zef
Center for
Development Research
University of Bonn

ZEF-Discussion Papers on
Development Policy No. 246

Miguel Almanzar and Maximo Torero

Media Coverage and Food Commodities: Agricultural Futures Prices and Volatility Effects

Bonn, October 2017

The **CENTER FOR DEVELOPMENT RESEARCH (ZEF)** was established in 1995 as an international, interdisciplinary research institute at the University of Bonn. Research and teaching at ZEF address political, economic and ecological development problems. ZEF closely cooperates with national and international partners in research and development organizations. For information, see: www.zef.de.

ZEF – Discussion Papers on Development Policy are intended to stimulate discussion among researchers, practitioners and policy makers on current and emerging development issues. Each paper has been exposed to an internal discussion within the Center for Development Research (ZEF) and an external review. The papers mostly reflect work in progress. The Editorial Committee of the ZEF – DISCUSSION PAPERS ON DEVELOPMENT POLICY includes Joachim von Braun (Chair), Christian Borgemeister, and Eva Youkhana. Chiara Kofol is the Managing Editor of the series.

Miguel Almanzar and Maximo Torero, Media Coverage and Food Commodities: Agricultural Futures Prices and Volatility Effects, ZEF – Discussion Papers on Development Policy No. 246, Center for Development Research, Bonn, October 2017, pp. 62.

ISSN: 1436-9931

Published by:
Zentrum für Entwicklungsforschung (ZEF)
Center for Development Research
Genscheralle 3
D – 53113 Bonn
Germany
Phone: +49-228-73-1861
Fax: +49-228-73-1869
E-Mail: zef@uni-bonn.de
www.zef.de

The author[s]:
Miguel Almanzar, International Food Policy Research Institute (IFPRI). Contact: M.Almanzar@cgiar.org
Maximo Torero, World Bank Group. Contact: mtorero@worldbank.org

Acknowledgements

The views and opinions expressed in this article are those of the authors and do not reflect the official position of the World Bank or IFPRI.

We thank Per Pinstrup-Andersen and two anonymous referees for helpful comments on an initial draft which were of great assistance and improved the quality of this version of the paper.

Abstract

We examine how media coverage of fluctuations in the price of agricultural commodities affects these prices and their volatility. We develop a unified empirical framework to analyze the media's effects on both returns and volatility using insights from the literature. We use daily prices of futures contracts for soybeans, hard wheat, soft wheat, rice, and maize, complemented by a unique dataset that follows a comprehensive set of global media outlets and uses an algorithm to determine sophisticated relationships among phrases in a news article which signal an increase or decrease in the price of those four commodities.

We find price effects that are economically important in size. Our estimates imply a net increasing effect of media coverage on the price of these four commodities; these effects are mostly concentrated in 2012 and from 2015 onwards, meaning that these effects are important in periods of both high and low prices. Across commodities, the price effects are concentrated in soybeans and maize. We find robust evidence that media coverage decreases volatility for these agricultural commodities on average for the period we study. The effects on volatility balance each other, with decreasing price coverage decreasing the variance of returns and increasing price coverage increasing the variance of returns of futures contracts of these commodities; however, the increase is than the decrease. Our results suggest that media coverage increases periods of normal volatility and decreases periods of excessive volatility.

These results point to the potential of using media coverage to bring attention to price surges and to decrease volatility during food crises or times when there is above-normal volatility. The dynamics between the price of agricultural commodities and media coverage may help prevent knee-jerk policy reactions by discouraging market overreaction, encouraging market stability, and promoting food security. They highlight crucial role of providing appropriate information as fast as possible so media coverage and reflects the fundamentals that drive food commodity prices.

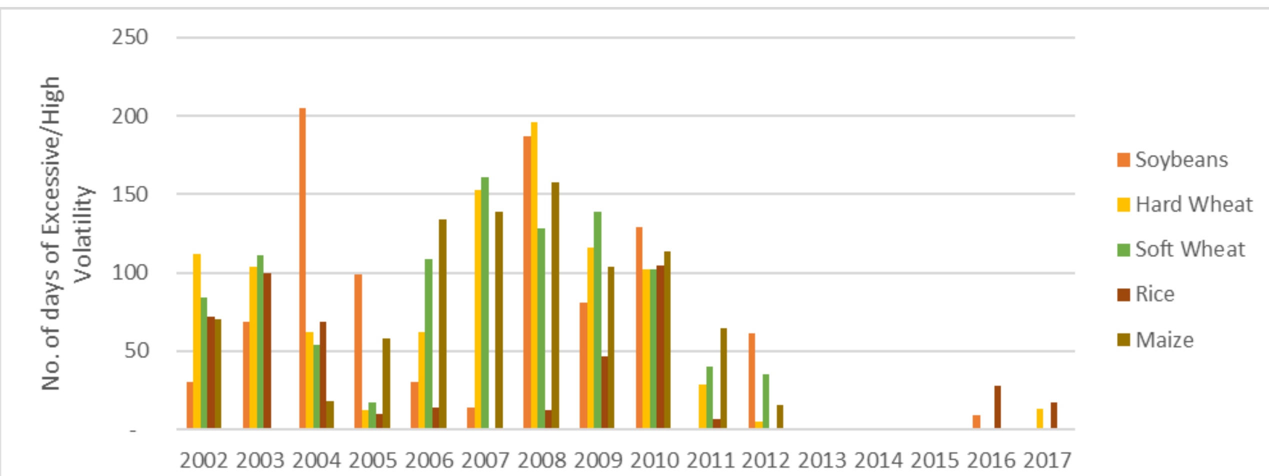
Keywords: sentiment analysis, textual analysis, agricultural food commodities, price spikes, volatility, media coverage, market efficiency

JEL codes: G13, G14, Q11, Q14, Q18

1. Introduction

The world faces a new food economy that likely involves both higher and more volatile food prices. After the food price crisis of 2007–2008, food prices started rising again in June 2010, with international maize prices doubling by the beginning of 2011 and international wheat prices increasing by a factor of 1.5 by May 2011. The peak came in February 2011, in a spike that was even more pronounced than that of 2008, according to the Food Price Index of the Food and Agriculture Organization (FAO) of the United Nations. Although the food price spikes of 2008 and 2011 did not reach the heights of the 1970s and although food prices have been decreasing in the last few years, price volatility — the amplitude of price movements over a particular period — was at the highest t level in the past 50 years during the 2007-2008 period. This volatility has affected wheat and maize prices in particular. For soft wheat and maize, for example, there were over 100 days of high or excessive price volatility annually between December 2006 and December 2010, according to a measure of price volatility recently developed at the International Food Policy Research Institute (IFPRI). Figure 1 shows the evolution of this volatility measure for soybeans, soft and hard wheat, rice, and maize from 2002 to July 2017 using the volatility measures developed in Martins-Filho, Yao and Torero (2015, 2016). The figure highlights a decrease in volatility seen from 2011-2017.

Figure 1 Evolution of the Number of Days of High and Excessive Price Volatility



Source: Authors' own calculations

Note: This figure shows the results of a model of the dynamic evolution of daily returns based on historical data going back to 1954. A period characterized by extreme price variation (volatility) is a period of time in which we observe a large number of extreme positive returns. An extreme positive return is defined to be a return that exceeds a value of return with 5 percent probability.

High and volatile food prices are two different phenomena with distinct implications for consumers and producers. High food prices may harm poorer consumers because they need to spend more money on their food purchases and therefore may have to cut back on the quantity or the quality of the food they buy or economize on other needed goods and services (Torero, 2012, 2016). For food producers, higher food prices could raise their incomes — but only if they are net sellers of food, if increased global prices feed through to their local markets, and if the price developments on global markets do not also increase their production costs.

Price volatility also has significant effects on both food producers and food consumers. Greater price volatility can lead to greater potential losses for producers because it implies price changes that are larger and faster than what producers can adjust to. Uncertainty about prices makes it more difficult for farmers to make sound decisions about how and what to produce. For example, which crops should they plant? Should they invest in expensive fertilizers and pesticides? Should they pay for high-quality seeds? Without a good idea of how much they will earn from their products, farmers may become more pessimistic in their long-term planning and may dampen their investments in areas that could improve their productivity[1]. By reducing supply, such a response could also lead to higher prices, which in turn would hurt consumers (Torero, 2012).

In rural areas, the line between food consumers and food producers is blurry because many households both consume and produce agricultural commodities. Therefore, if prices become more volatile and these households reduce their spending on seeds, fertilizer, and other inputs, this may affect the amount of food available for their own consumption. Even if households are net sellers of food, producing less and having less to sell will reduce their household income and thus still affect their consumption decisions.

As highlighted in the 2011 Global Food Policy Report (IFPRI, 2012), increased price volatility over time can also generate larger profits for investors, drawing new players into the market for agricultural commodities. Increased price volatility may thus lead to increased — and potentially speculative — trading that in turn can exacerbate price swings further.

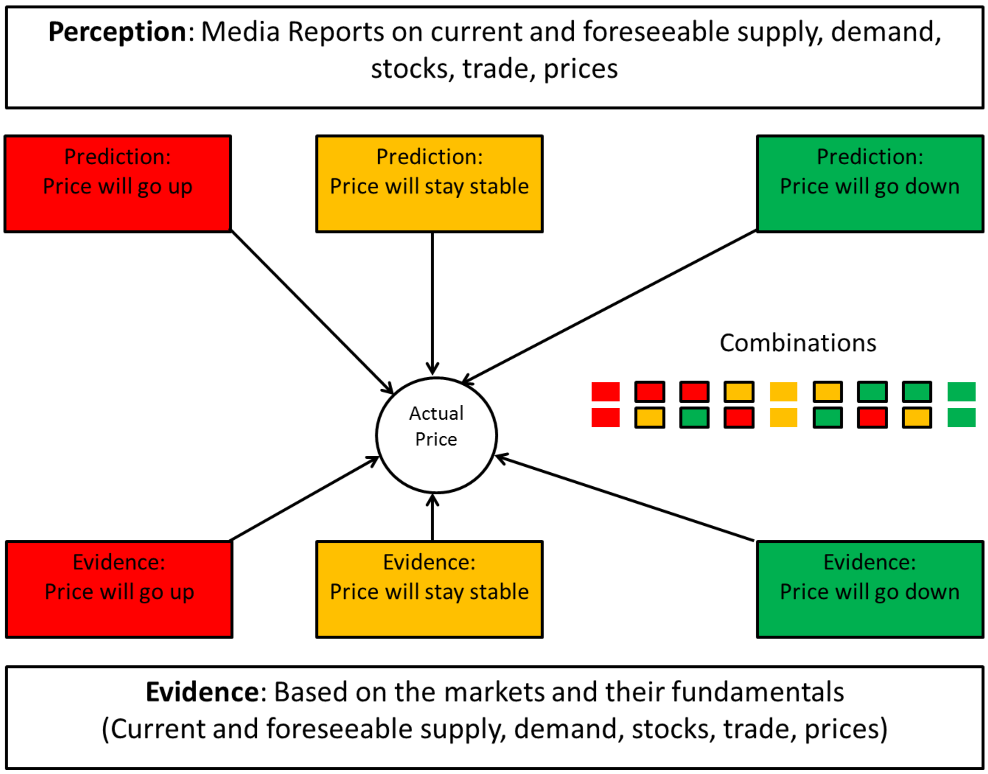
There has been long standing and increasing interest among researcher on the relationship between media and its effects on different financial tools and markets in general markets. Smith, van Ravenswaay, and Thompson (1988) and Carter and Smith (2007) estimate the effects on milk and corn prices after news of contamination of the milk and corn supply. Pruit (1987) on the prices of agricultural futures for commodities produced near Chernobyl after the nuclear accident. More recently in the financial economics literature, Tetlock (2007) and Ahmad et al. (2016) showing that not only coverage and tone of the media predicts changes in stocks even if

[1]The positive relationship between price volatility and producers‘ expected losses can be modeled in a simple profit maximization model assuming producers are price-takers. However, there is no uniform empirical evidence of the behavioral response of producers to volatility.

they are not necessarily related with “new” information[2]. Given the importance of both prices and price volatility, this paper explores how the attention paid by the media to the movement of food prices affect those prices and their volatility. We examine the effects of media coverage of commodity price increases and decreases on the price of the agricultural commodity covered. The mechanisms we advance in our analysis are depicted in Figure 2. For each commodity, there are **evidence-based market fundamentals,** like current and foreseeable supply, demand, stocks, trade, and current prices; these fundamentals allow us to predict the price realization for the specific commodity. There are three clear “possible futures” based – with margins of error – on this evidence: prices will either (1) go up, (2) stay stable, or (3) go down. On the other side, there is the **perception portrayed through media reports**, which - in an ideal world - would just amplify the experts’ opinion on “possible futures.” The actual or realized price then can reflect nine combinations based on the evidence and the media perception forecast. There are three combinations in which price realization based on market fundamentals and the reporting on these price developments in the media is identical. In these cases, the marginal effect of media should be minimal as efficient market prices should reflect the available information (Fama, 1970). On the other hand, the other six combinations, in which evidence and perception differ (where, for example, all market fundamentals show that prices will stay stable or even fall, but the media reports that prices will increase) could be a case in which the media can have a significant influence on food prices.

[2] A brief literature review that includes examples in the agricultural economics and financial economics literature is presented in Section 3.

Figure 2 Mechanisms of Effects of Media on Prices



In this paper, we examine how media coverage of changes in the price of agricultural commodities affects these prices and their volatility. The overarching goal is to quantify the size of these effects (if any) and to gauge the extent to which the media can be used as a policy tool to mitigate the negative effects of price spikes and volatility in agricultural commodities. We contribute to the literature by providing evidence of the relationship between media coverage and agricultural commodities' prices and volatility. First, we develop a unified empirical framework to analyze the effects on both returns and volatility using insights from the literature on the analysis of information in financial markets and compare the results from our model to other more common models in the literature. Second, we use a unique dataset to construct a measure of media coverage to estimate the effect of the intensity of media coverage on the price dynamics in these markets. The dataset follows a comprehensive set of global media outlets and uses an algorithm to determine sophisticated relationships in phrases in a media article which signal an increase or decrease in price.

We find price effects that are economically important in size. Our estimates imply a ***net increasing effect of media coverage on the price of these commodities***; these effects are mostly concentrated in 2012 and from 2015 on, implying that these effects are important in periods of both high and low prices. Across commodities, the price effects are concentrated in soybeans and maize. The findings that more news reports of increases and decreases in prices reinforce price movements in the direction of the report strengthen the case that increased media coverage in some periods of food crises can exacerbate price spikes.

We also find robust evidence that media coverage decreases volatility for these agricultural commodities on average for the period we study. The effects on volatility balance each other, with ***decreasing price coverage decreasing the variance of returns*** and ***increasing price coverage increasing the variance of returns*** but to a lesser extent than decreasing price. Coverage of increased prices increase variance by 3 percent, while and coverage of price decreases increases variance by 4.6 percent in our preferred model. Overall, in the period we study and across the different models, the evidence points to a ***decreased volatility effect due to media coverage*** for these commodities. Finally, our results suggest that media coverage **increases periods of normal volatility** and **decreases periods of excessive volatility**. This points to the potential of using media coverage to bring attention to price surges and, at the same time, to decrease volatility during food crises or times when there is above-normal volatility.

The rest of the paper is divided into eight sections, including the introduction. Section 2 presents the background and examples of the effects of media coverage, Section 3 presents a short literature review of the effects of media and information on prices; Section 4 presents our empirical framework and describes our estimation strategy; Section 5 describes the data used; Section 6 presents the results on the effects of media coverage on prices and volatility; Section 7 presents additional robustness checks, and, Section 8 concludes.

2. Background

As background, we present two examples: the Russian wheat export ban of 2010 and the increase in maize and soybean prices in the summer of 2012 due to a drought in the US Midwest.

In 2010, the global media overreacted to the news of Russia's wheat export ban, failing to explain that global wheat production and stocks were sufficient to compensate for the loss of Russian wheat. Moreover, every piece of news tracking the ban from August to October 2010 — even the US Department of Agriculture's better-than-expected projection that the world would harvest only 5 percent less wheat than the previous year — seemed to elicit a spike. The number of media articles on the price of wheat rose significantly between August and October 2010. The average quarter for 2010 had 122 articles mentioning that wheat prices were increasing; however, the quarter from August to October 2010 had 210 articles, i.e. 72 percent higher[3]. From those articles in the period of August-October 2010, 82 percent of articles mentioning wheat said that the price of wheat was going to increase. This figure was 69 percent in the previous quarters in 2010, a 12 percentage-point difference.

Table 1 shows that among the major reasons for the price increases reported in the media were the fires in Russia, with 62 percent of the articles referencing wheat price increases due to disasters. Note that even though global inventories and stocks were sufficient and were significantly higher than in the 2008 crisis, 25 percent of the articles reference a price increase due to low inventories from low production and stocks. Finally, only 7 percent of articles referred to policies, such as export bans, which were in fact the actual major reason for the increase in prices. This lack of information on global production led governments around the world to engage in panic buying, which exacerbated the situation and pushed prices up further.

Another example in which the media overreacted to conditions in the market was the increase in global maize and soybean prices in the summer of 2012. Prices skyrocketed during this period, and experts feared that price increases would continue unabated due to ongoing dry weather in the US Midwest. The US corn crop was hard-hit by the drought conditions, which began in May 2012 and stunted crops in the crucial pollination phase. While US government officials argued that an increase in maize acreage would offset the drop in yields, agricultural and trade analysts feared that the length and severity of the drought could continue to have a substantial impact on prices. However, as seen in Figure 3, the prices of maize experienced a decrease in the months

[3] To analyze media articles, we use Sophic Intelligence Software, which is built on the Biomax BioXMä Knowledge Management Suite. Each day, global food- and commodity-related news articles are loaded into Sophic Intel for linguistic analysis and semantic object network mapping. Sophic Intel generates wiki reports and heatmaps based on terms and phrases found in press articles that influence commodity price volatility and food security.

following October 2012. Similarly, soybean prices experienced sharp spikes in the summer of 2012, seeing their highest levels in nearly four years. This jump in prices was caused by a combination of dry weather throughout the US and South America, decreased acreage in the US in favor of more profitable maize, record levels of Chinese imports, and a subsequent rapid rate of stock disappearance.

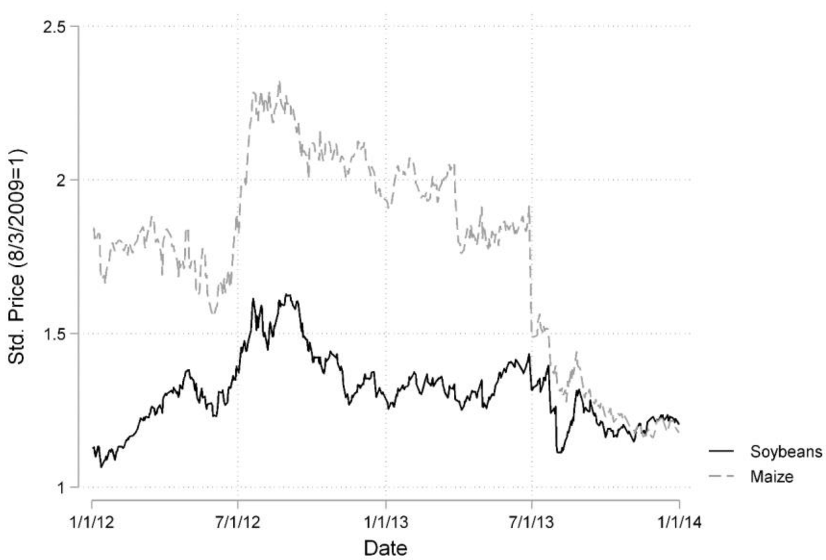
Table 1 Analysis of Media articles referring to wheat prices during the Russian export ban

Reason referred in media article	Reference to prices going up	
	All of 2010	Aug - Oct 2010
Financial	42	10
Inventories	99	40
Policies	37	12
Disasters and Civil Effects	159	101
Total of references to prices going up	337	163
Total articles	585	288

Source: Authors' calculations

Note: The periods correspond to the following dates: All 2010 - refers to January 1, 2010 to December 31st, 2010; and Aug-Oct 2010 - refers to 1st of August 2010 to October 31st, 2010. Some examples for each category are: (a) financial: domestic food price, expectations, expected prices, futures markets, trade barrier, trading volume; (b) inventories: production, domestic production, domestic supply, emergency reserves, storage, supply, surplus; (c) policies: export ban, export quota, food security, import quota, price controls, taxes; and (d) disasters and civil effects: drought, earthquake, famine, fire, flood, riots.

Figure 3 Daily Prices of Soybeans and Maize January 2012 to January 2014



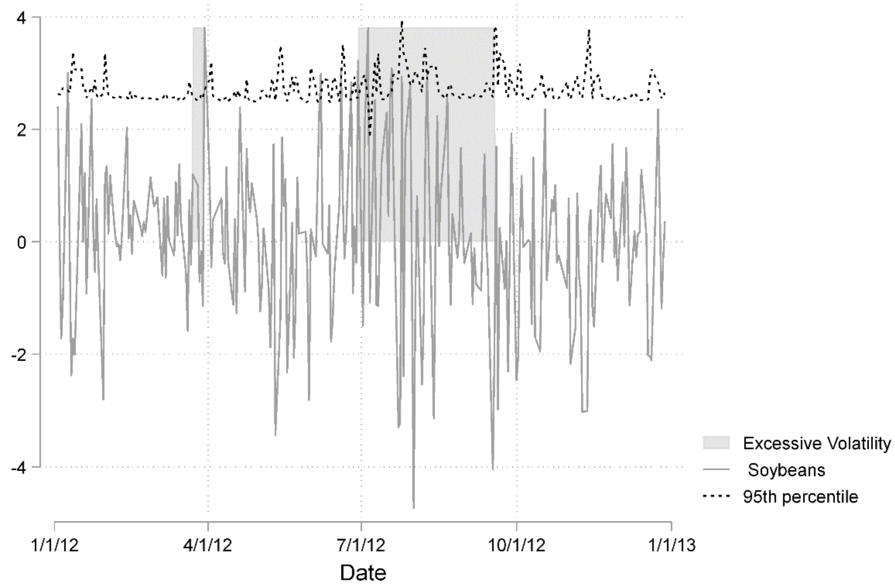
Source: Authors' own calculations using data from the Chicago Board of Trade

In the case of volatility, the period shows significantly higher volatility, with realized returns above the forecasted 95th percentile returns in several instances. There was a spell of excessive volatility in the price of soybeans and a shorter period of excessive volatility in maize prices at the begining of the price spike.

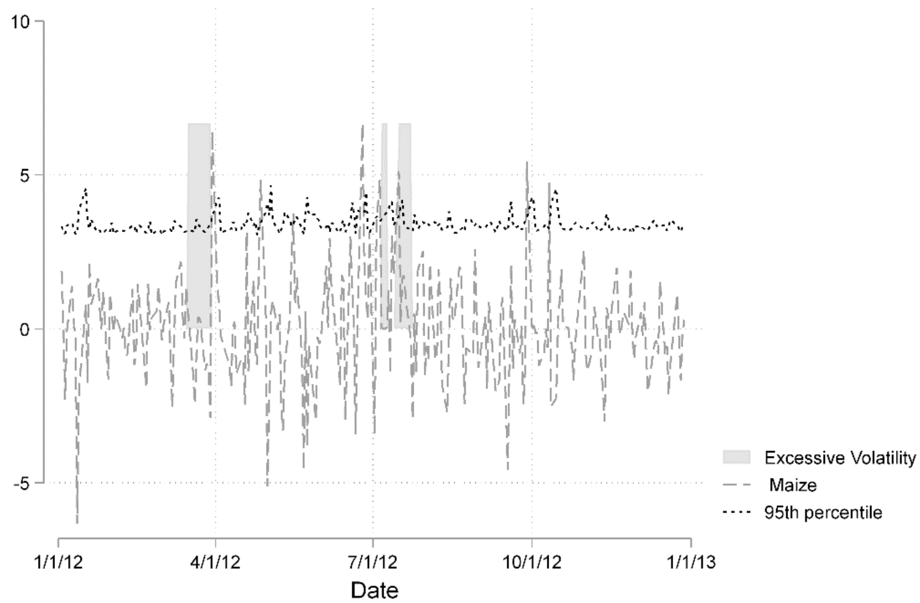
During this period, media coverage of maize and soybean prices was consistently high, as seen in Figure 5, with the majority of articles relating to increased prices between July and October of 2012.

Both of these examples show that it is important for policymakers to not react with knee-jerk policies such as stockpiling and export restrictions. While such policies may appease the population of a particular country or region, they can have devastating consequences for global food prices and food security.

Figure 4 Soybeans and Maize Price Volatility, January – December 2012



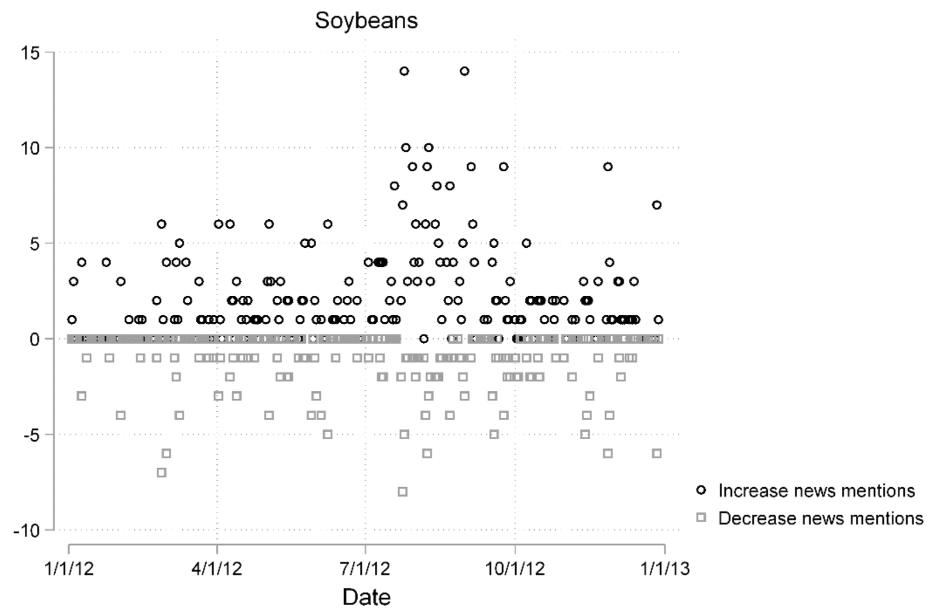
Source: Authors' own calculations using data from the Chicago Board of Trade and volatility procedure in Martins-Filho, Yao and Torero (2016)

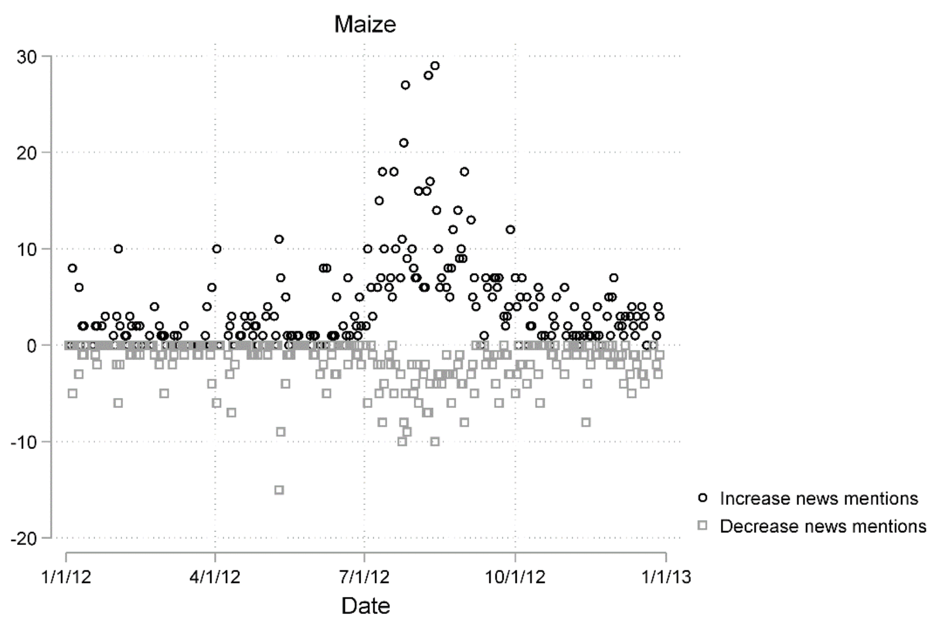


Source: Authors' calculations using data from the Chicago Board of Trade volatility procedure in Martins-Filho, Yao and Torero (2016)

Note: An abnormality occurs when an observed return exceeds the 95 percent conditional quantile — that is, a value of return that is exceeded with low probability (5 percent).

Figure 5 Media Coverage Mentioning Maize and Soybean Prices, 2012





3 Previous literature: Media, information and prices

The effect of information shocks on markets has a long history in economics. The efficient market hypothesis in its simplest form purports that market prices should 'fully' reflect available information. Generally, the tests of this hypothesis are of the semi-strong form, which investigate whether prices efficiently adjust to the information that is available (Fama, 1970). These tests exploit the variation induced by informational events, such as stock splits, policy announcements, dividend information, etc., essentially comparing abnormal prices to the dates of informational events. An economically sensible version of the efficient market hypothesis, where we see price reactions depending on the tradeoff between the benefits and cost of acting on additional information, fares well with the data in the empirical literature (Fama, 1970, 1991).

The effects of informational events on futures[4] prices have been studied by various authors, and thus differences in methodology and in what is meant by 'informational event' abound. Rucker et al. (2005) estimate the effect of several types of events (periodic, aperiodic, and irregular) on lumber futures prices to help shed light on the volatility of lumber prices. They find that periodic and aperiodic events are absorbed quickly compared to irregular events. Their study is not a test of market efficiency, however, since the authors do not exploit variation in timing of the events but are rather interested in the structural aspects of the market response to the types of events studied.

Other studies explore how unexpected news events affect market prices. Pruit (1987) studies the effects of the Chernobyl nuclear accident on the agricultural futures commodity prices produced in the Chernobyl area. He exploits the evolution of the news in the days surrounding the accident and finds that the commodities studied experienced a short-lived increase in volatility and that prices were affected as the efficient market hypothesis would predict. Continuing with the effects of adverse events, Carter and Smith (2007) estimate the effect of news concerning the contamination of the corn supply on the price of corn. They find that prices were affected and that the negative effect persisted for at least a year.

Another vein of study explores the effects of news on recalls and food safety on the prices of affected products. McKenzie and Thomsen (2001) find that red meat recalls due to contamination and food safety information negatively affects beef prices, but this transmission does not occur across all margins. Specifically, they find that farm-level prices do not respond to this information. In a similar study, Schlenker and Villas-Boas (2009) explore the effects that information on mad cow disease had on purchases and futures prices. They find that future prices were negatively

[4] A futures contract is represented as an agreed-upon price of a commodity or financial instrument to be delivered on a future date.

affected by the discovery of the first mad cow and that information that is not "news (for example, a talk show host just providing the information available on mad cow disease and thus just bringing attention to the issue) had half of the effect of the initial announcement of the discovery. Smith, van Ravenswaay, and Thompson (1988) study the impact of milk contamination on consumer demand and find that media coverage had an impact on demand for milk, with negative media coverage having larger impacts. These studies show that media coverage can have large impact on food prices, regardless of whether the information takes the form of official 'news' or is just bringing attention to the issues and propagating information regardless of its accuracy.

In the finance literature, there has been an increased interest in the relationship between media coverage and financial markets. Fang and Peress (2009) who find that media coverage can decrease information frictions and affect prices in in the stock market. They find that stocks with no media coverage have a premium when compare to those with high media coverage and that this effect is more pronounced for small stocks and stocks with high volatility. Tetlock (2007) who examines how qualitative information in media coverage in a popular column of the Wall Street Journal affects the stock market. Using the fraction of negative words in a popular news column about the stock market he finds that media pessimism predicts decreases in market prices but not from media reflecting market fundamentals or volatility. Tetlock, Saar-Tsechansky and Macskassy (2008) extend this work using individual firms' stock returns and a larger set of financial news stories on these firms. They find that negative words about a firm or the sentiment in media content can reflect hard-to-quantify aspects of firms' fundamentals that are quickly incorporated in the stock prices. More recently, Ahmad et al. (2016) arrive to a similar conclusion using data from US firms returns to construct a measure of negative media "tone." They find that media impacts are sometimes quickly reversed, while at other times they endure; this signals that media content and analysis can sometimes just be 'noise' but can other times contain relevant information or news. Engelberg and Parsons (2011) estimate the effect of local media coverage on the behavior of investors after earnings announcements and find that it is strongly related to whether the local paper covers the announcement and that the effect depends the timing of local reporting.

Media coverage of price changes can be a signal of volatility in a market. Given the extreme prices in food commodities that we observed during 2011-2012, the issue of the effect of media coverage on food commodity prices is increasingly important. News reports of food price increases and decreases do not always provide 'new' information to markets, as these news articles are reporting the tendencies of the price series as they occur. However, focusing attention on the dynamics of prices can serve as a signal of other underlying issues and could reinforce the tendency by updating the beliefs of both investors and consumers. Exaggeration or downplay of the importance of price increases by the media can cause welfare losses, given that agents will make decisions based on information that may not reflect the true nature of the pricing process.

4 Empirical framework

4.1 Price Effects

To quantify the effect of media coverage of price increases and decreases on food prices and volatility, we depart from a simple market model that accounts for global trends in market fundamentals and focus on the variation that is not explained by current market conditions.

A dynamic panel with fixed effects regression of price levels to account for serial correlation in prices is not consistent for a panel with small N or for small cross-sectional units (number of commodities, in this case; Nickell, 1981). Other methods that rely on a large time series component have been developed; in particular, instrumental variables using a general method of moments (IV-GMM) can be used to estimate these types of models (Anderson and Hsiao, 1981; Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998). However, these methods suffer from weak instrument problems, and the number of instruments grows with the time series (T) component of the panel (Han and Phillips, 2010). We use insights from these methods to estimate the impact of our media variables in the inherently dynamic structure of the price data.

The equation we estimate is:

(1) $p_{i,t} = \alpha_i + \theta p_{i,t-1} + \mu \cdot UP_{it} + \gamma \cdot DOWN_{it} + \beta \cdot X_t + \varepsilon_{it}$

where:

i= Soybeans, Hard Wheat, Soft Wheat, Rice, and Maize

$t = 1 \ldots T$ denotes the time periods

p_{it} is the log price returns (to account for the unit root in the price level)

α_i is a commodity-specific intercept (fixed effect)

UP_{it} is the number of 'increase in price of *i* news for day *t*

$DOWN_{it}$ is the number of 'decrease in price' of *i* news for day *t*

X_t is a matrix of market variables at date t

ε_{it} is a random error term

We assume that the news variables are predetermined or sequentially exogenous - that is, that $E[\varepsilon_{it}|X_t, Up_{i,t-k}, Down_{i,t-k}] = 0 \; for \; k = 1 \ldots t$, which allow us to use moment restriction to obtain a GMM-IV estimator. This equation might also contain lags of the regressors and/or additional lags of the dependent variable, but it captures the essential feature of the model that we want to estimate, namely, a dynamic effect of media coverage on the price level for which the speed of adjustment is governed by the coefficient of lagged price level[5].

The sequential exogeneity assumption implies that the regressors are uncorrelated to past and present values of the error term. It does not rule out correlation between the regressors and the individual effect. Lagged price levels will be correlated by construction with the fixed effect and with the lagged error term, but it may also be correlated with contemporaneous ε if ε is serially correlated, which is not ruled out by the sequential exogeneity assumption. Thus, the lagged dependent variable is effectively an endogenous explanatory variable in the equation with respect to both α_i and ε.

To derive the moment conditions to estimate the parameters in (1), we follow the Arellano-Bond-Blundell-Bover procedure and difference (Δ) equation (1) to obtain:

(2) $\Delta p_{i,t} = \theta \Delta p_{i,t-1} + \mu \cdot \Delta UP_{it} + \gamma \cdot \Delta DOWN_{it} + \beta \cdot \Delta X_t + \Delta \varepsilon_{it}$

From the sequential exogeneity assumption, we can see that $p_{i,t-2-k}$, $\Delta p_{i,t-1-k}$, $Up_{i,t-1-k}$, $Down_{i,t-1-k}$, $\Delta Up_{i,t-1-k}$, $\Delta Down_{i,t-1-k}$ for $k = 0 \dots t-1$ are orthogonal to error term.

We can derive the moments conditions to estimate (2) based on the lag levels and/ or the lag differences and recuperate the parameters from (1). The moment conditions for the parameters of interest are, for $k = 1 \dots K$:

(3) Lag Difference Instruments:

$$\mathrm{E}\left[\Delta p_{i,t-1-k} \cdot \varepsilon_{it}\right] = 0, \mathrm{E}[\Delta UP_{it-k}\, \varepsilon_{it}] = 0 \; and \; \mathrm{E}[\Delta Down_{t-k} \cdot \varepsilon_{it}] = 0$$

(4) Lag Instruments:

$$\mathrm{E}\left[p_{i,t-2-k} \cdot \varepsilon_{it}\right] = 0, \mathrm{E}[UP_{it-k} \cdot \varepsilon_{it}] = 0 \; and \; \mathrm{E}[Down_{t-k} \cdot \varepsilon_{it}] = 0$$

This restricted IV-GMM estimator based on these moment conditions, $\nu = [\theta, \mu, \gamma]$ is consistent as $T \to \infty$, as long as $E[Z_{it}\varepsilon_{it}] = 0$, since it retains its time series with the regressors being predetermined, where Z_{it} is the matrix of instruments, i.e the market controls, lagged differences, lag levels.

The moment conditions assume that for a set of k values, the lags or the lag differences in the price and the news variables are uncorrelated with the errors at time t. In theory, we could use all past differences; however, this would worsen the weak instrument problem that is inherent

[5] To avoid issues of cointegration of commodity prices and exchange rates, we use the return to the market (exchange) variables, which are stationary.

in this assumption, as the number of available instruments increases with T. This kind of procedure was devised for small T and large N panels (Arellano and Bond, 1991), which is not the case in our study. In our case, we have a large T, and OLS and fixed effects maintain their consistency. In addition, under the market efficiency hypothesis, sequential exogeneity is a plausible assumption, as all the information in the news variables in the time periods before t should be already reflected in the price of the commodity. The parameter identified with the IV-GMM estimators from equation (2) is the effect of the intensity of media coverage (measured by the number of articles mentioning increases and decreases in prices). The OLS estimates of the media variables (1) are consistent under our assumptions, and the restricted IV-GMM estimates exploits past variation in the media coverage in the number of articles mentioning price changes (lag instruments) or the changes in coverage in the past days (lag differences instruments), perhaps better signaling when the change in price is first realized.

The long-term effects of these variables can be computed by: $\check{\mu} = \frac{\mu}{1-\theta}$ and $\check{\gamma} = \frac{\gamma}{1-\theta}$. We cluster the standard errors by date and allow for auto-correlated (AR1) common disturbances and arbitrary heteroskedasticity, using a truncated kernel as recommended in Thompson (2009). This allows standard errors to adjust for the possibility that the errors have the following form:

$$\varepsilon_{it} = \varphi \cdot \psi_{t-1} + d_t + \epsilon_{it}$$

4.2 Volatility Effects

To explore the effects of media coverage on price volatility, we estimate the following model using the residual from the previous estimations (following Ohlson and Penman, 1985; Dubofsky, 1991). This model follows the financial economics literature to estimate a daily excess return (e_{it}) after controlling for market conditions:

(5) $e_{it}^2 = \rho e_{it-1}^2 + \nu \cdot UP_{i,t} + \psi \cdot DOWN_{i,t} + \pi_{it}$

where

$e_{it} = p_{it} - \hat{\theta} p_{it-1} - \hat{\alpha}_i - \hat{\mu} \cdot UP_{it} - \hat{\gamma} \cdot DOWN_{it} - \hat{\beta} X_t$, is the residual from the regression in (1)

π_{it} is an error term

UP_{it} is the number of 'increase in price of i news for day t

$DOWN_{it}$ is the number of 'decrease in price' of i news for day t

In addition, using this model, we compute 30-day rolling variance estimates to estimate the effects of the media variables on a smoother approximation of the variance (related to realized volatility in the asset pricing literature, Andersen and Benzoni, 2008).

Let $s_{i,t}^2 = \frac{\sum_{k=1}^{30} e_{i,t-k}^2}{30}$ and we estimate a smoothed version of (5). Namely,

$$(6)\ s_{i,t}^2 = \rho {\hat{s}_{i,t-1}}^2 + \nu \cdot UP_{i,t} + \psi \cdot DOWN_{i,t} + \pi_{it}$$

In both the price and the volatility estimations, we present the OLS and restricted IV-GMM estimates for sensitivity and robustness. Since we have a long panel time series dimension and a few cross-sectional units, we use procedures that require large T and include commodity-fixed effects to account for persistent commodity shocks and allow a flexible specification of the error terms to allow for persistent common shocks. Our procedures exploit the long-time series aspect of the data and allow for a flexible data-generating process for the error term, uncovering the causal or structural relationship between media coverage and prices in this dynamic pricing framework.

5 Data Sources and Description

We use various data sources to estimate the impact of media coverage on agricultural futures markets. The first is daily futures price data from the Chicago Board of Trade for futures of maize, soft wheat, soybean, and rice and from the Kansas City Board of Trade (prior to its close in 2015) for hard wheat. The future contract prices selected are the closest to maturity each day (contracts expiring between one and three months). To avoid double-counting of futures contracts, only a single contract on each commodity is used in a given day.

We augment these data with market variables, including the SP500 index, obtained from the St. Louis Federal Reserve Bank, and the daily exchange rates between the US dollar and the currencies of major participant countries/regions in the agricultural commodity markets, obtained from the Federal Reserve Bank. Exchange rates are included for Australia, the European Union, Brazil, Canada, China, India, Mexico, and Thailand, in addition to the nominal broad dollar index and the nominal major currencies dollar index.

The informational event variable, or the media coverage measure, is constructed from a list of global news sources and an algorithm that relates words in the articles with signs of increasing prices and decreasing prices. Every day, IFPRI monitors a comprehensive set of RSS (Really Simple Syndication) feeds drawn from global media outlets via Google news. A total of 31 feeds related to global food prices and food security are monitored; these feeds include search strings such as "food prices," "food crisis," "agricultural development," "commodity prices," "price of maize," "price of wheat," "price of rice," "price of soybean," etc. Articles are tagged if they are about:

1. global food security or food prices,
2. ongoing national, regional, or global food crises,
3. prices (international, regional, and national) or crop conditions of major agricultural commodities (wheat, corn, soybeans, and rice),
4. oil prices,
5. agricultural trade (export bans, import or export forecasts, etc.), or
6. agricultural/food policy research, such as new IFPRI reports.

At the end of each day, tagged articles are converted into .txt files and saved using the format "title_month_day_year.txt." The ".txt" files for each day are then uploaded into the IFPRI Food Security Analysis System, a tool built by Sophic Systems Alliance. This tool uses linguistic and semantic object network-mapping algorithms to analyze the relationships between key terms found in each article. When articles are uploaded each day, the tool mines the complete database of articles for a select set of key words.

The system provides daily reports analyzing movement (increases or decreases) in agricultural commodities prices. These reports provide a count of the number of articles each day with "up" or "down" movements for each commodity by analyzing the text within the articles and looking at phrases in the articles that influence commodity price volatility and food security.

We use a list of key words, synonyms, and relations functions to determine an "up" or "down" movement within our database of articles. For example, an article containing the words "soybean" and "surge" would denote an "up" movement in soybean prices; if the soybean "up" report on a given day is listed as "5", this means that on that day, of the articles uploaded, we found five occurrences (or mentions) containing words suggesting a rise in soybean prices[6].

The system can detect more sophisticated relationships to determine whether the phrase in the article is an "up" or a "down." Namely, we use four categories of informational events that can be related to an increase or decrease in prices: **Financial, Inventories, Polices, and Disasters.** For example, a phrase that indicates that inventories are decreasing is related to an "up" mention, while a mention in the financial category that the price of maize is over-valuated is related to a "down" mention.

We use these "up" and "down" variables to measure the intensity of media coverage of a price increase. Articles that are published on weekends and holidays, when the market is closed, are moved to the next day the market is open. With these data, we construct a panel of five agricultural commodities: S*oybean, Soft Wheat, Hard Wheat, Rice, and Soybeans*. The data span the period from August 3, 2009 to July 28, of 2017. In "market time," we obtain 1,940 days for each commodity, or 9,700 observations in the panel.

Using these series, we construct daily log-differences or returns for futures, defined as $p_{it} = 100 * \ln(\frac{P_{i,t}}{P_{i,t-1}})$, where P_{it} is the closing price for commodity *i* on day *t*. The price series for the different commodities is presented in Figure 6, where we standardize the price to the initial day in the analysis. The figure shows that all the agricultural commodities have seen both large increases and large decreases in the study period; maize, wheat, and soybeans experienced big spikes in mid-2012 and rice saw large spikes after January 2011. All the commodities prices have decreased after a spike in 2014 and are at similar levels at the beginning and the end of the study period.

[6] More details on the sources and keywords used in the media analysis is available in: "Appendix 1 - Details of Media Data and Sources".

Figure 7 shows the returns for each commodity; we can also see the volatility of the returns at different points in the study period, as the band in the figure widens at some points.

Table 2 shows the summary statistics for the variables of interest in in the analysis, and Table 3 shows the descriptive statistics for the market variables that are used as covariates in the analysis. During the study period, the prices of these commodities varied considerably; on average, prices increased by 11 percent. On average, the price of soybeans increased by 7 percent above the initial price in August 2009, while average maize prices increased by 36 percent. Only rice had an average price below its initial August 2009 price, with an average decrease of 4 percent during the study period. The returns are, on average, negative for all commodities except maize and are below the returns for the SP500 index and the major currencies dollar index during the study period, which were 0.047 percent and 0.009 percent, respectively. The average *daily* returns across commodities ranges from -0.009 percent for hard wheat to 0.003 percent for maize. We bring attention to the higher volatility in commodities returns, as evidenced by the larger standard deviation when compared with the exchange rates returns and the SP500 index. The largest negative return in the sample is for maize, at -25.23 percent in a day, while the biggest gains in returns are for wheat, with 13.12 percent in a day. This is compared to the SP500 Index, which ranges from -6.90 percent to 7.34 percent during this period.

As for the media coverage variables, we see that the commodity that is most active in the media sources we track is wheat, with 5.95 mentions of increases or decreases on average per day during the study period, followed by maize with 5.49 mentions on average. For both commodities, the mentions in the media tend to be about prices increasing, with 3.44 mentions per day for wheat and 3.09 mentions per day for maize during this period.

Table 2 Summary Statistics: Prices and Media coverage

	Mean	Std. Dev.	Median	Min	Max	N
Soybeans						
Std. Price (8/3/2009=1)	1.07	0.20	0.99	0.78	1.63	1,940
Returns	-0.004	1.51	0.00	-17.70	6.46	1,939
Increase news mentions	1.95	2.70	1.00	0.00	28.00	1,940
Decrease news mentions	1.58	2.31	1.00	0.00	18.00	1,940
All news mentions	3.53	4.61	2.00	0.00	42.00	1,940
Hard Wheat						
Std. Price (8/3/2009=1)	1.08	0.27	1.08	0.63	1.70	1,940
Returns	-0.009	1.83	0.00	-8.00	9.21	1,939
Increase news mentions	3.44	4.02	2.00	0.00	51.00	1,940
Decrease news mentions	2.51	3.13	1.00	0.00	23.00	1,940
All news mentions	5.95	6.42	4.00	0.00	61.00	1,940
Soft Wheat						
Std. Price (8/3/2009=1)	1.08	0.24	1.04	0.66	1.71	1,940
Returns	-0.007	2.02	-0.03	-9.09	13.12	1,939
Increase news mentions	3.43	4.08	2.00	0.00	51.00	1,940
Decrease news mentions	2.51	3.13	1.00	0.00	24.00	1,940
All news mentions	5.94	6.48	4.00	0.00	70.00	1,940
Rice						
Std. Price (8/3/2009=1)	0.96	0.16	0.99	0.67	1.35	1,940
Returns	-0.005	1.49	-0.06	-6.57	7.58	1,939
Increase news mentions	1.81	2.67	1.00	0.00	26.00	1,940
Decrease news mentions	1.26	1.96	0.00	0.00	25.00	1,940
All news mentions	3.07	4.18	2.00	0.00	37.00	1,940
Maize						
Std. Price (8/3/2009=1)	1.36	0.41	1.14	0.84	2.32	1,940
Returns	0.003	1.86	0.00	-25.23	10.36	1,939
Increase news mentions	3.09	3.75	2.00	0.00	29.00	1,940
Decrease news mentions	2.40	3.08	1.00	0.00	23.00	1,940
All news mentions	5.49	6.35	3.00	0.00	42.00	1,940

Table 3 Summary Statistics of Market Variables

	Mean	Std. Dev.	Median	Min	Max	N
Return Nominal Broad Dollar Index	0.008	0.321	0.000	-1.99	1.74	1,939
Return Nominal Major Currencies Dollar Index	0.009	0.450	0.012	-2.41	2.48	1,939
Return AUSTRALIA -- SPOT EXCHANGE RATE US$/AU$	-0.003	0.748	0.000	-4.46	5.08	1,939
Return EURO -- SPOT EXCHANGE RATE	-0.011	0.608	0.000	-2.67	3.06	1,939
Return BRAZIL -- SPOT EXCHANGE RATE, BRL	0.028	0.955	0.000	-5.30	8.67	1,939
Return CANADA -- SPOT EXCHANGE RATE, CANADIAN $/US$	0.008	0.565	0.000	-2.90	3.37	1,939
Return CHINA -- SPOT EXCHANGE RATE, YUAN/US$ P.R.	-0.001	0.142	0.000	-1.02	1.82	1,939
Return INDIA -- SPOT EXCHANGE RATE, RUPEES/US$	0.016	0.520	0.000	-3.76	3.03	1,939
Return MEXICO -- SPOT EXCHANGE RATE, PESOS/US$	0.016	0.738	-0.007	-4.55	7.01	1,939
Return THAILAND -- SPOT EXCHANGE RATE, BAHT/US$	-0.001	0.293	0.000	-1.98	1.45	1,939
Return SP500	0.047	0.984	0.049	-6.90	7.34	1,939

Note: Differences in the observations across measures are due to missing observations when the market was close or on a holiday.

Figure 8 presents the distribution of the media coverage variable with mentions of increases in prices in the positive Y-axis and mentions of decreases in the negative Y-axis[7]. As seen previously, the most activity in news coverage is for wheat and maize, with considerable spikes in the number of mentions of price increases. Compare this to the media mentions for rice, which are less volatile in the study period, as evidenced by the standard deviation of these variables in Table 2.

[7] To better visualize the data, we exclude observations with more than 20 mentions. However, all observations are included in the analysis.

Figure 6 Agricultural Futures Prices for Commodities Prices, Standardized (8/3/2009=1)

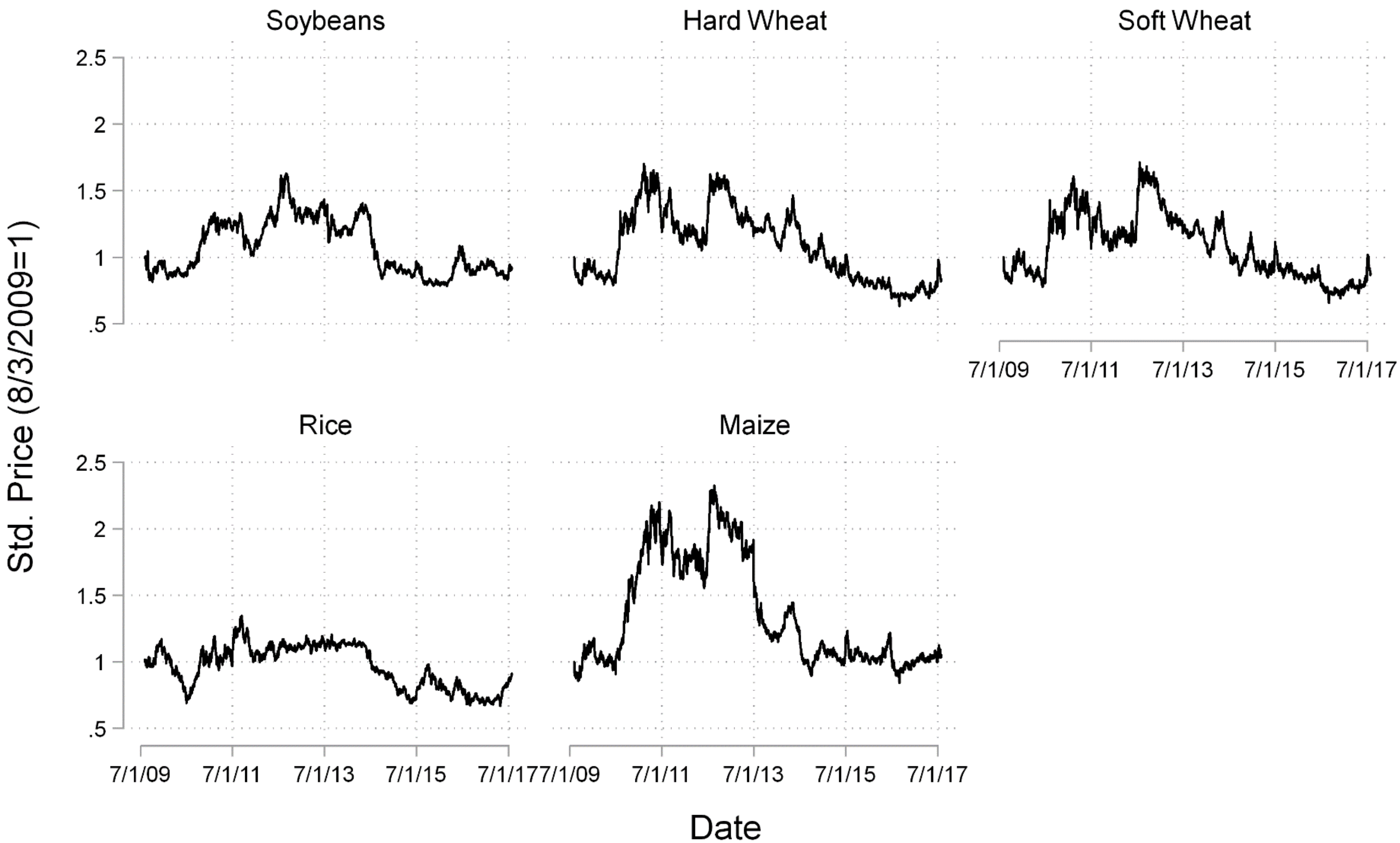


Figure 7 Agricultural Futures Prices for Commodities Returns

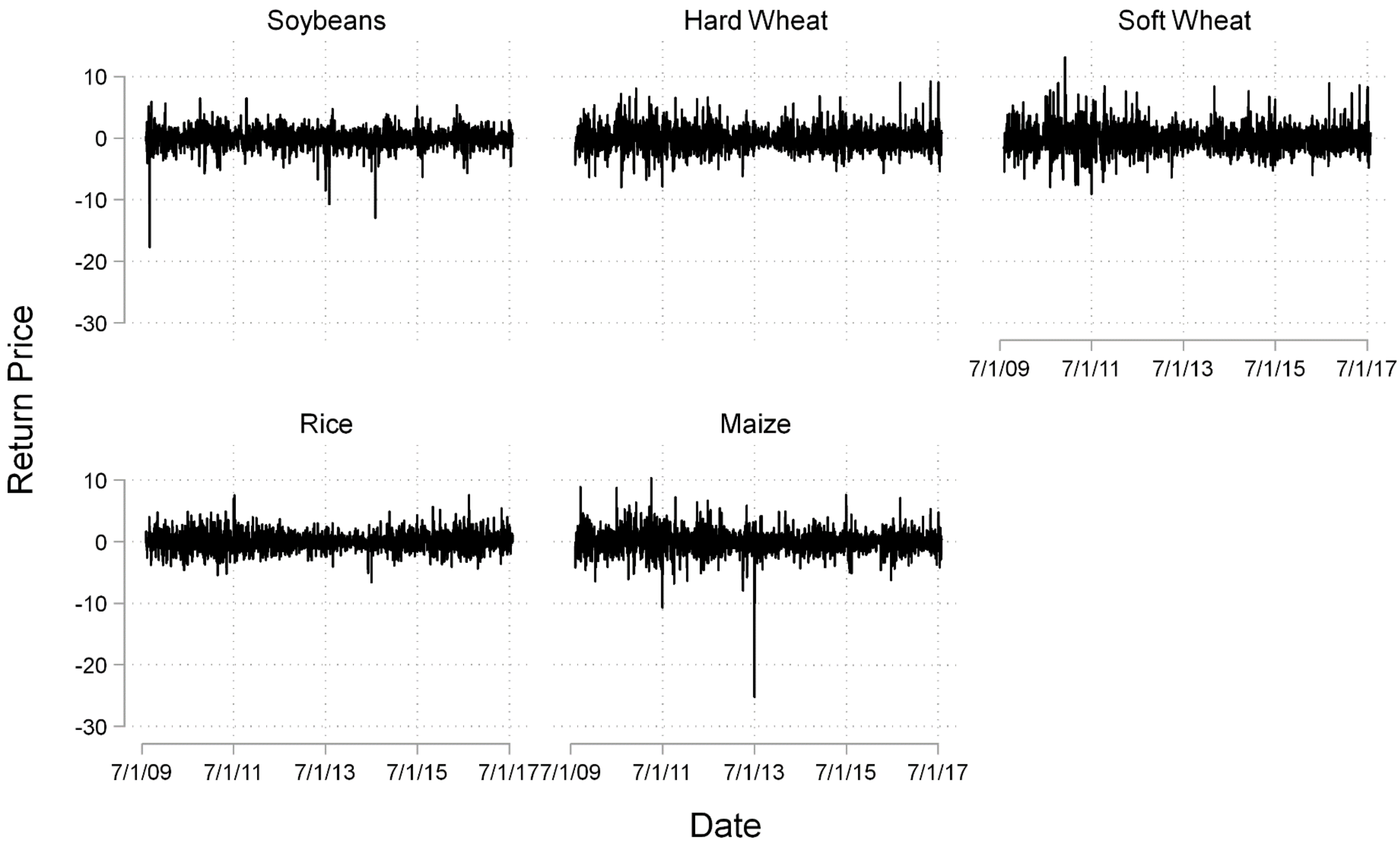
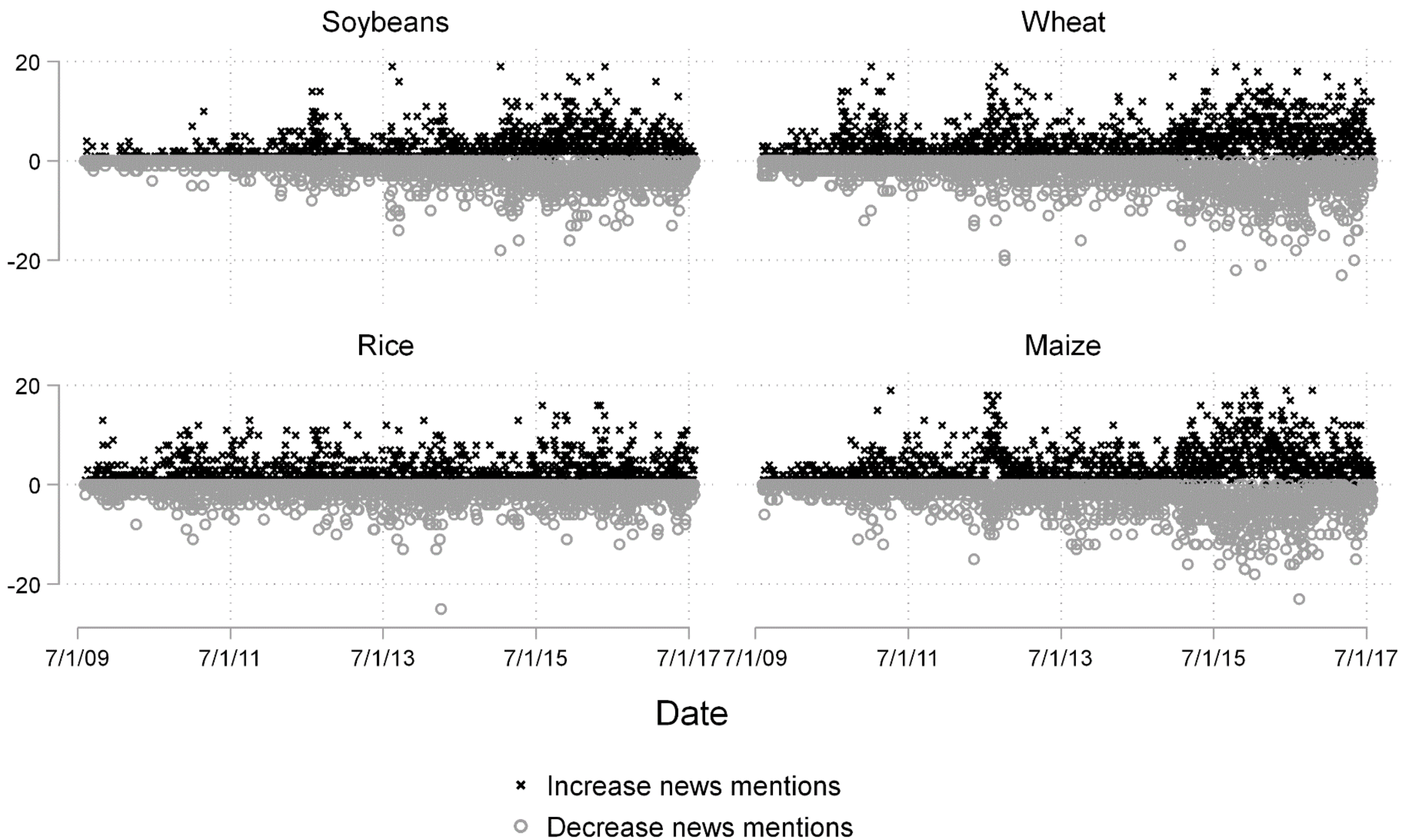


Figure 8 Media Coverage of Price Changes, Ups and Downs



6 Results

This section presents the OLS estimates of equation (1) and the IV-GMM estimates of equation (2) (the differenced equation) using the lags of the media variables and the dependent variable as instruments in one set of results and the lag-differences instruments in another. In addition, for each set of results, we include sequential commodity effects, market controls, and calendar effects. The calendar effects consist of indicator variables for day of the week, week in the year, month in the year, and year. Standard errors are presented in brackets and are clustered by date. We use heteroskedastic and autocorrelation-consistent (HAC) standard errors that are robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation in the panel, following Driscoll and Kray (1998)[8]. For the results using the IV-GMM estimator, we present the F statistic for weak identification test (Cragg-Donald or Kleibergen-Paap) of the excluded instruments and label this in the tables as "K-P rk Wald F."[9]

As is customary when analyzing futures price time series, we use the rate of return series in our analysis. We present results for the Dickey Fuller unit root tests for the panel in Table 4 for the log prices and the difference in log-prices or returns. The panel version of the test assumes a common autocorrelation parameter and relies on large T asymptotic theory. The tests provide evidence that the log-price levels have a unit root and that the returns (first difference) are stationary. This accounts for serial correlation in the price series; the tests show that we cannot reject the null of a unit root when we use the log-prices and that we reject the null with the log differences or the returns, with and without a trend in both cases.

Table 4 Augmented Dickey Fuller Panel Unit Root Tests

Prices	**Obs**	**Commodities**	**Periods**	**Chi-2**	**Df**	**P-Value**
No-Trend	9,700	5	1,940	9.66	10	0.471
Trend	9,700	5	1,940	10.22	10	0.421
Returns	**Obs**	**Commodities**	**Periods**	**Chi-2**	**Df**	**P-Value**
No-Trend	9,695	5	1,939	360.44	10	0.000
Trend	9,695	5	1,939	360.44	10	0.000

Note: Tests conduct unit-root tests for each panel individually, and then combine the p-values from these tests to produce an overall test. Assumes that T tends to infinity

[8] This variance-covariance estimator is a large-T estimator, and we used a truncated kernel.

[9] Lag variables are denoted using $L.X_t = X_{t-1}$ and difference using $D.X_t = \Delta X_t$, it follows that $LD.X_t = \Delta X_{t-1}$ and $LD2.X_t = \Delta X_{t-2}$

6.1 Price Effects of Media coverage

We now turn to the discussion of the return equations (1) and (2). The effects of media coverage on the prices of agricultural commodities are shown in Table 5; the dependent variable is the difference in log of the price or returns for each commodity. In columns (1) – (4), we present the results for equation (1), which are the OLS estimates of the returns equation. These estimates are consistent but could be biased downward if the media coverage variables on the current day only reflect the price levels in the past, which are included in the differencing operation to obtain the returns and the inclusion of the lag returns in the estimation.

Measurement errors in the media variables could also cause attenuation bias - for example, if important media outlets that traders in the commodities exchanges read are not included in the RSS we track to create these variables. These caveats to the OLS estimates are addressed with the IV-GMM estimates in columns (5) to (12), which show the results of the IV estimation for equation (2), the differenced equation. Column (5)-(8) use five lags in the news variables and lag two to five of the returns as 14 instruments for the lag difference in returns and the difference in the news variables. Columns (9)-(12) use five lag differences of the news variables as instruments instead of the lags. These estimates address both previous concerns under our sequential exogeneity assumption. These estimates are based on the variation on media coverage, and the returns in the previous days affect the media coverage in the current day.

The baseline model is given in column (1), where we include only the media coverage of increases and decreases in prices. We estimate this via OLS, adding controls from one column to the other. The estimate is robust to inclusion of commodity effects, market controls, and calendar effects. The OLS estimates indicate that media coverage of increasing prices increases returns by 0.031 percent across commodities after controlling for market condition and time-invariant unobservable differences across commodities. Media coverage of decreasing prices decreases returns further. One mention in the articles related to a decrease in prices decreases returns by 0.046 percent. This evidence suggests that media attention tends to accentuate price movements, with the acceleration effect for price decreases being larger than the acceleration effect for price increases. The persistence parameter suggests that the effects are present only for current price levels and that there is no transition to future periods.

In columns (5) through (12), we see that the estimates using the IV-GMM estimator from equation (2) are larger. The F statistics for the excluded instruments are large, showing that both sets of instruments are strong. Under both moment conditions, the estimates show that media coverage tends to accelerate price movements, with one mention of price increases further increasing prices between 0.059 percent and 0.061 percent on that day; one mention of price decreases drives further decreases of between 0.089 percent and 0.093 percent.

From these estimates, we can gather some of the dynamics between prices and media coverage; media coverage of price movements and the fundamentals that affect them through financial reasons, inventories, trade policies, and disasters accelerate the price movement in the same direction. Note that we include increases and decreases in price news, not the total number of mentions, to capture the effect of media coverage in a single parameter. While they are not necessarily the inverse of each other, they strongly correlated and in our case, the correlation is positive; we see articles pointing to both increases and decreases in prices on the same day. This is because, for example, different media outlets reflect the opinion of different experts who might put more emphasis on one aspect (say, inventories) over another (say, their perception that a contract is overvalued). Indeed, we find that the opposite effects highlight the importance of introducing increases and decreases separately in the model, as the reaction to increases and decreases in prices can be very different[10].

The price effects we find are economically important in size. The median return on the SP500 Index during the period we study is 0.049 percent per day; when we compare our estimates to this measure, we conclude that media coverage following increases in prices in the form of one mention is just 94.3 percent of the media return on the SP500, while for decreases, our returns are 20.9 percent above of the median return for the SP500. At the average level of mentions in the media observed in the period, the effects are similar. Using the global mean for media mentions of price changes (2.74 for increases and 2.05 for decreases), the estimates imply a 0.12 percent effect on returns. This shows that following the **intensity of media coverage of price changes in the futures markets for these commodities can open opportunities to reallocate risk** in investment portfolios, which can be beneficial to investors, and that media reports influence prices even if those reports are contradictory, reflecting the different information that different sources might have.

To conclude this section, we explore heterogeneity across commodities and time periods.[11] Figure 9 shows the OLS estimates (and confidence intervals), while Figure 10 shows the IV-GMM estimates with the lag difference instruments for each year in the data[12] and the aggregate estimate across years. The effects we observe are mostly concentrated in 2012 and from 2015 on. Figure 10 shows that the 2012 and 2015 effects were larger for both media coverage

[10] We thank a reviewer for bringing attention to the difference between media coverage in general (in a single parameter) and the asymmetric responses that increases and decreases of prices in the market can elicit in the media.

[11] We present both the OLS and the IV-GMM estimates in the figures. We note that the individual estimates use less data and that the strength of the instruments is lower for these estimations; thus, they are less precise.

[12] We include the data from 2009 and 2010 in one group and the data from 2016 and 2017 in another, since fewer observations in 2009 and 2017 would be a low power test of these effects. The tables with the coefficients on which the figures are based are available upon request from the authors.

measures but were not statistically different from each other or from the aggregate effect we find. The distribution of the effects across time shows that these media effects are important in periods of both high volatility (2012) and low volatility (post-2014), regardless of there is an increasing or a decreasing trend.

Figure 11 shows the OLS estimates and Figure 12 shows the IV-GMM that explores heterogeneity of the effects across commodities. This shows that media coverage has a large effect on soybeans prices but an insignificant effect on rice prices. The IV-GMM estimates for mentions of price increases is only significant at the 90 percent confidence level for soft wheat and maize, while those for mentions of decreases in prices are significant at conventional levels for soybeans, soft wheat, and maize.

Table 5 Media coverage Effects of Returns: OLS estimates of Returns and IV-GMM Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS Estimates of Eq. (1)				IV-GMM Estimates of Eq. (2)- IV Lags				IV-GMM Estimates of Eq. (2) - IV Lag-Diff			
Increase news mentions	0.043	0.043	0.04	0.038								
	[0.0087]***	[0.0087]***	[0.0087]***	[0.0087]***								
Decrease news mentions	-0.05	-0.05	-0.047	-0.045								
	[0.011]***	[0.011]***	[0.011]***	[0.011]***								
D.Increase news mentions					0.046	0.046	0.046	0.047	0.046	0.046	0.045	0.045
					[0.016]***	[0.016]***	[0.015]***	[0.015]***	[0.018]***	[0.018]***	[0.017]***	[0.016]***
D.Decrease news mentions					-0.059	-0.059	-0.063	-0.059	-0.058	-0.058	-0.06	-0.059
					[0.017]***	[0.017]***	[0.017]***	[0.017]***	[0.017]***	[0.017]***	[0.017]***	[0.018]***
LD.Log Price	0.023	0.023	0.018	0.00061								
	[0.016]	[0.016]	[0.016]	[0.016]								
LD2.Log Price					0.029	0.029	0.026	0.024	0.028	0.028	0.025	0.023
					[0.024]	[0.024]	[0.023]	[0.023]	[0.024]	[0.024]	[0.023]	[0.023]
Commodity Effects	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Market Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Calendar Effects	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
K-P rk Wald F					121.50	124.20	125.00	99.40	79.30	79.30	79.30	80.70
Observations	9,690	9,690	9,690	9,690	9,670	9,670	9,670	9,670	9,670	9,670	9,670	9,670

Standard errors in second row

HAC-Standard Errors (in brackets) and Statistics robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation. Clustered on date.

Column (1)-(4) is the OLS estimate of the returns equation. Columns (5)-(12) show the results of the IV estimation for the differenced equation. Column (5)-(8) use 5 lags in the news variables and lag 2 to 5 of the returns as instruments for the lag difference in returns and the difference in the news variables. Columns (9)-(12) use 5 lag differences of the news variables as instruments instead of the lags. The K-P rk Wald F is the F statistic for weak identification test (Cragg-Donald or Kleibergen-Paap) of the excluded instruments.

* p<0.10, ** p<0.05, *** p<0.01

Figure 9 OLS Estimates of Media Coverage Effects by Year

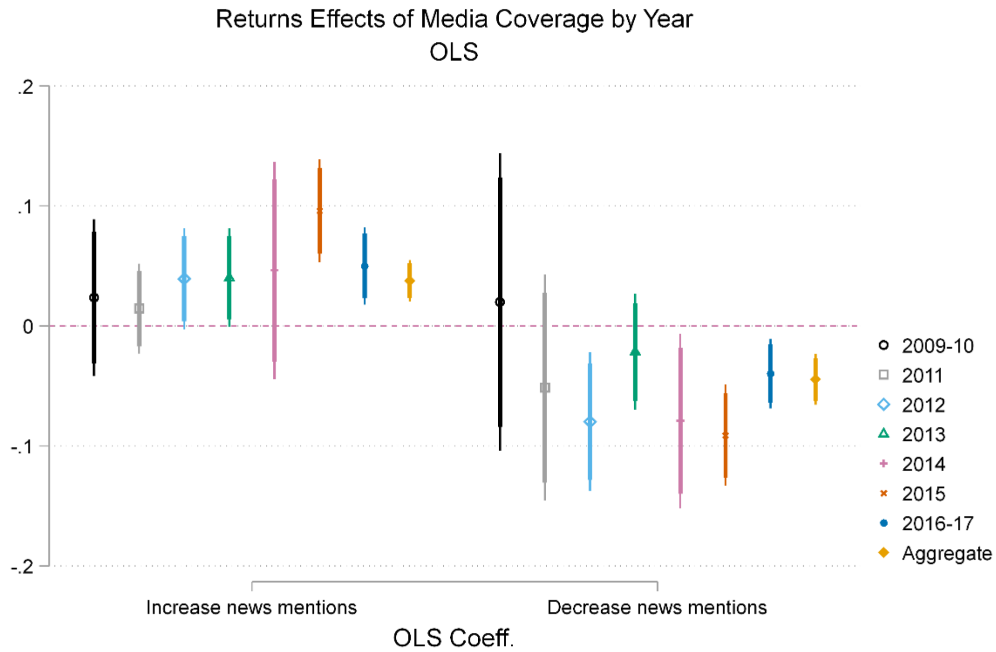


Figure 10 IV-GMM Estimates of Media Coverage Effects by Year: Lag-Difference IV

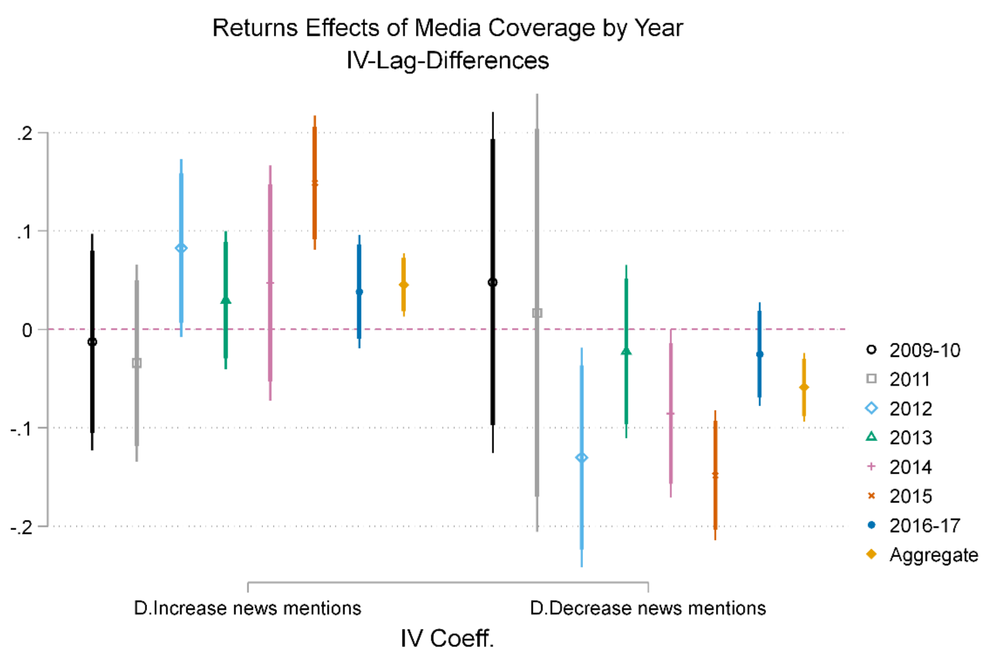


Figure 11 OLS Estimates of Media Coverage Effects by Commodity

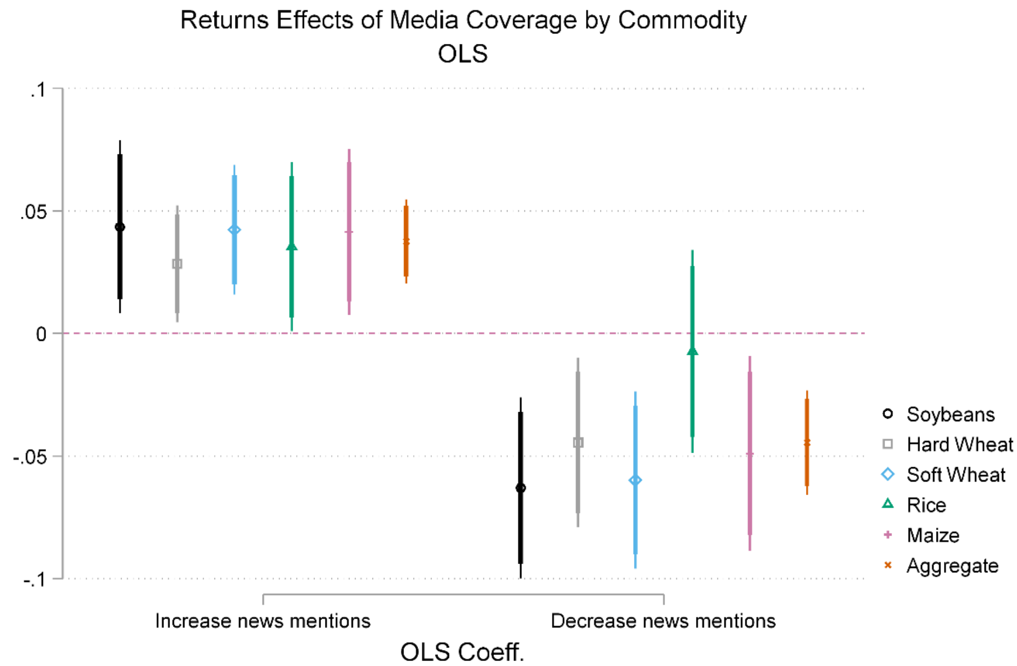
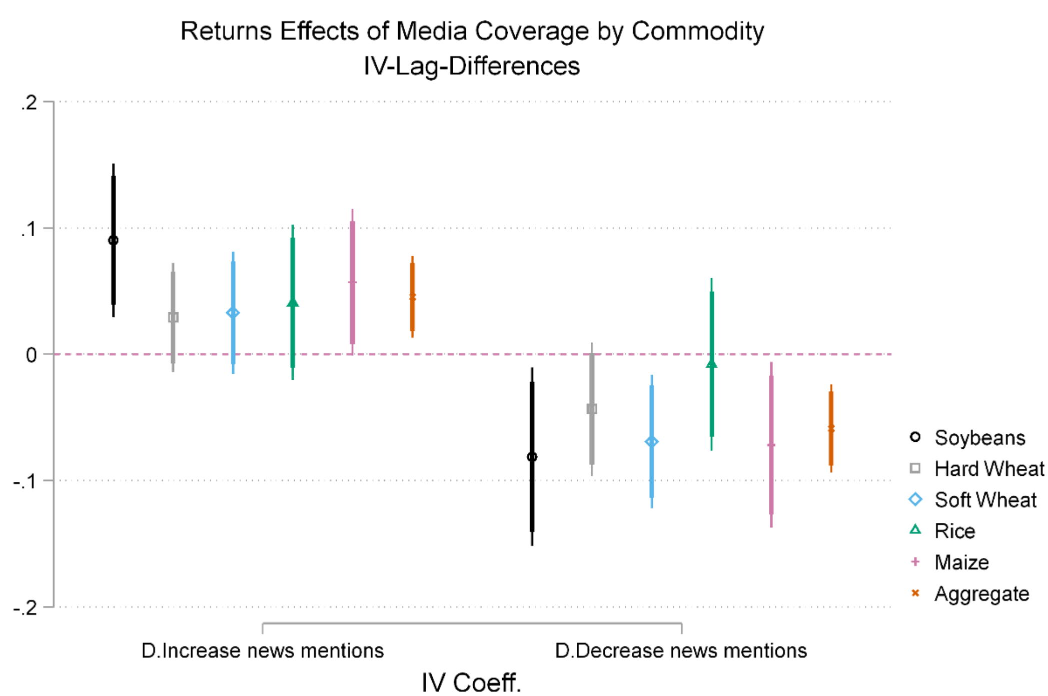


Figure 12 IV-GMM Estimates of Media Coverage Effects by Commodity: Lag-Difference IV



6.2 Volatility Effects of Media Coverage

The differences in volatility will be examined using several different methods. The first **(a)** is a straightforward F-test testing for difference of variances (or standard deviations); the second **(b)** is an OLS estimation and IV-GMM estimation of equation (5) that uses the residuals from equation (1) to compute excess returns based on the media and market variables; the third **(c) is a** model that uses a 30-day rolling variance estimate from the previous model; the fourth **(d)** presents the effects of media using the canonical GARCH model and using volatility indicators derived from non-parametric extreme quantile model.

(a) Difference in Variance: F-test

The ratio of the estimated variance of the prices (or returns) on days with news mentions relative to days with no news mentions, $\frac{\sigma^2_{no-news}}{\sigma^2_{news}}$ is distributed as an F-statistic under the null hypothesis of equal variances. We examine a possible volatility difference between days with and without news using this standard F-test for difference of variance on each type of day. We want to know whether volatility agricultural futures prices and returns is higher or lower on days when there is media coverage indicating ups or downs.

Table 6 shows the results of these tests for each type of coverage, namely days on which there is up news, down news, and any type of news for the prices studied. The null hypothesis is that the ratio is equal to one, and the alternatives are given in the column headers. The test for the returns suggests that ***there is lower volatility of returns on the days on which there is media coverage.*** The volatility of returns tends to be lower, observing that we reject the null in most cases under the alternative hypothesis that the ratio in greater than one, so that the variance on days on which we find no mentions of decreases or increases in prices is greater. This is strongly the case for soybeans and maize, but is not as clear-cut for the other commodities, highlighting the importance of performing a more sophisticated test and exploring heterogeneity across commodities.

Table 6 F-Test for the difference in Variance of Returns on News and Non-News Days

	Obs	Ratio<1	Ratio>1	Ratio≠1	F-Stat	SD-No News	SD-News
Soybeans							
Increase Event	1,939	1.000	0.000	0.000	1.38	1.66	1.41
Decrease Event	1,939	1.000	0.000	0.000	1.29	1.61	1.42
Any Event	1,939	1.000	0.000	0.000	1.27	1.63	1.45
Hard Wheat							
Increase Event	1,939	0.177	0.823	0.354	0.93	1.78	1.84
Decrease Event	1,939	0.941	0.059	0.117	1.11	1.89	1.79
Any Event	1,939	0.292	0.708	0.583	0.95	1.79	1.84
Soft Wheat							
Increase Event	1,939	0.444	0.556	0.889	0.99	2.01	2.02
Decrease Event	1,939	0.998	0.002	0.005	1.21	2.15	1.95
Any Event	1,939	0.827	0.173	0.346	1.08	2.08	2.01
Rice							
Increase Event	1,939	0.066	0.934	0.132	0.91	1.45	1.52
Decrease Event	1,939	0.657	0.343	0.685	1.03	1.50	1.48
Any Event	1,939	0.153	0.847	0.306	0.93	1.46	1.51
Maize							
Increase Event	1,939	1.000	0.000	0.000	1.55	2.15	1.73
Decrease Event	1,939	1.000	0.000	0.000	1.74	2.18	1.65
Any Event	1,939	1.000	0.000	0.000	1.66	2.23	1.73

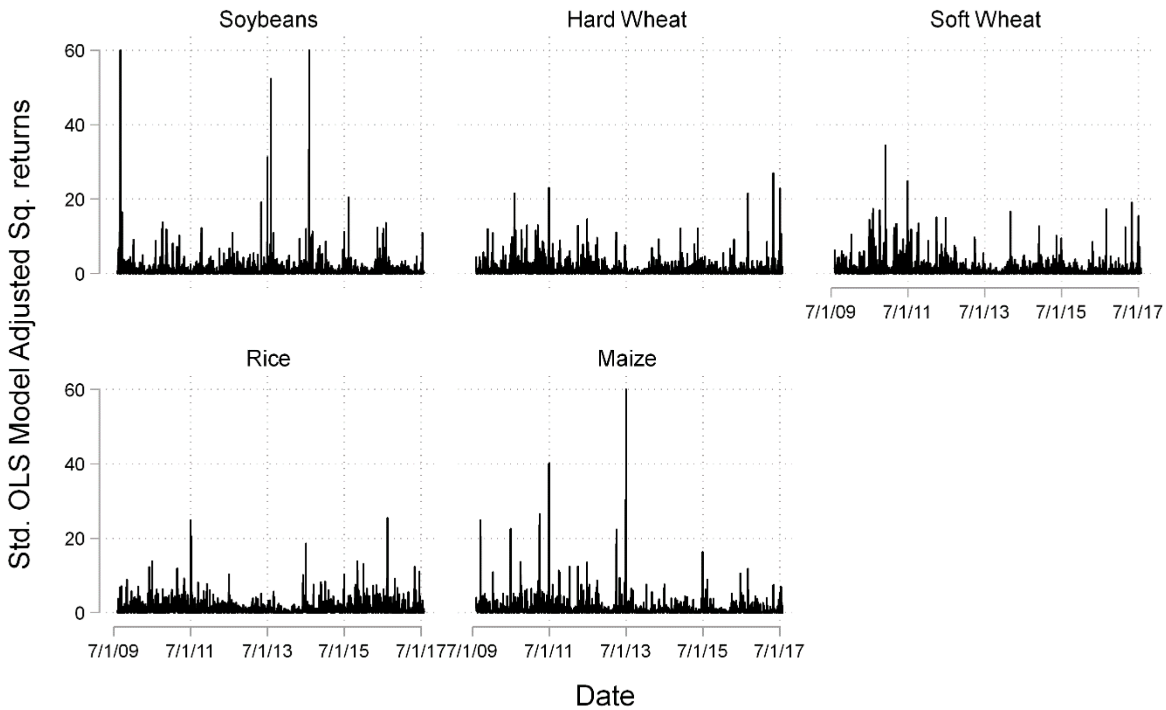
(b) Market Model of Excess Returns

Given the different effects across commodities we find using the simple F-test, we proceed with the second approach. This approach consists of estimating equation (1) and using the residuals to compute the excess returns, taking into account the market variables and direct effects of media coverage. These residuals are standardized and are then squared to serve as a proxy for the daily variance or volatility; it is this variation that remained unexplained in the first regression.[13] The evolution of this variable across time is presented in Figure 13, where we can observe the differences in volatility in different periods of time.

To start, we present a graphical analysis of the residuals, given that this simple test might not reflect the heterogeneity in volatility due to the intensity of media coverage. Creating dichotomous groups that agglomerate a day with one mention of price increases with a day with 10 mentions of price increases might give the impression that media coverage is positively correlated with volatility, when it could also be the opposite. Figure 14 shows that the excess returns squared are smaller when there is more attention (both for increases and decreases). The residuals around zero are higher than when we have more media mentions. This graphical evidence initially points to lower volatility when media coverage is more intense.

[13] We estimate the residuals only, including the market variables, and then estimate the residuals with both the market variables and the media variables. The results are qualitatively similar. These residuals are essentially standardized market-adjusted returns.

Figure 13 Volatility Time Series: Standardized Residuals Squared, OLS

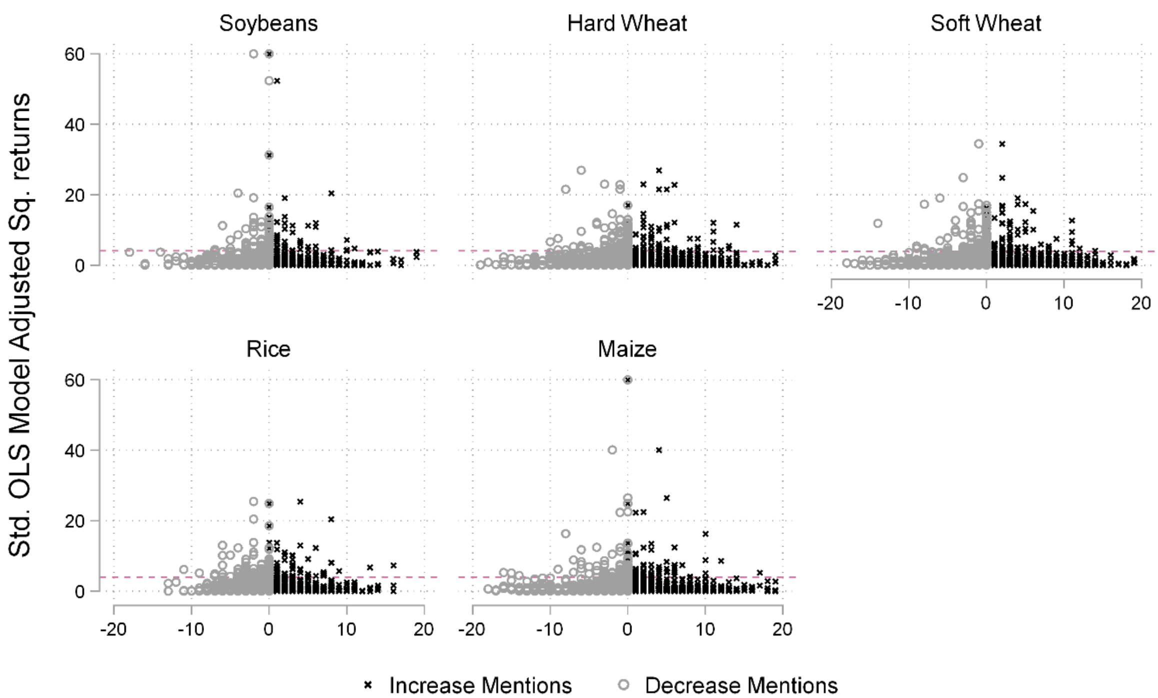


Note: Residuals truncated at 60 in the graph for visualization. All observations are included in the regression analysis.

Table 8 shows the estimations that put the figures in a regression framework, as described in equation (5)[14]. Columns (1) to (5) show the estimates based on the OLS regression. These estimates are robust to the inclusion of covariates, with the effects being symmetrical. Media mentions of price increases raise volatility by 0.03, or 3 percent of the variance of excessive returns, while the reverse is true for mentions of price decreases. Given that the mean and medians of the number of mentions in the media is higher for price increases, the estimates suggest that media coverage increases volatility on average and that these effects are important in periods in which prices are increasing. The results in columns (5) to (9) show larger effects - 3 percent of the variance for the mentions of price increases and 4.6 percent of the variance for mentions of price decreases, when using the lag differences as an instrument. The estimates using the lags of the media variable as instruments are similar in column (9).

[14] We also computed the excess residuals excluding the media variables from equation (5) and calculated these estimates (simple market adjusted residuals) and the results are qualitative similar in size and significance.

Figure 14 Volatility vs. Intensity of Media Coverage: Standardized Residuals Squared, OLS



Note: A line at 4 is presented for guidance. The standardized residuals above 2 are considered excessive and our measure squares this residual. Residuals truncated at 60 and news outside of the interval [-20,20] are excluded from the graph for visualization. All observations are included in the regression analysis.

In Figure 15, we estimate monthly effects of media coverage on volatility for the study period. The figure shows the in-sample prediction for each month in the data; that is, for each day, we calculate $\nu \cdot UP_{i,t}$ and $\psi \cdot DOWN_{i,t}$ from (5) and get the average for the month. Similarly, we calculate the net effect, calculating $\nu \cdot UP_{i,t} + \psi \cdot DOWN_{i,t}$ and computing the monthly average of this measure. These figures give us the evolution of the effects during the study period. In addition, we shade the months with excessive volatility, which we define as a month with more than 5 percent of the daily standardized residuals squared above the critical value of 4. Table 7 shows the net effect of media coverage for each commodity and each year. Figure 15 and Table 7 show that after 2011, the effect of the media on volatility increased, particularly for maize and wheat, and that for this period, the net effect is negative. This means that media coverage **decreased** volatility in the prices of these commodities by 0.99 percent of the estimated variance.

Intuitively, news of both decreasing and increasing prices could increase volatility because it generates trading and, hence, volatility. Under this assumption, news of decreasing prices reduces leverage, thus reducing risk and volatility. In our case, ***news of decreased prices decreases the variance*** and ***news of increased news increase the variance,*** but to a lesser degree than news of decreased prices; this implies a ***net decrease effect of media coverage on volatility*** for these commodities in the study period.

Table 7 Average Media Coverage Effects: % Of Variance of Return, Std. Residuals Sq.

	Soybeans	Hard Wheat	Soft Wheat	Rice	Maize
2009-10	0.2%	2.2%	2.2%	1.2%	0.0%
2011	0.4%	4.0%	3.9%	1.8%	1.8%
2012	1.8%	2.6%	2.8%	1.0%	3.6%
2013	-1.0%	-1.7%	-1.7%	-2.2%	-2.5%
2014	-1.9%	-1.9%	-1.8%	-2.6%	-3.2%
2015	-4.6%	-5.6%	-5.6%	-0.7%	-5.9%
2016-17	-3.2%	-5.2%	-5.2%	-0.7%	-3.8%

Table 8 Volatility Effect of Media: Std. Residuals Squared OLS Model Adjusted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS Estimates of Eq. (5)				IV-GMM Estimates of Differenced Eq. (5)- IV Lag-Diffs				IV Lags
Increase news mentions	0.022	0.021	0.02	0.03					
	[0.010]**	[0.010]**	[0.011]*	[0.010]***					
Decrease news mentions	-0.049	-0.049	-0.048	-0.03					
	[0.012]***	[0.013]***	[0.013]***	[0.014]**					
D.Increase news mentions					0.04	0.04	0.039	0.03	0.034
					[0.017]**	[0.017]**	[0.017]**	[0.018]*	[0.017]**
D.Decrease news mentions					-0.047	-0.047	-0.047	-0.046	-0.049
					[0.022]**	[0.022]**	[0.022]**	[0.022]**	[0.020]**
L.Std. OLS Model Adjusted Sq. returns	0.029	0.029	0.03	-0.02					
	[0.017]*	[0.017]*	[0.017]*	[0.016]					
LD.Std. OLS Model Adjusted Sq. returns					0.016	0.016	0.015	-0.0035	-0.0033
					[0.011]	[0.011]	[0.011]	[0.016]	[0.016]
Commodity Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Market Controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes
Calendar Effects	No	No	No	Yes	No	No	No	Yes	Yes
K-P rk Wald F					81.5	81.4	81.4	83.9	103.5
Observations	9,685	9,685	9,685	9,685	9,665	9,665	9,665	9,665	9,665

Standard errors in second row

HAC-SE (in brackets) and Statistics robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation. Clustered on date.
Colums (1)-(4) have the squared residuals of previous models defined in the text; (5)-(8) shows the estimated parameter from the differenced equation and the instruments are the 5 lags differences in the news variables and lags 2 to 5 of the volatility measure. In Column (9) the instruments are the 5 lags in the news variables and lags 2 to 5 of the volatility measure. The K-P rk Wald F is the F statistic for weak identification test (Cragg-Donald or Kleibergen-Paap) of the excluded instruments.

* p<0.10, ** p<0.05, *** p<0.01

Figure 15 Volatility Effects across Time from Std. Residuals Squared OLS Model

(c) 30-Day Rolling Variance Model

We use the previous model to estimate a 30-day variance for each of the series and estimate the effect of the media coverage variables on the variance of returns; the estimated variance is presented in Figure 16, while the estimated media effects are shown in Table 9. The results are similar to what we find previously, using the standardized residuals squared. In addition, these results show that the media effects persist somewhat through the innovation or error term used in this equation. The effects accounting for this persistence are similar; for example, for mentions of price increases, the effect is $\left(\frac{0.046}{1-0.04}\right) = 0.048$, or 4.08 percent of the estimated variance.

The estimates suggest media coverage increases volatility on average and that these effects are important in periods in which prices are increasing. In Figure 17, we present the monthly effects for each month in the data. For each day, we calculate $\nu \cdot \frac{UP_{i,t}}{\sigma^2_{30-day}}$ and $\psi \cdot \frac{DOWN_{i,t}}{\sigma^2_{30-day}}$ and get the average for the month; similarly, we calculate the net effect adding both effects. These figures give us the evolution of the effects during the study period using this model. For this model, we identify (shade) the months with excessive volatility, which we define as months in which more than 5 percent of the days during the 30-day rolling variance estimates were above the 95 percent quantile of the variance distribution. The figure shows that this model predicts fewer months of excessive volatility than the previous model and that after 2015, the net effect of media on volatility has increased. Using this model in Table 10 to calculate the net effect by year, **the net effect suggests that media coverage has increased volatility in the prices of these commodities** by 1 percent of the variance of returns on average during the study period, with differences across commodities. The results show that these effects vary over time, with the end of the study period seeing larger effects from media coverage variables and a small positive effect, on average, using this model.

Table 9 Volatility Estimates: 30 Day Moving Variance from Squared Residuals

	(1)	(2)	(3)	(4)
	OLS Estimates of Eq. (6)- 30-day Moving Variance			
Increase news mentions	0.064	0.042	0.041	0.046
	[0.013]***	[0.013]***	[0.013]***	[0.0098]***
Decrease news mentions	-0.099	-0.12	-0.12	-0.047
	[0.015]***	[0.015]***	[0.015]***	[0.012]***
L.OLS Model Adjusted Sq. returns	0.054	0.05	0.05	0.04
	[0.010]***	[0.0085]***	[0.0085]***	[0.0054]***
Commodity Effects	No	Yes	Yes	Yes
Market Controls	No	No	Yes	Yes
Calendar Effects	No	No	No	Yes
Observations	9,685	9,685	9,685	9,685

Standard errors in second row

HAC-SE (in brackets) and Statistics robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation. Clustered on date. Colums (1)-(4) have as a dependent variables the 30-day moving average of squared residuals of a regression of returns against from the OLS model.

* p<0.10, ** p<0.05, *** p<0.01

Table 10 Average Media Coverage Effects: % Of Variance of Return, 30-day Rolling Variance Model

	Soybeans	**Hard Wheat**	**Soft Wheat**	**Rice**	**Maize**
2009-10	0.3%	1.0%	0.7%	1.4%	0.4%
2011	1.1%	2.6%	1.8%	1.6%	1.3%
2012	2.4%	2.3%	2.2%	2.1%	2.6%
2013	0.6%	0.6%	0.6%	0.0%	0.4%
2014	1.0%	0.3%	0.3%	-1.3%	-0.1%
2015	1.8%	0.3%	0.4%	0.8%	1.2%
2016-17	1.1%	1.4%	1.1%	1.3%	2.8%

Figure 16 Volatility Time Series: 30-day Rolling Variance

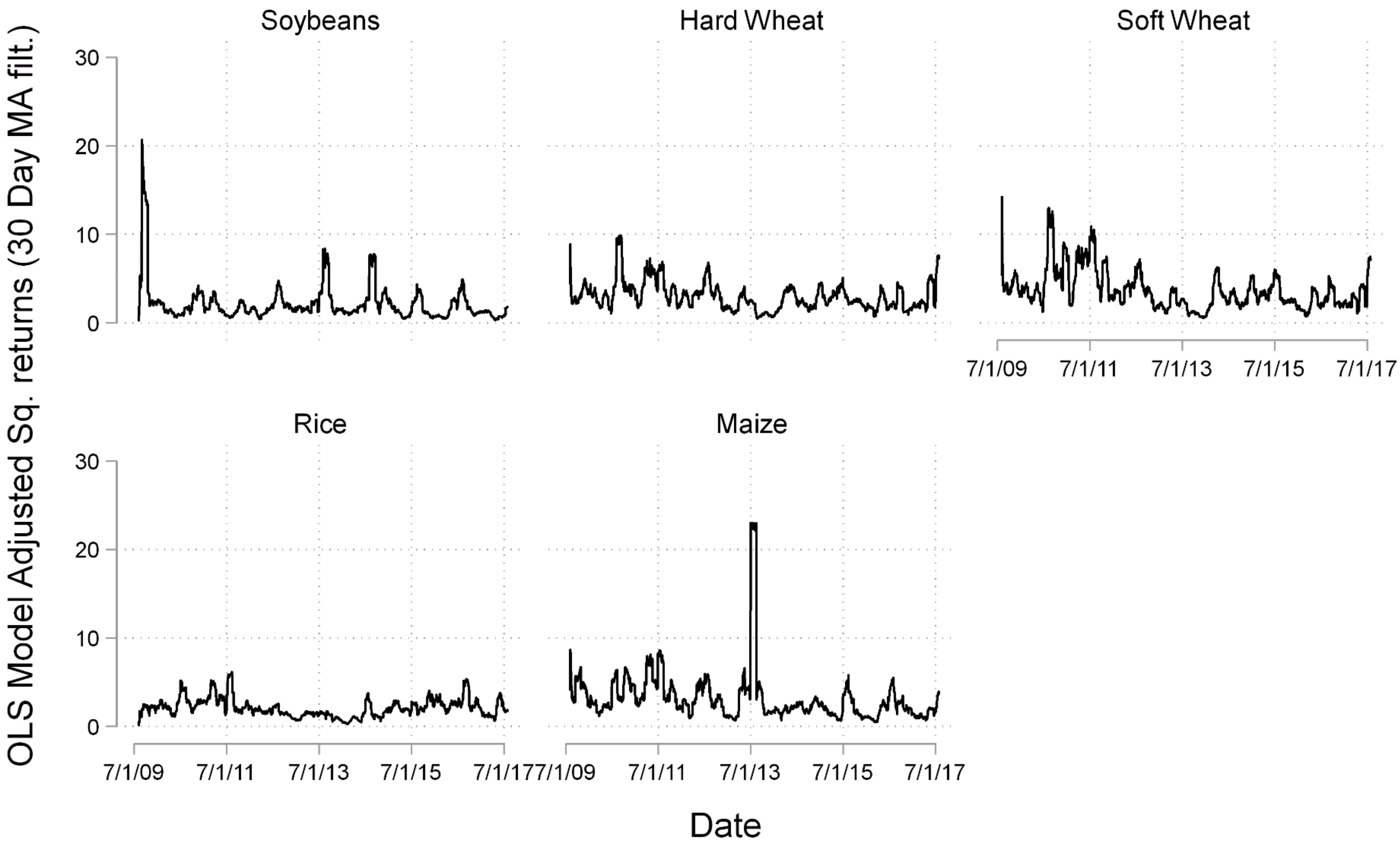
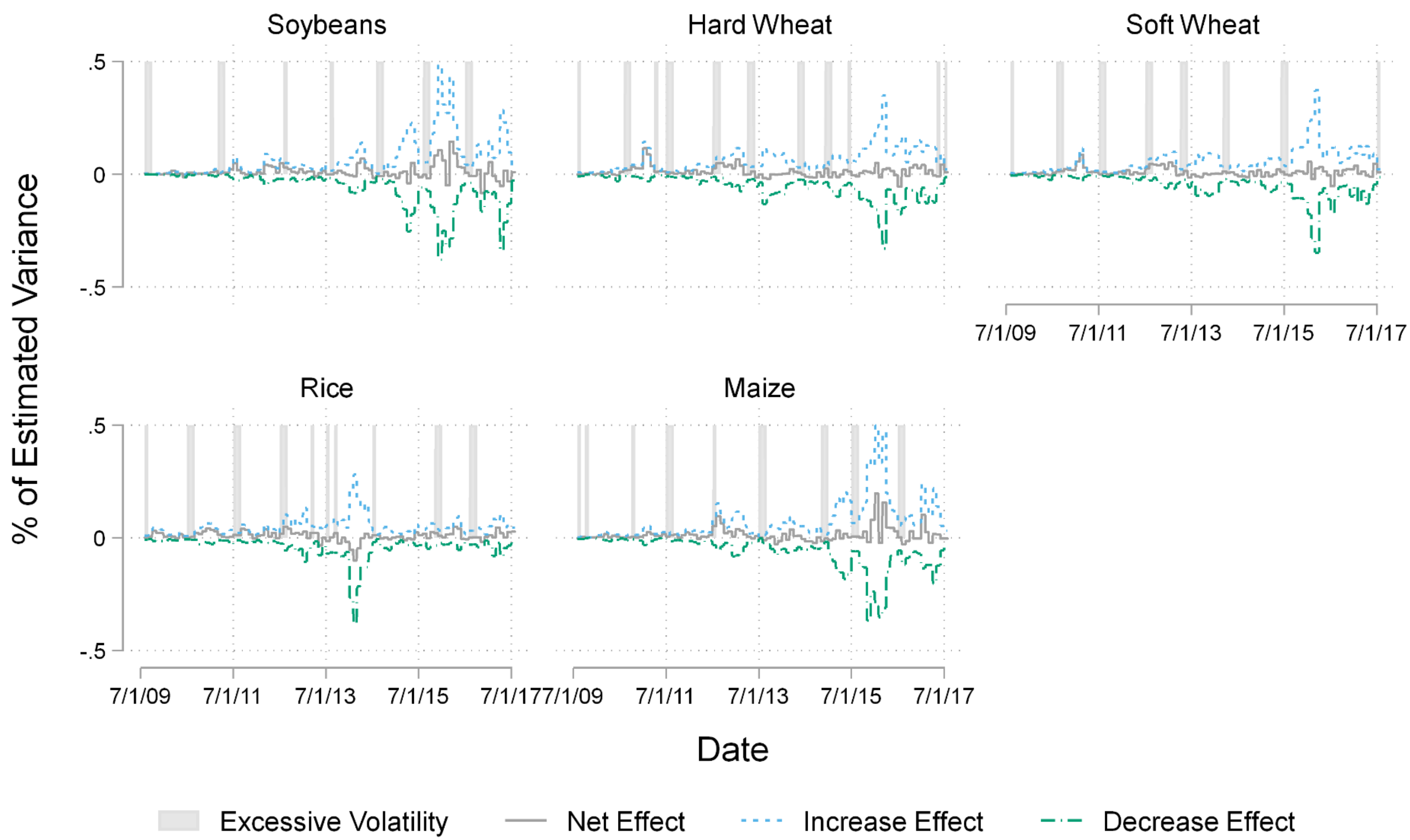


Figure 17 Volatility Effects across Time from 30-day Rolling Variance Model



(d) Other Models of Volatility

We are also interested in how our previous estimates compare to other models or ways of estimating used in the literature. In this section, we estimate a GARCH model that jointly estimates the effects of media on returns and the persistence and reaction of the markets to news; we use a nonparametric extreme quantile model combined with extreme value theory to estimate higher-order quantiles of the return series, which allows us to classify returns as extremely high or not extremely high and to explore how the media variables affect these indicators.

GARCH Model

Given the effects we find for both the returns and the volatility of returns, we explore how the effects of the media variables are affected when we jointly estimate a model of returns and when the variance of the error term from equation (1) depends on past errors and the past volatility of returns. We estimate a Generalized Autoregressive Conditional Heteroskedasticity (GARCH), first developed by Bollerslev (1986), in which the variance each day depends on one lag of the error in (1) and one lag of the variance itself. Namely, the variance at time t, σ_t^2 is:

$$\sigma_t^2 = \kappa + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot \sigma_{t-1}^2$$

The estimates from the pooled GARCH equations are presented in Table 11. The GARCH reaction parameter (α) is about 0.045, which, for daily data, implies a relatively stable market; the persistence parameter (β) indicates a slowly adjusting process, i.e. high persistence. These results are similar to what we find in Table 9 with the 30-day rolling variance equation. Table 12 shows the GARCH equations for each commodity. The estimates of the GARCH effects indicate that the **maize market is the most volatile during this period**, with a lower persistence parameter around 0.77 and a higher reaction at 0.18. The effects of the media coverage variable tend to be larger for mentions of decreasing prices (except for rice), which is **similar to what we found in the previous models**.

Table 11 Volatility Effects with GARCH (1,1) Conditional Variance

	(1)	(2)	(3)
	GARCH (1,1) Estimates of Eq. (1)		
Log Price			
Increase news mentions	0.048	0.049	0.048
	[0.0064]***	[0.0064]***	[0.0066]***
Decrease news mentions	-0.053	-0.052	-0.051
	[0.0073]***	[0.0074]***	[0.0080]***
Constant	-0.029	-0.015	0.12
	[0.020]	[0.033]	[0.082]
ARCH			
ε_{t-1}^2	0.043	0.044	0.045
	[0.0020]***	[0.0020]***	[0.0022]***
σ_{t-1}^2	0.95	0.95	0.95
	[0.0019]***	[0.0020]***	[0.0024]***
Constant	0.031	0.031	0.032
	[0.0025]***	[0.0025]***	[0.0028]***
Commodity Effects	No	Yes	Yes
Calendar Effects	No	No	Yes
Observations	9,695	9,695	9,695

Standard errors in second row. Column (1) with a constant, in (2) and (3) we add commodity indicators and calendar effects.
Calendar Effects are indicators for year, month and week of the calendar year.
* p<0.10, ** p<0.05, *** p<0.01

Table 12 Volatility Effects with GARCH (1,1) Conditional Variance by Commodity

	(1)	(2)	(3)	(4)	(5)
	Soybeans	**Hard Wheat**	**Soft Wheat**	**Rice**	**Maize**
Returns					
Increase news mentions	0.057	0.041	0.057	0.035	0.04
	[0.016]***	[0.014]***	[0.016]***	[0.015]**	[0.014]***
Decrease news mentions	-0.078	-0.047	-0.063	-0.0026	-0.041
	[0.021]***	[0.019]**	[0.020]***	[0.021]	[0.017]**
Constant	-0.28	-0.29	0.079	0.26	0.24
	[0.16]*	[0.24]	[0.22]	[0.17]	[0.19]
ARCH					
ε_{t-1}^2	0.088	0.047	0.034	0.024	0.18
	[0.010]***	[0.0070]***	[0.0051]***	[0.0042]***	[0.021]***
σ_{t-1}^2	0.89	0.94	0.96	0.97	0.77
	[0.013]***	[0.0095]***	[0.0056]***	[0.0057]***	[0.024]***
Constant	0.061	0.059	0.034	0.014	0.27
	[0.015]***	[0.017]***	[0.0087]***	[0.0044]***	[0.046]***
Calendar Effects	Yes	Yes	Yes	Yes	Yes
Observations	1,939	1,939	1,939	1,939	1,939

Standard errors in second row. GARCH (1,)stimates for each commodity indicated in the header.

Calendar Effects are indicators for year, month and week of the calendar year.

* p<0.10, ** p<0.05, *** p<0.01

Non-parametric Extreme Quantile Model and Extreme Value Theory

In this section, we use the estimator described in Martins-Filho, Yao and Torero (2016) to estimate conditional quantiles for log returns of futures prices. This model draws on extensive research into returns on agricultural commodity prices to explain when such price fluctuations and jumps are abnormally high given past observations on prices. It uses a Non-Parametric Extreme Quantile (NEXQ) with extreme value theory to estimate higher-order quantiles of the return series, allowing for classification of any realized return as high, extremely high, or normal based on daily returns since 2001.

One or two extremely high returns do not necessarily indicate a period of excessive volatility in this model; excessive volatility is identified based on a statistical test of the number of times extreme values occur in a window of consecutive 60 days.[15] Thus, a period of time characterized by extreme price volatility occurs when we observe a large number of extreme positive returns; that is, a value of return that exceeds the 95 percent conditional quantile.

We use this model to estimate the effects of media coverage on the volatility measures identified in this model. First, we create an indicator for the days characterized by extreme price volatility, defined as days when the return is above the 95 percent conditional quantile predicted by the NEXQ model. We use this indicator and the categorization of the period given by the model (as normal, high, or excessive volatility).

The classification of days in normal, high, or excessive volatility follows the results from a statistical test on the probability that we would observe k days of extreme returns in a 60-day window. The probability that we observe k days of extreme price returns (returns above the 95 percent quantile, as explained in the definition of excessive price volatility) in a period of 60 consecutive days is defined as:

$$Prob(Excessive\ Days = k) = \binom{60}{k}(0.05)^k(0.95)^{60-k}$$

Using this, we define three categories:

- Excessive Volatility: If the probability value is less than or equal to 2.5 percent, the null hypothesis that violations (i.e. days of extreme price returns) are consistent with expected violations is highly questionable, meaning that we are in a period of an excessive number of days of extreme price returns relative to that expected by the model.

[15] For comparison, in the previous sections we counted the number of days in a calendar month on which the returns exceeded a threshold and classified that month as one with excessive volatility. In this model, the 60-day rolling window and the statistical test allows for arbitrary periods of volatility, which can be short (4-5 days) or long (2-3 months).

- High volatility: If the probability value is bigger than 2.5 percent and less than or equal to 5 percent, the null hypothesis that violations are consistent with expectations is questionable at a low level, meaning that we are in a period of moderate number of days of extreme price returns relative to that expected.
- Normal volatility: If the probability value is bigger than 5 percent, we accept the null hypothesis that violations are consistent with expectations, meaning that the number of extreme price returns is consistent to what is expected from the model.

For example, if in a period of 60 days we observe 10 days with returns above the predicted 95 percent quantile, we compute $Prob(Excessive\ Days = 10) = \binom{60}{10}(0.05)^{10}(0.95)^{50} = 0.001$; this is a very unlikely event, so we characterize the last day in the window as a day with excessive volatility. Intuitively, in a 60-period window, if we see fewer than six days with returns above the predicted 95 percent quantile, that day is characterized as normal; if we see more than six days with above-normal returns, the day is characterized as having high or excessive volatility. Table 13 shows the proportion of days in the study period that were in each category. The model identifies 10 percent of days in the study period as having high or excessive volatility.

Table 13 Proportions of days with Normal, High, and Excessive Volatility

	Normal	High	Excessive
Soybeans	89%	5%	6%
Hard Wheat	90%	5%	6%
Soft Wheat	89%	5%	6%
Rice	90%	7%	3%
Maize	88%	4%	8%

Table 14 shows the effects of media coverage on the probability of observing extremely high returns (higher than the forecast the 95 percent conditional quantile). The estimates in columns (1) to (3) imply that media mentions of increased prices increase the probability of having a highly abnormal return by 0.095 percentage points, while mentions of decreased prices decrease that probability by 0.13 percentage points. For example, 10 mentions of prices increasing would further increase the probability of a rare event from 5 percent (having a return above the 95 percent conditional quantile) to 5.95 percent; in the case of 10 mentions of price decreases, that probability would decrease from 5 percent to 3.7 percent. In columns (4) to (6), we estimate the effect of the media variables depending on the volatility category of the period. In periods of normal volatility, mentions of price increases decrease the probability of having excessive returns; in periods of excessive volatility, such mentions of increased prices increase that probability by 1.18 percentage points. These estimates suggest that news of price increases during normal periods do not provoke rushed or automatic reactions in the market, but they do prompt such reactions when volatility is excessive. In the case of mentions of decreased prices, we only find significant effects for periods of normal volatility. During periods of normal volatility, news of decreasing prices decreases the probability of having higher returns, as one would expect if the media is reporting on the price decrease based on market fundamentals.

Using the predictions of this model, we calculate the number of continuous days in the current level of volatility for each day. For example, the variable takes a value of 30 on a day of normal volatility if there have been 30 days of low volatility since the last instance of high or excessive volatility; the variable takes a value of 10 on a day of high volatility if here have been 30 days of high volatility since the last instance of low or excessive volatility, etc.

Table 15 shows the results for this variable. The estimates in columns (1) to (3), which do not differentiate between the level of volatility, show non-robust effects. In columns (4) to (6), we see that the only media variables that matters is the mention of decreased prices during periods of excessive volatility. The estimate suggests that one mention of decreased prices during periods of high volatility can decrease the longevity of the excessive volatility days by 77 days. The increase in news mentions variables effects are not robust to the inclusion of covariates. Together, they show that media coverage **increases periods of normal volatility** and **decreases periods of excessive volatility**, implying that media coverage decreases volatility based on this model.

Table 14 Ext. Value Volatility: Prob. of return above the 95th conditional percentile

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-parametric Extreme Quantile Model					
Increase news mentions	0.077	0.074	0.095			
	[0.044]*	[0.044]*	[0.040]**			
Decrease news mentions	-0.22	-0.22	-0.13			
	[0.050]***	[0.049]***	[0.038]***			
Normal # Increase news mentions				-0.084	-0.087	-0.055
				[0.020]***	[0.020]***	[0.016]***
High # Increase news mentions				1.11	1.11	1.07
				[0.79]	[0.79]	[0.74]
Excessive # Increase news mentions				1.21	1.21	1.18
				[0.45]***	[0.45]***	[0.45]***
Normal # Decrease news mentions				-0.088	-0.084	-0.034
				[0.019]***	[0.019]***	[0.016]**
High # Decrease news mentions				-0.63	-0.6	-0.62
				[1.14]	[1.14]	[1.08]
Excessive # Decrease news mentions				-0.41	-0.41	-0.48
				[0.89]	[0.89]	[0.88]
Commodity Effects	Yes	Yes	Yes	Yes	Yes	Yes
Market Controls	No	Yes	Yes	No	Yes	Yes
Calendar Effects	No	No	Yes	No	No	Yes
Observations	9,700	9,695	9,695	9,521	9,516	9,516

Standard errors in second row

HAC-SE (in brackets) and Statistics robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation. Clustered on date. Columns (1)-(3) shows the estimates for the news variables with different covariates. Columns (4)-(6) shows the estimates for the days identified as normal (0), high (1), excessive (2).

* p<0.10, ** p<0.05, *** p<0.01

Table 15 Ext. Value Volatility: Spell or an indicator of days the model identifies with abnormal returns

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-parametric Extreme Quantile Model					
Increase news mentions	-2.74	-1.92	-5.01			
	[8.08]	[7.31]	[7.51]			
Decrease news mentions	73.6	72.5	4.81			
	[7.32]***	[7.77]***	[6.19]			
Normal # Increase news mentions				5.73	6.46	-1.34
				[8.11]	[7.41]	[6.58]
High # Increase news mentions				-35.6	-31.5	-20.1
				[19.5]*	[19.2]	[13.2]
Excessive # Increase news mentions				-39.1	-39.3	-17.9
				[10.4]***	[11.1]***	[19.7]
Normal # Decrease news mentions				68.5	67.4	3.94
				[7.74]***	[8.16]***	[6.84]
High # Decrease news mentions				-86.4	-93	-11
				[24.8]***	[29.3]***	[12.1]
Excessive # Decrease news mentions				-58.7	-57.8	-77.5
				[14.8]***	[15.4]***	[26.4]***
Commodity Effects	Yes	Yes	Yes	Yes	Yes	Yes
Market Controls	No	Yes	Yes	No	Yes	Yes
Calendar Effects	No	No	Yes	No	No	Yes
Observations	9,521	9,516	9,516	9,521	9,516	9,516
Standard errors in second row				5.73	6.46	-1.34

HAC-SE (in brackets) and Statistics robust to both arbitrary heteroskedasticity and arbitrary common autocorrelation. Clustered on date. Columns (1)-(3) shows the estimates for the news variables with different covariates. Columns (4)-(6) shows the estimates for the days identified as normal (0), high (1), excessive (2).
* p<0.10, ** p<0.05, *** p<0.01

7 Robustness Checks

In the main tables presented in this paper, we included estimates with variables that predict the market fundamentals in order to demonstrate that the relationship controlling for the price effects is what one might expect. In this section, we present some robustness checks of the estimates presented previously and address the possibility of completely endogenously determined news and prices.

It could be that some outside event influences both news and price direction across the panel, and it can be very difficult to controls for this type of relationship. To address this possibility, we provide Granger causality tests between the returns and the media variables to see if there is a predictable direct time link. Table 16 shows the results for these tests. The test in panel A cannot reject the null that the media coverage variables do not granger-cause prices; this is the stronger test of our sequential exogeneity assumption. This evidence implies that old news in the media do not directly affect returns after we control for the previous day's return; that is, old news articles do not provide additional information that affects prices. In panel B and C, we see the test for the news variables; panel B rejects the null hypothesis suggesting that increases in mentions are affected by previous returns and by the decrease in news mentions of the previous day. In panel C, we also reject the null hypothesis of joint granger-causality between returns and increased news and decreased news. Note that this is exactly what lies behind our IV-GMM estimation strategy: first, that news mentions on the current day are important for price determination (supported by panel A; it is reasonable to estimate (1) via OLS) and second, that news in the current day are affected by news in the previous days (both panel B and C) and by the price level in the previous days (panel B, and panel C when considering the joint test).

Table 16 Granger Causality from Homogeneous Panel Vector Auto-regression (VAR)

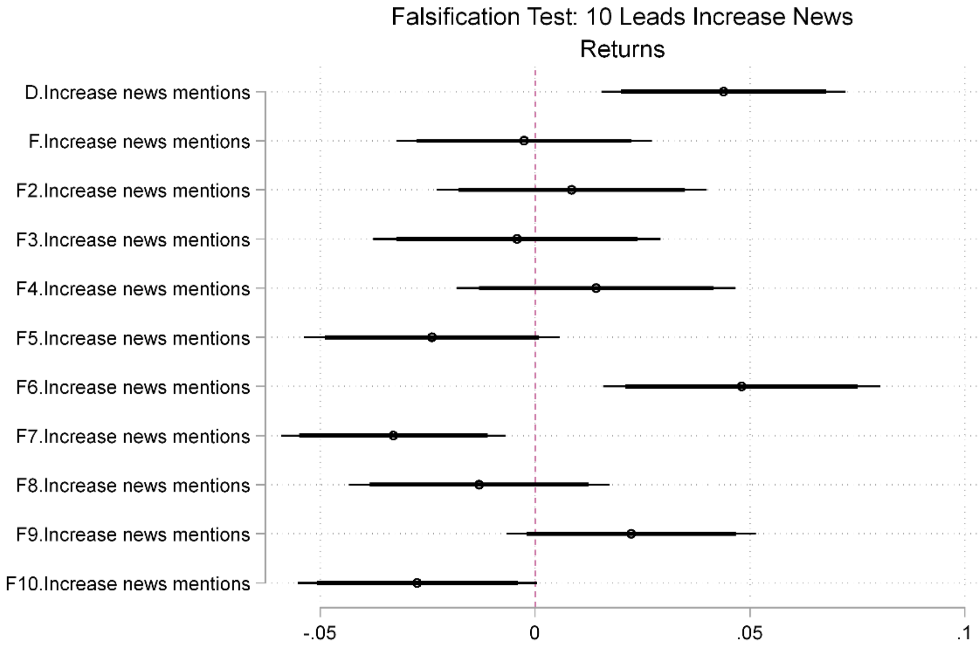
	chi2	df	Prob > chi2	Conclusion
Panel A: Return Price				
Increase news mentions	0.47	1	0.493	Cannot Reject
Decrease news mentions	0.97	1	0.324	Ho: News do not granger
ALL	1.39	2	0.498	cause return
Panel B: Increase news mentions				
				Reject
Return Price	5.79	1	0.016	Ho: Returns and Decrease
Decrease news mentions	6.73	1	0.009	News do not granger cause
ALL	13.01	2	0.001	Increase News
Panel C: Decrease news mentions				
Return Price	1.58	1	0.209	Reject
Increase news mentions	22.94	1	0.000	Ho: Returns and Increase
				News do not granger cause
ALL	26.41	2	0.000	Decrease News

Note: VAR for granger causality test fits a multivariate panel regression of each dependent variable on lags of itself and on lags of all other dependent variables using generalized method of moments (GMM)

In Figure 18, we present the estimates from the difference equation including leads of the media coverage variables. The idea is that news in the future should not be strongly related to prices on the current day and that any relationship should decrease as media reports occur farther in the future. In the figure, we include 10 leads of the increase and decrease mentions variables. The figure shows that the estimate of the media effects remains significant for both increase and decrease mentions and that the leads of these variables are not significant in most cases.[16] In Figure 19, we present the test for volatility using the standardize residuals (as in section (b) of the volatility results) and the IV-GMM estimate of the difference equation. The figure shows that the estimate of the media effects on volatility remains significant for both increase and decrease mentions variables and that the 10 leads of these variables are not significant.

[16] Only lead six and seven in the increase news mentions estimate are significant, which could be due to some persistence or a long spell in increase news. We estimate the model including up to 20 leads, and the estimates for the leads after 10 were not significant.

Figure 18 Falsification Tests for Returns Results: Leads of Media coverage



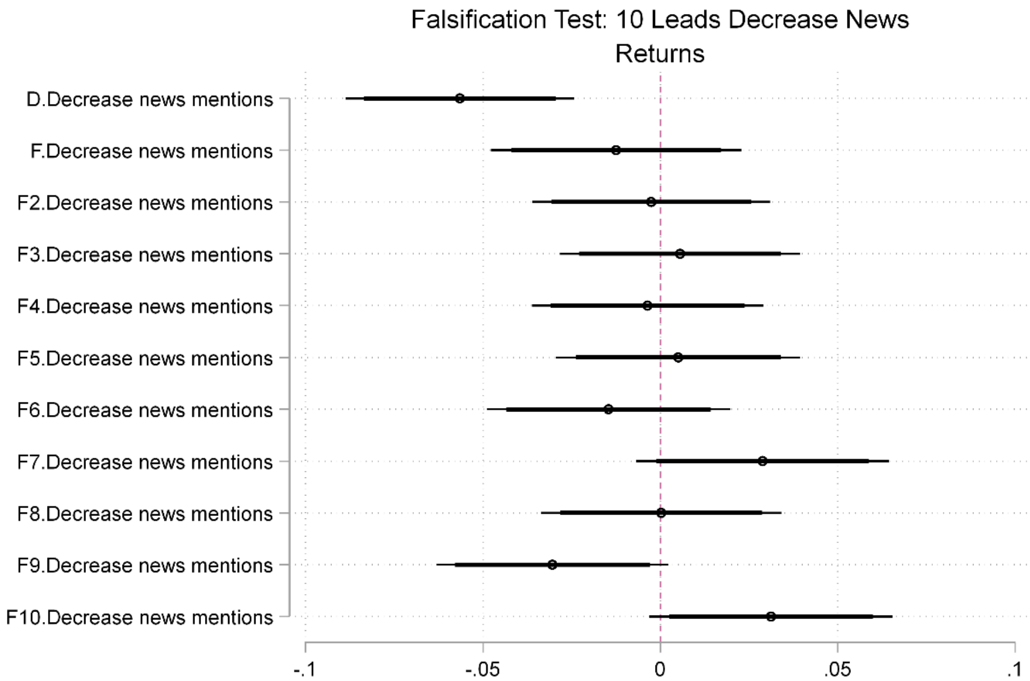
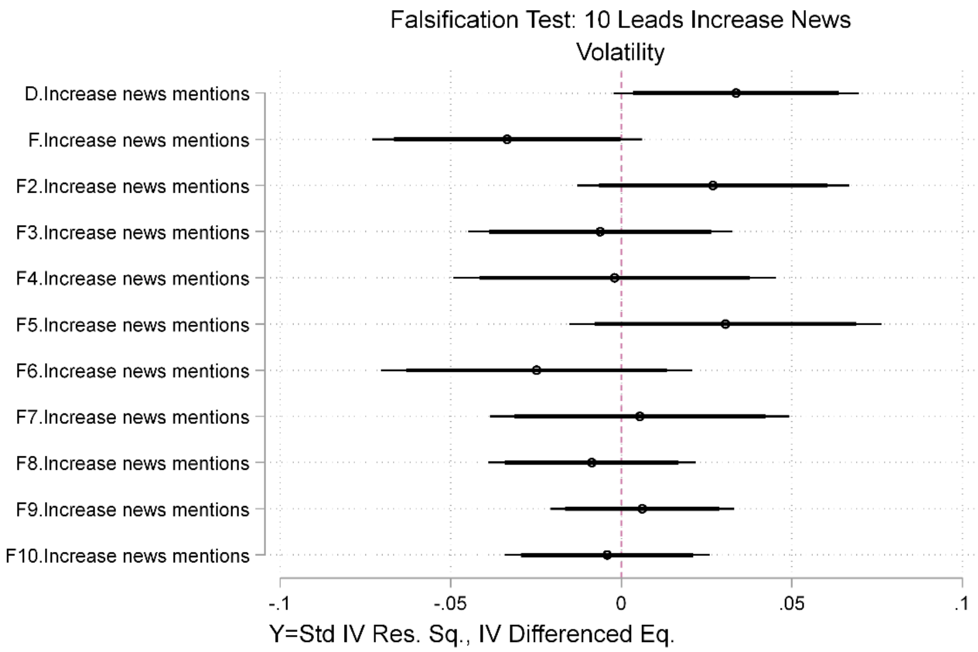
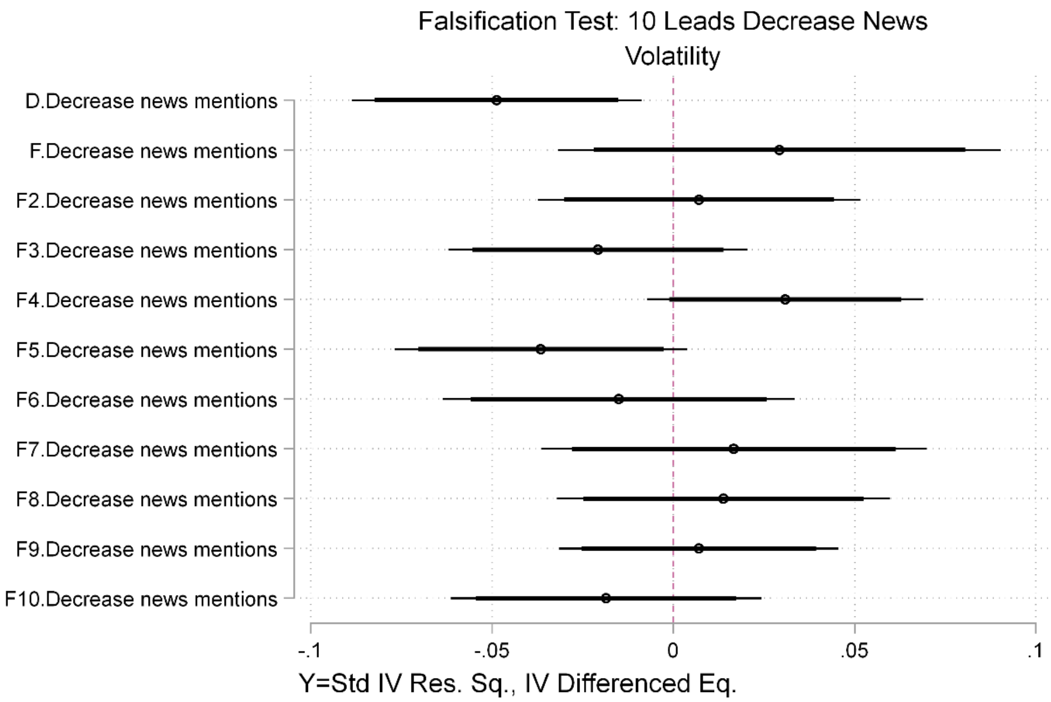


Figure 19 Falsification Test for Volatility Results: Leads of Media coverage





8 Conclusions

At times, the daily global news can be inundated with stories of rising food prices and accompanying rises in poverty and hunger. Droughts in China, Russia, and the US Midwest of the US have at times driven food prices, specifically commodity prices, up around the world. In addition to the food security challenges posed by rises in the price of staple commodities, such as wheat and maize, one of the major factors threatening global food security remains extreme price fluctuations and the observed political and market overreaction that normally follows. Policymakers are now faced with decisions regarding the appropriate response to increases in food prices and food price volatility.

In 2007-2008, volatile food prices led many major food producers to impose knee-jerk reactions, such as export restrictions. In 2012, restrictions on exports from large exporting countries placed more pressure on commodity prices, dramatically affecting consumers worldwide. Such an impact is magnified for poor consumers, who spend a large portion of their incomes on food. While commodity prices have decreased in recent years, the lessons from these events can better prepare policymakers to respond to price increases in the future.

This paper draws on extensive research into returns on agricultural commodity prices to explain how media coverage of price fluctuations in food commodity prices affects these prices and their volatility. We estimate the effects of media coverage on these two important aspects in agricultural commodities futures markets. We develop a unified empirical framework to analyze the effects of media coverage on both returns and volatility, using insights from the literature on the analysis of information in financial markets; we also compare the qualitative results from our model to other more common models in the literature.

We use a unique dataset to construct a measure of media coverage, and we uncover several interesting impacts of media coverage of varying intensity on the price dynamics in these markets. The data follows a comprehensive set of global media outlets and uses an algorithm to determine sophisticated relationships in phrases in a media article which signal an increase or decrease in price.

We find price effects that are economically important in size. When compared to the daily return of the SP500 Index during the study period, our estimates suggest that media coverage about increases in prices of agricultural commodities can account for 94.3 percent of the return on the SP500; for coverage of decreases in prices, a reallocation of market positions following the media coverage could lead to returns 20.9 percent above of the median return for the SP500. At the means for the study period, the estimates imply a 0.12 percent effect on returns, signaling ***a net increasing effect of media coverage on the price of these commodities***. We explore heterogeneity across commodities and time periods and find that the effects observed were mostly concentrated in 2012 and from 2015 on, signifying that these effects are important in periods of both high and low prices. Across commodities, the price effects were concentrated in soybeans and maize.

The finding that increased media coverage of price movements reinforces those movement in the direction of the news reports strengthens the case that increased media coverage during periods of food crises can exacerbate price spikes. In addition, our findings suggest that media coverage that follows market fundaments and indicates that prices should decrease in the near future can serve as a policy tool to soften the increase in the price of food commodities.

We find robust evidence that, on average, media coverage decreases volatility for the studied agricultural commodities during the study period. The effects on volatility balance each other, with ***coverage of decreasing prices decreasing the variance*** of returns and ***coverage of increasing prices increasing the variance of returns*** but to a lesser extent than the effect of the decreasing price coverage. Coverage of increased prices increases the variance by 3 percent, while and coverage of decreasing prices decreases the variance by 4.6 percent. While using the 30-day rolling variance model suggests that media coverage has **increased** volatility, on net, in the period we study and across the different models we estimate, the evidence points to a ***decrease in volatility effect due to media coverage*** for these commodities. Using a canonical GARCH model, we find effects that are similar to those using our empirical framework.

Using a data-driven, non-parametric model to identify periods of normal, high, and excessive volatility, we find that media coverage of increases in prices increases the probability of having a highly abnormal return by 0.095 percentage points, while coverage of decreases in price decreases that probability by 0.13 percentage points. We also find that the context in which the coverage occurs matters. During periods of normal volatility, media coverage of increases in price decreases the probability of having excessive returns, promoting stability in the market. In periods of excessive volatility, coverage of increased prices increases the probability of excessive returns, promoting more unstable or spiky markets. The results from this model suggest that media coverage **increases periods of normal volatility** and **decreases periods of excessive volatility**.

We concluded with evidence that the time link between price movements and media coverage supports our estimation strategy and that the results are not driven by spurious correlations.

Our estimates are consistent with efficiently functioning markets in which the media helps to process complex information that might not be reflected in objective or quantitative measures of market fundamental. They highlight crucial role of providing appropriate information as fast as possible so media coverage reflects the fundamentals that drive food commodity prices and not investor or trader speculation. From these estimates, we can better understand the dynamics between prices and media coverage, and this deeper understanding may help prevent rushed and automatic policy reactions by discouraging market overreaction, encouraging market stability, and promoting food security.

References

Ahmad, K., Han, J., Hutson, E., Kearney, C., and Liu, S., 2016. "Media-expressed negative tone and firm-level stock returns", Journal of Corporate Finance, Elsevier, vol. 37(C), pages 152-172.

Andersen, T. G. and Benzoni, L., 2008. "Realized volatility". In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.): Handbook of Financial Time Series, 554-575. Springer, New York.

Anderson, T. W. and Hsiao, C., 1981. "Estimation of Dynamic Models with Error Components", Journal of the American Statistical Association, 76, 598-606.

Arellano, M. and Bond, S., 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", Review of Economic Studies, 58, 277-297.

Arellano, M. and Bover, O., 1995. "Another Look at the Instrumental-Variable Estimation of Error-Components Models", Journal of Econometrics, 68, 29-51.

Baum, C.F., Schaffer, M.E., Stillman, S. 2010. "ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression." http://ideas.repec.org/c/boc/bocode/s425401.html

Blundell, R. and Bond, S., 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", Journal of Econometrics, 87, 115-143.

Bollerslev, T., 1986. "Generalized autoregressive conditional heteroskedasticity", Journal of Econometrics, 31, issue 3, p. 307-327.

Carter, C. A. and Smith, A., 2007. "Estimating the Market Effect of a Food Scare: The Case of Genetically Modified StarLink Corn", The Review of Economics and Statistics, MIT Press, vol. 89(3), pages 522-533, 07.

Driscoll, J.C. and Kraay, A. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data", The Review of Economics and Statistics. Vol. 80, No. 4, pp. 549-560.

Dubofsky, D. A., 1991. "Volatility Increases Subsequent to NYSE and AMEX Stock Splits," Journal of Finance, 46, 421-431.

Engelberg, J.E. and Parsons, A., 2011. "The Causal Impact of Media in Financial Markets", Journal of Finance, 66, issue 1, p. 67-97.

Fama, E. F., 1970. “Efficient capital markets: A review of theory and empirical work”, Journal of Finance 25, 383-417.

Fama, E. F., 1991. “Efficient Capital Markets: II.” The Journal of Finance, 46: 1575–1617.

Fang, L., and Peress, J., 2009. “Media Coverage and the Cross-Section of Stock Returns”. The Journal of Finance, Vol. 64, No. 5, pp. 2023-2052 Ahmad, K., Han, J., Hutson, E., Kearney, C., and Liu, S., 2016. "Media-expressed negative tone and firm-level stock returns", Journal of Corporate Finance, Elsevier, vol. 37(C), pages 152-172.

Han, C. and Phillips, P. C. B., 2010. "GMM Estimation For Dynamic Panels With Fixed Effects And Strong Instruments At Unity," Econometric Theory, Cambridge University Press, vol. 26(01), pages 119-151, February.

International Food Policy Research Institute (IFPRI), 2012. 2011 Global food policy report. Washington, D.C.: International Food Policy Research Institute (IFPRI)

Martins-Filho, C., Yao, F., and Torero, M. 2015. "High-Order Conditional Quantile Estimation Based on Nonparametric Models of Regression," Econometric Reviews, Taylor & Francis Journals, vol. 34(6-10), pages 907-958, December.

Martins-Filho, C., Yao, F., and Torero, M. 2016. “Nonparametric Estimation of Conditional Value-At-Risk and Expected Shortfall Based on Extreme Value Theory.” Econometric Theory, 1-45.

McKenzie, A.M., and Thomsen, M.R., 2011. "The Effect of E. Coli 0157:H7 on Beef Prices." Journal of Agricultural and Resource Economics 26(2001):431-44.

Nickell, S. (1981): “Biases in Dynamic Models with Fixed Effects”, Econometrica, 49, 1417-1426.

Ohlson, J. A., and Penman, S. H., 1985, “Volatility Increases Subsequent to Stock Splits: An Empirical Aberration,” Journal of Financial Economics, 14, 251-266.

Pruitt, S. W., Tawarangkoon, W. and Wei, K. C. J., 1987. “Chernobyl, commodities, and chaos: An examination of the reaction of commodity futures prices to evolving information”, Journal of Futures Markets, 7: 555–569.

Rucker, R. R., Thurman, W.N., and Yoder, J. K., 2005. "Estimating the Structure of Market Reaction to News: Information Events and Lumber Futures Prices", American Journal of Agricultural Economics, Agricultural and Applied Economics Association, vol. 87(2), pages 482-500

Schlenker, W., and Villas-Boas, S. B., 2009. "Consumer and Market Responses to Mad Cow Disease," American Journal of Agricultural Economics, Agricultural and Applied Economics Association, vol. 91(4), pages 1140-1152.

Smith, M. E., van Ravenswaay, E. O., and Thompson, S. R., 1988, “Sales loss determination in food contamination incidents: an application to milk bans in Hawaii,” American Journal of Agricultural Economics Vol. 70, No. 3 (Aug., 1988), pp. 513-520

Tetlock, P. C., 2007. “Giving Content to Investor Sentiment: The Role of Media in the Stock Market.” The Journal of Finance, 62: 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M. and Macskassy, S., 2008. “More Than Words: Quantifying Language to Measure Firms' Fundamentals.” The Journal of Finance, 63: 1437–1467.

Thompson, S. B., 2011. “Simple formulas for standard errors that cluster by both firm and time”, Journal of Financial Economics, Volume 99, Issue 1, January 2011, Pages 1-10

Torero, M., 2012. “Food prices: riding the rollercoaster.” In: 2011 Global food policy report. International Food Policy Research Institute, Washington, DC

Torero, M., 2016. “Consistency between theory and practice in policy recommendations by international organizations for extreme price and extreme volatility situations.” In: Food price volatility and its implications for food security and policy, eds. Matthias Kalkuhl, Joachim von Braun, and Maximo Torero. Chapter 19, pp. 45

Appendix 1 - Details of Media Data and Sources

Every day, we monitor a comprehensive set of Really Simple Syndication (RSS) feeds[17] drawn from global media outlets via Google news. A total of 31 feeds related to global food prices and food security are monitored; these feeds include search strings such as "food prices," "food crisis," "agricultural development," "commodity prices," "price of maize," "price of wheat," "price of oil," "price of rice," "price of soybean," etc. Stories are tagged if they are about: 1. global food security or food prices, 2. ongoing national, regional, or global food crises, 3. prices (international, regional, and national) or crop conditions of major agricultural commodities (wheat, corn, soybeans, and rice), 4. oil prices, 5. agricultural trade (export bans, import or export forecasts, etc.), or 6. agricultural/food policy research.

At the end of each day, all starred articles are converted into .txt files and saved using the format "title_month_day_year.txt." The ".txt" files of the day are then uploaded into the IFPRI Food Security Analysis System, a tool built by Sophic Systems Alliance, called Sophic Intelligence Software. This software, which is built on the Biomax Knowledge Management Suite, uses linguistic and semantic object network-mapping algorithms to analyze the relationships between key terms found in each article. When articles are uploaded each day, the tool mines the complete database of articles for a select set of key words. Sophic Intelligence Software generates a detail analysis of the text within the articles and look at phrases in the articles that influence commodity price volatility and food security.

Using the list of key words to determine an "up" or "down" movement within our database of articles, the software identifies how many times phrases occur in the articles. The categories that the software mines in the text in the articles are based on four categories or functions:

a. **Financial**: domestic food price, expectations, expected prices, futures markets, hedge, hedging, interest rate, international food price, monetary policy, rates, speculation, trade, trade barrier, trading volume;

b. **Inventories**: corn production, domestic production, domestic supply, emergency reserves, maize production, reserves, rice production, storage, supply, surplus, and wheat production;

c. **Policies**: export ban, export quota, food security, import quota, import restrictions, price controls, and taxes; and

d. **Disasters and civil effects**: drought, earthquake, famine, fire, flood, frost, hurricane, nutrition, plague, poverty, riots.

[17] Also called web feeds, RSS is a content delivery vehicle. It is the format used to syndicate news and other web content. When it distributes the content, it is called a feed.

Within each of these categories, the text is scanned for occurrences of keywords that suggest changes in prices: rise, reduce, collapse, grow, lower gain, shrink, etc. For example, an article containing the words "soybean" and "surge" would denote an "up" movement in soybean prices; if the soybean "up" report on a given day is listed as 5, this means that on that day, 5 articles contained words suggesting a rise in soybean prices. Daily, the system provides reports analyzing movements (increases -ups - or decreases - downs) in commodity prices. These reports provide a count of the number of times the articles mentions "up" or "down" movements for each commodity each day. We use these "up" and "down" variables to measure the intensity of media coverage of a price change.

The measures of media coverage are obtained by monitoring a comprehensive set of RSS feeds drawn from global media outlets via Google news. A non-exhaustive list of sources of these feeds is:

Details of Media Data and Sources		
ABC	Fox Business	Pakistan Daily Times
AFP	Futures Magazine	Politico
Agriculture.com	Ghana News Agency	Reuters
Agrimoney.com	Hindu Business Line	RTT News
All Africa	Huffington Post	San Francisco Chronicle
Arab News	Independent Online	The Australian
Associated Press of Pakistan	Indian Express	The Guardian
Barron's	Inside Futures	The Seattle Times
Bloomberg	Kuwait Times	Time Magazine
Business Day	Los Angeles Times	Times of India
Business Standard	NASDAQ	UK Telegraph
China Daily	New York Times	UN News Centre
CNBC	Newstime Africa	Wall Street Journal
Economic Times	NPR	Washington Post
Food World News	Pakistan Business Recorder	Weekly Times Now

A total of 31 feeds related to global food prices and food security are monitored; these feeds include search strings such as:

Keywords	
AGOA	Food security
Agricultural/food policy research	Global food security
Agriculture development	Import or export forecasts
Climate change	National, regional, or global food crisis
Commodity Prices	Oil world
Ethanol subsidies	Price of maize or maize prices or maize export
Export bans	Price of oil or oil prices or oil
FAO	Price of rice or rice prices or rice export
Food crisis	Price of soybean or soybean prices or soybean export or soybean
Food prices	Price of wheat or wheat prices or wheat export