.

We would like to stress that the presence of a link in the network depends on the set of commodities considered. We replicated the analysis on the smaller set of the three commodities gasoline, ethanol and corn, and obtained similar patterns as in the 10-commodity network. In this smaller network, the expected links between corn and ethanol and between ethanol and gasoline became clearly visible. The links are different in the larger network because (i) we visualize only the 15% largest volatility spillovers (ii) a volatility spillover is defined as the proportion of volatility forecast error variance that may be attributed to a shock in the volatility of another commodity; the magnitude of this proportion depends on the set of commodities considered.

4.3. Forecast accuracy

For a given forecast horizon h, we forecast the vector of log volatilities $\hat{\boldsymbol{y}}_{t+h}^{(t)} = [\hat{y}_{t+h,1}^{(t)}, \ldots, \hat{y}_{t+h,J}^{(t)}]$ from the data in the time window $[t-W+1, t]$. The forecast is obtained from the VAR model, with autoregressive coefficients $\hat{\boldsymbol{B}}$ estimated by the t-lasso, the Gaussian lasso and the conventional Least Squares (LS) estimator. We compare these three estimators in terms of forecast accuracy by computing the mean absolute forecast error

$$\text{MAFE} = \frac{1}{N-h-W+1}\sum_{t=W}^{N-h}\text{MAFE}_t \quad \text{with} \quad \text{MAFE}_t$$

$$= \frac{1}{J}\sum_{j=1}^{J}|\hat{y}_{t+h,j}^{(t)} - y_{t+h,j}|,$$

where N is the sample size, J the number of time series, W the window width and h the forecast horizon. Hence, MAFE_t is the average of the absolute forecast errors of the J series, computed for each time window. The MAFE then averages over the time windows. The smaller the value of the mean absolute forecast error, the more accurate the log-volatility forecasts.

Table 2
Mean absolute forecast error (left) and mean squared forecast error (right) for the t-lasso, the Gaussian lasso and the LS for forecast horizons $h = 1, 5, 20$ and 60 from a 220-day rolling window.

	t-Lasso	Gaussian lasso	LS
Mean absolute forecast error			
$h=1$	1.871	1.896	2.418
$h=5$	1.959	2.056	2.589
$h=20$	2.280	2.636	2.820
$h=60$	2.626	2.865	2.871
Mean squared forecast error			
$h=1$	6.527	6.763	10.253
$h=5$	7.291	8.117	11.470
$h=20$	9.579	12.043	13.215
$h=60$	11.979	13.654	13.696

Results. Table 2 reports the MAFE for the three estimators for different forecast horizons, namely $h=1$, $h=5$, $h=20$ and $h=60$, that is one day, one week, four weeks and three months, respectively. The t-lasso attains the best value of the MAFE for all forecast horizons. Using the Diebold–Mariano test (Diebold and Mariano, 1995), we find the difference in forecast accuracy with the other estimators to be significant (p-values < 0.01) in all cases. For forecast horizon $h=1$, the t-lasso gives a relative improvement in MAFE of 29% over the LS, while there is only a small gain over the Gaussian lasso (1.3%). The relative advantage of the t-lasso over the Gaussian lasso grows for forecast horizons $h=5$ (5%) and $h=20$ (16%). For long horizons (i.e. $h=60$), the t-lasso still performs best: at 3 months horizon, it outperforms both the LS and the Gaussian lasso. Overall, for short and long horizons the t-lasso attains significant and relevant gains in forecast accuracy. The superior forecast performance of the t-lasso is confirmed when the Mean Squared Forecast Error (MSFE) is used instead of the MAFE, see Table 2. The MSFE for the t-lasso is significantly lower than for its competitors (p-values < 0.01 in all cases).

Fig. 4 reports the evolution of the MAFE_t (for $h=20$) for the t-lasso (blue solid line), the Gaussian lasso (red dashed line) and Least Squares (gray dotted line) as a function of t, the end point of each

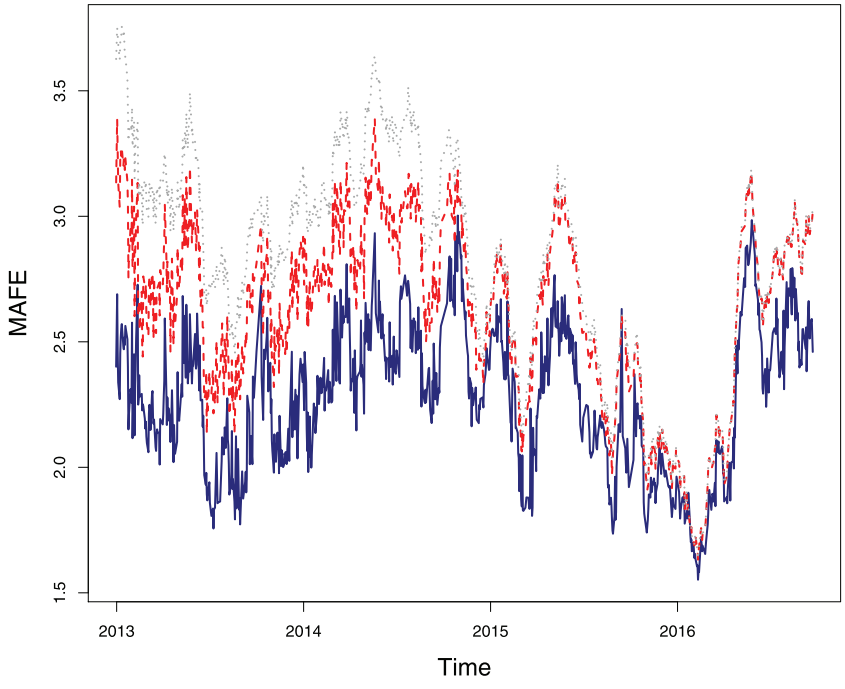

Fig. 4. Mean absolute forecast error for forecast horizon $h=20$ for the t-lasso (solid, blue), the Gaussian lasso (dashed, red) and the Least Squares (dotted, gray). The horizontal axis represents the ending date of a 220-day time window.

Table 3
Mean absolute estimation error for the four estimators and different degrees of freedom ν.

ν	LS	Gaussian lasso	t-Lasso	
			ν fixed	ν estimated
1	0.638	0.187	0.088	0.089
2	0.135	0.096	0.088	0.088
3	0.107	0.090	0.088	0.089
5	0.098	0.090	0.088	0.089
10	0.095	0.091	0.089	0.090
∞	0.093	0.091	0.091	0.091

time window. The t-lasso attains a lower mean absolute forecast error than the Gaussian lasso in all but five time windows and the difference in forecast accuracy is confirmed to be significant in all these time windows by the Diebold–Mariano test.

The difference between the three estimators reduces in 2015, when both the MAFE_t and the volatility spillover index (Fig. 2) attain the lowest values. In times of high incertitude – i.e. high MAFE_t and high volatility spillover index – the t-lasso attains a better forecast performance than the Gaussian lasso and the LS.

5. Simulations

We study the performance of the *t-lasso* by means of a simulation experiment, comparing it with the Gaussian lasso and the least squares estimator. The t-lasso we propose estimates the degrees of freedom ν of the t-distribution. In this simulation experiment we can compare it with a t-lasso where ν is fixed at its true value, and does not need to be estimated along.

Data generating process. We simulate from a VAR of order $P=2$ with $J=10$ time series. The dimensions of the VAR are in line with the empirical study made in Section 3. The data generating process is:

$$\boldsymbol{y}_t = \boldsymbol{B}_1\boldsymbol{y}_{t-1} + \boldsymbol{B}_2\boldsymbol{y}_{t-2} + \boldsymbol{e}_t,$$

for $P+1 \leq t \leq T=100$. The autoregressive coefficient matrices are sparse: B_1 and B_2 have their non-zero elements (0.4 for B_1 and 0.2 for B_2) on the main diagonal and on the first row. The error terms $\boldsymbol{e}_t$ follow a multivariate Student $t_\nu(\boldsymbol{0}, \boldsymbol{\Psi})$, with $\nu \in \{1, 2, 3, 5, 10, \infty\}$. In the special case $\nu=1$, the distribution is a multivariate Cauchy distribution, whereas for $\nu \to \infty$ the distribution is a multivariate normal. The ijth entry of $\boldsymbol{\Psi}$ is $\psi_{i,j} = 0.1^{|i-j|}$, such that inverse error covariance matrix $\boldsymbol{\Omega}$ is a band matrix. We take $S=1000$ simulations runs.

Performance measure. The different estimators are compared in terms of their estimation accuracy. To evaluate the estimation accuracy, we use the mean absolute estimator error

$$\text{MAEE}(\boldsymbol{B}, \hat{\boldsymbol{B}}) = \frac{1}{S}\frac{1}{PJ^2}\sum_{S=1}^{S}\sum_{i,j=1}^{J}\sum_{p=1}^{P}|\hat{B}_{p,ij}^{(s)} - B_{p,ij}|,$$

where $\hat{B}_{p,ij}^{(s)}$ is the ij^{th} entry of the estimate $\hat{\boldsymbol{B}}_p^{(s)}$ in simulation run s.

Simulation results. Table 3 reports the MAEEs for different values of the true degrees of freedom ν. The t-lasso with ν fixed at the true value always achieves the best MAEE. It is very closely followed by the t-lasso with ν estimated. There is no considerable loss in estimation accuracy when estimating ν. Both t-lasso estimators perform significantly better than the Gaussian lasso (the difference in estimation accuracy is tested with a paired t-test, all p-values < 0.01). For instance, the improvement in estimation accuracy is of 53% for $\nu=1$ and of 9% for $\nu=2$. The margin by which the t-lasso estimators outperform the Gaussian lasso decreases for larger degrees of freedom. In particular, for $\nu=\infty$ there is no difference between the t-lasso estimators and the Gaussian lasso. Indeed, as $\nu \to \infty$, a multivariate t-distribution converges to a multivariate normal. The LS suffers from the large number of parameters to estimate, given the length of the time series. The LS is not competitive with the lasso type estimators when estimating large VAR models.

Finally, Fig. 5 shows the frequencies of the estimated degrees of freedom by the t-lasso, for the different settings $\nu \in \{1, 2, 3, 10\}$. The estimated degrees of freedom are closely centered around the true value (vertical red line). The average (averaged over all simulation runs) estimated degrees of freedom are 0.94 (for $\nu=1$), 2.10 (for $\nu=2$), 3.16 (for $\nu=3$) and 11.14 ($\nu=10$). The variance of the estimated degrees of freedom increases for larger values of ν. By repeating this simulation experiment for larger values of the length of the time series T, one sees that the bias of the estimator of ν shrinks to zero, as well as its variance, indicating consistency.

6. Conclusions and policy implications

We first discuss our main findings related to the research questions stated in Section 1.

Q1: Volatility spillovers between energy, biofuel and agricultural commodities are detected and visualized using networks. In particular, we find evidence of bidirectional volatility spillovers between energy and agricultural commodities, regardless of the fact of being biofuel crops or not, as in (Rezitis, 2015).
Q2: Using a rolling window approach, we could see that volatility spillovers experience large swings over time. We observe that volatility spillovers become weaker when energy prices drop. This is in line with Hassler and Sinn (2016) and Chang and Su (2010), who observe weaker substitution among standard fuels and biofuels if the energy prices are low.
Q3: Due to the spikes typical of commodity volatility, the distribution of the innovations of the VAR model is fat-tailed. We capture this fat-tailedness by proposing the t-lasso, suited for estimating high-dimensional VAR models. The t-lasso procedure includes estimation of the degrees of freedom and was not considered before in the literature, up to our knowledge. The t-lasso is an extension of the (Gaussian) lasso, since it coincides with the latter if the estimated degrees of freedom tend to infinity. We do find an improvement in forecast accuracy (see Section 4.3) and precision (see Section 5) using it.

Further research could extend the proposed methodology to allow for asymmetric distributions of the VAR error terms using for instance a multivariate skewed t-distribution (Fernández and Steel, 1998), a skewed generalized-t (Theodossiou, 1998) or a generalized hyperbolic density (Eberlein and Keller, 1995). Note that in our empirical application, graphical displays did not reveal any strong deviation from symmetry of the residuals.

The proposed t-lasso is useful for other studies by practitioners, decision-makers and investors. Among other ones, the analysis of systemic risk, that is the risk that an individual firm default might threaten the stability of an entire market (Nucera et al., 2016), seems a natural application. After the 2007–2009, great importance has been given to the modeling of the spillovers among firms as a measure of systemic risk: typically, one would study a large number of firms (Hautsch et al., 2015), and model the non-normal innovation distribution caused by extreme downturns in the equity market (Engle et al., 2015). The proposed t-VAR approach would clearly address both issues, and its results would be highly interpretable in the form of risk networks.

Our results help policy makers and investors in quantifying and describing the nature of the volatility spillovers among energy, biofuel and agricultural commodities. Policy makers could benefit from our analysis to quantify the potential impact of a proposed

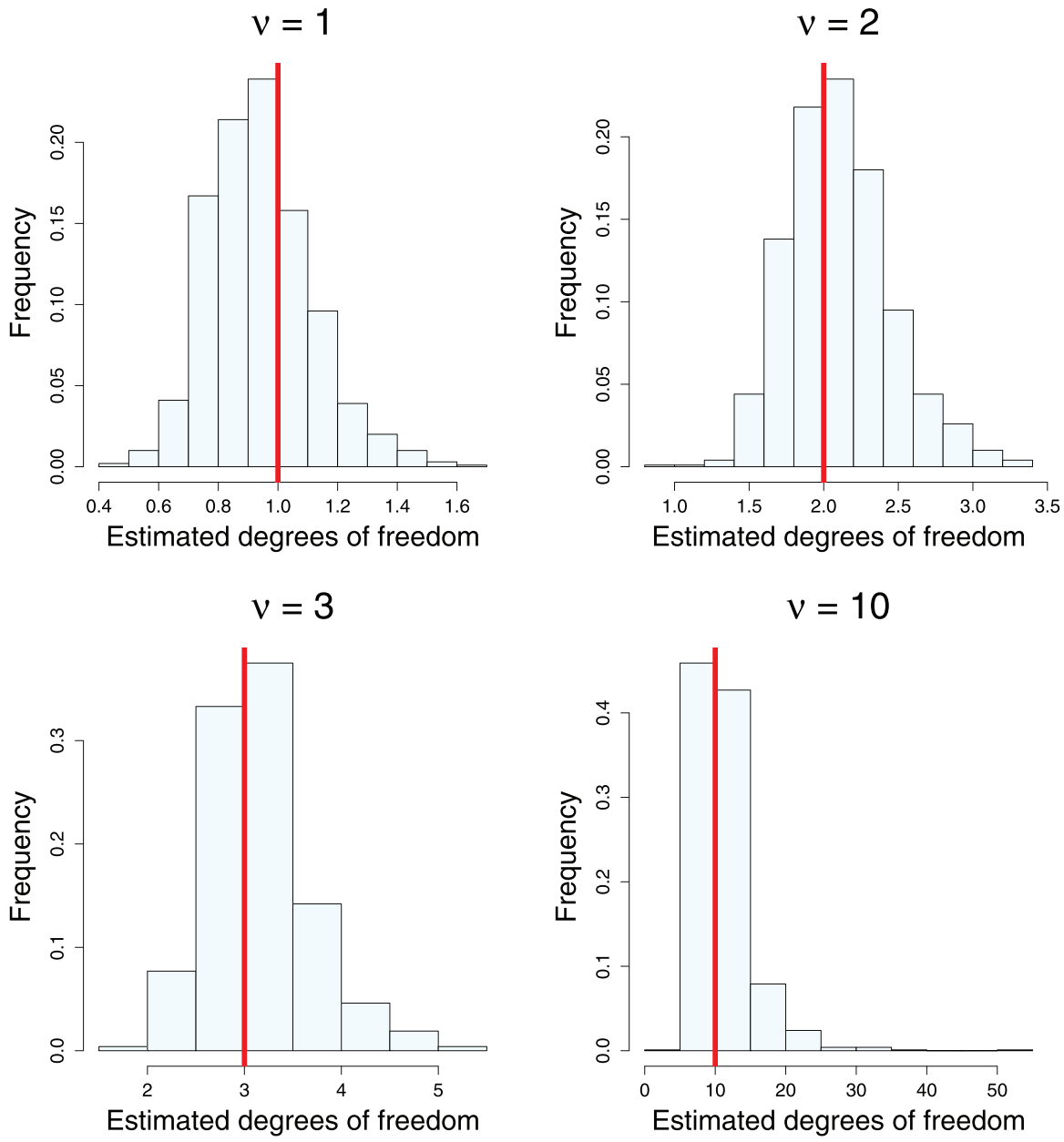

Fig. 5. Frequency of the estimated degrees of freedom, for setting $\nu = 1, 2, 3, 10$ in the simulation study. The true value of the degrees of freedom is indicated by the vertical red line (dark gray on a gray scale).

policy: for instance, a decision that would cause an increase in the energy market volatility would have strong spillovers to the biofuel and agricultural sectors in times of high energy prices, while it would generate weaker volatility spillovers in times of low energy prices. The good forecasting performance of the t-lasso is helpful for getting more precise volatility forecasts for short term financial investors and for long term institutional investors.

Acknowledgements

We thank the reviewers for their thorough review and highly appreciate their comments and suggestions which substantially improved the quality of the manuscript. We thank Piet Sercu and Kris Boudt for the useful comments that greatly improved the manuscript. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 832671.

Appendix A. Algorithm for the t-lasso

Consider the VAR model (2), with the t-distributed innovations. We first assume that the degrees of freedom ν is known and discuss the EM (Expectation Maximization) algorithm, following Finegold and Drton (2011). However, in practice ν is not known and needs to be estimated. To this end, we use the ECM (Expectation Conditional Maximization) algorithm as in Liu and Rubin (1995).

The Student t-distributed random vector $\boldsymbol{e}_t$ can be written as scale mixture of a multivariate normal $N(\boldsymbol{0}, \boldsymbol{\Psi})$ and a random variable τ_t distributed as a Gamma $\Gamma(\nu/2, \nu/2)$. One has

$$\boldsymbol{e}_t | \tau_t \sim N\left(\boldsymbol{0}, \frac{\boldsymbol{\Psi}}{\tau_t}\right) \tag{9}$$

and

$$\tau_t|\boldsymbol{e}_t \sim \Gamma\left(\frac{\nu+J}{2}, \frac{\nu+\boldsymbol{e}_t'\boldsymbol{\Omega}\boldsymbol{e}_t}{2}\right)$$

(e.g., Kotz and Nadarajah, 2004, p. 2). The expected value of the latter Gamma distribution equals

$$\hat{\tau}_t = E[\tau_t|\boldsymbol{e}_t] = \frac{\nu+J}{\nu+\boldsymbol{e}_t'\boldsymbol{\Omega}\ \boldsymbol{e}_t}. \tag{10}$$

A.1 EM algorithm with known ν

We treat τ_t as a hidden variable and want to estimate $\boldsymbol{B}$ and $\boldsymbol{\Omega}$. In the E-step, we compute the conditional expectation of the hidden variables according to (10). These expected values are put on the diagonal of an $N \times N$ diagonal matrix $\boldsymbol{\tau}$. In the M-step, we solve the optimization problem

$$\begin{aligned}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\Omega}}|\boldsymbol{\tau}) \quad &= \underset{\boldsymbol{B},\boldsymbol{\Omega}}{arg\,min} \frac{1}{2N}\text{tr}\left[(\boldsymbol{\tau}^{1/2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}))\boldsymbol{\Omega}(\boldsymbol{\tau}^{1/2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{B}))'\right] \\ &\quad -\frac{1}{2}\log|\boldsymbol{\Omega}| \\ &\quad +\lambda\sum_{i,j=1}^{J}\sum_{p=1}^{P}|B_{p,ij}| + \gamma\sum_{i\neq j}^{J}|\omega_{ij}|,\end{aligned} \tag{11}$$

where the conditional normality (9) of the error terms, given the hidden variables, is used. This M-step corresponds exactly to computing the Gaussian lasso (3) for weighted observations, with weights $\hat{\tau}_t, \quad t = 1, \ldots, N$. Algorithm 1 gives the details of this EM algorithm.

Algorithm 1. *t*-Lasso with known ν

- $\boldsymbol{Y}, \boldsymbol{X}$, degrees of freedom ν.
- $\hat{\boldsymbol{\Omega}}^{(0)} = \mathbf{I}_J$, $\hat{\boldsymbol{B}}_p^{(0)} = \boldsymbol{I}_J$ for $1 \le p \le P$, $\boldsymbol{e}_t^{(0)}$ the tth row of $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}^{(0)}$ for $t = 1, \ldots, N$.
- Iterate the following steps for $m = 0, 1, 2, \ldots$:
  - ○ Recompute the weights
    $$\hat{\tau}_t = \frac{\nu+J}{\nu+(\boldsymbol{e}_t^{(m)})'\hat{\boldsymbol{\Omega}}^{(m)}\boldsymbol{e}_t^{(m)}}.$$
  - ○ Compute $\hat{\boldsymbol{B}}^{(m+1)}$ and $\hat{\boldsymbol{\Omega}}^{(m+1)}$ using the Gaussian lasso with inputs $\mathbf{Y} = (\hat{\boldsymbol{\tau}})^{1/2}\boldsymbol{Y}$, and $\mathbf{X} = (\hat{\boldsymbol{\tau}})^{1/2}\boldsymbol{X}$. Here, $\hat{\boldsymbol{\tau}}$ is the diagonal matrix having $\hat{\tau}_t$ on its diagonal. Let $\boldsymbol{e}_t^{(m+1)}$ be the tth row of $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}^{(m+1)}$.
- Iterate until the relative change in the value of the objective function in (11) in two successive iterations is smaller than the desired accuracy.
- $\hat{\boldsymbol{B}} = \hat{\boldsymbol{B}}^{(m+1)}$ and $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}^{(m+1)}$.

Gaussian lasso. Algorithm 1 for the *t*-lasso with known degrees of freedom requires computation of the Gaussian lasso defined in (3). We follow Rothman et al. (2010) and first solve (3) for the autoregressive parameter $\boldsymbol{B}$ conditional on the inverse error covariance matrix $\boldsymbol{\Omega}$, using a coordinate descent algorithm (see Friedman et al., 2007). Secondly, we solve for $\boldsymbol{\Omega}$ conditional on $\boldsymbol{B}$, using the Graphical lasso algorithm of Friedman et al. (2008). These two steps are implemented in `R` using the `grplasso` and `glasso` packages, respectively. We iteratively repeat the two steps until convergence of the objective function is reached.

When solving for $\boldsymbol{B}|\boldsymbol{\Omega}$, we use a grid of regularization parameters λ and search for the optimal one minimizing the Bayesian Information Criterion (BIC)

$$\text{BIC}_\lambda = -2\log L_\lambda + \text{df}_\lambda \log(N),$$

where $\log L_\lambda$ is the estimated likelihood, i.e. the first term in (3), using regularization parameter λ, and df_λ is the number of non-zero components of $\hat{\boldsymbol{B}}$. Likewise, when solving for $\boldsymbol{\Omega}|\boldsymbol{B}$ we use a grid of regularization parameters γ and search for the optimal one minimizing the BIC

$$\text{BIC}_\gamma = -2\log L_\gamma + \text{df}_\gamma \log(N),$$

where df_γ is the number of non-zero lower diagonal elements of $\hat{\boldsymbol{\Omega}}$.

A.2 ECM algorithm with unknown ν

For unknown degrees of freedom ν, the M-step is replaced by two constrained maximizations (CM) steps, where first $(\boldsymbol{B}, \boldsymbol{\Omega})$ and then ν are estimated. An E-step is introduced before each CM-step, such that the weights are recomputed twice in each iteration. This result is a multi-cycle version of the EM algorithm. In the first CM-step, we take ν as fixed, and compute the residuals from the VAR model estimated with the Gaussian lasso. In the second CM-step, we fit a multivariate *t*-distribution with unknown degrees of freedom to these residuals, exactly as in Liu and Rubin (1995), and keep the estimate of ν. Algorithm 2 presents the complete ECM algorithm.

Algorithm 2. *t*-Lasso with unknown ν

- $\boldsymbol{Y}, \boldsymbol{X}$.
- $\hat{\boldsymbol{\Omega}}^{(0)} = \mathbf{I}_J$, $\hat{\boldsymbol{B}}_p^{(0)} = \boldsymbol{I}_J$ for $1 \le p \le P$, $\boldsymbol{e}_t^{(0)}$ the tth row of $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}^{(0)}$, for $t = 1, \ldots, N$
  $\hat{\nu}^{(0)} = 1000$.
- Iterate the following steps for $m = 0, 1, 2, \ldots$:
  - ○ Recompute the weights:
    $$\hat{\tau}_t = \frac{\hat{\nu}^{(m)}+J}{\hat{\nu}^{(m)}+(\boldsymbol{e}_t^{(m)})'\hat{\boldsymbol{\Omega}}^{(m)}\boldsymbol{e}_t^{(m)}}.$$
  - ○ Compute $\hat{\boldsymbol{B}}^{(m+1)}$ and $\hat{\boldsymbol{\Omega}}^{(m+1)}$ using the Gaussian lasso with inputs $\mathbf{Y} = (\hat{\boldsymbol{\tau}})^{1/2}\boldsymbol{Y}$, and $\mathbf{X} = (\hat{\boldsymbol{\tau}})^{1/2}\boldsymbol{X}$. Here, $\hat{\boldsymbol{\tau}}$ is the diagonal matrix having $\hat{\tau}_t$ on its diagonal.
    Let $\boldsymbol{e}_t^{(m+1)}$ be the tth row of $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{B}}^{(m+1)}$.
  - ○ Recompute the weights:
    $$\hat{\tau}_t = \frac{\hat{\nu}^{(m)}+J}{\hat{\nu}^{(m)}+(\boldsymbol{e}_t^{(m+1)})'\hat{\boldsymbol{\Omega}}^{(m+1)}\boldsymbol{e}_t^{(m+1)}}.$$
  - ○ Compute $\hat{\nu}^{(m+1)}$ with a one-dimensional search minimizing
    $$\begin{aligned}&-\varphi\left(\frac{\nu}{2}\right) + \log\left(\frac{\nu}{2}\right) + \frac{1}{N}\sum_{t=1}^{N}\left(\log(\hat{\tau}_t) - \hat{\tau}_t\right) + 1 \\ &+\frac{1}{N}\left[\varphi\left(\frac{\nu+J}{2}\right) - \log\left(\frac{\nu+J}{2}\right)\right] = 0.\end{aligned} \tag{12}$$
    Eq. (12) has been derived in Liu and Rubin (1995, Eq. (30)); φ stands for the derivative of the log-Gamma function.
- Iterate until the relative change in the value of the objective function in (11) in two successive iterations is smaller than the desired accuracy.
- $\hat{\boldsymbol{B}} = \hat{\boldsymbol{B}}^{(m+1)}$, $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}^{(m+1)}$ and $\hat{\nu} = \hat{\nu}^{(m+1)}$.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.eneco.2019.104555.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., Heiko, E., 2001. The distribution of realized stock return volatility. J. Finan. Econ. 61 (1), 43–76.

Arouri, M., Jouini, J., Nguyen, D., 2011. Volatility spillovers between oil prices and stock sector returns: implications for portfolio management. J. Int. Money Finance 30 (7), 1387–1405.

Barbaglia, L., Wilms, I., Croux, C., 2016. Commodity dynamics: a sparse multi-class approach. Energy Econ. 60, 62–72.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc.: Ser. B Stat. Methodol. 64 (2), 253–280.

Beckman, J., Borchers, A., Jones, C.A., 2013. Agriculture's supply and demand for energy and energy products. USDA-ERS Econ. Inform. Bull., 112.

Bubák, V., Kočenda, E., Žikeš, F., 2011. Volatility transmission in emerging European foreign exchange markets. J. Bank. Finance 35 (11), 2829–2841.

Callot, L.A.F., Kock, A.B., Medeiros, M.C., 2017. Modeling and forecasting large realized covariance matrices and portfolio choice. J. Appl. Econ. 32 (1), 140–158.

Caporin, M., Velo, G.G., 2015. Realized range volatility forecasting: dynamic features and predictive variables. Int. Rev. Econ. Finance 40, 98–112.

Chang, T.H., Su, H.M., 2010. The substitutive effect of biofuels on fossil fuels in the lower and higher crude oil price periods. Energy 35 (7), 2807–2813.

Chkili, W., Hammoudeh, S., Nguyen, D., 2014. Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. Energy Econ. 41, 1–18.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Finan. Econ. 7 (2), 174–196.

Corsi, F., Mittnik, S., Pigorsch, C., Pigorsch, U., 2008. The volatility of realized volatility. Econ. Rev. 27 (1–3), 46–78.

Davis, R., Zang, P., Zheng, T., 2016. Sparse vector autoregressive modeling. J. Comput. Graph. Stat. 25 (4), 1077–1096.

Derimer, M., Diebold, F.X., Liu, L., Yilmaz, K., 2018. Estimating global bank network connectedness. J. Appl. Econ. 33 (1), 1–15.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Business Econ. Stat. 13 (3), 253–263.

Diebold, F.X., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. Econ. J. 119 (534), 158–171.

Diebold, F.X., Yilmaz, K., 2012. Better to give than to receive: predictive directional measurement of volatility spillovers. Int. J. Forecast. 28 (1), 57–66.

Diebold, F.X., Yilmaz, K., 2015. Financial and Macroeconomics Connectedness: A Network Approach to Measurement and Monitoring. Oxford University Press, New York, US.

Ding, P., 2016. On the conditional distribution of the multivariate *t* distribution. Am. Stat. 70 (3), 293–295.

Du, X., Yu, C., Hayes, D., 2011. Speculation and volatility spillover in the crude oil and agricultural commodity markets: a Bayesian analysis. Energy Econ. 33 (3), 497–503.

Eberlein, E., Keller, U., 1995. Hyperbolic distributions in finance. Bernoulli 1, 281–299.

Engle, R., Jondeua, E., Rockinger, M., 2015. Systemic risk in Europe. Rev. Finance 19 (1), 145–190.

Fernández, C., Steel, M.F., 1998. On Bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. 93, 359–371.

Finegold, M., Drton, M., 2011. Robust graphical modeling of gene networks using classical and alternative *t*-distributions. Ann. Appl. Stat. 5 (2A), 1057–1080.

Friedman, J., Hastie, T., Höfling, H., Tibshirani, R., 2007. Pathwise coordinate optimization. Ann. Appl. Stat. 1 (2), 302–332.

Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9 (3), 432–441.

Gelper, S., Wilms, I., Croux, C., 2016. Identifying demand effects in a large network of product categories. J. Retail. 92 (1), 25–39.

Hafner, C.M., Herwartz, H., 2006. Volatility impulse responses for multivariate GARCH models: an exchange rate illustration. J. Int. Money Finance 25 (5), 719–740.

Hasanov, A., Poon, W., Al-Freedi, A., Heng, Z., 2018. Forecasting volatility in the biofuel feedstock markets in the presence of structural breaks: a comparison of alternative distribution functions. Energy Econ. 70, 307–333.

Hassler, J., Sinn, H.W., 2016. The fossil episode. J. Monetary Econ. 83, 14–26.

Hassler, U., Rodrigues, P.M.M., Rubia, A., 2016. Quantile regression for long memory testing: a case of realized volatility. J. Finan. Econ. 14 (4), 693–724.

Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical Learning with Sparsity: The Lasso and Generalizations. CRC Press.

Hautsch, N., Schaumburg, J., Schienle, M., 2015. Financial network systemic risk contributions. Rev. Finance 19 (2), 685–738.

Kang, S., Yoon, S., 2009. Forecasting volatility of crude oil markets. Energy Econ. 31, 119–125.

Karali, B., Ramirez, O., 2014. Macro determinants of volatility and volatility spillover in energy markets. Energy Econ. 46, 413–421.

Knittel, C.R., Pindyck, R.S., 2016. The simple economics of commodity price speculation. Am. Econ. J.: Macroecon. 8 (2), 85–110.

Kotz, S., Nadarajah, S., 2004. Multivariate t Distributions and Their Applications. Cambridge University Press, Cambridge, UK.

Lanne, M., Nyberg, H., 2016. Generalized forecast error variance decomposition for linear and nonlinear multivariate models. Oxf. Bull. Econ. Stat. 78 (4), 595–603.

Levin, A., Lin, C.F., Chu, C.S.J., 2002. Unit root tests in panel data: asymptotic and finite-sample properties. J. Econ. 108 (1), 1–24.

Liu, C., Rubin, D.B., 1995. ML estimation of the t distribution using EM and its extensions, ECM and ECME. Stat. Sin. 5, 19–39.

Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer, Heidelberg, Germany.

Martens, M., van Dick, D., 2007. Measuring volatility with the realized range. J. Econ. 138 (1), 181–207.

McAleer, M., Medeiros, M.C., 2008. Realized volatility: a review. Econ. Rev. 27 (1–3), 10–45.

Miao, H., Ramchander, S., Wang, T., Yang, D., 2017. Influential factors in crude oil price forecasting. Energy Econ. 68, 77–88.

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

Nucera, F., Schwaab, B., Koopman, S.J., Lucas, A., 2016. The information in systemic risk rankings. J. Emp. Finance 38 (A), 461–475.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Business 53 (1), 61–65.

Pesaran, H.H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. Econ. Lett. 58 (1), 17–29.

Qu, H., Duan, Q., Niu, M., 2018. Modeling the volatility of realized volatility to improve volatility forecasts in electricity markets. Energy Econ. 74, 767–776.

Rezitis, A.N., 2015. The relationship between agricultural commodity prices, crude oil prices and US dollar exchange rates: a panel VAR approach and causality analysis. Int. Rev. Appl. Econ. 29 (3), 403–434.

Rothman, A.J., Levina, E., Zhu, J., 2010. Sparse multivariate regression with covariance estimation. J. Comput. Graph. Stat. 19 (4), 947–962.

Schmidt, D., Makalic, E., 2017. Robust lasso regression with Student-*t* residuals. Conference Paper in Lecture Notes in Computer Science, http://dx.doi.org/10.1007/978-3-319-63004-5-29.

Serra, T., 2011. Volatility spillovers between food and energy markets: a semiparametric approach. Energy Econ. 33, 1155–1164.

Serra, T., Zilberman, D., 2013. Biofuel-related price transmission literature: a review. Energy Econ. 37, 141–151.

Serra, T., Zilberman, D., Gil, J.M., Goodwin, B.K., 2010. Price transmission in the US ethanol market. In: Handbook of Bioenergy Economics and Policy. Springer, pp. 55–72.

Shu, J., Zhang, J.E., 2006. Testing range estimators of historical volatility. J. Futures Markets 26 (3), 297–313.

Theodossiou, P., 1998. Financial data and the skewed generalized *t* distribution. Manag. Sci. 44, 1650–1661.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc.: Ser. B Stat. Methodol. 58 (1), 267–288.

Ziel, F., Weron, R., 2018. Day-ahead electricity price forecasting with high-dimensional structures: univariate vs. multivariate modeling frameworks. Energy Econ. 70, 396–420.