TI 2016-006/III
Tinbergen Institute Discussion Paper

Are the S&P 500 Index and Crude Oil, Natural Gas and Ethanol Futures related for Intra-Day Data?

Massimiliano Caporin[1]
Chia-Lin Chang[2]
Michael McAleer[3]

[1] *University of Padova, Italy;*
[2] *National Chung Hsing University, Taiwan;*
[3] *National Tsing Hua University, Taiwan; Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands; Complutense University of Madrid, Spain.*

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Are the S&P 500 Index and Crude Oil, Natural Gas and Ethanol Futures Related for Intra-Day Data? *

Massimiliano Caporin

Department of Economics and Management "Marco Fanno"
University of Padova, Italy

Chia-Lin Chang

Department of Applied Economics
Department of Finance
National Chung Hsing University, Taiwan

Michael McAleer

Department of Quantitative Finance
National Tsing Hua University, Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute, The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid, Spain

Revised: February 2016

* For financial support, the second author wishes to thank the National Science Council, Taiwan, and the third author wishes to acknowledge the Australian Research Council and the National Science Council, Taiwan.

Abstract

The energy sector is one of the most important in the world, so that time series fluctuations in leading energy sources have been analysed widely. As the leading energy commodities are traded on international stock exchanges, the analysis of the fluctuations in stock and financial derivatives prices and returns have also been investigated extensively in recent years. Much of the empirical analysis has concentrated on using daily, weekly or monthly data, with little research based on intra-day data. The paper analyses the relationships among the S&P 500 Index and futures prices, returns and volatility of three leading energy commodities, namely crude oil, natural gas and ethanol, using intra-day data. The detailed analysis of intra-day temporal aggregation in examining returns relationships and volatility spillovers across the equity and energy futures markets, and the effects of overnight returns, volume, realized volatility, asymmetry, and spillovers across the four financial markets, leads to interesting and useful results for decision making and hedging strategies. The empirical results relating to alternative models of mean and variance feedback and asymmetry for intra-daily returns, asymmetry and volatility spillovers, and dynamic conditional correlations and covariances, show that the relationships between the stock market and alternative energy financial derivatives, specifically futures prices and returns, can and do vary according to the trading range, whether daily or overnight effects are considered, and the temporal aggregation and time frequencies that are used.

Keywords: Trading range, Intra-day prices and returns, S&P 500 Index, Crude oil futures, Natural gas futures, Ethanol futures, Overnight returns, Overnight volume, Overnight realized volatility, Asymmetry, Spillovers.

JEL Codes: C22, C32, C58, G12, G15.

1. Introduction

As the energy sector is one of the most important in the world in terms of its contribution to world trade, income and employment, it is not surprising that the time series fluctuations in leading energy sources, such as oil, natural gas and ethanol have been analysed widely in terms of economics and finance, thereby leading to significant research in energy economics and energy finance. Given the recent emphasis on the development of green energy, in which agricultural products, especially sugar cane and corn, have been used to produce bio-ethanol, the emphasis on agricultural finance has also not been surprising. As the leading energy and agricultural commodities are traded on international stock exchanges, the analysis of the fluctuations in stock and financial derivatives prices and returns, as well as the accompanying dynamic volatility to analyse risk and to develop hedging strategies, have also been investigated extensively in recent years.

In this context, the returns, volatility and volatility spillovers (namely, the delayed effect of a returns shock in one financial/energy/agricultural asset on the subsequent volatility or covolatility in another financial/energy/agricultural asset), among alternative energy commodities across different markets have been analysed using a variety of univariate and multivariate returns and volatility models, alternative data frequencies, and different data sets. Given the recent interest and emphasis in bio-fuels and green energy, and the various agricultural products that can be used to produce bio-ethanol, there is a topical and developing literature on the spillovers between energy and agricultural markets (see Chang et al. (2015) for a recent crticial review of the literature that connects the energy and agricultural sectors).

Much of the preceding literature has used data frequencies such as one day, week or month to analyse the relationships and spillovers across different financial/energy/agricultural markets, with few if any examining the relationships among such financial, energy and agricultural assets using intra-day data. The advantages of using intra-day data include the ability ot examine the effects of intra-day temporal aggregation in examining returns relationships and volatility spillovers across different financial/energy/agricultural markets, such as 5-minute, 15-minute, 45-minute frequencies, as well as the effects of overnight returns, volume, realized volatility, asymmetry, and spillovers, in comparison with the daily frequency.

The purpose of the paper is to analyse the relationships among the S&P 500 Index and futures prices, returns and volatility of three leading energy commodities, namely crude oil, natural gas and ethanol, using intra-day data. This will lead to a detailed analysis of intra-day temporal aggregation in examining returns relationships and volatility spillovers across the equity and energy futures markets, and examine the effects of overnight returns, volume, realized volatility, asymmetry, and spillovers across the four financial markets.

The plan of the remainder of the paper is given as follows. Section 2 examines energy futures and the stock market index, including alternative daily and overnight measures and an empirical analysis of three energy futures, namely crude oil, natural gas and ethanol, and the S&P 500 Index, and alternative daily and overnight measures of returns, volume and realized volatility. Section 3 analyses alternative models of mean and variance feedback and asymmetry for intra-daily returns, asymmetry and volatility spillovers, dynamic conditional correlations and covariances, and empirical analysis. Some concluding comments are given in Section 4.

2. Energy Futures and Stock Market Index

2.1 Alternative Measures of Three Energy Futures and S&P 500 Index

In this paper we consider the time series behaviour, especially the correlations among, three important energy-related futures and the equity stock market index. All data are obtained from TickData.[1] We consider the futures of WTI Crude Oil, Natural Gas, and Ethanol, and the S&P500 equity market index. Ethanol futures are traded at CME/CBOT[2], while Natural Gas and Crude Oil futures are traded at NYMEX.

The futures data and the equity market index have different time coverage within the TickData database. Nevertheless, in order to perform analyses on a common sample, we restrict attention to a four-year period starting in January 2010 and ending in December 2013. The choice of sample period is limited by the availability of Ethanol futures data from January 2010.

[1] See www.tickdata.com for details on the database coverage of the TickData company, and for further information on data accessibility.
[2] The selected contract is not the most highly traded, but data for Ethanol Platts futures are not available through TickData.

The S&P500 data are available at the 1-minute frequency and have been pre-filtered by TickData to remove cancelled trades, misalignment of data with the previous or subsequent trades, and bad ticks. All these corrections have been validated with third-party sources. In addition, TickData time series take into account the effects of corporate actions. The available dataset for the equity index includes, for each 1-minute interval, the open-high-low-close prices. Empty intervals, if present, have been filled by using the last available prices, so that a zero return would be induced.

The management of futures data is more complex. In fact, on any given day, several futures are traded on the same underlying asset, and differ in the maturity date. However, for the purposes of this paper, a continuous time series is required. Consequently, from the TickData dataset we extract a continuous time series which is based on the front future contract (that is, the nearest contract which, in our case, is also the most actively traded). In addition, as several futures reach maturity within our sample, a roll method to the next future is required.

The Automatic Roll method provided by TickData through its TickWrite software is used. The automatic roll method performs the transition between the current front future and the next future on the basis of the daily volume.[3] Finally, in order to remove rollover gaps across contracts, we perform a backward ratio adjustment on the prices. For each 1-minute interval, we obtain the open-high-low-close prices, together with the volume traded within the minute, and the number of ticks recorded in the minute. Prices are adjusted for rollovers, and are also corrected and validated for the same type of errors and problems that can affect the equity index data. The volume data and tick counts refer only to the front contract adopted in the construction of the continuous time series.

A further aspect deserves attention, namely the trading hours of futures contracts, as opposed to the trading hours for the equity market. Taking as a reference the New York Stock Exchange, the equity market has trading beginning at 9:30 AM and ending at 4 PM Eastern US time. However, for the futures contracts, we can have trading expressed in either Eastern or Central US time. For reasons of simplicity, we use Eastern US time. In Eastern US time, trading for Natural Gas and Crude Oil futures lasts for almost the entire day in electronic form, with a break from 5:15 PM to 6PM. Trading starts on Sunday at 6 PM and ends on Friday at 5:15 PM. For both contracts, open outcry starts at 9 AM and ends at 2:30 PM, excluding week-ends.

[3] See the user manual of TickWrite software, which is available from the TickData website, for additional details.

The trading hours of Ethanol futures are sensibly shorter, and change during the sample period. In fact, up to mid-2012, electronic trading had a night session from 6 PM to 7:15 AM from Sunday to Friday, and a daily session (contemporaneous to the open outcry session) from 9:30 AM to 1:15 PM, Monday to Friday. In 2012, electronic trading was extended, beginning at 5 PM and ending at 2 PM on the following day, from Sunday to Friday. The daily session with open outcry was also extended from 9:30 AM to 2 PM, as of June 25, 2012.[4]

The various changes pose a relevant challenge to the joint analysis of the energy futures and equity data. We opt for what will be viewed as a simple but reasonable choice: when we consider the estimation of models involving data associated with the four assets (Natural Gas, Crude Oil, Ethanol, and the S&P500 index), we will restrict the time span to the range 9:30 AM to 1:15 PM, which lasts for 3 hours and 45 minutes. We are aware that relevant information may well be excluded. However, as will be argued below, information arising from the trading activity that occurs outside the previous range will be collapsed into an indicator or into additional variables that will be incorporated into the empirical analysis.

With the 1-minute data related to the assets of interest, we can perform a number of preliminary analyses to evaluate the evolution of the trading activity during the day. This can be monitored, for instance, by the intra-daily volume, tick count, and volatility. Intuitively, we can expect significantly different levels of volume and tick counts when comparing ethanol futures with the futures on light crude oil and natural gas.

In addition, as the ethanol market liquidity is lower than the liquidity on oil and gas, we can expect a large number of 1-minute intervals with zero volumes and zero returns. Such preliminary analyses suggests aggregation of the available 1-minute data to that of a 5-minute frequency. This leaves a total of 288 intra-day observations for each 24-hour day. At this stage, we will focus on the entire day as the trading of financial futures is allowed even during the night, as was noted above.

[4] The trading time has been further modified and aligned to that of Crude Oil and Natural Gas, as reported in the CME website (last access date, July 2014).

For each five-minute interval, we recover the number of tick counts, volume, close-to-close returns,[5] and a suitable proxy for volatility, as measured by squared returns. The graphical evaluation of the raw time series is, however, not particularly informative. More interesting issues arise when we average across days. For a given measure (such as returns or volumes), denoted by $m_{i,t}$, where *i* refers to the five-minute interval during the day, and *t* is the daily time index, the 5-minute average value is given as:

$$\overline{m}_{i,\cdot} = \frac{1}{T}\sum_{t=1}^{T} m_{i,t} \;. \qquad (1)$$

In the sample, we have a total of 1461 days, including holidays and Saturdays, which is the only day of the week where there are no trades for energy-related commodities. Where tick counts and returns do not provide insightful results, the analysis of volume and volatility can provide more useful informative.

2.2 Empirical Analysis of Three Energy Futures and S&P 500 Index

Figures 1 to 3 report the average traded volume for Ethanol, Crude Oil, and Natural Gas futures. The shaded areas in the graphs highlight the range 9:30 AM to 1:15 PM. The S&P 500 index is not present as the volume for the index is not included in our database. The Ethanol plot (Figure 1) shows clearly visible three different spikes associated with the opening of the outcry/daily negotiation (9:30 AM), and with the closing of the outcry/daily trading session. For the latter, we have two spikes, at 1:15 PM and at 2PM, as the closing time was modified in June 2012. We also note that the volume is extremely high at closing, namely about 10 times larger than during the morning, and five times larger than at opening. Trades occurring before 9:30 AM and after 2 PM, that is, during the night or the electronic sessions, is sparse and is concentrated in the last months of our sample.

For Natural Gas and Crude Oil futures (Figures 2 and 3, respectively), we observe a similar pattern, with spikes in the volume observed at 9 AM, 10:30 AM and 2:30 PM. The first and third spikes are

[5] At the five-minute frequency, the close-to-close return is computed as the log difference between the last price (close price) of two consecutive 5-minutes intervals. For the first trading period after a market closure (for instance, at 6PM – the market is closed from 5:15 PM to 6 PM - for Natural Gas, or after week-ends), we consider the closing price of the last trading period.

associated with the beginning and end of the open outcry trading sessions. We associate the 10:30 AM spike with the closing of the European markets.[6]

A different picture emerges when we examine the volatility proxy (that is, squared returns). These are graphically represented in Figures 4 to 7, which includes the S&P500 index. In the Ethanol case (Figure 5), volatility is high at the beginning of the daily trading session, that is, at around 9:30 AM. However, what is interesting is that the highest spike occurs at 5 PM when the night session begins. This could be associated with the reaction of Ethanol prices to news released after 2 PM, or inferred by traders and operators from other energy commodity trading prices that are observed from 2 PM to 5 PM.

Overall, the Ethanol 5-minute volatility (as measured by squared returns) shows a pattern that is similar to that of equities, and is generally called an Inverse-J shape, which involves peaking at the opening, decreasing during the morning, with a minimum at around 12 PM, and then recovering towards the close of the day session.

Moving to Crude Oil and Natural Gas (Figures 6 and 7, respectively), we observe increases in the volatility at around the opening of the outcry session (9 AM), and at its close (2:30 PM). In addition, we have a further spike at around 10:30 AM, which is striking in the Natural Gas case. Notably, for both Natural Gas and Crude Oil, we observe a further spike at 6 PM, when the market re-opens after the weekday evening break.

As for Ethanol, we associate this spike with the typical movement observed at the equity market opening; for energy commodities, we observe two "opening" spikes, one in the morning at the opening of the outcry session, and one in the evening. These two spikes are very clear for Natural Gas, while for Crude Oil the behaviour is more erratic. For both commodities, the Inverse-J shape is distorted, but its fundamental movement might be observed if we focus on the morning up to 2 PM, and discard the 10:30 AM spike. For the S&P 500 index, we observe the known effect with the inverse-J shape for the 5-minute volatility.

[6] We note the double spike is not associated with data management issues associated with daylight saving time. Had this been the case, a double spike would have been observed at the closing, specifically at 3:30 PM.

In order to proceed to further empirical analysis that will consider joint modelling of the three commodities and the equity index, we restrict attention to a subset of the daily trading activity that lasts from 9:35 AM to 1:15 PM.[7] With such a choice, in principle, we exclude a relevant amount of information, in particular, for Crude Oil and Natural Gas, whose trading range lasts for almost the entire day. In order to recover the information from trades (in terms of price, volume, and volatility) executed from 1:15 PM for a given day, up to 9:30 AM of the following trading day, we create a set of "overnight" indicators. Although these indicators are not strictly "overnight" as they monitor trades occurring during the traditional daily trading sessions, they will nevertheless have the same impact as the overnight returns, as in Gallo (2001).

2.3 Alternative Measures of Returns, Volume and Realized Volatility

The three measures we consider which, by construction, have a daily frequency, are the following, where it is noted that the overnight return could span more than one day as $t+1$ might be a holiday or be located during week-ends:

(1) **overnight returns**: computed as the log-price difference between the price observed at 1:15 PM of a given trading day t, and the opening price observed at 9:30 AM of the subsequent trading day;
(2) **overnight volume**: computed as the sum of all trades that occur between 1:15 PM of day t and 9:30 AM of the subsequent trading day;
(3) **overnight realized volatility**: computed as the sum of all the squared 5-minute returns observed in the range 1:15 PM of day t up to 9:30 AM of the next trading day.

In addition to the construction of overnight-related variables, for purposes of defining a common sample, we exclude some days when either the Ethanol futures or the equity market was not trading.

After the selection of the daily trading range and the exclusion of specific days, the overall sample reduces to a total of 997 trading days. For these specific days, we have 5-minute data for all of the variables of interest, corresponding to 45 observations per day. The data can also be aggregated into lower frequencies, namely 15 minutes, for 15 observations each day, or to 45 minutes, for 5 observations each day, or to the daily frequency.

[7] The first 5-minute interval is marked at 9:35 AM as it refers to the 5-minute period ending at 9:35 AM.

Tables 1 to 5 report some descriptive statistics across the various frequencies for both the equity index and the three energy-related commodity futures. We note some well-known stylized facts. The average returns are small, positive for the equity index, and negative for all the energy-related commodities. Such behaviour is stable across the different frequencies. The standard deviations of the energy commodities are always higher than those of the index, coherently with the different nature of the series (that is, the commodities are associated with futures contracts, while the equity is a spot market index), with the perception that commodities are more risky than equities.

It is notable the Ethanol futures have volatility that is smaller than that of the Natural Gas contract. However, such a result might be affected by both the number of zeros present in the series and the limited trading activity on the contract, and corresponds to the limited movements that are observed in the Ethanol price. We also note the increase in the kurtosis when moving to higher frequencies. However, the levels of excess kurtosis are sensibly different: the Natural Gas results are similar to those of the S&P500 index, despite being lower, but those of Crude Oil futures are quite different, with excess kurtosis, which does not increase exponentially in moving to higher frequencies, but instead increases slowly, reaching a level of 5.4 at the 5-minute frequency.

On the contrary, the Ethanol kurtosis explodes, but this is a by-product of the infrequent trading of the Ethanol futures contract. This is confirmed by the large amount of zeros, which we detect at the highest frequencies (specifically, 75% of the series is comprised of zero values at the 5-minute frequency, and about 45% at the 15-minute frequency). Asymmetry in increasing (in absolute terms) with an increase in frequency, and is negative, coherently with what is observed on single stocks, with the exception of Natural Gas, which is characterized by positive asymmetry. Again, for Ethanol, we observe the largest values.

The volume time series levels are influenced by the nature of the future contracts. In fact, the Crude Oil contract unit is 1,000 barrels, while for Natural Gas the contract is 10,000 million British thermal units, and for Ethanol, the contract size is 29,000 gallons. Nevertheless, we observe the small volume of Ethanol, and the large number of transactions for both Natural Gas and Crude Oil.

The overnight series show some differences compared with the daily time series, which are associated with a trading range starting at 9:30 AM and ending at 1:15 PM. As shown in Table 5, overnight

Ethanol futures returns have positive asymmetry, and the volume is quite high for Crude Oil, but with very large dispersion. In addition, we observe how the Realized Variance (RV) is comparable to the overnight dispersion of the returns.[8]

3. Mean and Variance Feedback and Asymmetry for Intra-daily Returns

3.1 Models of Energy Futures and Equity Intra-daily Returns

We first discuss the model for the analysis of futures and equity intra-daily returns. Denote by $Y_{i,t} = \left[R_{i,t}\ E_{i,t}\ O_{i,t}\ G_{i,t} \right]'$ the day *t* interval *i* returns for the variables of interest, namely the S&P 500 index, $R_{i,t}$ Ethanol futures, $E_{i,t}$, Light Crude Oil futures, $O_{i,t}$, and Natural Gas futures, $G_{i,t}$. The "day" can contain a different number of intervals, depending on the frequency adopted (namely, 5, 15 or 45 minutes). If the data are daily, the indication of the interval becomes redundant and can easily be avoided.

The mean dynamics follow a VARMAX-type model:

$$Y_{i,t} = \mu + \Phi_1 Y_{i-1,t} + \Phi_2 Y_{i,t-1} + \mathrm{B}_i X_{i,t-1} + \varepsilon_{i,t} \ , \tag{2}$$

where $Y_{i-1,t}$ represents a standard VAR(1) term that captures the short-term serial correlation that is commonly observed for high-frequency data (see Cont (2001), and Dacorogna et al. (2001), among others); if *i=1*, the term refers to the last observation of day *t-1*; $Y_{i,t-1}$ is a VAR term associated with the interval *i* returns of the previous day (for example, this corresponds to lag 45 for 5-minute data); and $X_{i,t-1}$ is a matrix of lagged exogenous explanatory variables.

[8] In order to compare values, we must take into account that the figures reported in Table 5 for the mean and variance are multiplied by 100, and that the RV columns report Realized Variances. Thus, we take the RV column mean, divide it by 100, take the square root, and then multiply by 100 for compatibility with the Returns Standard Deviation.

The matrix X_t is partitioned into four elements, each of which contains a set of asset-specific variables, that is, $X_{i,t} = \left[X_{i,t}^R \; X_{i,t}^E \; X_{i,t}^O \; X_{i,t}^G \right]'$. Each asset-specific element $X_{i,t}^j$ contains the following indicators:

- previous interval traded volume: $v_{i-1,t}^j$;
- overnight traded volume: $onv_{i,t}^j = onv_t^j$;
- overnight returns: $or_{i,t}^j = or_t^j$;
- overnight realized volatility: $orv_{i,t}^j = orv_t^j$.

Note that, despite being indexed at time *t*, the overnight variables are included in the information set at time *t-1* as they monitor market activity up to the opening of day *t*. Thus, the information set at time *t-1* contains information available up to the opening of the markets on day *t*. The overnight variables have a daily frequency, as highlighted above, and volume is not available for the equity index.

In order to allow for a different impact of the overnight variables across the intra-daily observations, the coefficient matrix, B_i, is interval specific. In such a way, we might have overnight variables that impact with different coefficients across all intervals during the day, or might impact in the first part of the day (as one can easily obtain by means of a set of zero restrictions), or might even have an impact for the first observation of the day. The last of these is empirically the most relevant structure, leading to matrix, B_i, with just two designs, namely one for the first interval of the day, and a second for the other intervals of the day.

In such a case, the model can be recast in the following alternative representation:

$$Y_{i,t} = \mu + \Phi_1 Y_{i-1,t} + \Phi_2 Y_{i,t-1} + \mathrm{B}_{vol} V_{i,t-1} + \mathrm{B}_{ove} \bar{X}_{t-1} D_{i,t}^1 + u_{i,t}, \tag{3}$$

where $V_{i,t-1}$ contains only the previous interval volume of the three Energy-related commodities; B_{vol} is a parameter matrix containing 12 coefficients (namely the volume of, say, crude oil impacts on all the commodities, as well as on the equity index); $\bar{X}_{t-1}$ contains the overnight variables of the energy commodities and the equity index (namely, 11 variables, as the equity index volume is not available), and is constant across intra-daily intervals; $D_{i,t}^{1}$ is a dummy variable, taking the value 1 for the first interval of the day, and zero otherwise; and, finally, B_{ove} is a matrix of coefficients that allow interactions across all the overnight variables and each endogenous (that is, dependent) variable.

In order to evaluate the possible asymmetric impact of past returns and overnight returns, we generalize the model, as follows:

$$\begin{aligned} Y_{i,t} = \mu + \Phi_1 Y_{i-1,t} + \Phi_2 Y_{i,t-1} \mathrm{B}_{vol} V_{i,t-1} + \mathrm{B}_{ove} \bar{X}_{t-1} D_{i,t}^{1} + \Upsilon_1 \left[I\left(Y_{i-1,t} < 0\right) Y_{i-1,t} \right] + \Upsilon_2 \left[I\left(Y_{i,t-1} < 0\right) Y_{i,t-1} \right] \\ + \Upsilon_3 \left[I\left(W_t < 0\right) W_t \right] D_{i,t}^{1} + u_{i,t} \end{aligned} \tag{4}$$

where *I(.)* is an matrix indicator variable that takes the following form for a generic *k*-dimensional vector *a*:

$$I\left(a<0\right) = \begin{bmatrix} I\left(a_1<0\right) & 0 & \cdots & 0 \\ 0 & I\left(a_2<0\right) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I\left(a_k<0\right) \end{bmatrix} \tag{5}$$

so that it has zero elements outside the main diagonal, while on the diagonal it assumes a value of 1 when the argument is true, and zero otherwise; the vector W_t contains the overnight returns of the four assets; and the parameter matrices, Υ_1, Υ_2, and Υ_3, monitor the additional impacts of negative returns and overnight returns. Note that asymmetric terms can be also introduced in specification (2). Although leverage is not permitted, given the specification of the volatility models (for further details, see McAleer (2014)).

In order to verify the relevance of the various components of the models introduced above, we will estimate five different specifications, as follows: (i) a model with only the VAR component (M1) corresponding to equation (2), with the parameter matrix B_i restricted to have zero coefficients; (ii) model M1 extended with the introduction of exogenous variables, lagged volume and overnight variables, with the latter having an impact over the entire daily trading range (M2), corresponding to equation (2); (iii) model M2, with the overnight variables impacting only on the first interval of the day (M3), corresponding to equation (3); (iv) M2 extended with asymmetric terms (M4), that the model in (2) with the inclusion of asymmetric impact as in (4); and (v) M3 extended with asymmetric terms (M5), that is, the model in equation (4).

All the models are estimated using the least squares method, with robust Newey-West HAC standard errors due to the presence of residual heteroskedasticity. Moreover, we estimate the models for the various intra-daily frequencies previously considered, namely at 5, 15 and 45 minutes, as well as for daily data. In the case of daily data, the models M3 and M5 are not considered.

3.2 Empirical Results

A summary of the empirical results is reported in Tables 6 to 9, while detailed tables reporting all the estimated coefficients are available upon request. What emerges from the mean estimation is that statistically significant coefficients are few, relative to the total number of coefficients included in each model.[9] This is somewhat to be expected as financial returns are being modelled. We also note that the asymmetry components are statistically significant (at least for some of the coefficients) when the model allows for the impact of overnight variables for the whole trading session. On the contrary, asymmetry is not present if we have an impact of the overnight asymmetry solely on first return of the day.

However, when we test for the joint significance of asymmetry behaviour by means of Wald-type tests of models M4 and M5, the result is opposite, and asymmetry seems to be more relevant when the models include the impact of overnight measures only on the first return of the day. Such an empirical result, which is slightly surprising, might occur as a by-product of the number of coefficients included in the models, in which case the joint test could be more powerful than the significance of individual coefficients. Overall, the empirical results show some evidence of the presence of asymmetric impacts of both overnight variables and past returns.

In addition, serial correlation is relevant at the intra-daily level, as can be seen for the Wald test for model M1, but not at the daily level, as can be seen in Table 9. This confirms the relevance of serial

[9] In each table we report the number of coefficients included in each estimated model

correlation for the intra-daily data and its limited, or virtually absent, impact at the daily frequency. The overnight variables and volume are relevant drivers of information, despite the number of significant coefficients being very small.

Nevertheless, we also note that the joint significance of the coefficients, as derived from the Wald tests, is more pronounced if we restrict the impact of the overnight variables for the first interval of the day (as can be seen in a comparison of the Wald tests for models M2 and M3). This suggests that the overnight information impacts mostly during the first interval, which is an expected result that confirms previous studies, but it can also be influenced by the way we define the daily trading range and the overnight variables.

3.3 Asymmetry and Volatility Spillovers

We now analyse the conditional variances and their relation with spillovers of shocks and asymmetry. In this case, we follow an approach similar to that of Billio and Caporin (2010) and to the VARMA-GARCH family that was proposed by Ling and McAleer (2003) as a generalization of the CCC model of Bollerslev (1990), and subsequently extended to include asymmetry, as in McAleer et al. (2007, 2008, 2009).

When dealing with models that include asymmetry and volatility spillovers, the number of parameters can increase quickly, thereby making estimation infeasible (which is widely known as the "curse of dimensionality"). In terms of volatility spillovers and co-volatility spillovers across different financial assets, Chang et al. (2015) give three new definitions of volatility spillovers, specifically full volatility, full covolatility spillovers, and partial covolatility spillovers, and evaluate the leading alternative multivariate models in terms of the new definitions,

For the data considered in the paper, the model should also take into account the periodic evolution of intra-daily volatility, which is a known feature of the data (see Andersen and Bollerslev (1997), and Dacorogna et al. (2001), among others). In order to control for the increase in the number of parameters, we filter out the periodic evolution in a preliminary step. Therefore, for each mean innovation, $u_{i,t}^{j}$, $j = R, E, O, G$, we assume that the periodic pattern is a multiplicative component that leads to:

$$u_{i,t}^{j} = p_{i,t}^{j}\varepsilon_{i,t}^{j}, \;\; j = R, E, O, G\,. \tag{6}$$

We recover the periodic pattern with the following linear model:

$$\ln\left(u_{i,t}^{j}\right)^{2} = \ln\left(p_{i,t}^{j}\right)^{2} + \ln\left(\varepsilon_{i,t}^{j}\right)^{2} = \sum_{l=1}^{k}\delta_{l}^{j}D_{i,t}^{l} + \ln\left(\varepsilon_{i,t}^{j}\right)^{2} \tag{7}$$

where $D_{i,t}^{l}$ are dummy variables for each intra-daily interval, and *k* is the number of intervals within a day. Thus, we take a simplified approach to remove the periodic behaviour. Alternative methods could be used (see, among others, Andersen and Bollerslev (1997) and Boudt et al. (2011)). After the periodic pattern has been estimated, we can easily filter it out from the mean innovations, use the estimates of the series, $\varepsilon_{i,t}$, and then analyse asymmetric and spillover effects by estimating, for example, a multivariate GARCH model, of which several alternatives are available.

3.4 Dynamic Conditional Correlations and Covariances

The model we consider focuses on the conditional variance, $\Sigma_{i,t}$, of the innovations, $\varepsilon_{i,t}$. For the second-order moment, we use the following dynamic structure:

$$\Sigma_{i,t} = D_{i,t}R_{i,t}D_{i,t} \tag{8}$$

where $D_{i,t}$ is the diagonal matrix of conditional volatilities, such as $D_{i,t} = diag\left(\left[\sigma_{i,t}^{R}\ \sigma_{i,t}^{E}\ \sigma_{i,t}^{O}\ \sigma_{i,t}^{G}\right]'\right)$, and $diag(a)$ is a matrix operator that creates a diagonal matrix, with the vector *a* on the main diagonal. Moreover, $R_{i,t}$ is assumed to be a dynamic conditional correlation matrix, although the caveats given in McAleer et al. (2008) should be borne in mind in interpreting the estimates from $R_{i,t}$ as conditional correlations.

If the conditional correlations can be estimated consistently, the conditional covariances, as defined by equation (8), can also be estimated consistently. As the definition in equation (8) can be rewritten to express the conditional correlations in terms of the conditional covariances, if the conditional covariances can be estimated consistently, the conditional correlations can also be estimated consistently.

We then model the log-conditional variances, $H_{i,t} = \left[\ln\left(\sigma_{i,t}^{2,R}\right)\ \ln\left(\sigma_{i,t}^{2,E}\right)\ \ln\left(\sigma_{i,t}^{2,O}\right)\ \ln\left(\sigma_{i,t}^{2,G}\right)\right]'$, as follows:

$$\begin{aligned} H_{i,t} &= \mu + diag\left(\boldsymbol{\beta}_1\right) H_{i-1,t} + diag\left(\boldsymbol{\beta}_2\right) H_{i,t-1} + \boldsymbol{\alpha}_1 \tilde{\varepsilon}_{i-1,t} + \boldsymbol{\alpha}_2 \tilde{\varepsilon}_{i,t-1} + \boldsymbol{\gamma}_1 \left[I\left(\varepsilon_{i-1,t} < 0\right) \tilde{\varepsilon}_{i-1,t} \right] \\ &\quad + \boldsymbol{\gamma}_2 \left[I\left(\varepsilon_{i,t-1} < 0\right) \tilde{\varepsilon}_{i,t-1} \right] + \boldsymbol{\delta}_1 D_{i,t}^1 ORV_t + \boldsymbol{\delta}_2 D_{i,t}^1 \left[I\left(W_t < 0\right) ORV_t \right] \end{aligned} \tag{9}$$

where $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, $\boldsymbol{\delta}_1$, *and* $\boldsymbol{\delta}_2$ are parameter matrices,

$$\tilde{\varepsilon}_{i,t} = \left[\ln\left(\left(\varepsilon_{i,t}^R\right)^2\right)\ \ln\left(\left(\varepsilon_{i,t}^E\right)^2\right)\ \ln\left(\left(\varepsilon_{i,t}^O\right)^2\right)\ \ln\left(\left(\varepsilon_{i,t}^G\right)^2\right) \right]', \tag{10}$$

and

$$ORV_t = \left[\ln\left(orv_t^R\right)\ \ln\left(orv_t^E\right)\ \ln\left(orv_t^O\right)\ \ln\left(orv_t^G\right) \right]'. \tag{11}$$

As previously discussed, the last component in equation (11), despite being indexed with time *t*, is known before "day" *t*, as it contains variables observed up to the opening. Moreover, the indicator variables, as well as the dummy variable pre-multiplying the overnight volatility, have the same structure as those adopted for the mean dynamics.

The model provides dynamics for the log-variances, which is similar to Bordignon et al. (2007, 2009). The innovations are not given by variance standardized residuals, as in the EGARCH model of Nelson (1991), but rather by log-squared innovations. The model shares the same advantage of EGARCH in excluding positivity restrictions to model the parameters. In addition, as for EGARCH, the model allows for asymmetry, which is defined as the possibly different impacts on volatility of positive and negative shocks of equal magnitude, but not leverage, which was defined in Black (1976), based on the debt-equity ratio (see McAleer and Hafner (2014) and McAleer (2014) for caveats regarding the signs of the coefficients in EGARCH).

The model in equation (9) shares some similarities with the model in Billio and Caporin (2010) as it does not allow for volatility spillovers. In fact, the log-conditional variances of a given asset are a function of only the past log-conditional variances of the same asset. This is due to the diagonality restriction imposed on the parameter matrices that pre-multiply the lagged log-conditional variances. On the contrary, the shocks of the other assets can be relevant, and their effects can depend on the signs allowing for asymmetry, as in McAleer et al. (2007, 2009).

The diagonality assumption adopted for the GARCH part of the model, as highlighted in Billio and Caporin (2010), allows parameter estimation on the basis of marginal univariate likelihoods, but at the cost of reduced efficiency. Nevertheless, we believe the loss in efficiency is acceptable in light of the sensible computational advantages. In fact, the full model, including the four variables, has 108 parameters. By moving to univariate models, the parameter number decreases to 27, which is still a large number, but with computationally feasible estimation. We also note that the model is similar to the models developed in McAleer et al. (2007, 2009) with respect to the introduction of asymmetry for the variables that are under consideration.

The model includes an "exogenous" variable, namely the overnight realized volatility. In fact, this is a natural information driver for volatility (as well as its asymmetric impact), given the movements occurring outside the trade range under consideration. Note that the asymmetric impact is driven by the sign of the overnight returns. Moreover, asymmetry and overnight volatility impact only on the first observation of the day, given the empirical evidence that has already been presented for the conditional mean.

3.5 Estimation

Estimation is performed at the univariate level, and the parameter restrictions are not needed to ensure the positivity of the conditional variances. However, the parameters are restricted to ensure stationarity. The constraints can be recovered by considering the ARMA representation of the model. We start from the dynamic equation of a single conditional variance, as follows:

$$\begin{aligned}\ln\left(\sigma_{i,t}^{2}\right) &= \omega+\beta_{1}\ln\left(\sigma_{i-1,t}^{2}\right)+\beta_{2}\ln\left(\sigma_{i,t-1}^{2}\right)+\alpha_{1}\ln\left(\varepsilon_{i-1,t}^{2}\right)+\alpha_{2}\ln\left(\varepsilon_{i,t-1}^{2}\right)+\gamma_{1}\ln\left(\varepsilon_{i-1,t}^{2}\right)I\left(\varepsilon_{i-1,t}<0\right) \\ &\quad+\gamma_{2}\ln\left(\varepsilon_{i,t-1}^{2}\right)I\left(\varepsilon_{i,t-1}<0\right)+\boldsymbol{\delta}'X_{i,t}\end{aligned} \tag{12}$$

where $X_{i,t}$ is a set of exogenous variables. Note that the innovations also include the other asset shocks, which are considered as exogenous variables. The innovations, $\varepsilon_{i,t}$, are given as products of the conditional volatility and the variance standardized shocks, $z_{i,t}$, with $\varepsilon_{i,t} = \sigma_{i,t} z_{i,t}$.

Upon taking squared logs, we have:

$$\ln\left(\varepsilon_{i,t}^2\right) = \ln\left(\sigma_{i,t}^2 z_{i,t}^2\right) = \ln\left(\sigma_{i,t}^2\right) + \ln\left(z_{i,t}^2\right). \tag{13}$$

Moreover, given that the distribution of $\ln\left(z_{i,t}^2\right)$ is non-standard, we define its unconditional mean as $\tau = E\left[\ln\left(z_{i,t}^2\right)\right]$. By taking the latter into account, we can define a martingale difference sequence, $\nu_{i,t} = \ln\left(\varepsilon_{i,t}^2\right) - \tau - \ln\left(\sigma_{i,t}^2\right)$, that can be used to recover the ARMA representation of the model in equation (12), which is given as follows:

$$\phi(L)\ln\left(\varepsilon_{i,t}^2\right) = \omega + \left(1 - \beta_1 - \beta_2\right)\tau + \left(1 - \beta_1 L - \beta_2 L^k\right)\nu_{i,t} + \boldsymbol{\delta}' X_{i,t}, \tag{14}$$

where

$$\phi(L) = 1 - \left(\alpha_1 + \gamma_1 I\left(\varepsilon_{i-1,t} < 0\right) + \beta_1\right) - L\left(\alpha_2 + \gamma_2 I\left(\varepsilon_{i,t-1} < 0\right) + \beta_2\right)L^k, \tag{15}$$

and k is the number of intervals for a given day.

Therefore, under symmetry of the density of $\varepsilon_{i,t}$, the expectation of the indicator variables equals ½, so that:

$$\phi(L) = 1 - \left(\alpha_1 + \gamma_1 \frac{1}{2} + \beta_1\right)L - \left(\alpha_2 + \gamma_2 \frac{1}{2} + \beta_2\right)L^k.$$

Stationarity is associated with the roots of $\phi(L)$ lying outside the unit circle, thereby indirectly imposing the constraints on the parameters.

3.6 Empirical Analysis

Tables 10 to 13 report the estimated parameters for the four assets and the four different frequencies, while Figure 8 contains an example of the estimated conditional variance patterns. For each (asset, frequency) combination, we estimate three different models, namely: (1) a baseline specification that does not include other asset shocks, overnight realized volatility (H1), or asymmetry; (2) a model that adds to H1 asymmetry with respect to the asset shocks and ORV (H2); and (3) the most general specification with asymmetry and all the other asset shocks (H3).

We first focus on the high frequency results, namely Tables 10 to 12. It is noted that the GARCH coefficients are highly significant for lag 1, with point values decreasing in some cases when the data frequency is decreased towards daily data. This is particularly evident for Crude Oil, mainly at the 45-minute frequency. The GARCH coefficient capturing the "daily" effect has a minor relevance, and its significance decreases with the data frequency. We also note that some GARCH coefficients for lag are occasionally greater than 1. This does not necessarily imply non-stationarity as the condition is acting on the roots of the polynomial in equation (15), and not on each single coefficient.

In considering the ARCH coefficients, which reflects the short-run persistence of returns shocks on volatility, there is confirmation of the observations reported for the GARCH component of the model, namely they are statistically significant at lag 1, irrespective of the frequency, and the significance of the "daily" coefficients decreases with the data frequency. In addition, as is commonly observed in empirical GARCH models, the size of the ARCH coefficients is much smaller than those of the GARCH coefficients. The previous comments are valid irrespective of the estimated model, namely H1, H2 or H3.

Consider now the shock spillovers, namely those monitored by the coefficients $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, which are located in the off-diagonal terms. These coefficients, which are included only in model H3, are in most cases not significant. The largest number of significant coefficients is observed at the 5-minute frequency. The coefficient sizes are small, in some cases negative and, in general, are lower (in absolute terms) than the values obtained for the diagonal coefficients (namely, the standard ARCH components). Thus, we conclude that the shock spillovers across the assets have limited impacts across the intra-daily frequencies.

An important result of the empirical analysis is associated with the impact of the overnight volatility. Such an exogenous variable is statistically significant for all assets and for all data frequencies for model H1 (namely, the specification without asymmetry and shocks and asymmetry spillovers). The coefficients are always positive, thereby implying that the overnight volatility increases the daily volatility level. This is a somewhat expected outcome as this variable conveys the information from the close of the previous day's trading range to the opening of the daily trading range.

When asymmetry is included in specification H2, it is observed that the positive impact on the intra-day volatility for the S&P index is mainly associated with the overnight volatility for the negative overnight returns. On the contrary, the overnight volatility that is matched with positive overnight returns leads to a decrease in the intra-daily volatility for higher frequencies. The results of Ethanol suggest a compensation effect between positive and negative overnight returns across the various frequencies. For Crude Oil, most relevant are the negative overnight returns, leading to an increase in the intra-daily volatility.

Somewhat surprisingly, for Natural Gas the most relevant effects arise from the positive overnight returns, where the volatility leads to an increase in intra-daily volatility, while negative overnight returns do not have such an impact. The latter results might be explained by examining the estimates for model H3, which include the cross-impacts of overnight volatility. We observe that the most relevant role is now played by the S&P 500 overnight volatility that enters (for both positive and negative overnight returns) into the dynamics of the energy commodities in many cases, with a negative coefficient for positive overnight returns, and a positive coefficient for negative overnight returns. Therefore, it can be concluded that the overnight volatility has a relevant role for the intra-daily dynamics of energy commodities, with a fundamental role afforded by the equity market variable.

Finally, we focus on the shock spillovers and asymmetry across the equity market and energy commodities. The baseline specification is model H2, with only asymmetry but not any cross-impacts. Across the intra-daily frequencies, we note that the asymmetric coefficients (as included in the matrices $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$) are, in most cases, statistically significant, positive at lag 1, and negative at the "daily" lag. These results imply that the previous intra-day period sign matters, and that negative returns have larger impacts, as compared with positive returns.

The introduction of asymmetry also explains the few cases of insignificant ARCH coefficients at lag 1. The "daily" lag has the opposite effect, a kind of mean-reversion applied to the volatility dynamics. Nevertheless, we note that the coefficients are, as is typical in empirical (G)ARCH models, very small. When spillovers for both the shocks and asymmetry are taken into account, the previous results are confirmed, with a positive effect from the previous period and a negative effect from the "daily" lag. Moreover, many cross-asset effects are present, but all with very small magnitudes. In addition, the relevance of the asymmetry and shock spillovers is sensibly reduced at the 45-minute frequency, which indicates that the relation is present at quite high data frequencies.

The last empirical finding might also be associated with the somewhat unusual results at the daily frequency (see Table 13). In fact, for this frequency, the estimated parameters are, in some sense, non-standard, with smaller GARCH effects, higher ARCH coefficients, and an extremely large impact of the overnight variables.

This empirical result might be explained by the approach taken for the construction of the trading range, which was driven by the need for creating a common foundation for the spillovers. The spillovers can be detected at higher frequencies, and the GARCH parameters have standard values owing to the persistence of the intra-daily volatility patterns. When moving to lower frequencies (such as the 45-minute interval) or the daily frequency, the role of the excluded trading periods becomes fundamental, thereby leading to an extremely large impact of the overnight volatility.

As a final check on the variance standardized series, $\eta_{i,t} = \varepsilon_{i,t}\sigma_{i,t}^{-1}$, we fit the dynamic conditional correlation (DCC) model of Engle (2002), and the Asymmetric DCC model of Cappiello et al. (2006). The results, which are not reported but are available upon request, show the persistence of conditional correlations. As stated previously, the caveats given in McAleer et al. (2008) should be borne in mind in interpreting the estimates from these two models as conditional correlations.

Table 14 reports the unconditional correlations observed over the various frequencies on the variance standardized residuals (using conditional variances from model H3). We first note that the only correlation assuming a somewhat large value is that between Crude Oil futures and the stock market index. This is understandable given the interaction between the oil price, economic growth or the economic cycle, and the subsequent relation of each with financial markets and the financial cycle.

On the contrary, all the other correlations take very small values. Moreover, by increasing the sampling frequency, the conditional correlations decrease, again with the exception of the Oil futures and S&P 500 index, where we note an increase from the daily to the intra-daily values. The latter depends on the construction of the trading range, and is associated with the odds results on modelling daily conditional covariances.

In moving to conditional modelling, as already mentioned, we obtained parameter estimates for the DCC model, which indicates high persistence, with data that are close to being integrated (see also Aielli, 2013). Figure 9 reports the time evolution of conditional correlations at the 15-minute frequency. Similar patterns can be obtained for other intra-day frequencies. Again, the caveats given in McAleer et al. (2008) should be borne in mind in interpreting the estimates from DCC as conditional correlations.

Several relevant elements are noted. At first, the only correlation providing high value is, as expected, given Table 14, namely for the S&P 500 index and Crude Oil futures. The correlation shows a decreasing trend, a pronounced decrease, followed by a recovery, with a minimum in February 2011. Somewhat differently, the dynamic correlations involving Ethanol futures all oscillate around zero, while the conditional correlations between Natural Gas and the S&P or Crude Oil are positive, but are nevertheless quite low.

The last two empirical findings suggest that the estimated persistence could be driven by the stronger relation between Crude Oil futures and the stock market. This is confirmed by the fit of more general models allowing for asymmetry in the correlation dynamics and/or for correlation-specific parameters (as in Cappiello et al. (2006)). However, the estimation of these models provide non-standard results and convergence problems that are due to the joint presence of constant and dynamic correlations.

Such empirical findings would suggest, as a possible alternative, the estimation of alternative systems of bivariate models. Nevertheless, with the focus of this paper on the evaluation of spillovers and asymmetry across variables, bivariate specifications would be expected to have minor relevance, and are thereby not considered.

4. Concluding Remarks

The purpose of the paper was to analyse the relationships among the S&P 500 Index and futures prices, returns and volatility of three leading energy commodities, namely crude oil, natural gas and ethanol, using intra-day data. The analysis led to a detailed analysis of intra-day temporal aggregation in examining returns relationships and volatility spillovers across the equity and energy futures markets, and examined the effects of overnight returns, volume, realized volatility, asymmetry, and spillovers across the four financial markets.

The paper examined the time series fluctuations in three energy futures, namely crude oil, natural gas and ethanol, and the stock market index, including alternative measures of returns, volume and realized volatility at the daily frequency.

The empirical results relating to alternative models of mean and variance feedback and asymmetry for intra-daily returns, asymmetry and volatility spillovers, dynamic conditional correlations and covariances, showed that the relationships between the stock market and alternative energy financial derivatives, specifically futures prices and returns, could and did vary according to the trading range, whether daily or overnight effects were considered, and the temporal aggregation and time frequencies that were used.

References

Aielli, G.P. (2013), Dynamic conditional correlations: On properties and estimation, Journal of Business and Economic Statistics, 31, 282-299.

Andersen, T.G. and Bollerslev, T. (1997), Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance, 4, 115-158.

Billio, M., and Caporin, M. (2010), Market linkages, variance spillovers and correlation stability: empirical evidences of financial market contagion, Computational Statistics and Data Analysis, 54-11, 2443-2458.

Black, F. (1976), Studies of Stock Price Volatility Changes, Proceedings of the Business and Economics Section of the American Statistical Association, 177-181.

Bollerslev, T. (1990), Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach, Review of Economic and Statistics 72, 498-505.

Bordignon, S., Caporin, M., and Lisi, F. (2007), Generalised Long Memory GARCH models for intradaily volatility, Computational Statistics & Data Analysis, 51-12, 5900-5912.

Bordignon, S., Caporin, M., and Lisi, F. (2009), Periodic Long Memory GARCH models, Econometric Reviews, 28, 60-82.

Boudt, K, Croux, C., and Laurent, S. (2011), Robust estimation of intraweek periodicity in volatility and jump detection, Journal of Empirical Finance, 18, 353-367.

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

Chang, C.-L., Li, Y.-Y., and McAleer, M. (2015), Volatility spillovers between energy and agricultural markets: A critical appraisal of theory and practice, Tinbergen Institute Discussion Paper, TI 2015-077/III.

Cont, R. (2001), Empirical properties of asset returns: stylized facts and statistical issues, Quantitative Finance, 1, 223-236.

Dacorogna, M.M., Gencay, R., Muller, A.U., Olsen, R.B., and Pictet, O.V. (2001), An Introduction to High Frequency Finance, Academic Press, San Diego.

Engle, R.F. (2002), Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional hereoskedasticity models, Journal of Business and Economic Statistics, 20, 339-350.

Gallo, G.M. (2001), Modeling the impact of overnight surprises on intra-daily volatility, Australian Economic Papers, 40(4), 567-580.

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

McAleer, M. (2014), Asymmetry and leverage in conditional volatility models, Econometrics 2(3), 145-150.

McAleer, M., Chan, F., Hoti, S. and Lieberman, O. (2008), Generalized autoregressive conditional correlation, Econometric Theory, 24(6), 1554-1583.

McAleer, M., Chan, F., and Marinova, D. (2007), An econometric analysis of asymmetric volatility: theory and application to patents, Journal of Econometrics 139, 259-284.

McAleer, M. and Hafner, C. (2014), A one line derivation of EGARCH, Econometrics, 2(2), 92-97.

McAleer, M., Hoti, S. and Chan, F. (2009), Structure and asymptotic theory for multivariate asymmetric conditional volatility, Econometric Reviews, 28, 422-440.

Nelson, D.B. (1991), Conditional heteroskedasticity in asset returns: A new approach, Econometrica, 59, 347-370.

Table 1: 5-minute Descriptive Statistics

	S&P	Ethanol		Crude Oil		Natural Gas	
	returns	returns	volume	returns	volume	returns	volume
Mean	0.00084	-0.00065	2.88	-0.00010	2357.92	-0.00300	952.41
St. Dev.	0.12225	0.20871	7.58	0.15479	1820.53	0.25377	1176.06
Min.	-0.02617	-0.11940	0	-0.01656	0	-0.04890	0
Max.	0.02677	0.08970	202	0.01626	34407	0.05915	31133
Asymmetry	-0.351	-4.301	5.485	-0.155	3.011	0.315	7.645
Exc. Kur.	74.826	396.763	49.890	5.382	18.769	47.756	98.559
No. zeros	234	33585	29431	1753	1	3043	14
No. obs.	44865						

Note: Mean and St. Dev. multiplied by 100 for returns and RV (= Realized Variances).

Table 2: 15-minute Descriptive Statistics

	S&P	Ethanol		Crude Oil		Natural Gas	
	returns	returns	volume	returns	volume	returns	Volume
Mean	0.00252	-0.00196	8.63	-0.00029	7073.75	-0.00899	2857.23
St. Dev.	0.20873	0.35563	14.53	0.26451	4625.80	0.42893	2882.46
Min.	-0.03048	-0.11940	0	-0.02656	501	-0.04911	0
Max.	0.03534	0.08970	322	0.01715	72589	0.06601	44366
Asymmetry	-0.128	-2.690	3.797	-0.238	2.204	0.170	5.133
Exc.Kur.	35.075	142.410	29.993	4.208	9.649	21.101	39.166
No. zeros	0	6825	5229	3	0	13	4
No. obs.	14955						

Note: Mean and St. Dev. multiplied by 100 for returns and RV (= Realized Variances).

Table 3: 45-minute Descriptive Statistics

	S&P	Ethanol		Crude Oil		Natural Gas	
	returns	returns	volume	returns	volume	returns	Volume
Mean	0.00755	-0.00589	25.88	-0.00088	21221.25	-0.02697	8571.68
St. Dev.	0.36535	0.59427	27.50	0.45108	11410.25	0.73309	6458.57
Min.	-0.04203	-0.11940	0	-0.04202	1808	-0.05066	0
Max.	0.03297	0.08567	405	0.01948	95323	0.06496	63092
Asymmetry	-0.204	-1.698	2.469	-0.350	1.322	0.214	2.943
Exc. Kur.	15.140	51.922	13.613	3.552	2.730	7.873	11.895
No. zeros	0	669	386	0	0	1	1
No. obs.	4985						

Note: Mean and St. Dev. multiplied by 100 for returns and RV (= Realized Variances).

Table 4: Daily Descriptive Statistics

	S&P	Ethanol		Crude Oil		Natural Gas	
	returns	returns	volume	returns	volume	returns	Volume
Mean	0.00038	-0.00029	129.41	-0.00004	106106.23	-0.00135	42858.40
St. Dev.	0.00827	0.01375	74.33	0.01007	37014.83	0.01616	18654.90
Min.	-0.04132	-0.13226	11	-0.04097	17510	-0.07160	9970
Max.	0.03602	0.09771	894	0.03404	269954	0.09709	143511
Asymmetry	-0.311	-0.726	2.118	-0.299	0.807	0.079	1.276
Exc. Kur.	2.695	13.039	12.852	1.216	1.227	2.518	2.140
No. zeros	0	0	0	0	0	0	0
No. obs	997						

Note: Mean and St. Dev. multiplied by 100 for returns and RV (= Realized Variances).

Table 5: Overnight Descriptive Statistics

	S&P		Ethanol			
	returns	RV	returns	volume	RV	
Mean	0.011	0.002	0.134	27.23	0.023	
St. Dev.	0.524	0.006	1.011	45.42	0.241	
Min.	-0.026	0.000	-0.043	0	0.000	
Max.	0.028	0.001	0.103	376	0.059	
Asymmetry	-0.258	13.418	1.717	2.386	19.362	
Exc. Kur.	4.213	231.063	15.286	8.321	415.625	
No. zeros	135	1035	16470	25830	25335	
	Crude Oil			Natural Gal		
	returns	Volume	RV	returns	volume	RV
Mean	-0.024	101813	0.018	0.005	36806	0.034
St. Dev.	1.273	344712	0.019	1.652	13855	0.027
Min.	-0.069	0	0.000	-0.077	0	0.000
Max.	0.046	391342	0.002	0.076	160567	0.004
Asymmetry	-0.315	2.168	5.507	0.134	1.862	5.169
Exc. Kur.	1.540	10.583	47.144	1.435	8.445	56.293
No. zeros	135	45	45	360	45	45

Note: Mean and St. Dev. multiplied by 100 for returns and RV (= Realized Variances).

Table 6: 5-minute frequency results

	M1	M2	M3	M4	M5
No. of significant intercepts	1	0	1	0	0
No. of significant AR terms (diagonal)	5	6	4	3	0
No. of significant AR terms (off-diagonal)	5	5	5	3	0
No. of significant Vol. Change coefficients		2	2	3	0
No. of significant Overnight coefficients		8	8	6	1
No. of significant Asymmetry coefficients				4	0
No. of coefficients in the system	36	92	92	140	140
Wald	104.57	183.39	625.39	54.70	71.56
P.value	0.00	0.00	0.00	0.24	0.02

Note: M1 refers to the model including only the VAR(1) and VAR(45) lags; M2 adds to M1 the exogenous variables (lagged volume change and overnight variables); M3 adds to M1 the exogenous variables with the overnight variables affecting only the first interval of the day; M4 adds to M2 the asymmetry coefficients; M5 adds to M3 the asymmetric coefficients; the Wald test reported in the two bottom lines is a joint significance test for the VAR(1) and VAR(5) coefficients for M1, a joint significance test on the exogenous variables for M2 and M3, and a joint significance test of the asymmetry coefficients in M4 and M5.

Table 7: 15-minute frequency results

	M1	M2	M3	M4	M5
No. of significant intercepts	1	0	1	0	0
No. of significant AR terms (diagonal)	2	2	2	2	0
No. of significant AR terms (off-diagonal)	0	0	0	2	0
No. of significant Vol. Change coefficients	---	1	1	0	0
No. of significant Overnight coefficients	---	8	8	5	3
No. of significant Asymmetry coefficients	---	---	---	3	0
No. of coefficients in the system	36	92	92	140	140
Wald	58.66	175.11	474.77	43.94	69.74
P.value	0.00	0.00	0.00	0.64	0.02

See Note to Table 6.

Table 8: 45-minute frequency results

	M1	M2	M3	M4	M5
No. of significant intercepts	1	0	1	0	0
No. of significant AR terms (diagonal)	2	3	1	1	0
No. of significant AR terms (off-diagonal)	3	3	3	3	0
No. of significant Vol. Change coefficients	---	1	1	0	0
No. of significant Overnight coefficients	---	9	6	5	5
No. of significant Asymmetry coefficients	---	---	---	1	0
No. of coefficients in the system	36	92	92	140	140
Wald	81.80	196.85	358.22	35.77	69.28
P.value	0.00	0.00	0.00	0.90	0.02

See Note to Table 6.

Table 9: Daily frequency results

	M1	M2	M4
No. of significant intercepts	1	0	0
No. of significant AR terms (diagonal)	1	0	1
No. of significant AR terms (off-diagonal)	0	0	0
No. of significant Vol. Change coefficients		0	0
No. of significant Overnight coefficients		7	6
No. of significant Asymmetry coefficients			1
No. of coefficients in the system	20	76	108
Wald	12.48	288.56	31.73
P.value	0.71	0.00	0.67

Note: M1 refers to the model including only the VAR(1) and VAR(45) lags; M2 adds to M1 the exogenous variables (lagged volume change and overnight variables); M4 adds to M2 the asymmetry coefficients; the Wald test reported in the bottom two lines is a joint significance test for the VAR(1) and VAR(5) coefficients for M1, a joint significance test of the exogenous variables for M2, and a joint significance test of asymmetry in M4.

Table 10: 5-minute frequency results

	H1				H2				H3			
	SP	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**
ω	**0.032***	**0.016***	**0.026***	**0.057***	**0.028***	**0.001***	**0.015***	**0.044***	**0.026***	**0.000***	**0.013***	**0.039***
β_1	**0.955***	**0.982***	**0.951***	**0.856***	**0.945***	**0.981***	**0.932***	**0.845***	**0.944***	**1.003***	**0.924***	**0.841***
β_2	**0.016***	0.008	**0.022***	**0.097***	**0.020***	**0.013***	**0.035***	**0.102***	**0.021***	**-0.003***	**0.045***	**0.105***
α_{1R}	**0.029***				**0.025***				**0.024***	**0.001***	0.000	0.000
α_{1E}		**0.004***				**0.002***			0.001	**0.001***	0.002	0.003
α_{1O}			**0.019***				**0.013***		-0.001	**0.001***	**0.012***	0.000
α_{1G}				**0.038***				**0.034***	**-0.002***	0.000	0.001	**0.034***
α_{2R}	**-0.005***				0.000				0.000	0.000	-0.001	0.000
α_{2E}		**-0.001***				**-0.002***			-0.001	**-0.001***	**-0.004***	**0.005***
α_{2O}			0.000				**0.004***		**-0.002***	**-0.001***	**0.004***	-0.001
α_{2G}				-0.002				-0.001	**-0.002***	-0.001	**-0.004***	-0.001
δ_{1R}	**1.425***				**-0.825***				**-2.258***	**-0.835***	**-2.177***	-0.170
δ_{1E}		**0.049***				**0.104***			-0.018	0.003	**-0.039***	-0.021
δ_{1O}			**0.631***				0.057		**0.362***	**0.137***	**0.250***	**-0.787***
δ_{1G}				**0.332***				**0.428***	**0.156***	**0.039***	-0.019	**0.645***
γ_{1R}					**0.016***				**0.014***	-0.001	**0.006***	0.002
γ_{1E}						**0.007***			0.001	**0.005***	0.001	0.002
γ_{1O}							**0.023***		**0.008***	**0.003***	**0.021***	0.001
γ_{1G}								**0.020***	-0.001	**-0.003***	**0.004***	**0.020***
γ_{2R}					**-0.008***				**-0.006***	-0.001	0.001	**-0.009***
γ_{2E}						**-0.002***			0.000	**-0.005***	0.000	0.000
γ_{2O}							**-0.007***		0.001	-0.001	**-0.005***	0.002
γ_{2G}								**-0.005***	0.002	**0.004***	0.002	**-0.006***
δ_{2R}					**2.160***				**1.911***	**0.557***	**1.682***	0.261
δ_{2E}						**-0.098***			**0.034***	-0.001	0.024	0.014
δ_{2O}							**0.501***		**0.418***	**-0.283***	**0.591***	0.163
δ_{2G}								**-0.286***	-0.049	**-0.097***	0.054	**-0.222***
LL	-14747.4	-21535.1	-19000.2	-19935.7	-14517.3	-20811.6	-18691.9	-19852.5	-14408.3	-20549.8	-18603.5	-19800.4
LR vs H1					0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LR vs H2									0.00	0.00	0.00	0.00

Note: The parameters are defined according to equation (9): ω is the intercept; β_1 and β_2 refer to the one-period and daily GARCH effects, respectively; α_{1i} and α_{2i} refer to the one-period and daily ARCH effects, respectively; i = R,E,O,G, corresponding to S&P 500, Ethanol, Oil and Gas, respectively; δ_{1i} is the impact of overnight volatility and δ_{2i} the overnight asymmetry; γ_{1i} refers to one-period asymmetry, and γ_{2i} refers to daily asymmetry. Over columns, H1 identifies the baseline model with just the overnight volatility impact, while H2 adds asymmetry, with both models excluding spillovers, while H3 includes spillovers. The last two rows contain likelihood ratio tests between the three nested models, H3 being the most general. Bold stars denote statistical significance at the 1% level.

Table 11: 15-minute frequency results

	H1				H2				H3			
	SP	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**
ω	**0.032***	**0.068***	**0.011***	**0.030***	**0.021***	**-0.001***	**0.004***	**0.021***	**0.021***	**0.003***	0.002	**0.018***
β_1	**0.956***	**0.945***	**0.969***	**0.827***	**0.948***	**1.026***	**0.993***	**0.800***	**0.939***	**0.988***	**0.992***	**0.814***
β_2	0.011	**-0.039***	0.014	**0.134***	**0.019***	**-0.026***	0.000	**0.156***	**0.024***	0.012	0.001	**0.146***
α_{1R}	**0.033***				**0.024***				**0.024***	0.001	0.000	0.001
α_{1E}		**0.013***				**0.006***			**0.004***	**0.007***	-0.001	**-0.006***
α_{1O}			**0.020***				**0.010***		**-0.003***	-0.003	**0.010***	**-0.014***
α_{1G}				**0.032***				**0.029***	0.002	0.003	**-0.005***	**0.029***
α_{2R}	**-0.009***				-0.003				0.000	-0.004	0.001	-0.001
α_{2E}		**0.009***				**-0.006***			0.001	**-0.007**	0.000	0.009
α_{2O}			**-0.010***				**-0.006***		-0.002	0.003	**-0.006***	0.008
α_{2G}				**-0.008***				**-0.006***	-0.004	-0.005	0.005	**-0.008***
δ_{1R}	**1.261***				**-1.054***				**-2.096***	-0.296	**-0.697***	**-1.572***
δ_{1E}		**0.070***				0.005			-0.008	**0.011***	-0.002	0.032
δ_{1O}			**0.298***				**-0.056***		**0.416***	**0.159***	0.087	0.021
δ_{1G}				**0.354***				**0.327***	0.043	**0.104***	0.014	**0.382***
γ_{1R}					**0.023***				**0.020***	**0.008***	**0.011***	**0.014***
γ_{1E}						**0.010***			0.001	**0.012***	0.002	0.002
γ_{1O}							**0.019***		0.007	**0.011***	**0.015***	**0.020***
γ_{1G}								**0.019***	**-0.009***	**-0.010***	0.000	**0.013***
γ_{2R}					**-0.013***				**-0.012***	**-0.006***	**-0.010***	**-0.012***
γ_{2E}						**-0.010***			0.000	**-0.013***	-0.002	-0.004
γ_{2O}							**-0.016***		-0.001	-0.006	**-0.013***	-0.004
γ_{2G}								-0.007	0.004	**0.012***	0.001	**-0.008***
δ_{2R}					**2.434***				**2.451***	0.481	**0.976***	**1.484***
δ_{2E}						0.002			0.020	**-0.016***	-0.010	-0.041
δ_{2O}							**0.164***		0.210	**-0.374***	**0.133***	-0.257
δ_{2G}								0.042	-0.102	**-0.193***	-0.050	0.100
LL	-5264.41	-7083.14	-6447.81	-6845.28	-5145.52	-6797.81	-6363.55	-6830.40	-5096.78	-6679.42	-6326.65	-6784.87
LR vs H1					0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LR vs H2									0.00	0.00	0.00	0.00

See note to Table 10.

Table 12: 45-minute frequency results

	H1				H2				H3			
	SP	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**	**SP**	**ET**	**CL**	**NG**
ω	**0.026***	**0.091***	**0.014***	-0.003	0.005	**0.029***	-0.006	0.003	-0.006	**0.033***	**-0.027***	-0.006
β_1	**0.981***	**0.718***	**0.871***	**1.090***	**0.894***	**0.800***	**0.643***	**0.874***	**0.867***	**0.797***	**0.554***	**1.023***
β_2	-0.020	**0.178***	0.098	**-0.120***	0.053	**0.118***	**0.311***	0.079	0.069	**0.119***	**0.409***	-0.055
α_{1R}	**0.032***				**0.022***				**0.022***	**0.027***	-0.002	**0.008***
α_{1E}		**0.035***				**0.017***			-0.002	**0.015***	-0.011	-0.009
α_{1O}			**0.012***				0.010		-0.008	**-0.017***	0.004	0.009
α_{1G}				**0.010***				**0.019***	0.003	0.002	0.006	**0.010***
α_{2R}	-0.009				-0.002				-0.001	0.005	0.003	-0.009
α_{2E}		0.014				0.010			0.003	0.010	-0.006	0.001
α_{2O}			0.004				0.004		-0.001	-0.018	0.000	-0.006
α_{2G}				-0.001				-0.002	-0.009	-0.001	0.002	-0.001
δ_{1R}	**1.056***				-0.439				**-1.143***	-0.930	**-3.696***	0.267
δ_{1E}		**0.047***				**0.144***			0.011	**0.141***	-0.003	-0.001
δ_{1O}			**0.248***				-0.099		**0.342***	0.231	0.105	**0.228***
δ_{1G}				**0.206***				**0.225***	0.067	-0.123	-0.046	**0.194***
γ_{1R}					**0.041***				**0.038***	0.020	**0.057***	-0.003
γ_{1E}						**0.048***			0.005	**0.048***	0.013	0.002
γ_{1O}							**0.014***		0.017	-0.008	0.000	**0.020***
γ_{1G}								0.007	-0.016	-0.016	0.010	-0.001
γ_{2R}					-0.011				-0.008	0.000	**-0.030***	0.006
γ_{2E}						**-0.019***			0.010	**-0.020***	-0.013	-0.008
γ_{2O}							**0.019***		-0.005	0.008	**0.023***	-0.009
γ_{2G}								-0.003	0.013	0.011	-0.001	0.002
δ_{2R}					**2.106***				**2.207***	-0.556	**3.141***	-0.553
δ_{2E}						**-0.145***			-0.002	**-0.143***	-0.038	0.003
δ_{2O}							**0.483***		0.087	-0.118	**0.729***	**-0.213***
δ_{2G}								0.106	-0.144	-0.046	0.056	0.053
LL	-2002.93	-2312.47	-2273.07	-2290.81	-1969.65	-2255.95	-2242.10	-2290.49	-1946.05	-2232.19	-2188.58	-2261.65
LR vs H1					0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.00
LR vs H2									0.00	0.00	0.00	0.00

See note to Table 10.

Table 13: Daily frequency results

	H1				H2				H3			
	SP	ET	CL	NG	SP	ET	CL	NG	SP	ET	CL	NG
ω	**-2.171***	**-1.723***	**-6.318***	**-3.643***	**-1.913***	**-2.297***	**-5.430***	**-3.058***	-3.450	**-2.372***	**-5.111***	**-3.403***
β_1	**0.755***	**0.717***	0.261	**0.633***	**0.797***	**0.654***	**0.373***	**0.687***	**0.621***	**0.584***	**0.372***	**0.657***
α_{1R}	**0.029***				**0.025***				0.021	0.013	0.021	-0.025
α_{1E}		**0.071***				**0.071***			0.010	**0.058***	0.019	**0.031***
α_{1O}			**0.064***				**0.055***		0.025	0.017	**0.044***	-0.003
α_{1G}				**-0.032***				**-0.032***	-0.003	0.031	0.004	**-0.039***
δ_{1R}	**2.519***				**2.587***				1.427	-1.020	-1.590	0.233
δ_{1E}		0.000				**0.176***			0.035	**0.184***	**-0.152***	**-0.106***
δ_{1O}			**0.898***				**1.376***		0.847	0.334	**1.477***	0.139
δ_{1G}				**0.716***				**0.536***	-0.160	**-0.259***	-0.005	**0.569***
γ_{1R}					**-0.021***				**-0.019***	**-0.016***	**-0.017***	0.001
γ_{1E}						0.001			-0.010	0.004	0.001	0.001
γ_{1O}							-0.006		0.007	0.005	-0.002	0.002
γ_{1G}								**0.017***	0.007	0.008	0.003	**0.015***
δ_{2R}					-0.189				-0.489	-1.219	1.039	-0.539
δ_{2E}						-0.186			-0.036	**-0.184***	0.012	**0.081***
δ_{2O}							**-0.663***		0.136	0.196	-0.626	0.034
δ_{2G}								0.043	0.116	0.047	-0.056	0.033
LL	4437.90	3814.79	4100.89	3651.62	4451.83	3831.16	4105.78	3652.09	4466.81	3843.12	4119.97	3661.75
LR vs H1					0.00	0.00	0.02	0.82	0.00	0.00	0.01	0.50
LR vs H2									0.04	0.16	0.06	0.37

See note to Table 10. It should be noted that the one-period and daily effects now collapse to a single coefficient.

Table 14: Unconditional correlations

	S&P	Ethanol	Crude Oil	S&P	Ethanol	Crude Oil
	Daily frequency			45-minute frequency		
Ethanol	0.0518			0.0305		
Crude Oil	0.3792	0.1112		0.4599	0.0301	
Natural Gas	0.0536	0.0715	0.1091	0.0651	0.0214	0.1029
	15-minute frequency			5-minute frequency		
Ethanol	0.0029			-0.0044		
Crude Oil	0.4536	0.0130		0.4542	0.0056	
Natural Gas	0.0530	0.0168	0.0810	0.0492	0.0101	0.0790

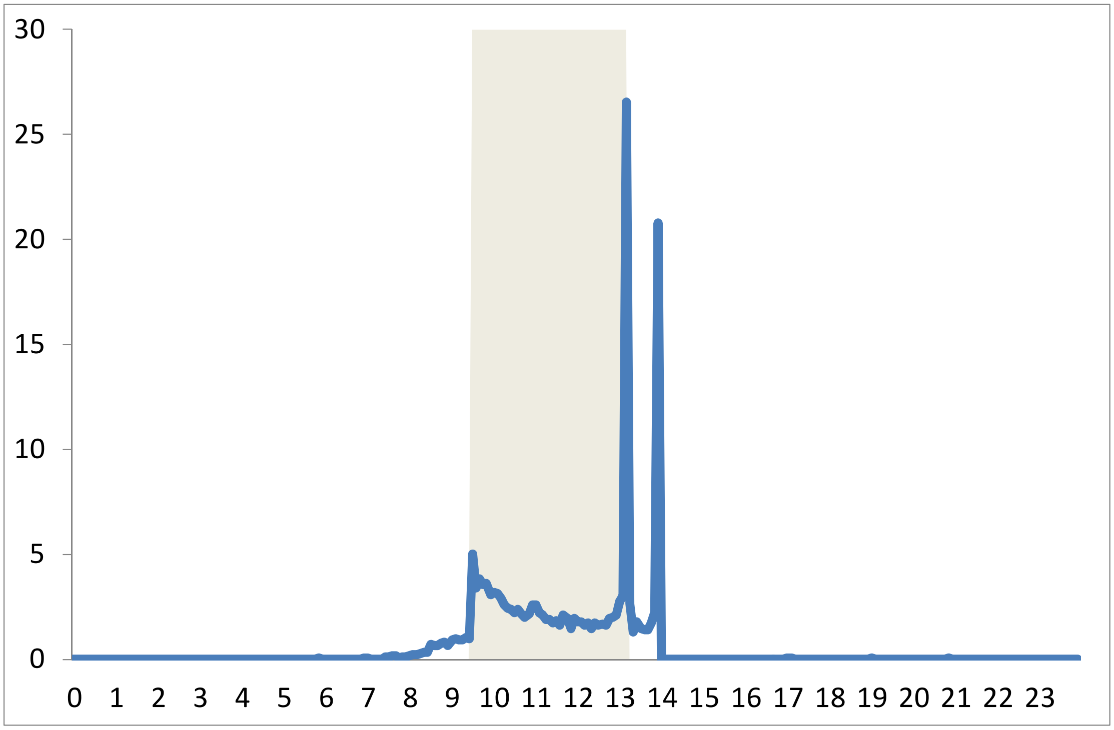

Figure 1: Average 5-minute ethanol futures volume (shaded area denotes 9:35-13:15)

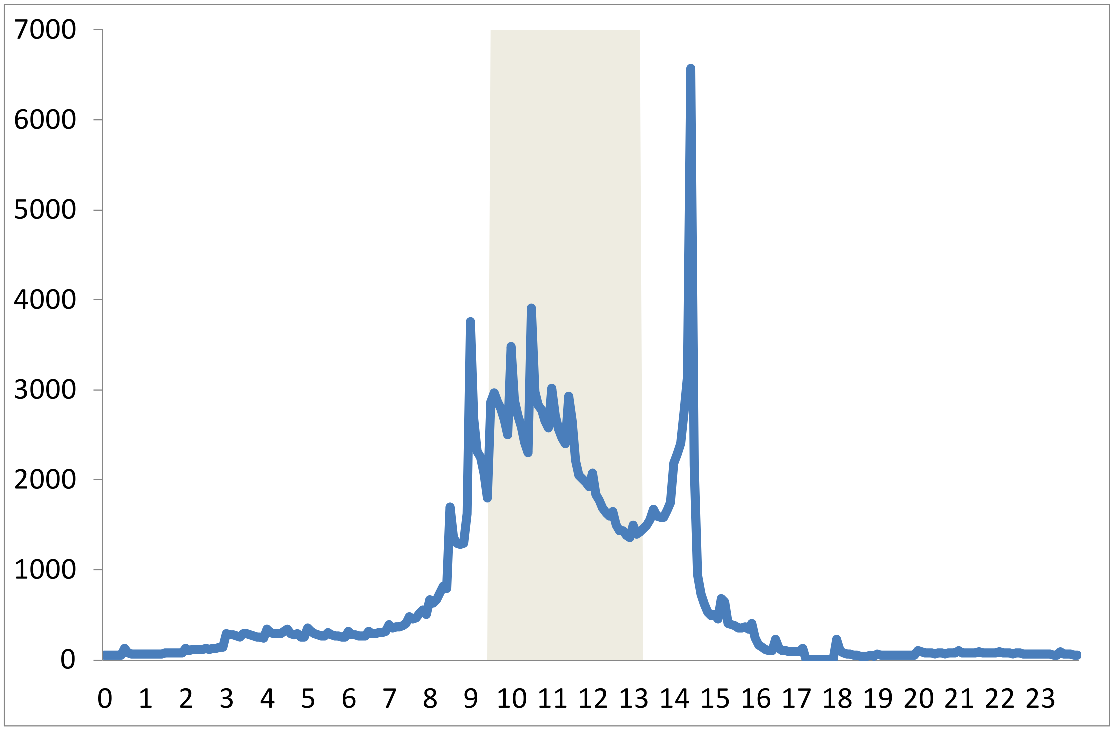

Figure 2: Average 5-minute crude oil futures volume (shaded area denotes 9:35-13:15)

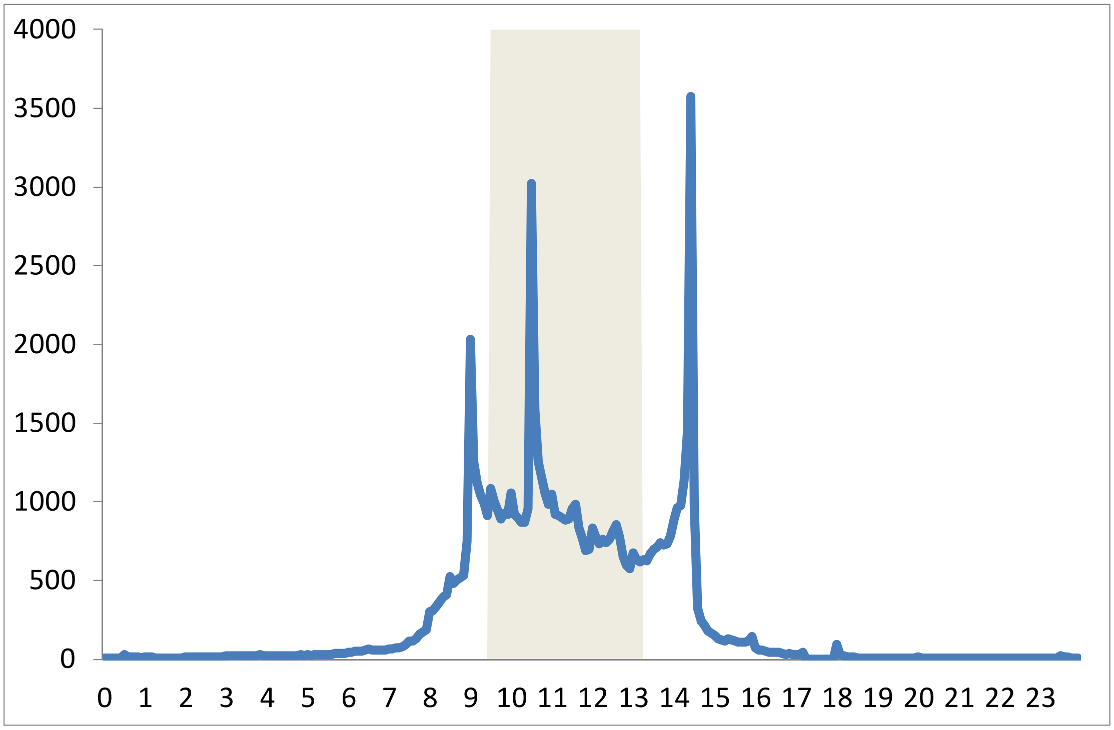

Figure 3: Average 5-minute natural gas futures volume (shaded area denotes 9:35-13:15)

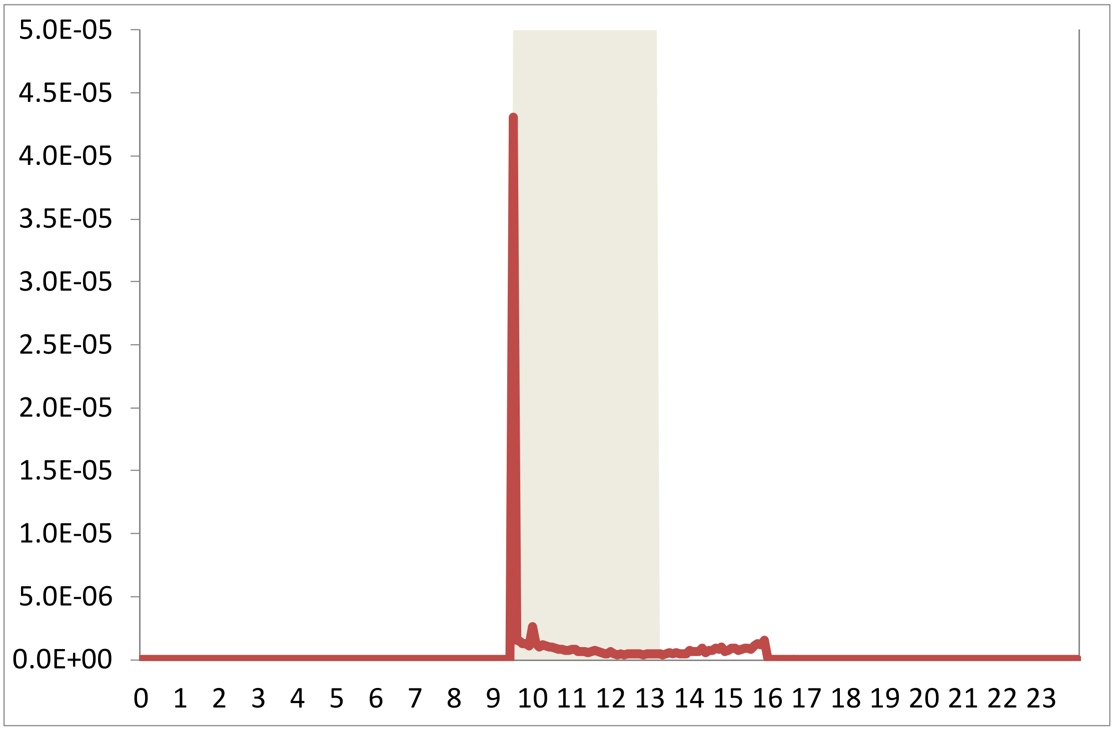

Figure 4: Average squared 5-minute S&P500 index returns (shaded area denotes 9:35-13:15)

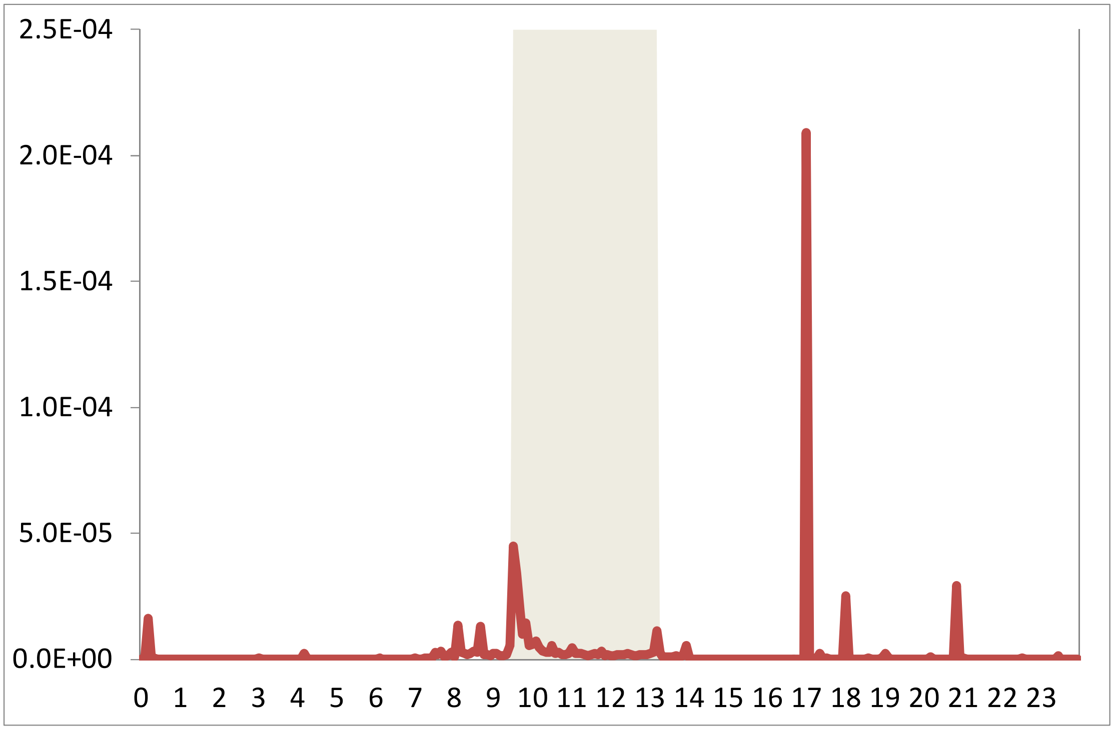

Figure 5: Average squared 5-minute ethanol futures returns (shaded area denotes 9:35-13:15)

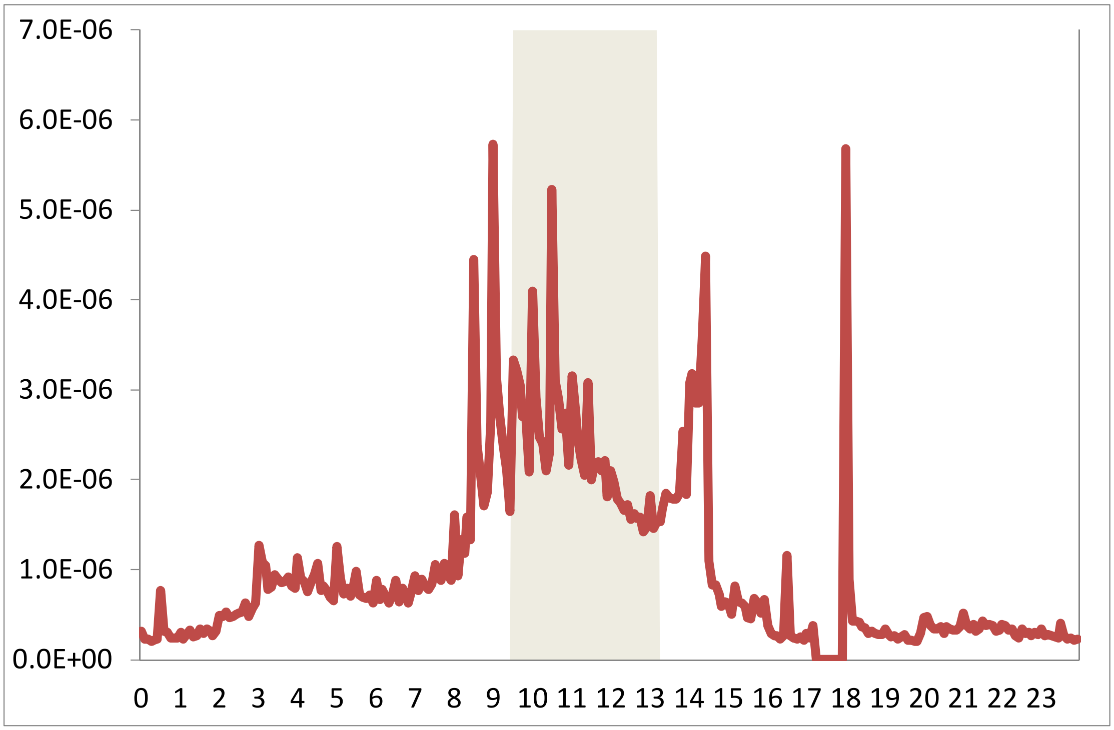

Figure 6: Average squared 5-minute crude oil futures returns (shaded area denotes 9:35-13:15)

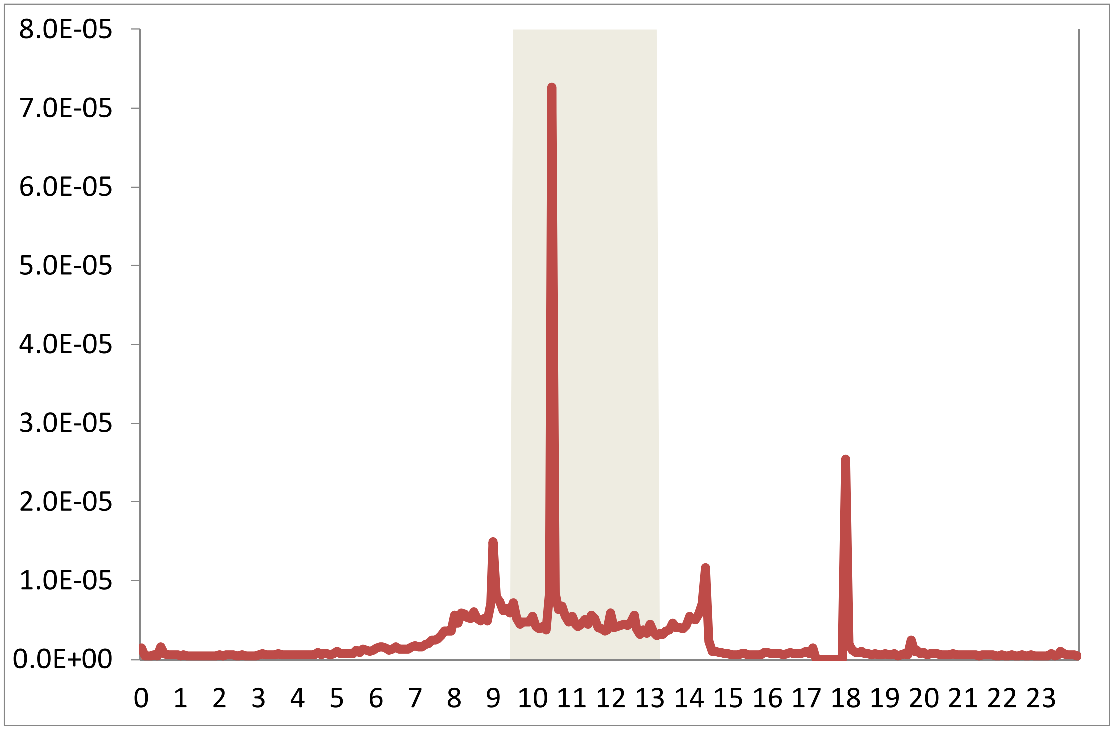

Figure 7: Average squared 5-minute natural gas futures returns (shaded area denotes 9:35-13:15)

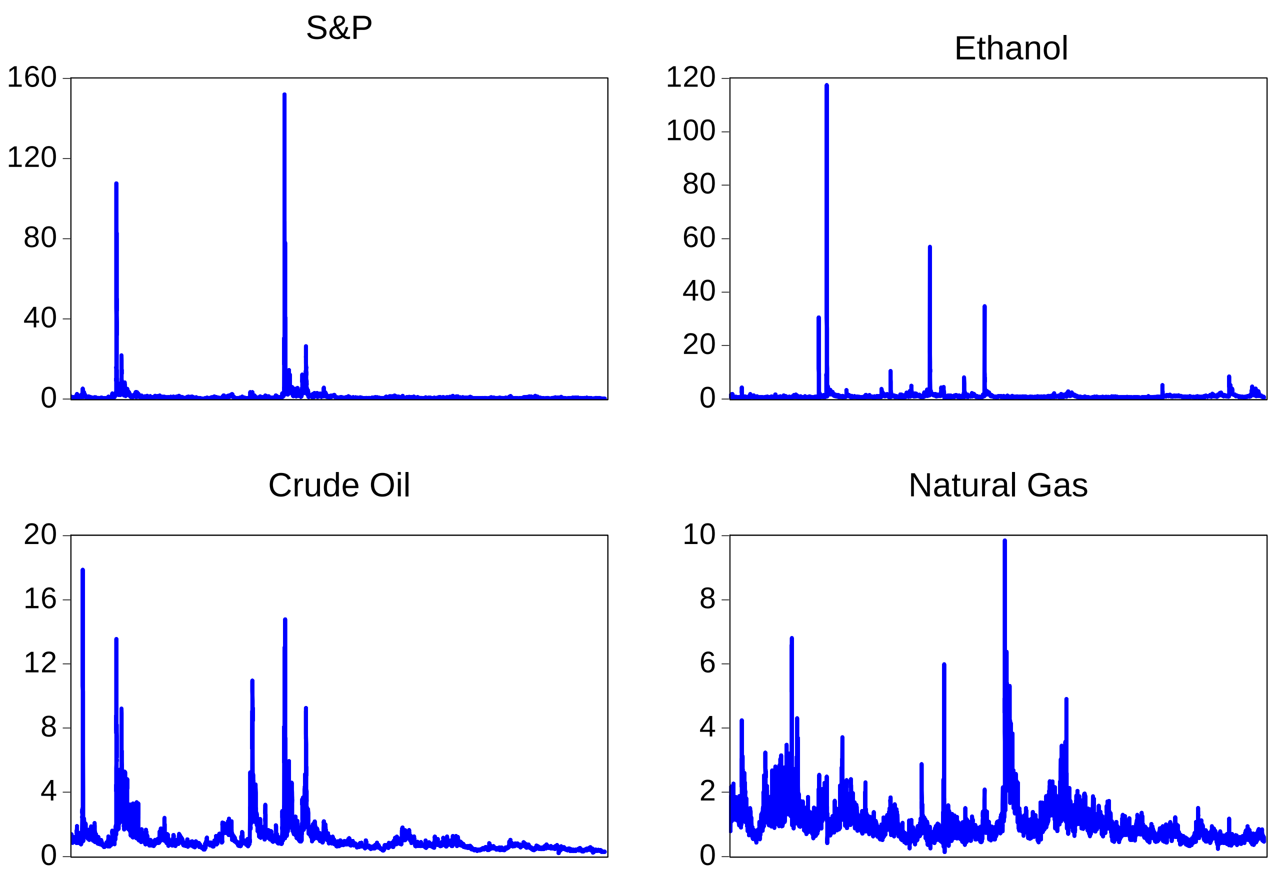

Figure 8: Conditional variances of model H3 in log-scale at the 15-minute frequency

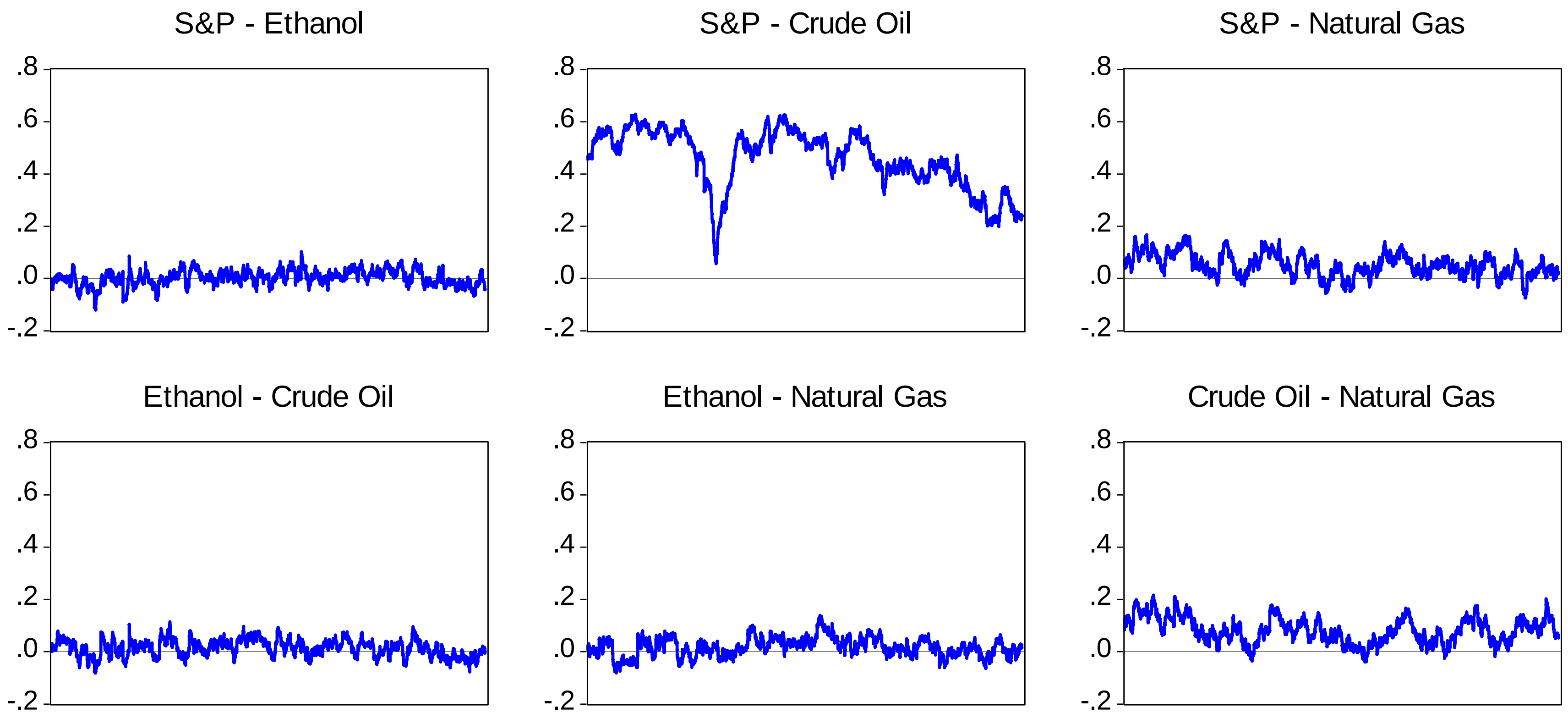

Figure 9: Conditional correlations of scalar DCC at the 15-minute frequency