Matteo Manera,[a] Marcella Nicolini,[b*] and Ilaria Vignati [c]

Futures price volatility in commodities markets: The role of short term vs long term speculation

Abstract: This paper evaluates how different types of speculation affect the volatility of commodities' futures prices. We adopt four indexes of speculation: Working's T, the market share of non-commercial traders, the percentage of net long speculators over total open interest in future markets, which proxy for long term speculation, and scalping, which proxies for short term speculation. We consider four energy commodities (light sweet crude oil, heating oil, gasoline and natural gas) and six non-energy commodities (cocoa, coffee, corn, oats, soybean oil and soybeans) over the period 1986-2010, analyzed at weekly frequency. Using GARCH models we find that speculation is significantly related to volatility of returns: short term speculation has a positive and significant coefficient in the variance equation, while long term speculation generally has a negative sign. The robustness exercise shows that: i) scalping is positive and significant also at higher and lower data frequencies; ii) results remain unchanged through different model specifications (GARCH-in-mean, EGARCH, and TARCH); iii) results are robust to different specifications of the mean equation.

JEL Codes: C32; G13; Q11; Q43.
Keywords: Commodities futures markets; Speculation; Scalping; Working's T; Data frequency; GARCH models

---

[a] University of Milan-Bicocca, Milan, and Fondazione Eni Enrico Mattei, Milan. E-mail: matteo.manera@unimib.it
[b] University of Pavia, Pavia, and Fondazione Eni Enrico Mattei, Milan. E-mail: marcella.nicolini@unipv.it
[c] Fondazione Eni Enrico Mattei, Milan. E-mail: ilaria.vignati@feem.it
*Corresponding author

The authors would like to thank Bahattin Buyuksahin, Marzio Galeotti, Alessandro Lanza, Eduardo Rossi, Giovanni Urga and seminar participants at FEEM, University of Milan-Bicocca, 51st Meeting of the EWGCFM and 1st Conference of the Research Centre for Energy Management (London, 2013), FEEM Workshop on "Oil price forecasts and trends" (Milan, 2013) and 6th International Workshop on Empirical Methods in Energy Economics EMEE (Ottawa, 2013) for helpful comments.

Futures price volatility in commodities markets: The role of short term vs long term speculation

1. Introduction

Financial markets have faced a number of significant changes in the last decade. Commodities' prices grew dramatically during the first years of the 2000s and speculators often have been alleged to influence their levels and drive their increases (Masters, 2008). A related issue is whether speculators' activity affects the volatility of futures prices. On the one hand, speculators increase market liquidity thus reducing price volatility. On the other hand, critics argue that an increasing trading volume, especially by speculators, positively affects volatility. While recent empirical analysis suggests that financial speculation generally does not influence the returns of commodities (e.g. Manera et al., 2013), the evidence of its impact on volatility is lagging behind.

Our paper fills this gap by investigating the role of short and long term speculation over the period 1986-2010 in ten futures markets. In particular, we contribute to the literature in at least three different directions. First, we use different measures of speculation. Second, we analyze data at different frequencies. Third, we adopt alternative specifications for the volatility of futures returns.

Traditionally, the literature has measured excess speculation by means of the Working's T (1960) index, which is based on the relative weight of speculators and hedgers in the market. Alternative measures, based on the distinction between hedgers and speculators, are the market share of non-commercial traders and the percentage of net long speculators over total open interest. All these measures require a classification of agents between the two categories, which is provided by the U.S. Commodity Futures Trading Commission (CFTC). The measures based on this classification refer to a specific type of speculation, position trading, which seeks profits over a sizeable time span.

We are interested in investigating also the role of short term speculators, i.e. scalpers and day traders, whose typical market actions are aimed at obtaining a profit from the small price gaps created by the bid-ask spread. Scalpers, as long term speculators, are not interested in contracts for their physical content, but trade paper contracts to gain a margin from small changes in prices. In this sense, the scalping variable proxies for short term speculation, which seeks immediate profits, and it differs from the other three indexes, which could be considered as proxies for long term speculation.

We test if these different measures significantly affect the volatility of commodities' prices using Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. We include macroeconomic controls in the mean equation, such as returns on the T-bill, the Standard & Poor' s 500 (S&P) returns and the junk bond yield, as well as a speculative index in the variance equation. We find that speculation is significantly related to price volatility in the period 1986-2010. More precisely, scalping has a positive and significant coefficient in the variance equation, suggesting that short term speculation actually increases the noise in the information formation process, thus positively affecting volatility. The other three indexes have a negative effect (when significant), thus suggesting that long term speculation does not destabilize prices. Our results are in line with previous evidence (Peck, 1981; Streeter and Tomek, 1992) and more recent contributions (Brunetti et al., 2011).

We test the robustness of these results moving along several dimensions. First, we investigate if our results are robust across different data frequencies, finding that the scalping index is always significant and positive also at higher and lower frequency of data. Second, our findings remain unchanged through more refined model specifications, such as GARCH-in-mean, threshold GARCH and asymmetric exponential GARCH. Third, we investigate if results are affected somehow by the correct specification of the mean equation. Focussing on crude oil, we find that the inclusion of controls for the demand, production and stocks does not affect the main findings.

The reminder of the paper is structured as follows: Section 2 discusses the relevant literature, Section 3 presents the data, Section 4 illustrates the econometric specification, while the results are presented and discussed in Section 5. Finally, Section 6 concludes.

2. Literature review

The way speculators can influence markets is the object of a vast literature. In principle, the presence of speculators (i.e. agents that buy or sell an asset because its price is expected to change) is fundamental to the efficient operation and stability of markets. As Smith (2009) points out, "Speculation is not price manipulation, but is sometimes used to exploit efforts to manipulate prices by other means. In such cases, it is the manipulation of prices that is objectionable, not speculation, per se" (p. 26). Thus, the role of speculators might be stabilizing or destabilizing and understanding their behaviour and how it affects returns and volatility is extremely important. This issue has been debated extensively in literature. On one side, some authors suggest that the participation of speculators, which are considered uninformed traders, lowers the quality of information in the futures market, and might have a destabilizing effect on prices, thus increasing volatility (Stein 1987). Hart and Kreps (1986) show that, even in a general equilibrium with optimizing speculators, prices can be destabilized. On the other, speculators are supposed to bring efficiency to price predictions, lowering volatility. In particular, Friedman (1953) suggests that rational speculation stabilizes prices, Powers (1970) shows that speculative activity of futures traders reduces the random component of price variation, while Cox (1976) suggests that speculation increases the information content of prices. More recently, Brunetti et al. (2011) find that speculative trading activity reduces volatility, Alquist and Gervais (2013) support the view that oil price increases are explained by a series of positive demand shocks emanating from emerging countries, whereas Manera et al. (2013) show that financial speculation is poorly significant in modelling commodity returns.

Speculative activity can be classified according to its time perspective, i.e. the length of time a position is held. When positions are held for a matter of seconds, we have scalping activity. If the positions are closed by the end of the trading session, we have day traders. Both actions aim at gaining immediate profits and might be defined as short term speculation. When positions are held for longer time periods we have long term speculation.

Working's T index is the most widely adopted measure for long term speculation in literature. It quantifies the excess of speculation relative to hedging based on position data provided by commitments of traders (COT) data from the U.S. Commodity Futures Trading Commission (CFTC). Recently, Till (2009), Sanders et al. (2010) and Sanders and Irwin (2013) have shown that speculative positions in energy and agriculture U.S. futures markets are not excessive relative to hedging activity. The market share of non-commercial traders on total open interest is used in Büyükşahin and Robe (2010) to show that the composition of traders in futures markets helps explain the linkages between equity and commodity returns. The authors find that hedge funds increase the equity-commodity return correlations, while swap dealers, index traders, commercial traders, etc., do not influence the correlations. As net long positions of traders are concerned, some authors (Brunetti et al., 2011; Medlock and Jaffe, 2009; Büyükşahin and Harris, 2011; Irwin and Sanders, 2012) adopt the difference between long and short positions held by non-commercial traders. Others adopt this difference relative to total open interest (Hedegaard, 2011) or relative to open interest held by non-commercial traders (Brunnermeier et al., 2008; Sanders et al., 2010).[1] Net long positions are usually employed because speculators go mostly long on futures contracts (they buy the risk of hedgers traders), hence this measure is considered a good proxy to detect non-commercial traders. Moreover, net long positions of speculators have increased in commodity markets after 2004, especially in the oil market (Irwin and Sanders, 2010; Khan, 2009; Medlock and

[1] This is the so called "speculative pressure" (see De Roon et al., 2000, and Sanders et al., 2004).

Jaffe, 2009), leading to allegations that these positions have pushed prices up (Medlock and Jaffe 2009).

Short term speculation refers to different phenomena, such as scalping and day trading. Scalping is an intraday activity, made up of instant transactions by traders which open and close contract positions within a very short period of time to realize profits (Cornell, 1981; Du et al., 2011; Peck, 1981; Working, 1967). Scalpers are typically intended as types of traders who dart in markets even hundreds of times a day to make profits: they "[…] stand willingly to buy a tick below the last trade or sell a tick above it" (Cornell, 1981, p. 305) and, again, "scalpers trade price ticks, holding a position for a matter of moments anticipating the last price change will be followed by an opposite price move" (Roswell and Purcell, 1992, p. 206). Scalping is generally proxied as the ratio of volume to open interest (Du et al., 2011; Leuthold, 1983; Peck, 1981; Streeter and Tomek, 1992).[2] Short term speculators are more likely to close a contract within a day than hedgers, whose orders are generally held for more than one day. For this reason, changes in daily trading volume over open interest might be primarily interpreted as a reflection of speculative activity. Streeter and Tomek (1992) consider monthly data on soybeans future prices and find a positive and significant impact of scalping on prices volatility. Chatrath et al. (1996) find a positive relationship between the ratio of volume to open interest and exchange rate volatility. Luu and Martens (2003) find a positive and significant relationship between the volume to open interest ratio and volatility in the context of the Mixture of Distribution Hypothesis (MDH) for S&P 500 Index future contracts.[3] Robles et al. (2009) investigate speculative activity in four agricultural future markets in the 2000s', finding that past changes in scalping index help forecast changes in the price of wheat and rice. Du et al. (2011) analyze the role of speculation in driving crude oil price spike of 2008. Adopting a stochastic volatility model with Merton jumps in the weekly returns on crude oil future prices from 1998 to 2009, they find that both scalping and Working's T index have a significant positive impact on price volatility.

Overall, previous research finds that long term speculation (proxied by Working's T index) has a negative impact on price variability, while short term (measured by the ratio of volume to open interest) has a positive impact (Peck, 1981; Roswell and Purcell, 1992; Streeter and Tomek, 1992). These studies, however, use monthly data and approximate volatility through the average daily price range. Conversely, our work is novel in several respects. First, it uses different measures of speculation. Second, it analyzes data at higher frequencies. Third, it adopts different specifications for the volatility of futures returns.

[2] Open interest is the total number of contracts not yet offset by a transaction.

[3] The MDH analyzes the relationship between trading activity and price volatility, assuming they are correlated as being influenced by the same information arrival process (see also Andersen, 1996, and Tauchen and Pitts, 1983).

3. Data description

We collect data of futures prices for four energy commodities (light sweet crude oil, heating oil, gasoline and natural gas) and six non-energy commodities (cocoa, coffee, corn, oats, soybean oil and soybeans).[4] Daily (5 days) data on futures prices[5] for each commodity are obtained from Datastream for the period 1986-2010.[6] Data on position traders are publicly available at weekly frequency from the U.S. Commodity Futures Trading Commission (CFTC).

We measure speculation using four different indexes: Working's T index, the market share of non-commercial traders, the ratio of net long speculators over total open interest and scalping. We first discuss the measures for long term speculation.

Working's T index proxies the excess of speculation relative to hedging activity. This index is calculated as the ratio of non-commercial positions to total commercial positions:

$$\begin{cases} 1+\dfrac{SS}{HS+HL} & if \quad HS \geq HL \\ 1+\dfrac{SL}{HS+HL} & if \quad HS < HL \end{cases} \tag{1}$$

where *SS* is the number of positions held by speculators who are short, *SL* is speculation long, *HS* is hedging short and *HL* is hedging long. It should be noted that the calculation of the Working's T index crucially depends on the classification of the market operators between hedgers and speculators. CFTC also provides data for "Non-Reportable" agents,[7] which are not classified into any of the two categories. However, open interest held by these subjects should be included in the computation of the index. Several rules to treat non-reportables are at hand. One could consider them as being all hedgers or, more likely, all speculators. Indeed, hedgers are generally known by CFTC and are less likely to be among non-reportables. We follow an intermediate approach, assuming that 70% of them are speculators and 30% are hedgers.

---

[4] All energy commodities are traded on the New York Mercantile Exchange, while non-energy commodities are traded on the New York Board of Trade (cocoa and coffee) and the Chicago Board of Trade (corn, oats, soybean oil and soybeans).

[5] We use the continuous futures price series, calculated by Thomson Financial. Those series start at the nearest contract month, which forms the first value for the continuous series and switches over on 1st day of new trading month.

[6] The detailed description of the variables is presented in Table A.1 in the Statistical Appendix available from the authors upon request.

[7] CFTC defines this category as follows: "The long and short open interest shown as Non Reportable Positions is derived by subtracting total long and short Reportable Positions from the total open interest. Accordingly, for Non Reportable Positions the number of traders involved and the commercial/non-commercial classification of each trader are unknown." (see http://www.cftc.gov/MarketReports/CommitmentsofTraders/ExplanatoryNotes/index.htm).

As proposed by Büyükşahin and Robe (2010), we compute the market share of non-commercial traders as the average of the long and short positions of all non-commercial (or speculators) traders on the total open interest in that market:

$$\frac{SL + SS}{2 * OI} \tag{2}$$

where *OI* is the total open interest. In treating non-reportable positions, we follow the same approach used for Working's T.

The last measure of long term speculation is the ratio of net long speculative positions over total open interest. As in Hedegaard (2011), it is defined as:

$$\frac{SL - SS}{OI} \tag{3}$$

where non-reportable are treated as discussed above. This is a measure of the extent to which speculators are long or short in aggregate: if it is positive (negative), speculators go long (short) in futures markets. We adopt index (3) for two reasons. First, it is a "relative" measure, hence it is directly comparable with the other indexes. Second, it is highly correlated (0.92) with the measure of "speculative pressure" (see footnote 1).

As for the measure of short term speculation we adopt the ratio of volume to open interest:

$$\frac{VO}{OI} \tag{4}$$

Daily data, sourced from Datastream, do not allow us to disentangle between scalping and day trading, but this measure is able to grasp both activities.

To control for macroeconomic factors we follow, among others, Chevallier (2009) and Manera et al. (2013) and we collect daily (5 days) data on Moody's Aaa and Baa corporate bond yield, 3-month Treasury bill and S&P 500 index over the period 01/02/1986 - 12/31/2010 from Federal Reserve Economic Data (FRED) provided by the Federal Reserve of St. Louis.[8] For all these series we consider weekly averages of the daily data: this is the highest frequency which allows to compare results on commodities returns among the four speculative measures we have adopted.

[8] From the Federal Reserve of Philadelphia we also have retrieved the Aruoba-Diebold-Scotti (ADS) index, which is a measure of real business condition (see http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index for further details).

Descriptive statistics of speculation indexes for weekly data are reported in Table 1.[9] Since futures prices contain a unit root,[10] to obtain stationarity we consider the return r_{it}, which is defined as $\log(P_{it} / P_{it-1})$, where P_{it} and P_{it-1} are the prices of commodity i at time t and $t-1$, respectively. The same transformation is applied to the macroeconomic variables, while the speculation indexes are stationary in levels and are not transformed.[11]

[TABLE 1 ABOUT HERE]

The first panel of Table 1 shows that the commodities with the average highest values of the speculation index measured by scalping are gasoline (0.339), crude oil (0.326) and soybeans (0.310), while the lowest mean value is that of cocoa (0.104). Moreover, on the whole sample scalping scores the maximum values of 1.002 (soybeans) and 0.939 (natural gas). In the second panel of Table 1 we observe that Working's T index ranges, on average, from 1.105 (gasoline) to 1.268 (soybeans), while its maximum value is larger than 1.5 (natural gas and oats). The panel presenting the statistics for the market share of non-commercial traders shows that the highest mean values are those of soybeans (0.362), oats (0.355) and corn (0.345) and the lowest is that of natural gas (0.187), indicating that there are more speculative traders in former markets. Finally, we present the descriptive statistics of net long speculative positions. On average, non-commercial are net long in aggregate, since all values are positive. Moreover, oats market has the highest mean value (0.282) and crude oil has the lowest (0.009). However, the minimum values reported reveal that speculators vary their positions over time going also short. These four measures of speculation lead to different results, but they all agree in identifying agricultural markets as those with more speculative activity.

---

[9] In the Statistical Appendix, which is available from the authors upon request, descriptive statistics are reported for all the variables of interest at different frequencies (Tables A.2.a, A.2.b and A.2.c).

[10] Figure A.1 in the Statistical Appendix reports the behaviour of future prices at daily frequency (the highest frequency available in data) over the time period considered. In each graph, the series show a non-stationary behaviour, as well as an evident spike in prices in 2008. See also the ADF tests in Tables A.2.a (and A.2.b, A.2.c) in the Statistical Appendix.

[11] When the ADF test indicates the presence of a unit root (see the first panel of Table 1), we control the associated p-value and, if it is close to 0.05, we differentiate the series to obtain stationarity.

4. The econometric specification

After testing the stationarity of all the series, we estimate a model where the returns of each commodity *i* at time *t* depend on two sets of explanatory variables, namely macroeconomic and speculative factors:

$$
\begin{aligned}
r_{it} &= \alpha_0 + \alpha_1 int\_rate_t + \alpha_2 junk\_bond\_yield_t + \alpha_3 S\&P_t + \alpha_4 speculation\_ST_{it} \\
&+ \alpha_5 speculation\_LT_{it} + \varepsilon_{it}
\end{aligned}
\tag{5}
$$

In equation (5) the macroeconomic factors are represented by the returns of 3-month Treasury bill (int_rate_t), the junk bond premium ($junk\_bond\_yield_t$), defined as the difference between Baa and Aaa corporate bond yield, and the returns of S&P 500 index ($S\&P_t$). The short term speculation variable ($speculation\_ST_{it}$) is represented by the scalping index for the market *i* at time *t*, considered alone or associated with one long term speculation variable ($speculation\_LT_{it}$) such as Working's T, share of non-commercial or net long speculators. The estimation period for all eleven markets spans from 1986:w1[12] to 2010:w52.

We first estimate the model using Ordinary Least Squares (OLS) and test for autoregressive conditional heteroskedasticity (ARCH) effects in the residuals. If these effects are present, we move to GARCH specification, including an autoregressive term of order *p*, AR(*p*), when necessary. We aim at evaluating if speculation directly affects the volatility of returns, thus we consider speculation variables as exogenous regressors in the variance equation of the GARCH models. Given the excess of kurtosis which is present in the data, we choose a conditional Student's t density distribution for the error terms. Therefore, we end up estimating a model where the conditional mean equation is:

$$
r_{it} = \gamma_0 + \gamma_1 int\_rate_t + \gamma_2 junk\_bond\_yield_t + \gamma_3 S\&P_t + \gamma_4 r_{it-1} + \varepsilon_{it} \tag{6.a}
$$

with an AR(*p*) error term if the null hypothesis of absence of residual autocorrelation is rejected by the data. The conditional variance is defined as:

$$
\sigma_{it}^2 = \alpha_0 + \sum_{j=1}^{p_i} \beta_j \varepsilon_{it-j}^2 + \sum_{j=1}^{q_i} \omega_j \sigma_{it-j}^2 + \delta speculation\_ST_{it} + \phi speculation\_LT_{it} \tag{6.b}
$$

[12] For natural gas and heating oil, the estimation sample starts form 1990:w14 and 1986:w22, respectively.

where the variance σ_{it}^2 of the regression model's disturbances is a linear function of lagged values of the squared regression disturbances, of its past value and of measures of speculation, p defines the order of the ARCH term, and q of the GARCH term. Values for p and q are chosen depending on the outcome of residual tests (ARCH-LM test and correlogram on squared residuals). Short run speculation is proxied with the scalping index, while long run speculation is modelled using the Working's T index, the market share of non-commercial traders, and net long speculative positions. In Section 5.1 we present the results at weekly frequency. To check the robustness of our findings, we replicate the same analysis on different data frequencies and using different econometric techniques in Section 5.2.

We choose not to consider a multivariate GARCH approach as the impact of speculative activity is likely to be mostly relevant on the price volatility of its own market only. Indeed, a preliminary analysis of the conditional correlations shows that these are generally low and suggests to keep an univariate approach. [13]

5. Results

5.1 Main results

Table 2 shows the results obtained when short term speculation alone is included in the variance equation.[14] We estimate the model using OLS and then test for ARCH effects using a standard Lagrange multiplier test (not reported). For all commodities reported in Table 2, this test suggests the presence of ARCH effects in the residuals of the estimated model. Thus, we move to a GARCH(p,q) specification. Generally, p=q=1 is the preferred lag order, but there are some exceptions like cocoa, coffee, and soybean oil, where we adopt ARCH(2,0), ARCH(3,0) and ARCH(1,0), respectively. Additionally, the Ljung-Box test (not reported) on the GARCH(p,q) model shows that the residuals contain autocorrelation up to order 1. Introducing an AR(1) term in the models generally removes autocorrelation. The variance equation shows that the ARCH (β) and GARCH (ω) terms are always statistically significant. In particular, the ARCH estimates are generally small (between 0.114 for gasoline and 0.250 for heating oil) and the GARCH estimates are generally high and close to one (see for example 0.824 in the gasoline equation). This indicates that a shock in the volatility series impacts on futures volatility over a long horizon. The only exceptions are represented by heating oil and oats, which show lower GARCH estimates. In the

[13] The constant conditional correlation matrix is available upon request.

[14] We also estimated a model with Working's T index, market share of non commercial traders and net long speculative positions in isolation. Long term speculation indexes have generally a negative sign or they are not significant. These results are not reported but they are available upon request.

mean equation the only variable which significantly affects the returns across the commodities is the S&P 500 index: when it is significant, it is always positive, suggesting that returns are pro-cyclical. The T-bill is poorly significant and positive (apart from cocoa) and junk bond yield is never significant (apart from gasoline where it is poorly significant).[15] The speculation index in the variance equation is always small and significant. In particular, it is always positive, indicating that an increase of short term speculation, proxied by an increase in the ratio of volume to open interest, corresponds to an increase, although small, in the volatility of commodity futures returns.

[TABLE 2 ABOUT HERE]

The second set of estimates introduces the Working's T index in the variance equation, together with the scalping index. Results are presented in Table 3. The short term speculation variable remains positive and significant (apart from gasoline, where it is not significant). The Working's T index, instead, is generally negative and significant (only for cocoa it is not significant), meaning that speculation is associated with reduced volatility of commodities futures prices. This result is in line with the strand of literature which finds that long term speculation has the stabilizing effect of smoothing the price process (Brunetti et al. 2011). As far as the macroeconomic variables and the GARCH specification are concerned, we obtain similar results.

[TABLE 3 ABOUT HERE]

The third set of results considers the market share of speculators associated with the scalping index and is reported in Table 4. The market share of non-commercial traders exhibits a similar behaviour to the Working's T index, showing a negative coefficient in the variance equation. This is not surprising, given the high and positive correlation between these two measures.[16] We still find a positive impact of scalping on volatility.

[TABLE 4 ABOUT HERE]

The last set of estimates, reported in Table 5, presents the percentage of net long speculators, together with the scalping index. While for the mean equation we get results similar to the previous

[15] We also estimate the model with the ADS index in the mean equation. It is generally poorly significant, therefore we prefer the specification reported.
[16] Indeed, this high correlation prevents us from including the three different measures of speculation in one single specification.

models, we find some differences in the variance equation: scalping remains significant and positive across commodities, whereas net long speculative positions have mixed results. This index is generally not significant and, when it is significant, it is either positive (corn and soybeans equation) or negative (natural gas, cocoa and oats).

[TABLE 5 ABOUT HERE]

To sum up, the evidence shows that the scalping index has a positive and significant coefficient even when it is associated to other indexes. The evidence on the scalping index, which is usually employed to capture short term financial trading, suggests that this kind of phenomenon has a destabilizing effect on price volatility. However, if we consider long term speculation measures, we find that they generally have a negative impact on volatility, smoothing the price process.

5.2 Robustness analysis

In order to analyze if the main results vary under different conditions, we focus on three types of robustness checks: we extend the analysis adopting different data frequencies, we investigate whether the results are unaffected adopting alternative GARCH models and we check if different controls in the mean equation impact somehow on the results obtained in the variance equation.

5.2.1 Data frequency

We repeat the previous analysis at daily and monthly frequency to see if speculation indexes show a different impact on price volatility. While it is possible to replicate estimations at monthly level for each measure of speculation, we are forced to exclude from the daily analysis the Working's T index, the market share of non-commercial traders and the percentage of net long speculators over total open interest. Indeed, data to construct these indexes are available from CFTC only at weekly level (Büyükşahin and Robe 2010).[17] The results are discussed focussing on the model with scalping only in the variance equation.

Table 6 presents the scalping coefficients estimated at different frequencies. The monthly data do not have ARCH effects in the residuals of OLS estimation for a number of commodities and thus a GARCH(p,q) specification is no longer supported. We observe that scalping index maintains its sign and significance level across different frequencies (apart from the case of cocoa, where the coefficient loses significance at monthly frequency). The only difference is in the magnitude of

[17] Data on daily positions of traders are collected by CFTF, but they are not public.

coefficients which are smaller (greater) with data at daily (monthly) frequency. Nevertheless, they remain small and close to zero.[18]

[TABLE 6 ABOUT HERE]

5.2.2 Econometric specification

We repeat the previous analysis adopting alternative GARCH models to see if the results are influenced by the type of models employed. We estimate the GARCH-in-Mean (GARCH-M, see Engle et al. 1987), which introduces the conditional variance or standard deviation into the mean equation, the threshold ARCH (TARCH, see Zakoïan 1994), which allows the conditional standard deviation to depend upon the sign of the lagged innovations, and the asymmetric exponential GARCH (EGARCH, see Nelson 1991), which explicitly allows for asymmetries in the relationship between returns and volatility. We compare the results of the model with scalping index and Working's T in the variance equation.

Table 7 shows the results across different econometric specifications.[19] We can see that GARCH-M and TARCH have quite the same sign and significance of the GARCH model: scalping is positive and Working's T index remains generally negative. Moreover, in the GARCH-M estimation, we have found that the conditional variance (or standard deviation) added in the mean equation is generally not significant. This means that the estimated coefficient on the expected risk (the risk premium) has no influence on expected returns of commodities investments, i.e. there is no feedback from the variance to the mean. The asymmetric EGARCH model obtains larger coefficients but generally gets to the same results. Finally, the asymmetric models, TARCH and EGARCH, do not display significant asymmetric effects on conditional variance. We find some evidence of asymmetry for cocoa and soybeans although with an unexpected sign: bad news in futures markets decrease volatility. Overall, we might say that the leverage effect does not seem to be present.

[TABLE 7 ABOUT HERE]

[18] The complete set of estimation on daily and monthly data can be found in Tables A.4, A.5.a, A.5.b, A.5.c and A.5.d in the Statistical Appendix.

[19] The complete set of estimation on weekly data of GARCH-M models can be found in Tables A.6.a, A.6.b, A.6.c and A.6.d in the Statistical Appendix. Comparisons of different econometric estimations on every frequency of data and on every type of combination of variables in the variance equation can be found in Tables A.7, A.8.a, A.8.b, A.8.c, A.9.a, A.9.b, A.9.c and A.9.d in the Statistical Appendix.

5.2.3 Focus on crude oil

The macroeconomic controls we have used in our analysis might be not sufficient to model the economic cycle. Hamilton (2009), for example, suggests that economic fundamentals such as demand, supply and storage are more relevant in explaining crude oil returns. Thus, we focus on crude oil[20] and verify, at weekly frequency, how results on speculative indexes change when the mean equation is otherwise specified.

Table 8 presents six different specifications, one for each set of macroeconomic variables employed. The dependent variable in each equation is the crude oil return. The first four models include some controls specific for the oil market, i.e. data on demand, production and stocks, which are however poorly significant. The fifth model corresponds to the crude oil equation in Table 4. The scalping index is always positive, close to zero and significant at least at 5%, independently from the macro-variables' choice in the mean equation. The same happens for Working's T index, which is always negative, close to zero and statistically significant. Oil demand and stocks do not seem to significantly affect the returns. Only the production variable is significant, with an expected negative sign which does not have any impact on the results in the variance equation.[21]

[TABLE 8 ABOUT HERE]

We can conclude that the results on weekly data presented in Section 5.1 are invariant to changes in data frequency and econometric specification and that the choice of macroeconomic variables in the mean equation does not affect the results in the variance equation, which is the focus of our analysis.

6. Conclusions

This paper considers alternative measures of speculative activity and evaluates if there is a relationship between speculation and the volatility of commodity futures prices. We test this relationship using data for futures prices for four energy commodities (crude oil, heating oil, gasoline and natural gas) and six agricultural commodities (cocoa, coffee, corn, oats, soybean oil and soybeans) over the period 1986-2010 at weekly frequency. Short term speculation is measured by means of scalping, while long term speculation can be proxied by the Working's T index, the

[20] The focus on oil is motivated by the availability and the frequency of the oil data, which are not generally matched by other commodities.

[21] We implement the same robustness exercise at the monthly frequency and we obtain similar results. Estimates are presented in Table A.10 of the Statistical Appendix.

market share of non-commercial traders and the percentage of net long speculators over total open interest.

Our work brings fresh evidence in the literature under different respects. First, we distinguish between short term and long term measures of speculation. In the first category we consider the scalping index, while in the second we employ the most frequently adopted Working's T index, the market share of non-commercial traders and net long speculative positions. Second, we analyze if these different measures of speculation impact in a stabilizing or destabilizing way on price volatility, using more comprehensive econometric specifications than in previous studies (Peck, 1981; Streeter and Tomek, 1992). Finally, we run a robustness exercise to check if the main results are invariant to changes in data frequency, econometric specification and control variables in the mean equation.

In the econometric analysis commodity returns are modelled according to a GARCH(p,q) with an AR(1) term. Speculation indexes are included as exogenous variables in the variance equation of the models: short term speculation, i.e. the scalping index, and long term speculation indexes are jointly considered. Our estimation results suggest that, among macroeconomic factors, S&P 500 index is generally positive and significant and it is the most relevant control to explain commodity futures returns. We find that speculation significantly affects the volatility of returns, although in contrasting ways. The scalping index has a positive and significant coefficient in the variance equation, suggesting that short term speculation has a positive impact on volatility. The other three indexes have instead a negative effect (when significant), that is long term speculation does not destabilize prices (see, among others, Brunetti et al. 2011).

We evaluate if and how the main results change moving along several dimensions. In particular, we consider alternative data frequencies, finding that scalping index is always significant and positive also at higher and lower frequency of data. Moreover, the main results in the variance equation remain unchanged across different econometric models (such as GARCH-M, TARCH and asymmetric EGARCH). Finally, if we change the specification in the mean equation to include additional economic controls, the results in the variance equation are unaffected.

References

Alquist, Ron, Olivier Gervais. 2013. "The role of financial speculation in driving the price of crude oil" *The Energy Journal* 34 (3), 35-54.

Andersen, Torben G. 1996. "Return volatility and trading volume: An information flow interpretation of stochastic volatility" *Journal of Finance* 51 (1): 169-204.

Brunetti, Celso, Bahattin Büyükşahin, Jeffrey H. Harris. 2011. "Speculators, prices and market volatility", mimeo (available at SSRN: http://ssrn.com/abstract=1736737).

Brunnermeier, Markus K., Stefan Nagel, Lasse H. Pedersen. 2008. "Carry Trades and Currency Crashes" NBER Working Paper Series Num. 14473 (available at: http://www.nber.org/papers/w14473).

Büyükşahin, Bahattin, Jeffrey H. Harris. 2011. "Do Speculators Drive Crude Oil Futures Prices?" *The Energy Journal* 32 (2): 167-202.

Büyükşahin, Bahattin, Michel A. Robe. 2010. "Speculators, Commodities and Cross-Market Linkages" Working Paper Series (available at SSRN: http://ssrn.com/abstract=1707103).

Chatrath, Arjun, Sanjay Ramchander, Frank Song. 1996. "The Role of Futures Trading Activity in Exchange Rate Volatility" *The Journal of Futures Markets* 16 (5): 561-584.

Chevallier, Julien. 2009. "Carbon Futures and Macroeconomic Risk Factors: a View from the EU ETS" *Energy Economics* 31: 614-625.

Cornell, Bradford. 1981. "The Relationship between Volume and Price Variability in Futures Markets" *The Journal of Futures Markets* 1 (3): 303-316.

Cox, Charles C. 1976. "Futures trading and market information" *Journal of Political Economy* 84 (6): 1215-1237.

De Roon, Frans A., Theo E. Nijman, Chris Veld. 2000. "Hedging Pressure Effects in Futures Markets" *The Journal of Finance* 55 (3): 1437-1456.

Du, Xiaodong, Cindy L. Yu, Dermot J. Hayes. 2011. "Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis" *Energy Economics* 33: 497-503.

Engle, Robert F., David M. Lilien, Russell P. Robins. 1987. "Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model" *Econometrica* 55 (2): 391-407.

Friedman, Milton. 1983. “The case for flexible exchange rates” in Essays in Positive Economics, University of Chicago Press, Chicago, 157-203.

Hamilton, James D., 2009, “Understanding crude oil prices” *The Energy Journal* 30: 179-206.

Hart, Oliver D., David M. Kreps. 1986. “Price destabilizing speculation” *Journal of Political Economy* 94 (5): 927-952.

Hedegaard, Esben. 2011. “How Margins are Set and Affect Asset Prices” Job Market Paper.

Irwin, Scott H., Dwight R. Sanders. 2010. “The Financialization of Commodity Futures Markets or: How I Learned to Stop Worrying and Love the Index Funds” Working Papers Series (available at: http://ssrn.com/abstract=1699793).

Irwin, Scott H., Dwight R. Sanders. 2012. “Testing the Masters Hypothesis in commodity futures markets” *Energy Economics* 34: 256-269.

Khan, Mohsin S. 2009. “The 2008 Oil Price “Bubble”” Peterson Institute for International Economics Policy Brief 09-19.

Leuthold, Raymond M. 1983. “Commercial Use and Speculative Measures of the Livestock Commodity Futures Markets” *The Journal of Futures Markets* 3 (2): 113-135.

Luu, James C., Martin Martens. 2003. “Testing the Mixture-of-Distribution Hypothesis Using Realized Volatility” *The Journal of Futures Markets* 23 (7): 661-679.

Manera, Matteo, Marcella Nicolini, Ilaria Vignati. 2013. “Financial speculation in energy and agriculture futures markets: A multivariate GARCH approach” *The Energy Journal* 34 (3), 55-81.

Masters, Michael W. 2008. “Testimony of Michael W. Masters before the Committee on Homeland Security and Governmental Affairs United States Senate”, May 2008. http://hsgac.senate.gov/public/_files/052008Masters.pdf.

Medlock III, Kenneth B., Amy M. Jaffe. 2009. “Who Is In The Oil Futures Market and How Has It Changed?” James A. Baker III Institute for Public Policy, Rice University.

Nelson, Daniel B. 1991. “Conditional Heteroskedasticity in Asset Returns: A New Approach” *Econometrica* 59 (2): 347-370.

Peck, Anne E. 1981. “The adequacy of speculation on the wheat, corn, and soybean futures markets” *Research in Domestic and International Agribusiness Management* 2: 17-29.

Powers, Mark J. 1970. “Does futures trading reduce price fluctuations in the cash markets?” *American Economic Review* 60 (3): 460-464.

Robles, Miguel, Maximo Torero, Joachim von Braun. 2009. “When Speculation Matters” International Food Policy Research Institute, Issue Brief 57.

Rowsell, John B., Wayne D. Purcell. 1992. “Speculative Activity and Price Volatility in the Live Cattle Futures Market” Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL. http://www.farmdoc.uiuc.edu/nccc134.

Sanders, Dwight R., Keith Boris, Mark Manfredo. 2004. “Hedgers, funds, and small speculators in the energy futures markets: an analysis of the CFTC’s Commitments of Traders reports” *Energy Economics* 26: 425-445.

Sanders, Dwight R., Scott H. Irwin, Robert P. Merrin. 2010. “The Adequacy of Speculation in Agricultural Futures Markets: Too Much of a Good Thing?” *Applied Economic Perspectives and Policy* 32 (1): 77-94.

Sanders, Dwight R., Scott H. Irwin. 2013. “Measuring index investment in commodity futures markets” *The Energy Journal* 34 (3), 105-127.

Smith, James L. 2009. “World Oil: Market or Mayhem?” *Journal of Economic Perspectives*, 23 (3): 145-64.

Stein, Jeremy C. 1987. “Informational externalities and welfare-reducing speculation” *Journal of Political Economy* 95 (6): 1123- 1145.

Streeter, Deborah H., William G. Tomek. 1992. “Variability in Soybean Futures Prices: An Integrated Framework” *The Journal of Futures Markets* 12 (6): 705-728.

Tauchen, George E., Mark Pitts. 1983. “The price variability-volume relationship on speculative markets” *Econometrica* 51 (2): 485-505.

Till, Hilary. 2009. “Has There Been Excessive Speculation in the US Oil Futures Markets?” Working Paper, EDHEC - Risk Institute.

Working, Holbrook. 1960. “Speculation on Hedging Markets” *Food Research Institute Studies* 1:185-220.

Working, Holbrook. 1967. “Tests of a Theory Concerning Floor Traders” *Food Research Institute Studies*, Supplement to Vol. VII, reprinted in Anne E. Peck, editor, *Readings in Futures Markets*, Vol. I, Chicago Board of Trade, Chicago, 1978.

Zakoïan, Jean-Michel. 1994. “Threshold Heteroskedastic Models” *Journal of Economic Dynamics and Control* 18: 931-955.

Tables and figures

Table 1: Summary statistics for speculation indexes

	Obs	Mean	Std. Dev.	Min	Max	Unit Root Test
			SCALPING			
Gasoline	1299	0.339	0.091	0.083	0.656	-5.090***
Heating Oil	1279	0.283	0.071	0.107	0.579	-7.107***
Natural Gas	1079	0.191	0.080	0.025	0.939	-1.757
Crude Oil	1298	0.326	0.101	0.065	0.805	-2.373
Cocoa	1296	0.104	0.041	0.018	0.288	-4.732***
Coffee	1298	0.194	0.080	0.025	0.518	-3.296**
Corn	1299	0.174	0.065	0.015	0.428	-5.031***
Oats	1299	0.136	0.070	0.008	0.465	-4.317***
Soybean Oil	1297	0.208	0.071	0.038	0.443	-3.790***
Soybeans	1298	0.310	0.112	0.031	1.002	-10.605***
			WORKING'S T			
Gasoline	1299	1.105	0.046	1.036	1.386	-8.144***
Heating Oil	1297	1.154	0.051	1.050	1.340	-6.460***
Natural Gas	1079	1.128	0.083	1.021	1.517	-7.519***
Crude Oil	1298	1.140	0.039	1.051	1.278	-4.615***
Cocoa	1296	1.115	0.045	1.016	1.258	-8.787***
Coffee	1298	1.178	0.073	1.053	1.400	-6.806***
Corn	1299	1.250	0.047	1.146	1.401	-5.673***
Oats	1299	1.180	0.091	1.040	1.593	-6.250***
Soybean Oil	1297	1.183	0.065	1.051	1.364	-7.373***
Soybeans	1298	1.268	0.068	1.113	1.492	-8.080***
			SHARE NON-COMMERCIAL			
Gasoline	1299	0.213	0.048	0.097	0.448	-6.026***
Heating Oil	1297	0.253	0.054	0.149	0.441	-5.470***
Natural Gas	1079	0.187	0.061	0.048	0.475	-5.444***
Crude Oil	1298	0.217	0.044	0.110	0.364	-5.649***
Cocoa	1296	0.237	0.051	0.096	0.392	-5.351***
Coffee	1298	0.303	0.055	0.183	0.479	-5.398***
Corn	1299	0.345	0.038	0.250	0.434	-5.881***
Oats	1299	0.355	0.068	0.154	0.548	-5.395***
Soybean Oil	1297	0.291	0.053	0.153	0.409	-6.544***
Soybeans	1298	0.362	0.048	0.244	0.464	-7.949***
			NET LONG POSITIONS OF NON-COMMERCIAL OVER OPEN INTEREST			
Gasoline	1299	0.103	0.113	-0.176	0.407	-7.602***
Heating Oil	1279	0.077	0.082	-0.172	0.304	-9.089***
Natural Gas	1079	0.017	0.101	-0.226	0.268	-6.540***
Crude Oil	1298	0.009	0.071	-0.242	0.211	-8.248***
Cocoa	1296	0.093	0.154	-0.367	0.501	-6.331***
Coffee	1298	0.146	0.140	-0.226	0.485	-9.441***
Corn	1299	0.027	0.127	-0.291	0.279	-6.490***
Oats	1299	0.282	0.144	-0.136	0.620	-6.399***
Soybean Oil	1297	0.109	0.157	-0.240	0.519	-7.322***
Soybeans	1298	0.103	0.149	-0.318	0.402	-5.269***

Notes: Column "Unit Root Test" reports the Augmented Dickey-Fuller statistic for the null hypothesis that there is a unit root in the series. *, ** and *** denote significance at 10%, 5% and 1% levels, respectively.

Table 2: Estimates of univariate GARCH models – Scalping as exogenous variable in the variance equation

		Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
Mean Equation	Tbill	0.035**	0.032***	0.025	0.018	-0.035***	-0.002	0.016	-0.006	0.016**	0.007
		(0.016)	(0.011)	(0.018)	(0.016)	(0.010)	(0.007)	(0.010)	(0.009)	(0.008)	(0.010)
	Junk Bond Yield	-0.053*	-0.020	0.055	-0.036	-0.018	-0.037	-0.014	0.000	-0.019	-0.026
		(0.031)	(0.028)	(0.035)	(0.027)	(0.025)	(0.027)	(0.020)	(0.028)	(0.019)	(0.018)
	S&P500	0.056	0.103**	0.107**	0.122**	0.081*	0.239***	0.069*	0.083*	0.106***	0.081**
		(0.052)	(0.046)	(0.049)	(0.051)	(0.044)	(0.048)	(0.036)	(0.044)	(0.034)	(0.033)
	AR(1)	0.190***	0.194***	0.204***	0.164***	0.182***	0.179***	0.203***	0.169***	0.235***	0.212***
		(0.028)	(0.031)	(0.033)	(0.029)	(0.027)	(0.026)	(0.029)	(0.029)	(0.029)	(0.030)
	Constant	0.002	0.002	0.000	0.002	0.000	-0.001	0.000	0.000	0.001	0.001
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Variance Equation	ARCH(1)	0.114***	0.250***	0.157***	0.126***	0.082**	0.049*	0.185***	0.171***	0.128***	0.171***
		(0.026)	(0.044)	(0.034)	(0.028)	(0.037)	(0.029)	(0.034)	(0.046)	(0.040)	(0.029)
	ARCH(2)					0.059*	0.064**				
						(0.033)	(0.029)				
	ARCH(3)						0.051*				
							(0.030)				
	GARCH(1)	0.824***	0.361***	0.764***	0.819***			0.752***	0.232***		0.797***
		(0.037)	(0.079)	(0.040)	(0.037)			(0.037)	(0.078)		(0.030)
	Scalping	4.21E-04**	0.003***	0.002***	2.98E-04*	0.011***	0.008***	0.001***	0.008***	0.002***	9.44E-05*
		(1.93E-04)	(0.001)	(0.001)	(1.62E-04)	(0.001)	(0.001)	(1.39E-04)	(0.001)	(3.14E-04)	(4.86E-05)
	Constant	-2.07E-05	-3.29E-04***	-1.25E-04***	-4.81E-06	-6.23E-05	-2.32E-04**	-4.85E-05**	-7.88E-05	8.54E-05	2.77E-06
		(5.87E-05)	(1.14E-04)	(2.93E-05)	(4.31E-05)	(8.92E-05)	(1.01E-04)	(2.34E-05)	(5.83E-05)	(5.68E-05)	(1.63E-05)
	ARCH+GARCH terms	0.938	0.611	0.921	0.945	0.141	0.164	0.937	0.403	0.128	0.968
	Test ARCH LM (F-stat)	1.156	0.006	0.991	0.018	0.096	1.101	0.038	0.153	0.135	1.974
	N. of Obs.	1297	1277	1077	1296	1294	1296	1297	1297	1295	1296

Notes: The error distribution is a Student's T. Standard errors in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 3: Estimates of univariate GARCH models – Scalping and Working's T as exogenous variables in the variance equation

		Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
Mean Equation	Tbill	0.034**	0.031**	0.021	0.017	-0.034**	-0.001	0.009	-0.006	0.014	0.009
		(0.016)	(0.012)	(0.017)	(0.016)	(0.014)	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)
	Junk Bond Yield	-0.053*	-0.022	0.056	-0.036	-0.026	-0.035	-0.011	-0.014	-0.026	-0.029
		(0.032)	(0.029)	(0.050)	(0.027)	(0.026)	(0.027)	(0.021)	(0.028)	(0.018)	(0.019)
	S&P500	0.047	0.087*	0.203***	0.111**	0.077*	0.237***	0.078**	0.078*	0.089**	0.084**
		(0.051)	(0.046)	(0.068)	(0.049)	(0.046)	(0.051)	(0.035)	(0.045)	(0.035)	(0.033)
	AR(1)	0.191***	0.192***	0.199***	0.164***	0.198***	0.176***	0.199***	0.177***	0.223***	0.212***
		(0.028)	(0.030)	(0.036)	(0.029)	(0.028)	(0.027)	(0.030)	(0.029)	(0.028)	(0.030)
	Constant	0.002	0.001	0.002	0.002	-0.000	-0.001	0.000	0.000	-0.000	0.001
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Variance Equation	ARCH(1)	0.105***	0.230***	0.181***	0.133***	0.063***	0.057**	0.231***	0.151***	0.090***	0.168***
		(0.025)	(0.042)	(0.042)	(0.029)	(0.017)	(0.027)	(0.048)	(0.043)	(0.033)	(0.031)
	ARCH(2)						0.051**				
							(0.024)				
	ARCH(3)						0.043*				
							(0.026)				
	GARCH(1)	0.824***	0.384***	0.574***	0.796***	0.914***		0.465***	0.234***		0.754***
		(0.039)	(0.075)	(0.072)	(0.040)	(0.022)		(0.074)	(0.083)		(0.035)
	Scalping	2.23E-04	0.003***	0.005***	0.001**	4.96E-04**	0.008***	0.002***	0.007***	0.003***	2.80E-04***
		(1.89E-04)	(0.001)	(0.001)	(2.23E-04)	(2.32E-04)	(0.001)	(3.23E-04)	(0.001)	(3.03E-04)	(7.14E-05)
	Working's T	-0.001*	-0.002***	-0.001***	-0.001*	3.31E-04**	-0.001	-0.002***	-0.002***	-0.002***	-3.51E-04***
		(2.97E0-4)	(0.001)	(2.31E-04)	(3.84E-04)	(1.47E-04)	(0.001)	(3.30E-04)	(3.56E-04)	(2.65E-04)	(1.01E-04)
	Constant	0.001*	0.002***	0.001***	0.001*	-3.90E-04**	0.001	0.002***	0.002***	0.002***	4.20E-04***
		(3.80E-04)	(0.001)	(2.73E-04)	(4.15E-04)	(1.70E-04)	(0.001)	(4.20E-04)	(4.31E-04)	(3.34E-04)	(1.26E-04)
	ARCH+GARCH terms	0.929	0.614	0.756	0.929	0.977	0.151	0.696	0.385	0.090	0.922
	Test ARCH LM (F-stat)	1.460	0.004	1.294	0.097	0.138	0.620	1.113	0.111	0.447	2.250
	N. of Obs.	1297	1277	1077	1296	1294	1296	1297	1297	1295	1296

Notes: The error distribution is a Student's T. Standard errors in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 4: Estimates of univariate GARCH models – Scalping and market share of non-commercial traders as exogenous variables in the variance equation

		Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
Mean Equation	Tbill	0.033**	0.030**	0.021	0.018	-0.032***	0.001	0.011	-0.003	0.014	0.010
		(0.016)	(0.013)	(0.014)	(0.016)	(0.010)	(0.008)	(0.010)	(0.011)	(0.011)	(0.010)
	Junk Bond Yield	-0.054*	-0.018	0.025	-0.036	-0.018	-0.037*	-0.008	-0.004	-0.025	-0.026
		(0.032)	(0.029)	(0.052)	(0.027)	(0.025)	(0.027)	(0.022)	(0.029)	(0.018)	(0.019)
	S&P500	0.042	0.077	0.217***	0.110**	0.081*	0.234***	0.072*	0.071	0.063**	0.081**
		(0.050)	(0.048)	(0.067)	(0.050)	(0.044)	(0.049)	(0.038)	(0.048)	(0.037)	(0.034)
	AR(1)	0.191***	0.192***	0.201***	0.165***	0.182***	0.181***	0.199***	0.169***	0.227***	0.211***
		(0.028)	(0.030)	(0.034)	(0.029)	(0.027)	(0.027)	(0.030)	(0.028)	(0.027)	(0.030)
	Constant	0.002	0.002	0.000	0.002	0.000	-0.001***	0.000	0.000	-0.000	0.001
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Variance Equation	ARCH(1)	0.108***	0.222***	0.168***	0.133***	0.082**	0.047*	0.184***	0.119***	0.069**	0.174***
		(0.025)	(0.043)	(0.038)	(0.029)	(0.038)	(0.028)	(0.045)	(0.039)	(0.030)	(0.031)
	ARCH(2)					0.059*	0.063**				
						(0.033)	(0.029)				
	ARCH(3)						0.050*				
							(0.029)				
	GARCH(1)	0.821***	0.342***	0.488***	0.790***			0.379***	0.207**		0.755***
		(0.037)	(0.076)	(0.036)	(0.041)			(0.095)	(0.089)		(0.036)
	Scalping	3.62E-04**	0.003***	0.005***	0.001**	0.011***	0.008***	0.002***	0.007***	0.004***	2.67E-04***
		(1.82E-04)	(0.001)	(0.001)	(2.14E-04)	(0.001)	(0.001)	(3.77E-04)	(0.001)	(3.10E-04)	(7.12E-05)
	Share Non-Commercial	-0.001**	-0.003***	-0.005***	-0.001**	-1.07E-04	-1.75E-04	-0.003***	-0.003***	-0.003***	-4.08E-04***
		(2.85E-04)	(0.001)	(4.47E-04)	(3.36E-04)	(7.42E-04)	(0.001)	(0.001)	(0.001)	(2.56E-04)	(1.42E-04)
	Constant	1.44E-04	0.001***	0.001***	1.03E-04	-3.82E-05	-1.66E-04	0.001***	0.001***	0.001***	1.22E-04**
		(1.01E-04)	(1.82E-04)	(1.78E-04)	(7.53E-05)	(1.92E-04)	(3.20E-04)	(2.13E-04)	(2.58E-04)	(1.04E-04)	(5.00E-05)
	ARCH+GARCH terms	0.929	0.564	0.656	0.924	0.141	0.160	0.563	0.326	0.069	0.929
	Test ARCH LM (F-stat)	1.455	0.018	1.008	0.040	0.091	1.173	0.506	0.009	0.707	2.731*
	N. of Obs.	1297	1277	1077	1296	1294	1296	1297	1297	1295	1296

Notes: The error distribution is a Student's T. Standard errors in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 5: Estimates of univariate GARCH models – Scalping and net long positions of non-commercial traders over open interest as exogenous variables in the variance equation

		Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
Mean Equation	Tbill	0.034**	0.031***	0.022	0.018	-0.036***	-0.003	0.016	-0.006	0.017**	0.006
		(0.016)	(0.011)	(0.019)	(0.015)	(0.009)	(0.007)	(0.010)	(0.010)	(0.008)	(0.010)
	Junk Bond Yield	-0.053*	-0.019	0.063	-0.036	-0.018	-0.030	-0.012	0.002	-0.020	-0.027
		(0.031)	(0.028)	(0.043)	(0.027)	(0.025)	(0.027)	(0.020)	(0.032)	(0.019)	(0.018)
	S&P500	0.056	0.103**	0.161**	0.119**	0.088**	0.242***	0.063*	0.095*	0.103***	0.082**
		(0.052)	(0.046)	(0.075)	(0.051)	(0.044)	(0.050)	(0.035)	(0.049)	(0.034)	(0.033)
	AR(1)	0.190***	0.194***	0.177***	0.164***	0.180***	0.175***	0.202***	0.175***	0.235***	0.213***
		(0.028)	(0.031)	(0.031)	(0.029)	(0.027)	(0.027)	(0.029)	(0.028)	(0.028)	(0.030)
	Constant	0.002	0.002	0.001	0.002	0.000	-0.001	0.000	-0.001	0.001	0.001
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Variance Equation	ARCH(1)	0.114***	0.250***	0.168***	0.124***	0.077**	0.048*	0.189***	0.107***	0.129***	0.167***
		(0.026)	(0.044)	(0.040)	(0.028)	(0.037)	(0.029)	(0.037)	(0.032)	(0.039)	(0.029)
	ARCH(2)			0.098***		0.061*	0.062**		0.051*		
				(0.037)		(0.033)	(0.029)		(0.030)		
	ARCH(3)			0.105***							
				(0.040)							
	GARCH(1)	0.824***	0.355***		0.820***			0.693***	0.175**		0.801***
		(0.037)	(0.080)		(0.037)			(0.045)	(0.068)		(0.029)
	Scalping	4.27E-04**	0.003***	0.015***	2.93E-04*	0.011***	0.008***	0.001***	0.006***	0.003***	8.86E-05*
		(2.15E-04)	(0.001)	(0.001)	(1.60E-04)	(0.001)	(8.00E-04)	(1.77E-04)	(0.001)	(3.27E-04)	(4.91E-05)
	Net Long Positions / OI	-8.48E-06	-2.15E-04	-0.002***	-1.51E-04	-0.001**	3.21E-04	2.36E-04***	-0.001***	2.43E-04	5.06E-05*
		(1.53E-04)	(4.04E-04)	(5.07E-04)	(1.62E-04)	(2.81E-04)	(3.51E-04)	(7.44E-05)	(2.42E-04)	(1.61E-04)	(2.81E-05)
	Constant	-2.21E-05	-2.92E-04**	-4.99E-04***	-2.74E-07	-8.70E-05	-1.98E-04**	-6.23E-05***	2.84E-04**	3.25E-05	-3.35E-07
		(6.35E-05)	(1.27E-04)	(4.67E-05)	(4.30E-05)	(9.22E-05)	(1.00E-04)	(2.38E-05)	(1.25E-04)	(6.41E-05)	(1.64E-05)
	ARCH+GARCH terms	0.938	0.605	0.371	0.944	0.138	0.110	0.882	0.333	0.129	0.968
	Test ARCH LM (F-stat)	1.168	0.010	0.435	0.042	0.034	1.074	7.02E-08	0.053	0.206	1.577
	N. of Obs.	1297	1277	1077	1296	1294	1296	1297	1297	1295	1296

Notes: The error distribution is a Student's T. Standard errors in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 6: Scalping coefficients estimated on different frequencies – GARCH models

	Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
Daily	5.21E-05***	7.39E-05***	0.001***	1.71E-04***	2.07E-05**	0.002***	6.81E-05***	0.003***	1.78E-05***	6.01E-06**
Weekly	4.21E-04**	0.003***	0.002***	2.98E-04*	0.011***	0.008***	0.001***	0.008***	0.002***	9.44E-05*
Monthly	-	0.019***	-	0.024***	3.77E-04	0.028***	-	-	-	0.004***

Notes: * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Weekly coefficients are the same as those in Table 3. Daily coefficients are obtained from the estimation on daily data of univariate GARCH(1,1) models for gasoline, heating oil, natural gas, corn, soybean oil and soybeans, univariate GARCH(2,0) models for coffee and oats, univariate GARCH(1,3) model for crude oil and univariate GARCH(2,1) for cocoa. Monthly coefficients are obtained from the estimation on monthly data of univariate GARCH(1,0) models for heating oil and crude oil, univariate GARCH(1,1) model for cocoa, univariate GARCH(2,0) models for coffee and soybeans (all monthly specifications have normally distributed error terms).

Table 7: Scalping and Working's T coefficients estimated with different GARCH models

		Gasoline	Heating Oil	Natural Gas	Crude Oil	Cocoa	Coffee	Corn	Oats	Soybean Oil	Soybeans
GARCH	Scalping	2.23E-04	0.003***	0.005***	0.001**	4.96E-04**	0.008***	0.002***	0.007***	0.003***	2.80E-04***
	Working's T	-0.001*	-0.002***	-0.001***	-0.001*	3.31E-04**	-0.001	-0.002***	-0.002***	-0.002***	-3.51E-04***
GARCH-M([a])([b])	Scalping	1.98E-04	0.003***	0.003***	0.001**	0.011***	0.007***	0.002***	0.007***	0.003***	2.73E-04***
	Working's T	-0.001**	-0.002***	-1.31E-04	-0.001*	0.002*	-0.001*	-0.002***	-0.002***	-0.002***	-3.35E-04***
TARCH([a])	Scalping	8.40E-05	0.003***	0.002***	0.001***	0.011***([c])	0.007***	0.002***	0.006***	0.003***	2.02E-04***([c])
	Working's T	-0.001**	-0.002***	-4.48E-04***	-0.001**	0.002**([c])	-0.001	-0.002***	-0.002***	-0.002***	-2.08E-04**([c])
EGARCH([a]) asymmetric	Scalping	0.037	2.561***	0.455**	0.159	10.377***([c])	7.943***	7.254***	7.087***	6.724***	0.434***([c])
	Working's T	-0.248	-1.428***	-0.012	-0.092	1.685([c])	-1.050	-7.455***	-2.235***	-3.853***	-0.515***([c])

Notes: * significant at 10% level, ** significant at 5% level, *** significant at 1% level. (a) GARCH-M, TARCH and EGARCH use in general the same lag structure as in the standard GARCH. The exceptions are: GARCH-M(2,0) and TARCH(2,0) for cocoa; EGARCH(2,0) for cocoa and corn; EGARCH(2,1) for soybean and gasoline. (b) For soybean oil only, the variance is positive and significant at 1% in the mean equation. (c) We find asymmetric effects on the conditional variance for cocoa and soybean only.

Table 8: Focus on crude oil - Estimates of univariate GARCH models – Scalping and Working's T index as exogenous variable in the variance equation

		WEEKLY – CRUDE OIL					
		(1)	(2)	(3)	(4)	(5)	(6)
Mean Equation	Demand	-0.011 (0.045)	-	-	0.013 (0.049)	-	0.007 (0.049)
	Production	-	-0.108** (0.046)	-	-0.110** (0.050)	-	-0.106** (0.050)
	Ending Stock	-	-	-0.129 (0.195)	-0.079 (0.196)	-	-0.098 (0.197)
	Tbill	-	-	-	-	0.017 (0.016)	0.016 (0.016)
	Junk Bond Yield	-	-	-	-	-0.036 (0.027)	-0.036 (0.027)
	S&P500	-	-	-	-	0.111** (0.049)	0.111** (0.049)
	AR(1)	0.163*** (0.029)	0.165*** (0.029)	0.161*** (0.029)	0.163*** (0.029)	0.164*** (0.029)	0.164*** (0.029)
	Constant	0.002 (0.001)	0.002 (0.001)	0.002 (0.001)	0.002 (0.001)	0.002 (0.001)	0.002 (0.001)
Variance Equation	ARCH(1)	0.131*** (0.028)	0.129*** (0.028)	0.132*** (0.028)	0.129*** (0.028)	0.133*** (0.029)	0.131*** (0.028)
	GARCH(1)	0.795*** (0.040)	0.800*** (0.039)	0.794*** (0.040)	0.799*** (0.040)	0.796*** (0.040)	0.799*** (0.040)
	Scalping	0.001*** (2.31E-04)	0.001** (2.25E-04)	0.001*** (2.34E-04)	0.001** (2.27E-04)	0.001** (2.23E-04)	5.30E-04** (2.20E-04)
	Working's T	-0.001** (3.91E-04)	-0.001* (3.82E-04)	-0.001** (3.95E-04)	-0.001** (3.85E-04)	-0.001* (3.84E-04)	-0.001* (3.78E-04)
	Constant	0.001** (4.22E-04)	0.001* (4.12E-04)	0.001** (4.26E-04)	0.001* (4.15E-04)	0.001* (4.15E-04)	0.001* (4.08E-04)
	ARCH+GARCH terms	0.926	0.929	0.926	0.928	0.929	0.930
	Test ARCH LM (F-stat)	0.011	0.016	0.019	0.021	0.097	0.120
	N. of Obs.	1296	1296	1296	1296	1296	1296

Notes: The error distribution is a Student's T. Standard errors in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Demand is the Weekly U.S. Refiner Net Input of Crude Oil in Thousand Barrels per Day (source: EIA); Production is the Weekly U.S. Field Production of Crude Oil in Thousand Barrels per Day (source: EIA); Ending Stock is the Weekly U.S. Ending Stocks of Crude Oil in Thousand Barrels (source: EIA).