Information Demand and Agriculture Commodity Prices

Massimo Peri*, Daniela Vandone**, and Lucia Baldi*
*Department of Agricultural, Food and Environmental Economics, Università degli Studi di Milano, Italy
**Department of Economics, Business and Statistics, Università degli Studi di Milano, Italy
massimo.peri@unimi.it, daniela.vandone@unimi.it, lucia.baldi@unimi.it

Abstract

This paper investigates the effect of information flow on corn futures price variability for the period January 2004 -July 2011. The theoretical framework is the Mixture Distribution Hypothesis, that posits a joint dependence of return volatility and information. The main contribution of this article is that we use two different proxy for the observed component of information flow that allow to separate the effect of supply (News) and demand (Internet Search Volume) of information. Empirical estimates highlight that: i) results support the MDH since observed volatility persistence appears to be related to the information flow; ii) variation in information demand has a significant effect on volatility of futures corn returns even controlling for variation in information supply and such result can be interpreted in light of behavioural finance.

Keywords: commodity prices, information, futures markets, Mixture Distribution Hypothesis, egarch.

JEL: C32, G13, G14, Q11

1 Introduction

Over the last few years exceptional shocks from a host of external sources have had a deep effect on agriculture; it is common felt that structural changes in global commodity markets have greatly contributed to rising prices, to amplify and perpetuate volatility. These fundamental trends toward higher prices have been a key lure for increased speculative activity on the major futures exchanges. While speculation is crucial to proper functioning markets, when we are in presence of exceptional shocks it can trigger a vicious cycle of rising fragility and increased volatility, with negative impacts on food security (Prakash and Gilbert, 2011). Financial investors are progressively looking towards investing in commodity derivatives, which allows for portfolio diversification benefits, that is reduction of risk for any given level of expected return. The enormous sums of money being poured into commodities has led to suspicion that behaviour in commodity exchanges is amplifying volatility and causing persistence in the high prices of many important foodstuffs. Such volatility is related to trading activity. It is well known that, from a theoretical perspective, price movements are caused by information. Accordingly to the Efficient Market Hypothesis (Fama 1985) shifts in investor demand for securities are completely rational and reflect reactions to public announcements that affect future growth rate of dividends, risk, tax, institutional reasons and other adjustment to news conveyed through the trading process itself. However, during years a voluminous literature have examined the intricacies between announcements, news and market activity and alternatives to the efficient markets approach have been developed. Such alternatives are based on the
assumption that some investors (generally called ‘retail investors’, ‘individual investors’ or ‘noise traders’, opposite to ‘institutional investors’) are not fully rational and their demand for risky assets seem to be a response to changes in expectations that are not completely justified by fundamental news, but based on beliefs, sentiments, pseudo-signals or popular models using inflexible trading strategies (Tersky & Kahneman, 1974). In theory, these demand shifts only matter if the judgment biases afflicting investors in processing information tend to be the same; conversely, if all investors trade randomly, their trades cancel out and there are no aggregate shifts in demand. Empirical studies and experiments in the fields of psychology and behavioural finance find evidence that biases tend to be the same and investors tend to make the same mistakes, so that shifts in the demand for stocks that do not depend on fundamental factors are likely to affect volatility and prices (Meier and Sprenger 2007; Kilborn 2005; Franken et al. 2008).

Within this framework the aim of the paper is to investigate the effect of information flow on price variability, taking into account trader’s characteristics. To study this relationship we follow the theoretical approach of the Mixture Distribution Hypothesis (MDH) (Clark, 1973; Epps and Epps, 1976), that posits a joint dependence of returns volatility and information. Those authors suggest the use of treated volume as a good proxy for information flow; conversely, following recent progress in the MDH theory (Kalev et al, 2004; Vlastalakis et al., 2010) in the empirical application we use two different proxy for the observed component of information flow that allow to separate the effect of supply and demand of information. For the information supply we use the number of news published on the major world newspaper, while for the information demand we use the Internet Search Volume (ISV), a recent variable developed on the basis of the study of Varian (2007); ISV is used in information study and also very appreciate in financial literature.

The main contribution of this article is that for the first time we test the effect of information flow on agricultural commodity markets (corn futures prices), and quantify the component of additional volatility related to information supply and demand. It is also the first time that the ISV variable is applied to the commodity market. Moreover, considering that information demand through the use of internet can be considered a proxy for information used by noise traders to decide their investments (Da et al. 2011), the analysis of such relationship within the theoretical framework of MDH enable us to draw some consideration about the role of retail traders on the market.

The article is organized as follows: Section 2 describes the theoretical framework. Section 3 presents the dataset used for the purpose of the study and a brief analysis of variables trend. Section 4 proposes the econometric methodology and section 5 presents the empirical results. Section 6 includes the discussion and final remarks.

2 Theoretical issues

The link between information flow and financial markets is well known to financial economists (Fama et al. 1965, French and Roll 1986). The widespread hypothesis is that measures of market activity – such as prices, return volatility and trading volume – are directly related to the rate of arrival of information in the market. From a market microstructure perspective, price movements are caused primarily by the arrival of new information and the process that incorporates this information into market prices. Theory suggests that variables such as the trading volume, the number of transactions, the bid-ask spread, or the market liquidity are related to the return volatility process.
On the empirical front, a relevant strand in the literature has documented a strong positive correlation between trading volume, return volatility and information. Following the work of Clark (1973), the empirical applications tend to follow the specification associated with the MDH, that provides an explanation to the observed link between volatility and trading volume by imposing a joint dependence of both volume and returns on a latent information process. Specifically, the MDH argues that the variance of returns at a given interval is proportional to the rate of information arrival, proxied by the trading volume (see, among other, Epps and Epps 1976, Tauchen and Pitts 1983, Harris 1987). A direct consequence of the MDH is that observed patterns in market activity, such as volatility persistence, are caused by the existence of the same patterns in information flow.

Based on the MDH, the empirical literature has exploring different approaches and variables to test the effect of information flow on volatility. Some studies substitute the “trading volume” with *ad hoc* data-set collecting the number of news on the topic investigated published by the newspaper, or other relevant magazines (Berry and Howe 1993, Mitchell and Mulherin 1994, Ederington and Lee 1993, Kalev et al. 2004). Just recently the wide dissemination of the *world wide web* and the subsequent development of internet utilities has made available a new set of information useful to investigate the consumer behavior and attitude. Among other, recent empirical application of the MDH use the information demand that has been available through “*google insight*” for search.

The underlying idea is that institutional investors generally rely on research studies and public news in order to valuate the fundamental value of the asset and properly take investment decisions, while retail traders, that do not have access to more sophisticated information or are not able to interpret them, rely more on information demanded through search in internet.

The internet search volume offers a unique opportunity to empirically study the impact of retail traders, indeed, as Barber and Odean (2008) underline, the easy access to information by internet has definitively changed the production, intermediation, dissemination and consumption of information in the financial industry. Taking advantage from this new framework, a new body of literature (Da et al., 2011; Vlastakis and Markellos, 2011, Bank et al., 2011) has analysed financial market dynamics by the use of ISV variable as a proxy for investors information demand. In particular Vlastakis and Markellos (2011) have demonstrated the good performance of ISV data as a measure of information demand in the analysis of stock market indexes.

In this article we analyze corn futures price volatility in order to empirically verify if an easier access to information by the use of internet has influenced food commodity (corn) price dynamics. Indeed, accordingly to Grossman and Stiglit (1980), who analyse a noisy rational expectations equilibrium, demand for information is increasing with the magnitude of noise. Moreover, Veldkamp (2006), revisiting the Grossmand and Stiglitz model, suggests that payoff volatility increases information demand, whereas information increases asset prices. Since information demand through the use of ISV is a proxy for information used by retail investors, the analysis of such relationship enable us to draw some consideration about the role of noise traders on the market.
3 Data

To quantify demand of information in corn price dynamic we use a novel source of information provided by Google, the largest search engine in the world. Google Insights\(^1\), a special utility of Google, provide the number of Internet search queries. It gives the search volume (Internet Search Volume - ISV) for specific keywords with a weekly frequency and as a value relative to the total number of search on Google in the corresponding time interval. One of the main problems using these data concerns the identification of keywords to perform the request. After many trials we identified the following search words: corn price and corn prices that represent the biggest volume of search and we sum the results from each of these searches, as in Da et al. (2011). Google insight provides also other information linked to the specific request, like the main categories involved, the “top searches” and “rising searcher”. Very interesting, the 25%-50% of the total research volume is linked to “Finance” category, the 10%-25% to “Business & Industrial” while only 0-10% to “Food & Drink”. This leads us to suppose that the key words used in the empirical study well describe the trader information demand. Such assumption also found support in the graphical visualization of series. Indeed ISV shows a dynamic very similar to that of corn prices reaching exceptional heights during mid-2008 and then subsequently declined with remarkable speed (see Figure 1).

For information supply data we use LexisNexis® Academic data base that collects the world’s most reputable news. In particular we consider all news published on major world newspapers containing the words corn price. In this case the searching engine capture all the articles that contain the combination of corn price (or corn prices) in the same phrases. Then we had rearranged the information collected building a weekly variable (named NEWS).

Futures prices are collected from DataStream and are denoted in U.S. cents per bushel. The serie is from CBOT and is considered as the leading benchmark price for international trade. Futures prices are relative to corn U.S. No.2 yellow. The futures contracts are rolled over to the next contract on the first business day of the contract month; this is the standard procedure in the literature since the nearby futures contract is highly liquid and the most active (Yang et al., 2001).

Figure 1 reports futures prices, information demand and information supply. Prices increased first during the beginning of 2007, then slipped slightly to be followed by a new, very strong increase. In 2008, corn prices went through a moment of particular impetus induced by the ethanol boom, which absorbed an increasing amount of production (about one-fifth of the previous harvest was used for the distillation of biofuel). This situation was intensified by the dry climate that reduced yields. However, in the following period, high maize prices gave way to a substantial increase in plantings and this, together with favourable weather conditions, boosted world output with an ensuing slight fall in prices.

As far as the variable ISV and the variable NEWS, the graph shows that during the period of sharp rise in futures prices - in 2008 and, later, from late 2010 onwards - the information flow variables show similar trends. The simple analysis of the chart already clearly highlights how these variables are interrelated. In other words, during the worst periods of the global food crises, information demand for corn price shows the same trend. The same occurs for the information supply variable.

\(^1\) See www.google.com/insights/search/
The following graphs (Figure 2) show the corn futures return, the squared return and the absolute return; similarly, the figures show that volatility is particularly strong during the above mentioned period of time. These three graphs show evidence of volatility clustering: low values of volatility followed by low values and high values of volatility followed by high values. All the data collected are weekly, span from January 2004 to July 2011 and are expressed in logarithmic form.

Source: CBOT, Google Insight and LexisNexis® Academic data base

**Figure 1.** Futures corn price, information demand variable (ISV) and information supply variable (NEWS)

**Figure 2.** Return of corn future
The descriptive statistics in Table 1 show the stylised facts that all the variables display a significant departure from normality as reflected by the high level of kurtosis and the Jarque-Bera test for normality. Part of this non-normality is caused by some outliers around the above mentioned food crises. Moreover, the level of skewness indicates that for what concerns futures returns there is a negative asymmetry, while for news and ISV the asymmetry is positive. All this evidences enable us to follow a GARCH approach.

<table>
<thead>
<tr>
<th>Table 1.</th>
<th>Descriptive statistics (absolute value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Futures return</td>
</tr>
<tr>
<td>Mean</td>
<td>1.09</td>
</tr>
<tr>
<td>Median</td>
<td>1.25</td>
</tr>
<tr>
<td>Maximum</td>
<td>99.75</td>
</tr>
<tr>
<td>Minimum</td>
<td>-93.50</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>21.61</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.25</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>7.66</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>369.41</td>
</tr>
<tr>
<td>Probability</td>
<td>0.000</td>
</tr>
<tr>
<td>Sum Sq. Dev.</td>
<td>187,797</td>
</tr>
<tr>
<td>Observations</td>
<td>403</td>
</tr>
</tbody>
</table>

Source: our elaboration on CBOT data, Google Insight and LexisNexis® Academic data base

4 Econometric methodology

In the empirical application we analyze the conditional volatility of futures corn price returns controlling for information (supply and demand). We model corn futures price return volatility in a GARCH framework constructing an EGARCH model to capture conditional volatility and possible leverage effects.

Our main focus in this work is to model volatility in presence of different information. However, a stylized fact of financial volatility is that volatility tends to be higher in a falling market than in a rising market. This asymmetric information impact on volatility is commonly referred to as the leverage effect (Black, 1976) that is bad news (negative shocks) tends to have a larger impact on volatility than good news (positive shocks).

The leverage effect can be incorporated into a GARCH model using exponential GARCH (EGARCH) model proposed by Nelson (1991).

\[
\log(\sigma_i^2) = \omega + \sum_{j=1}^q \alpha_j \left| \frac{\epsilon_{i-j}}{\sigma_{i-j}} \right| + \gamma \epsilon_i + \sum_{j=1}^q \beta_j \log(\sigma_{i-j}^2) \tag{1}
\]

where \(\alpha, \beta\) and \(\gamma\) are the constant parameters. The \(\alpha\) parameter represents the symmetric effect of the model whereas \(\beta\) measures the persistence in conditional volatility. The asymmetric effect of the volatility is expressed by \(\gamma\) coefficients and can be tested by the hypothesis that \(\gamma = 0\). In this case a positive return shock has the same effect on volatility as a negative return shock of the same amount. If \(\gamma > 0\) “good news” (or positive shocks) are more destabilizing than “bad news” (or negative shocks) whereas \(\gamma < 0\) “bad news” generate a larger impact on volatility.
One of the benefits of using this model is that the logarithmic form of the EGARCH(p,q) model certifies the non-negativity of the conditional variance without the need to constrain the model’s coefficients (Zivot and Wang, 2006). Another advantage is that only coefficients of GARCH term govern the persistence of volatility shock.

According to Lamoureux and Lastrapes (1990) the conditional variance of $\varepsilon_t$ shows patterns of time dependence when there is serial correlation in the news arrival processes under the hypothesis that each new information will set up a new price equilibrium. In other words it is possible that the true conditional variance is also function of other exogenous variables.

Kalev et al. (2004) argue that when we include a proxy for information flows into a conditional variance equation most of the observed volatility persistence would disappear. So equation (1) can be modified adding explanatory variables to conditional variance equation in a straightforward way (Zivot, 2008).

In our work we consider a EGARCH (1,1) model that can be adequate to obtain a good model fit for financial time series (Hansen and Lunde 2004). Thus the final model we use to model conditional variance of return of future corn price considering information variables as exogenous is given as follows:

$$\log(\sigma_t^2) = \omega + \alpha \left( \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) + \beta \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \lambda (ISV)_t + \delta (NEWS)_t \quad (2)$$

(ISV)$_t$ and (NEWS)$_t$ are set as exogenous variables in the conditional variance equation.

### 5 Empirical results

We consider four different EGARCH(1,1) models explained as follows: a) model without variables included in the variance equation; b) model with ISV variable included; c) model with NEWS variable included; d) model with both ISV and NEWS variables included $^2$. For each of these models we report results of estimates of mean equation, variance equation and the diagnostic tests to evaluate the adequacy of the fit.

Our empirical results reveal a significant impact of information flows on the conditional volatility of corn futures price returns.

The estimation of model a) without variance parameters (Table 2) shows that $\alpha$ and $\beta$ parameters are positive and significant, while the coefficient $\gamma$ is negative and significant at 0.10 level; such results are coherent with a correct specification of the EGARCH model.

Then we consider the other three models. From a statistical point of view, in all specifications $\alpha$ and $\beta$ coefficients are significant and positive, while $\gamma$ is significant and always negative. So we can assert that all variables considered produce a leverage effect, i.e. volatility is likely to be higher when previous shocks are negative.

Our results are in line with the MDH: by comparing estimates of volatility persistence before and after the inclusion of information variables, results show that the inclusion of information variables in the historical volatility models reduces volatility persistence significantly as in Kalev et al. (2004).

Specifically, a very interesting results is that when we consider ISV variable (model b) in the variance equation the $\beta$ coefficient that expresses the persistence in the volatility decreases

---

$^2$ To account also for trading activity we try to estimate model with Open Interest variable measured by the corn futures open interest collected from the U.S. Commodity Futures and Trading Commission (CFTC). We find OI variable be always not significant even when it is inserted with the other variables. So we decide to present the results only with the two variables ISV and NEWS.
from 0.729 to 0.540, that is, the information demand variable contributes to explain about 26% of the volatility. Moreover, $\gamma$ coefficient appears significant and higher compared to the basic model a); this means that search activity on Google intensifies the volatility produced by negative shocks instead of absorbing it.

When we include information supply as exogenous regressor in the variance equation (model c) the estimate of the $\beta$ coefficient is significant and equal to 0.677, and the persistence of volatility decrease only by 7% compared to the basic model.

In other words, if the models are estimated using only information demand variables or only information supply variables as variance regressors we can conclude that most of the reduction in persistence originated from information demand.

Such a result can be interpreted in light of behavioural finance, where studies consider information demand as an expression of noise trading (Da et al. 2011). The search of information on google by non-professionally (and less informed) traders can amplify volatility especially in case of negative shock, when investment decisions are more easily influenced by panic or irrational behavior.

When we consider both the information variables as regressors in the variance equation (model d) we can see that $\beta$ coefficient is still significantly different from zero, but the value is higher than the $\beta$ in the model with only the ISV variable. This implies information exogenous variables together can explain the persistence of corn futures return volatility in a less incisive way compared to information demand alone. This may be due to the different nature of the content of the information coming from google compared to information obtained from other sources. While ISV is a direct measure of individual attention NEWS measures public and official information. As argue Da et al. (2011) and Vlastakis and Markellos (2011), there are intricacies between supply and demand information. Their correlation is positive but low and there are evidence of bidirectional causality suggesting that news media information and google search could interact with each other.

Finally, based on standardized residual from the EGARCH(1,1) model we carried out a set of diagnostic tests to empirically support the specification of the different EGARCH models used. The ARCH Engle’s Lagrange multiplier (LM) tests the null hypothesis that there is no ARCH effects up to order in the residuals. Table 3 reports the test value and the corresponding p-value, as is clear from the table for all the model considered the Engle’s test indicates that there are no heteroskedasticity in the residuals. The Ljung-Box test statistic at lag $s$, $Q(s)$, tests the null hypothesis that there is no autocorrelation up to order $s$ for standardized residuals. Table 3 reports the $Q(20)$ and their $p$-values; the null hypothesis of no autocorrelation up to order 20 is not rejected for all the model considered. In the same table the results of the latter test is also confirmed by the $Q^2(20)$ that tests the standardized residuals squared.

Table 2.

<table>
<thead>
<tr>
<th>EGARCH(1,1) results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
</tr>
<tr>
<td>Without variance parameters</td>
</tr>
<tr>
<td>With ISV</td>
</tr>
<tr>
<td>With NEWS</td>
</tr>
<tr>
<td>With ISV and NEWS</td>
</tr>
</tbody>
</table>

Note: The results are based on the estimation of (2) model. This model is first estimated without any exogenous variables, then progressively includes ISV and NEWS variables. *, ** and *** denote statistical significance at 0.10, 0.05 and 0.001 level respectively.
Table 3. Diagnostic tests of EGARCH(1,1) models

<table>
<thead>
<tr>
<th></th>
<th>ARCH-LM</th>
<th>Q(20)</th>
<th>Q2(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without variance parameters</td>
<td>0.616</td>
<td>13.009</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>0.433</td>
<td>0.368</td>
<td>0.749</td>
</tr>
<tr>
<td>With ISV</td>
<td>0.633</td>
<td>13.553</td>
<td>4.924</td>
</tr>
<tr>
<td></td>
<td>0.426</td>
<td>0.33</td>
<td>0.96</td>
</tr>
<tr>
<td>With NEWS</td>
<td>0.185</td>
<td>12.776</td>
<td>5.364</td>
</tr>
<tr>
<td></td>
<td>0.668</td>
<td>0.386</td>
<td>0.945</td>
</tr>
<tr>
<td>With ISV and NEWS</td>
<td>0.420</td>
<td>13.268</td>
<td>5.031</td>
</tr>
<tr>
<td></td>
<td>0.517</td>
<td>0.35</td>
<td>0.957</td>
</tr>
</tbody>
</table>

Notes: ARCH(LM) is Lagrange multiplier test proposed by Engle (1982), Q(20) and Q2(20) denote the Ljung-Box test statistic on returns and squared returns. p-value in italics

6 Conclusion

These preliminary results allow us to better understand the cause of high prices volatility recorded by agricultural commodity prices during the latest years. Overall, our results support the MDH since the observed volatility persistence in our stock return data appears to be related to the demand and supply of information. Such result is itself interesting since no studies have so far applied the MDH to commodities markets. Moreover, results show that variations in information demand have a significant effect on corn futures volatility, and this effect is robust even controlling for variations in the supply of information. In our opinion this result is relevant since it can be interpreted in light of behavioural finance, where studies consider information demand as an expression of noise trading. The search of information on google by non-professionally (and less informed) traders can amplify volatility especially in case of negative shock, when investment decisions are more easily influenced by panic or irrational behavior.

References


