Center for Quantitative Economics

# Price Discovery in Agricultural Commodity Markets: Do Speculators Contribute?

Martin T. Bohl [†], Pierre L. Siklos [#], Martin Stefan [†] und Claudia Wellenreuther [†]

75/2018

[†] Department of Economics, University of Münster, Germany
[#] Department of Economics, Wilfrid Laurier University, Waterloo, Canada

wissen.leben
WWU Münster

# Price Discovery in Agricultural Commodity Markets: Do Speculators Contribute?

Martin T. Bohl [a] Pierre L. Siklos [b] Martin Stefan [c]
Claudia Wellenreuther [d]

October 10, 2018

**Abstract**

Previous literature on price discovery in commodity markets is mainly focused on the question whether the spot or the futures market dominates the price discovery process. Little attention, however, has been paid to the question how the price discovery process is affected by futures speculation. Using different measures for speculation and hedging and a new price discovery metric, the present study analyzes this relationship for various agricultural commodities. The results indicate that speculative activity generally reduces the level of noise in the futures market, while increasing the relative contribution of the futures market to the price discovery process.

**Keywords:** Commodity Markets, Futures Speculation, Price Discovery

[a]Corresponding author: Department of Economics, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Germany, Phone: +49 251 83 25005, E-mail: martin.bohl@wiwi.uni-muenster.de

[b]Department of Economics, Wilfrid Laurier University

[c]Department of Economics, Westfälische Wilhelms-Universität Münster

[d]Department of Economics, Westfälische Wilhelms-Universität Münster

# 1 Introduction

In the past decade the role of speculators in commodity futures markets has been subject to great debate. Public commentators, politicians and market participants feared that speculation could impair the functioning of futures markets. Two key functions of futures markets are risk transfer and price discovery. Concerning the transfer of risk, the importance of futures speculation is straightforward. Speculators provide liquidity to the market and help hedgers to find a matching counterparty to transfer price risk (Keynes 1923, Kaldor 1940). With regard to futures markets' price discovery function, the role of speculators is, however, less clear.

Previous research largely focused on determining whether price discovery occurs primarily in the futures or in the spot market. Across different agricultural commodities, the majority of studies concludes that futures markets dominate the price discovery process (see e.g. Garbade & Silber 1983, Zapata & Fortenbery 1996, Yang & Leatham 1999, Yang et al. 2001). This is in line with the theoretical argument of Grossman (1977) that futures markets are an important place for gathering and exchanging information. However, little attention has been paid to the role of speculation in affecting futures markets' price discovery function. This article sets out to fill this gap in the literature.

Price discovery is defined as the "efficient and timely incorporation of the information implicit in investor trading into market prices" (Lehmann 2002, p. 259). While "efficient" means the relative absence of noise, "timely" refers to how quickly a price series reflects new information about an asset's fundamental value (Putniņš 2013). It seems reasonable to expect that speculation can affect both the level of noise and the processing of new information. Therefore, we follow Yan & Zivot (2010) and Putniņš (2013) who establish a new way of measuring price discovery which specifically allows disentangling these two elements of price discovery. Their approach combines two commonly used measures of price discovery, namely the component shares developed by Garbade & Silber (1983) and Gonzalo & Granger (1995) as well as the information shares of Hasbrouck (1995) and Lien & Shrestha (2009). Using a structural vector error correction model, Yan & Zivot (2010) show that the component shares primarily measure the avoidance of noise, while the information shares capture both the avoidance of noise and the processing of new information. The authors then combine these measures to subtract the noise avoidance component from the information shares to obtain a new metric of informational leadership, which Putniņš (2013) reformulates as information leadership shares.

This article studies the impact of speculation on information leadership in the price discovery process in four agricultural commodity futures markets, namely the markets for corn, soybean, cocoa and feeder cattle futures. Our analysis comprises two steps. First, price data are used to compute time-varying information leadership shares using rolling window estimation. Second, this price discovery metric is regressed on different measures of speculative and hedging activity including the total levels of speculation and hedging, as well as the level of excessive speculation as measured by Working's (1960) T. Here, we rely on trader positions data from the Commodity Futures Trading Commission's (CFTC) Commitment of Traders (COT) reports.

Our results indicate that futures speculation significantly enhances the price discovery function of the futures market. This holds for both the total level of speculation as well as the level of excessive speculation. Conversely, hedging activity is found to reduce the futures market's ability to process new information. This result supports the intuitive idea that hedgers intend to lock in prices for the future, while speculators seek to execute profitable trades by using new information. Hence, speculators are informed traders and important market participants who allow futures markets to fulfill their key functions.

The remainder of this paper is structured as follows. Section 2 outlines the methodology used to measure price discovery and how it is affected by futures speculation, while section 3 describes the data used in this study. The results of this paper are presented in section 4 which are subsequently tested for their robustness. Lastly, section 5 concludes.

# 2 Methodology

To study the impact of futures speculation on price discovery, a two-step approach is adopted. Based on a simple cost-of-carry model, we first compute time-varying metrics of the relative contributions of the spot and futures markets to the price discovery process. Thereafter, a linear regression analysis is used to investigate how the price discovery function of the futures market is influenced by different measures of speculative and hedging activity.

## 2.1 Cost-of-carry model

In a simple cost-of-carry model, $S_t$ denotes the spot price of a commodity in period $t$ and $F_t^T$ the price of a futures contract with delivery in period $T$. To carry the commodity from $t$ to $T$, one incurs the cost of carry $(r + c - y) \cdot (T - t)$, where $r$ is the interest rate, $c$ denotes storage and transportation costs and $y$ is the convenience yield. Spot and futures market are naturally linked by arbitrage. Given that all arbitrage opportunities are exploited it must hold that

$$F_t^T = S_t \cdot e^{(r+c-y)\cdot(T-t)} \ . \tag{1}$$

Taking logs on both sides of the equation yields

$$f_t^T = s_t + \theta \ . \tag{2}$$

Here, $f_t^T$ denotes the logarithm of the futures price and $s_t$ that of the spot price. The parameter $\theta$ comprises the cost-of-carry-rate and all other components of spot and future price differentials. To investigate this relationship empirically, the following equation is estimated:

$$f_t^T = \theta + \xi s_t + \varepsilon_t \; . \tag{3}$$

In this equation $\varepsilon_t$ resembles white noise, while the unbiasedness hypothesis (Engel 1996, Zivot 2000) implies that the parameter $\xi$ will be equal to one.

Consequently, the following cointegration relationship with a cointegration vector of $(1, -1)$ is obtained:

$$ec_{t-1} = f_{t-1} - \xi_0 - \xi_1 s_{t-1} \; . \tag{4}$$

Based on the Granger representation theorem (Engle & Granger 1987) the relationship between futures and spot prices can be expressed using the following vector error correction (VEC) model, where the operator $\Delta$ denotes first differences:

$$\Delta f_t = \gamma_f + \alpha_f ec_{t-1} + \sum_{k=1}^{K} \gamma_{fs,k} \Delta s_{t-k} + \sum_{q=1}^{Q} \gamma_{ff,q} \Delta f_{t-q} + \varepsilon_{f,t} \; , \tag{5a}$$

$$\Delta s_t = \gamma_s + \alpha_s ec_{t-1} + \sum_{k=1}^{K} \gamma_{ss,k} \Delta s_{t-k} + \sum_{q=1}^{Q} \gamma_{sf,q} \Delta f_{t-q} + \varepsilon_{s,t} \; . \tag{5b}$$

Parameters $\gamma_f$ and $\gamma_s$ are the equations' intercepts. The error-correction coefficients $\alpha_f$ and $\alpha_s$ reflect the speed of adjustment of the two price times series towards the long-run price equilibrium in response to short-run deviations of the system. Conversely, parameters $\gamma_{ij}$ captures the model's short run dynamics, i.e. how previous changes in market $j = f, s$ affect present changes in market $i = f, s$.

As shown by Stock & Watson (1988) the VEC model of equation (5) has the following vector moving average (VMA) representation:

$$\boldsymbol{p}_t = \boldsymbol{p}_0 + \boldsymbol{\Psi}(1) \sum_{i=1}^{t} \boldsymbol{\varepsilon}_i + \boldsymbol{\Psi}^*(L)\boldsymbol{\varepsilon}_t \ , \tag{6}$$

with

$$\boldsymbol{p}_t = (f_t \ \ s_t)' \ , \tag{7}$$

where $\boldsymbol{p}_0$ is a $(2 \times 1)$ vector of constants and the $(2 \times 2)$ matrix $\boldsymbol{\Psi}(1)$ is the moving average impact matrix. This matrix contains the cumulative VMA coefficients such that the term $\boldsymbol{\Psi}(1) \sum_{i=1}^{t} \boldsymbol{\varepsilon}_i$ summarizes the long-run impact of an innovation at time $t$ on each of the prices. Lastly, $\boldsymbol{\Psi}^*(L)$ is a matrix polynomial in the lag operator, $L$, where $\boldsymbol{\varepsilon}_t$ is a $(2 \times 1)$ error vector with variance-covariance matrix $\boldsymbol{\Omega}$. The cointegration relationship described in equation (4) implies that $\boldsymbol{\gamma}^T \boldsymbol{\Psi}(1) = 0$ and $\boldsymbol{\Psi}(1)\boldsymbol{\alpha} = 0$. Based on the parameters of the VEC and the VMA model, the literature has proposed various price discovery measures which are discussed below.

## 2.2 Measuring price discovery

As explained earlier, price discovery is defined as the efficient and timely incorporation of the information implicit in investor trading into market prices (Lehmann 2002). While "efficient" means the relative absence of noise, "timely" refers to how quickly a price time series reflects new information about an asset's fundamental value (Putniņš 2013). To quantify the relative contribution of each market to the price discovery process, the literature has proposed various price discovery metrics. However, as argued by Yan & Zivot (2010) and Putniņš (2013) these metrics capture two different aspects of the price discovery process, namely noise avoidance and actual price leadership in the sense of incorporating new information more quickly.

In this sense, a measure for the avoidance of noise is based on the contributions of Garbade & Silber (1983), Schwarz & Szakmary (1994) and Gonzalo & Granger (1995). The latter show how the price vector $\boldsymbol{p}_t$ can be decomposed into a permanent (common factor) component and a transitory (stationary) component. The authors propose two component shares ($CS_f$ and $CS_s$) which can be calculated from the VEC model's long-run adjustment coefficients as

$$CS_f = \frac{|\alpha_s|}{|\alpha_f| + |\alpha_s|} \quad \text{and} \quad CS_s = \frac{|\alpha_f|}{|\alpha_f| + |\alpha_s|}\,, \tag{8}$$

where it holds by construction that $CS_f + CS_s = 1$. It follows from equation (8) that the smaller a market's long-run adjustment coefficient, the greater its component share. Thus, if $\alpha_f = 0$, the futures market does not react to any deviations from the long-run equilibrium. Consequently, all adjustment to the equilibrium occurs in the spot market. In this case $CS_s = 0$ such that the spot market is said to make no contribution to the price discovery process, which is then entirely dominated by the futures market.

While these component shares only measure the markets' relative avoidance of noise, an alternative measure developed by Hasbrouck (1995) and Lien & Shrestha (2009) captures both noise avoidance and the processing of new information. Both argue for making use of the information contained in the variance-covariance matrix of the error term in equation (5), $\boldsymbol{\Omega}$, in order to account for short-term dynamics. For this purpose, the authors propose the following information shares

$$IS_f = \frac{\left(\left[\boldsymbol{\Psi F}\right]_f\right)^2}{\boldsymbol{\Psi\Omega\Psi'}} \quad \text{and} \quad IS_s = \frac{\left(\left[\boldsymbol{\Psi F}\right]_s\right)^2}{\boldsymbol{\Psi\Omega\Psi'}}\,, \tag{9}$$

where the $(1\times 2)$ matrix $\boldsymbol{\Psi}$ is either of the two identical rows $\boldsymbol{\Psi}_1$ or $\boldsymbol{\Psi}_2$ contained in $\boldsymbol{\Psi}$ and $\boldsymbol{F}$ is a transformed version of the error covariance matrix $\boldsymbol{\Omega}$.[1] To obtain $\boldsymbol{F}$, Hasbrouck (1995) originally advocates the use of a Cholesky factorization. However, depending on the orderings of the prices this procedure produces two information shares for each market, referred to as upper and lower bounds. Conventionally, averages of these upper and lower bounds are computed. But in the case of highly divergent bounds, a market's average information shares are rather weak in their explanatory power. For this reason, Lien & Shrestha (2009) propose using an eigen-decomposition of $\boldsymbol{\Omega}$ to obtain a unique matrix $\boldsymbol{F}$ which then circumvents the problem of the Hasbrouck bounds. Similar to the component shares, these information shares also sum to one and have an analogous interpretation regarding the dominance of either market, i.e. price discovery primarily occurs in the market with the higher information share.

Yan & Zivot (2010) and Putniņš (2013) note that the component shares as well as the information shares of Hasbrouck (1995) and the modified version by Lien & Shrestha (2009) are only able to accurately measure price discovery in the sense of "which market moves first" when both markets exhibit similar levels of noise. This finding is based on a structural cointegration model of Yan & Zivot (2010). The authors show that the component shares only capture the relative response of the system to transitory shocks, while the information shares measure responses to both transitory shocks and permanent shocks. Consequently, in the case of differing noise levels, the price discovery measures can be higher for two reasons. On the one hand, the price discovery measures

[1] Baillie et al. (2002) show that $\boldsymbol{\Psi}$ can directly be obtained from the VEC model in equation (5). In particular, the authors prove that

$$\boldsymbol{\Psi}(1) = \boldsymbol{\xi}_{\perp}\pi\boldsymbol{\alpha}'_{\perp} = \begin{pmatrix} \boldsymbol{\Psi} \\ \boldsymbol{\Psi} \end{pmatrix} = \begin{pmatrix} \Psi_1 & \Psi_2 \\ \Psi_1 & \Psi_2 \end{pmatrix} = \pi \begin{pmatrix} \delta_f & \delta_s \\ \delta_f & \delta_s \end{pmatrix} , \tag{10}$$

where $\delta_f$ an $\delta_s$ denote the orthogonal complements to the vector of error correction adjustment coefficients ($\alpha_{\perp}$). The scalar $\pi$ is an arbitrary constant resulting from the computation of the orthogonal complements, which cancels out when calculating the information shares.

are higher if a price time series processes new information more quickly. On the other hand they are also higher if the price time series contains less noise. Thus, when examining either of the price discovery measures individually, one will not be able to tell which of the two causes dominates.

In order to obtain a price discovery metric that exclusively measures how markets process new information, Yan & Zivot (2010) propose removing the "noise" component from the existing information shares by dividing the quotient of the information shares by the quotient of the component shares

$$IL_f = \frac{IS_f}{IS_s} \cdot \frac{CS_s}{CS_f} \quad \text{and} \quad IL_s = \frac{IS_s}{IS_f} \cdot \frac{CS_f}{CS_s}\,. \tag{11}$$

While originally unnamed, Putniņš (2013) refers to these two metrics as "information leadership" ($IL$). To ease interpretation and allow for straightforward comparisons between markets, the author suggests making use of the following normalization

$$ILS_f = \frac{IL_f}{IL_f + IL_s} \quad \text{and} \quad ILS_s = \frac{IL_s}{IL_f + IL_s}\,. \tag{12}$$

The resulting information leadership shares ($ILS$) add up to one, and thus have the same quantitative interpretation as the component and information shares.

In order to obtain a time varying measure of price discovery, this paper estimates equation (5) using a rolling window estimation. In the baseline version we select a window size of 250 trading days which corresponds to one trading year. We also conduct three robustness exercises. First, we decrease the window size to 200 trading days, then increase it to 375 and 500 trading days. While 200 days are roughly the minimum number of observations needed to meaningfully estimate a VEC model, 375 and 500 trading days correspond to one and a half respectively two whole trading years. Based on the resulting time-varying estimates for the vector of long-run coefficients $\boldsymbol{\alpha}$ and the variance-covariance matrix of residuals $\boldsymbol{\Omega}$, the three measures of price discovery are computed as explained above[2]. Comparing how futures speculation affects the different price discovery metrics $CS$, $IS$ and $ILS$ then allows disentangling the effects that speculation has on either the level of noise or the actual processing of new information in the futures market.

## 2.3 The impact of speculation on price discovery

To measure speculative activity, we rely on trader position data provided in the Commitment of Traders (COT) reports by the Commodity Futures Trading Commission (CFTC). The reports contain data on the end-of-day open interest of three different trader types for a variety of different commodity futures markets: First, there are commercial traders who can be classified as hedgers. The second category consists of non-commercial traders who are generally viewed as speculators. And third, there are traders whose positions fall below the CFTC's minimum reporting standards such that the nature of these traders' operations is unknown.

The literature has proposed various ways of measuring speculative activity using such trader position data. A straightforward measure which is e.g. used by Manera et al. (2016) is the total percentage share of open interest held by speculators, $S_t$. This measure is defined as

$$S_t = \frac{NCL_t + NCS_t + \alpha \cdot (NRL_t + NRS_t)}{2 \cdot MOI_t} \cdot 100 \,. \tag{13}$$

Long and short positions of non-commercial speculators are denoted by $NCS_t$ and $NCL_t$. Analogously, the positions of non-reporting traders are denoted by

[2] The detailed regression results of the rolling window estimations are available upon request.

$NRL_t$ and $NRS_t$. Lastly, $MOI_t$ resembles market open interest in week $t$, while $\alpha$ captures the fraction of speculative positions among non-reported positions. We follow Rutledge (1977) and Sanders et al. (2010) and allocate non-reported positions to hedging and speculative positions in the same proportions that are observed for reported positions.

Another common measure that specifically aims to capture excessive amounts of speculation is Working's (1960) T index, $W_t$. This index is defined as

$$W_t = \begin{cases} 1 + \dfrac{SS_t}{HS_t + HL_t}\,, & \text{if } HS_t \geq HL_t \\ 1 + \dfrac{SL_t}{HS_t + HL_t}\,, & \text{if } HS_t < HL_t \end{cases}\,, \tag{14}$$

where $SL_t$, $SS_t$, $HL_t$ and $HS_t$ represent the long and short positions of speculators and hedgers, respectively. These are given by

$$SL_t = NCL_t + \alpha \cdot NRL_t\ , \tag{14a}$$

$$SS_t = NCS_t + \alpha \cdot NRS_t\ , \tag{14b}$$

$$HL_t = CL_t + (1 - \alpha) \cdot NRL_t\ , \tag{14c}$$

$$HS_t = CS_t + (1 - \alpha) \cdot NRS_t\ . \tag{14d}$$

The idea behind this index is that any trade by a hedger will require an opposing trade by a speculator. If hedgers as a whole prefer taking a short position, i.e. $HS_t \geq HL_t$, then speculators must take a sufficiently large long position for the market to clear. If, however, speculators take a short position, $SS_t$, which exceeds the needs of hedgers, this can be regarded as a sign of excessive speculation. Consequently, Working's T has a lower bound of unity and increases with this speculative pressure. Note, however, that as highlighted by Fattouh et al. (2013) and Boyd et al. (2018) "excessive" is to be understood in a technical and not a normative manner. An index value above one only indicates that there are more speculative positions in the market than necessary to offset the positions of hedgers, not that market functioning is fundamentally impaired by speculative activity.

In addition to these two speculation measures we also consider a measure of hedging activity, which is computed in analogy to the total open interest share of speculators. But instead of using the positions of non-commercial traders, we now consider the positions of commercial traders. Hence, the total percentage share of hedgers, $H_t$, is defined as

$$H_t = \frac{CL_t + CS_t + (1-\alpha) \cdot (NRL_t + NRS_t)}{2 \cdot MOI_t} \cdot 100 \; . \tag{15}$$

Given these measures of speculation and hedging activity, $S_t$, $W_t$ and $H_t$, we are now able to analyze both how speculation and hedging affect the price discovery process. For this purpose the following linear regression model is estimated:

$$PD_t = \beta_0 + \beta_1 SPEC_t + \beta_2 VOL_t + \beta_3 BASIS_t + \eta_t \; , \tag{16}$$

where $PD_t$ is one of the measures of the futures market's contribution to price discovery in period $t$, i.e. $CS_{ft}$, $IS_{ft}$ or $ILS_{ft}$. Similarly, $SPEC_t$ is a measure of the dominance of either trader type, i.e. $S_t$, $W_t$, or $H_t$. $VOL_t$ represents the trading volume, while $BASIS_t$ denotes the spread between spot and futures prices. Lastly, $\eta_t$ denotes random disturbances.

When investigating the impact of speculation (hedging) on the component share, positive values for $\beta_1$ indicate that speculators (hedgers) increase the futures market's relative avoidance of noise, i.e. speculators (hedgers) reduce the amount of noise in the futures market. Whereas a positive impact of speculation (hedging) on the information share implies either an increase of new information and/or a reduction of noise in the market. Finally, a positive impact on the information leadership share suggests that speculation (hedging) amplifies the futures market's ability to impound new information.

# 3 Data

This study analyzes four agricultural commodity markets included in the CFTC's COT reports. The choice of these markets is motivated by the average trading volume of the markets' associated futures contracts. The analysis focuses on the two largest and two of the smallest contracts for which long-term CFTC data are available, i.e. the markets for corn, soybeans, cocoa and feeder cattle. Corn, soybeans and feeder cattle futures are traded at the Chicago Board of Trade (CBOT) which is now part of the CME group, while cocoa futures are traded at the Intercontinental Exchange (ICE).

Daily data for spot and futures prices and futures trading volume are for all of these commodities obtained from Thomson Reuters Datastream. The futures price time series consist of continuously compounded settlement price composites. Data on the open interest of the different trader types is taken from the CFTC's COT reports. Table 1 characterizes the selected futures contracts as well as the sample range for each market and provides descriptive statistics for prices, volume and the speculation variables. Data for corn and soybeans are available from 30 March 1982 until 29 June 2018, while data for cocoa and feeder cattle start from 1 November 1983 and 4 January 1993, respectively. All series end on 29 June 2018. However, for the later regression analysis, the data are limited by the availability of CFTC COT report data.

[ Table 1 about here. ]

As shown in table 1, corn has the highest average trading volume with over 132 thousand futures contracts traded per day. Feeder cattle has the lowest average daily trading volume with roughly 5 thousand futures contracts. The feeder cattle market features on average the highest values for total speculation ($S_t$) and excessive speculation ($W_t$), while cocoa features the lowest. Conversely, total hedging ($H_t$) is on average the highest in the cocoa market, while the feeder cattle market exhibits the lowest hedging levels.

[ Figure 1 about here. ]

Figure 1 displays the daily futures and spot prices for each of the four commodity contracts. In each panel, the futures price is indicated in black, while the spot price series is highlighted in gray. For all commodities, the graphs show a strong co-movement between spot prices and futures prices. Moreover, similar price developments can be observed for all commodities. Starting from moderate levels in the 1990s the commodities experience rapid price increases with subsequent declines around the mid 2000s.

[ Table 2 about here. ]

As indicated before, using a VEC model requires that the time series considered are cointegrated which in turn requires them to be integrated of order one. To test the price time series' order of integration we employ two different stationarity tests, namely the tests of Dickey & Fuller (1979, 1981) and Kwiatkowski et al. (1992), hereafter ADF and KPSS test. While the former assumes a unit root in the null hypothesis, the latter tests the null hypothesis that the time series are stationary. The test results for the log series as well as log first differences are reported in Table 2. The results of both tests show that the log series contain a unit root, while the log differences are stationary, indicating that the price time series are integrated of order one. In order to test for cointegration, Johansen's (1988, 1991) trace and eigenvalue tests are used. The results of the tests, also reported in Table 2, show for all commodities that spot and futures price are cointegrated as suggested by the cost-of-carry model.

# 4 Results and robustness

Figure 2 shows how the information leadership shares of the futures markets (in percent) evolve over time. When the share is above 50 % the futures market is dominant in the price discovery process. In this case, the futures market processes new information more quickly than the spot market. Alternating phases of price dominance by either market are observed for all commodities. However, the futures market generally dominates more often. In particular for the small markets, i.e. cocoa and feeder cattle, a particularly strong and persistent dominance of the futures market is visible as of 2010. Table 3 reports the futures markets' information leadership shares when calculated in a static version based on estimating equation (5) for the full sample. Moreover, the table lists the results of the average value of the information leadership shares in the time-varying case of the rolling window estimation, and also the percentage share of days where the futures market dominates. For all four commodities, it is observed that price discovery is generally dominated by the futures market.

[ Figure 2 and Table 3 about here. ]

Table 4 reports the results of regressing the different price discovery shares of the futures market on total speculation, $S_t$. For the component share, as outlined in Section 2.3, a positive coefficient estimate on $S_t$ indicates that speculators reduce the amount of noise in the market. For the information share a positive coefficient estimate implies that speculators reduce the level of noise and/or increase the ability of the futures market to process new information.

Lastly, the coefficient estimate of speculation in the regression with the information leadership share represents the net effect that speculators exert on the futures market's price discovery function.

The results show that in two markets, corn and cocoa, speculation reduces the level of the noise as indicated by the statistically significant positive coefficients of $S_t$ on $CS_t$. For all markets, the impact of speculation on the information share is always positive and higher than the impact on the component share. This implies that speculators contribute to the processing of new information in the market. The net effect of speculation on the relative speed at which the futures markets incorporates new information is given by the coefficient estimates of $S_t$ on $ILS_t$. The estimated parameters are positive and highly statistically significant for three of the four markets, namely corn, soybeans and cocoa. Only for feeder cattle total speculation has no statistically significant positive effect on price discovery.

[ Tables 4 to 6 about here. ]

Table 5 presents the results from repeating the earlier analysis but with a focus on excessive speculation as measured by Working' T, $W_t$. The findings are highly consistent with what has been reported above for total speculation. Again, the coefficient estimates of (excessive) speculation are positive and highly significant for the information leadership regressions. This implies that not only total speculation, but also excessive speculation, improves the futures markets ability to process new information. Note that with excessive speculation the coefficient estimates are also significantly positive for feeder cattle, which were insignificant when considering the total level of speculation.

The results regarding the influence of hedging on price discovery are displayed in Table 6. Except for cocoa, also hedgers significantly reduce the level of noise in the market. However, we observe for all four commodities that hedgers, unlike speculators, do not add to the markets' information processing ability. Intuitively, this can be explained by the idea that hedgers are interested in locking in futures prices, while speculators intend to profit from expected price changes. Hence, speculative traders have, compared to hedgers, a greater incentive to gather and interpret new information about fundamentals.

From an econometric point of view, one caveat of rolling window estimations is the arbitrary choice of the window size. Therefore, the earlier analysis is repeated using a variety of alternative window sizes, namely 200, 375 and 500 trading days. The former is arguably just large enough to have sufficient observations for estimating the VEC model. The latter two amount to one and a half and to two trading years, respectively. The results from repeating the baseline speculation regressions with the alternative window sizes are reported in Table 7.

[ Table 7 about here. ]

The findings are highly consistent across the different window sizes. Again, total as well as excessive speculation are generally found to significantly reduce the amount of noise and increase the speed of information processing in the futures market. Conversely, hedging activity is again found to have a negative impact on the price discovery function of the futures market, in the sense of quickly incorporating new information.

# 5 Conclusion

While numerous studies have investigated whether price discovery primarily occurs in the spot or the futures market, little work has been done to study the impact of futures speculation on the price discovery process. This paper has investigated how speculation affects the price discovery function of four agricultural commodity futures markets, namely the markets for corn, soybeans, cocoa and feeder cattle. As price discovery describes the timely and efficient incorporation of new information about fundamental values into prices, speculators can contribute to price discovery in two ways. First, they can reduce the amount of noise in the futures price. Second, speculators can add to the futures market's ability to process new information. For this purpose, a new measure of price discovery is used in this paper, namely the information leadership shares by Yan & Zivot (2010) and Putniņš (2013) which precisely distinguish between these two channels of impact. Alternative measures of speculative and hedging activity are then regressed on the time-varying information leadership shares of the futures markets.

Our findings suggest that speculation, both total and excessive, improves the futures markets' price discovery function in two ways. First, speculation reduces the level of noise incorporated in futures prices. Second, speculative activity increases the speed at which futures prices reflect new information about changes in market fundamentals. Conversely, hedgers are found to reduce the futures markets' ability to process new information. Therefore, speculators, in contrast to hedgers, appear to be more informed and trade with a stronger focus on potential changes of fundamental values. The results of this paper imply that efforts by regulators to curb futures speculation could have adverse consequences for the price discovery process in commodity markets.

# References

Baillie, R. T., Booth, G. G., Tse, Y. & Zabotina, T. (2002), 'Price discovery and common factor models', *Journal of Financial Markets* **5**(3), 309–321.

Boyd, N. E., Harris, J. H. & Li, B. (2018), 'An update on speculation and financialization in commodity markets', *Journal of Commodity Markets* **10**, 91–104.

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**(366), 427–431.

Dickey, D. A. & Fuller, W. A. (1981), 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica* **49**(4), 1057–1072.

Engel, C. (1996), 'The forward discount anomaly and the risk premium: A survey of recent evidence', *Journal of Empirical Finance* **3**(2), 123–192.

Engle, R. F. & Granger, C. W. J. (1987), 'Co-integration and error correction: Representation, estimation, and testing', *Econometrica* **55**(2), 251–276.

Fattouh, B., Kilian, L. & Mahadeva, L. (2013), 'The role of speculation in oil markets: What have we learned so far?', *The Energy Journal* **34**(3), 7–33.

Garbade, K. D. & Silber, W. L. (1983), 'Price movements and price discovery in futures and cash markets', *The Review of Economics and Statistics* **65**(2), 289–297.

Gonzalo, J. & Granger, C. (1995), 'Estimation of common long-memory components in cointegrated systems', *Journal of Business & Economic Statistics* **13**(1), 27–35.

Grossman, S. J. (1977), 'The existence of futures markets, noisy rational expectations and informational externalities', *The Review of Economic Studies* **44**(3), 431–449.

Hasbrouck, J. (1995), 'One security, many markets: Determining the contributions to price discovery', *The Journal of Finance* **50**(4), 1175–1199.

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control* **12**(2-3), 231–254.

Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica* **59**(6), 1551.

Kaldor, N. (1940), 'A symposium on the theory of the forward market: A note on the theory of the forward market', *The Review of Economic Studies* **7**(3), 196–201.

Keynes, J. M. (1923), *The Collected Writings of John Maynard Keynes - Vol. XII*, Cambridge University Press.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root', *Journal of Econometrics* **54**(1-3), 159–178.

Lehmann, B. N. (2002), 'Some desiderata for the measurement of price discovery across markets', *Journal of Financial Markets* **5**(3), 259–276.

Lien, D. & Shrestha, K. (2009), 'A new information share measure', *Journal of Futures Markets* **29**(4), 377–395.

MacKinnon, J. G. (1996), 'Numerical distribution functions for unit root and cointegration tests', *Journal of Applied Econometrics* **11**(6), 601–618.

MacKinnon, J. G., Haug, A. A. & Michelis, L. (1999), 'Numerical distribution functions of likelihood ratio tests for cointegration', *Journal of Applied Econometrics* **14**(5), 563–577.

Manera, M., Nicolini, M. & Vignati, I. (2016), 'Modelling futures price volatility in energy markets: Is there a role for financial speculation?', *Energy Economics* **53**, 220–229.

Putniņš, T. J. (2013), 'What do price discovery metrics really measure?', *Journal of Empirical Finance* **23**, 68–83.

Rutledge, D. J. S. (1977), 'Estimation of hedging and speculative positions in futures markets: An alternative approach', *Food Research Institute Studies* **16**(3), 205–211.

Sanders, D. R., Irwin, S. H. & Merrin, R. P. (2010), 'The adequacy of speculation in agricultural futures markets: Too much of a good thing?', *Applied Economic Perspectives and Policy* **32**(1), 77–94.

Schwarz, G. (1978), 'Estimating the dimension of a model', *The Annals of Statistics* **6**(2), 461–464.

Schwarz, T. V. & Szakmary, A. C. (1994), 'Price discovery in petroleum markets: Arbitrage, cointegration, and the time interval of analysis', *Journal of Futures Markets* **14**(2), 147–167.

Stock, J. H. & Watson, M. W. (1988), 'Testing for common trends', *Journal of the American Statistical Association* **83**(404), 1097–1107.

Working, H. (1960), 'Speculation on hedging markets', *Food Research Institute Studies* **1**(2), 185–220.

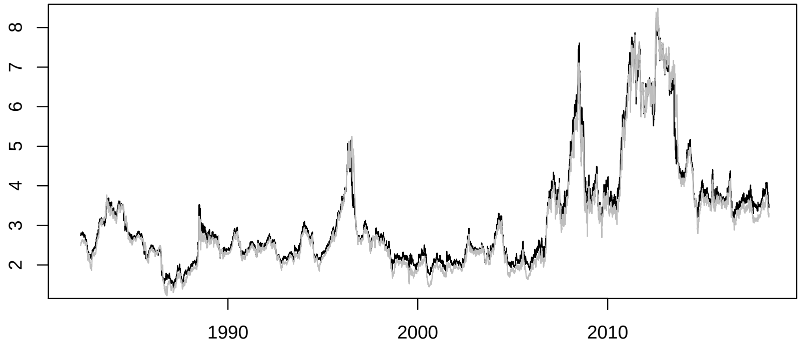
Yan, B. & Zivot, E. (2010), 'A structural analysis of price discovery measures', *Journal of Financial Markets* **13**(1), 1–19.

Yang, J., Bessler, D. A. & Leatham, D. J. (2001), 'Asset storability and price discovery in commodity futures markets: A new look', *Journal of Futures Markets* **21**(3), 279–300.

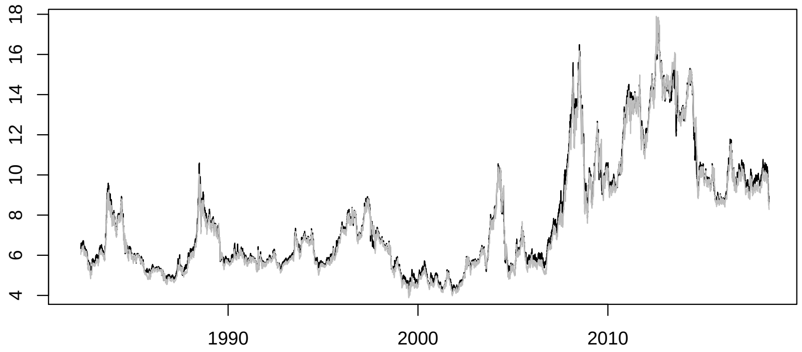
Yang, J. & Leatham, D. J. (1999), 'Price discovery in wheat futures markets', *Journal of Agricultural and Applied Economics* **31**(2), 359–370.

Zapata, H. O. & Fortenbery, T. R. (1996), 'Stochastic interest rates and price discovery in selected commodity markets', *Applied Economic Perspectives and Policy* **18**(4), 643–654.

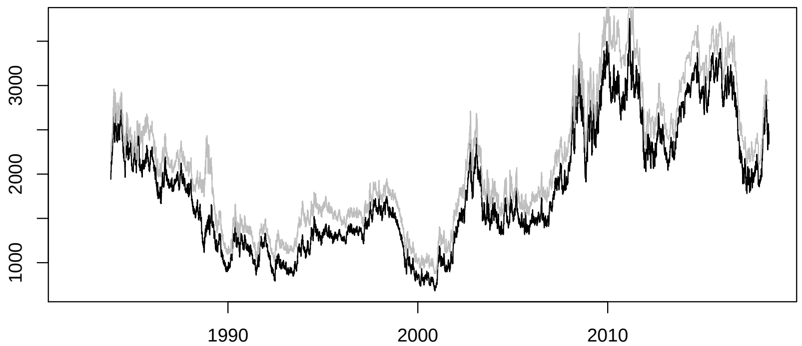
Zivot, E. (2000), 'Cointegration and forward and spot exchange rate regressions', *Journal of International Money and Finance* **19**(6), 785–812.



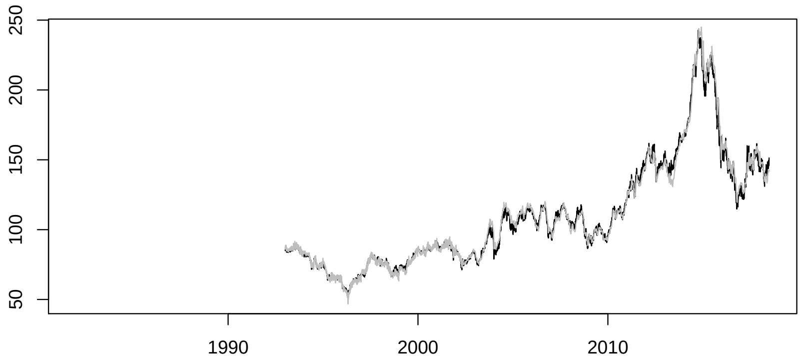

(a) Corn




(b) Soybeans



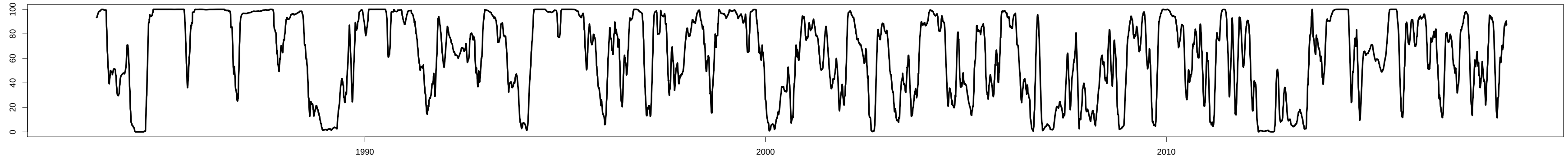

(c) Cocoa



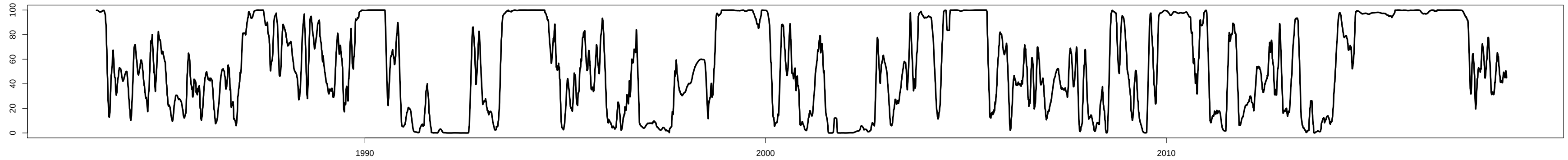

(d) Feeder Cattle

Figure 1: Futures and Spot Prices

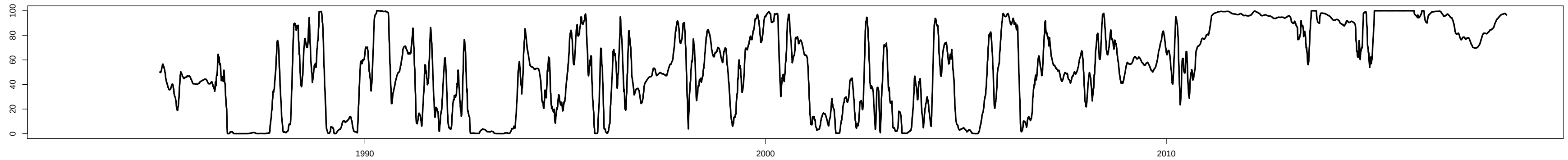
*Note:* Futures prices are depicted using black lines, spot prices are displayed in gray.




(a) Corn




(b) Soybeans




(c) Cocoa


(d) Feeder Cattle

Figure 2: Futures Market Information Leadership Shares (Monthly Moving Average)

Table 1: Descriptive Statistics

| Variable | Min. | Mean | Max. | St.dev. | Skew. | Kurt. |
|---|---|---|---|---|---|---|
| *CME No.2 Yellow Corn (5,000 bu), 30 Mar. 1982 - 29 June 2018* | | | | | | |
| Future | -0.25 | 0.00 | 0.10 | 0.02 | -0.65 | 16.89 |
| Spot | -0.12 | 0.00 | 0.11 | 0.02 | -0.29 | 7.43 |
| Volume | 0.19 | 132.76 | 1090.83 | 133.35 | 1.71 | 6.30 |
| Speculation | 19.99 | 27.11 | 39.98 | 3.66 | 0.93 | 3.82 |
| Working's T | 1.09 | 1.18 | 1.39 | 0.06 | 0.97 | 3.24 |
| Hedging | 48.74 | 61.38 | 72.97 | 5.46 | -0.40 | 2.37 |
| *CME No.1 Yellow Soybeans (5,000 bu), 30 Mar. 1982 - 29 June 2018* | | | | | | |
| Future | -0.14 | 0.00 | 0.08 | 0.01 | -0.58 | 8.42 |
| Spot | -0.17 | 0.00 | 0.08 | 0.02 | -0.68 | 9.25 |
| Volume | 0.10 | 86.78 | 804.24 | 83.82 | 1.65 | 6.54 |
| Speculation | 19.35 | 26.71 | 37.10 | 2.98 | 0.84 | 3.69 |
| Working's T | 1.07 | 1.16 | 1.35 | 0.05 | 0.76 | 3.57 |
| Hedging | 48.49 | 60.15 | 70.05 | 3.64 | -0.24 | 2.90 |
| *ICE Cocoa (10 mtr. t), 01 Nov. 1983 - 29 June 2018* | | | | | | |
| Future | -0.13 | 0.00 | 0.13 | 0.02 | -0.01 | 6.02 |
| Spot | -0.14 | 0.00 | 0.15 | 0.02 | -0.09 | 8.61 |
| Volume | 0.03 | 12.82 | 99.15 | 12.49 | 2.27 | 9.42 |
| Speculation | 7.18 | 22.19 | 37.56 | 5.66 | 0.03 | 2.53 |
| Working's T | 1.01 | 1.11 | 1.30 | 0.05 | 0.72 | 3.33 |
| Hedging | 50.57 | 70.48 | 88.81 | 9.28 | -0.16 | 1.85 |
| *CME Feeder Cattle (50,000 lbs), 04 Jan. 1993 - 29 June 2018* | | | | | | |
| Future | -0.06 | 0.00 | 0.06 | 0.01 | -0.11 | 5.92 |
| Spot | -0.09 | 0.00 | 0.09 | 0.01 | -0.03 | 9.13 |
| Volume | 0.00 | 4.99 | 29.67 | 3.82 | 1.73 | 6.59 |
| Speculation | 33.76 | 46.57 | 61.71 | 4.04 | 0.10 | 3.13 |
| Working's T | 1.27 | 1.50 | 1.89 | 0.09 | 0.47 | 4.04 |
| Hedging | 29.42 | 43.15 | 56.41 | 5.07 | 0.31 | 2.73 |

*Note:* The first two rows of each commodity panel report the summary statistics of futures and spot log returns. Trading volume is reported in thousands. The speculation, hedging variables and Working's T are computed based on CFTC data as outlined in Section 2.

Table 2: Unit Root and Cointegration Tests

| | | ADF test | | KPSS test | |
|---|---|---|---|---|---|
| Commodity | Variable | Logs | Log returns | Logs | Log returns |
| Corn | Future | -2.21 | -67.05$^{***}$ | 21.80$^{***}$ | 0.04 |
| | Spot | -2.18 | -94.69$^{***}$ | 31.80$^{***}$ | 0.04 |
| Soybeans | Future | -2.11 | -93.94$^{***}$ | 43.60$^{***}$ | 0.04 |
| | Spot | -2.07 | -97.52$^{***}$ | 42.90$^{***}$ | 0.04 |
| Cocoa | Future | -2.22 | -92.77$^{***}$ | 56.90$^{***}$ | 0.03 |
| | Spot | -1.94 | -71.12$^{***}$ | 38.30$^{***}$ | 0.03 |
| Feeder Cattle | Future | -1.08 | -54.22$^{***}$ | 8.53$^{***}$ | 0.06 |
| | Spot | -1.12 | -15.79$^{***}$ | 1.20$^{***}$ | 0.08 |

| | Johansen tests | | |
|---|---|---|---|
| Commodity | Null hypothesis | Trace | Eigenvalue |
| Corn | $H_0 : r = 0$ | 87.18$^{***}$ | 81.81$^{***}$ |
| | $H_0 : r = 1$ | 5.37 | 5.37 |
| Soybeans | $H_0 : r = 0$ | 178.83$^{***}$ | 174.12$^{***}$ |
| | $H_0 : r = 1$ | 4.70 | 4.70 |
| Cocoa | $H_0 : r = 0$ | 158.96$^{***}$ | 154.85$^{***}$ |
| | $H_0 : r = 1$ | 4.11 | 4.11 |
| Feeder Cattle | $H_0 : r = 0$ | 522.88$^{***}$ | 521.42$^{***}$ |
| | $H_0 : r = 1$ | 1.47 | 1.47 |

*Note:* The table reports for log prices and log returns the critical values of the ADF and the KPSS as computed by MacKinnon (1996) and Kwiatkowski et al. (1992), while the optimum lag-lengths have been determined using the Bayesian-Schwarz information criterion (1978). Moreover, the table displays for log prices the critical values of the Johansen (1988, 1991) trace and eigenvalue tests as computed by MacKinnon et al. (1999). The null hypothesis $r = 0$ tests for no cointegration, while the null hypothesis $r = 1$ tests for a cointegration rank of one. Statistical significance is reported as $^{*}$ for $p < 0.1$, $^{**}$ for $p < 0.05$ and $^{***}$ for $p < 0.01$.

Table 3: Futures Market Information Leadership Shares

| Commodity | Full Sample Static | Average Time-Varying | Relative Dominance |
|---|---|---|---|
| Corn | 87.98 | 62.93 | 64.39 |
| Soybeans | 55.01 | 53.77 | 51.77 |
| Cocoa | 57.71 | 54.03 | 55.64 |
| Feeder Cattle | 98.17 | 75.18 | 73.92 |

*Note:* The second column reports the futures market's information leadership share based on the full-sample VEC model. Values above 50 indicate that the futures market is dominant in the price discovery process. The third columns displays the average information leadership shares of the time-varying VEC model. For the same regressions, the rightmost column shows the percentage share of days where the futures market is dominant. Values greater than 50 indicate that the futures market dominates more often than the spot market.

Table 4: Results of Regressions with Total Speculation ($S_t$)

| | Corn | | | Soybeans | | | Cocoa | | | Feeder Cattle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS |
| $S_t$ | $1.562^{***}$ | $2.561^{***}$ | $2.091^{***}$ | $-0.751^{***}$ | $0.975^{***}$ | $2.948^{***}$ | $1.081^{***}$ | $1.653^{***}$ | $1.475^{***}$ | $-0.017$ | $0.090$ | $0.180$ |
| | (0.260) | (0.279) | (0.320) | (0.263) | (0.362) | (0.420) | (0.155) | (0.152) | (0.175) | (0.113) | (0.097) | (0.159) |
| $VOL$ | $-0.011^{*}$ | $-0.020^{***}$ | $-0.025^{***}$ | $-0.020^{***}$ | $0.011$ | $0.042^{***}$ | $0.621^{***}$ | $0.871^{***}$ | $0.810^{***}$ | $-0.833^{***}$ | $0.484^{***}$ | $3.113^{***}$ |
| | (0.006) | (0.007) | (0.008) | (0.008) | (0.011) | (0.012) | (0.061) | (0.059) | (0.068) | (0.122) | (0.105) | (0.173) |
| $BAS$ | $11.258^{**}$ | $22.899^{***}$ | $21.855^{***}$ | $5.449^{**}$ | $14.277^{***}$ | $13.926^{***}$ | $-0.068^{***}$ | $-0.072^{***}$ | $-0.021^{**}$ | $-0.557^{***}$ | $-0.394^{***}$ | $-0.348^{**}$ |
| | (4.847) | (5.194) | (5.955) | (2.750) | (3.776) | (4.388) | (0.008) | (0.008) | (0.009) | (0.125) | (0.107) | (0.176) |
| $Cons.$ | $6.320$ | $-17.612^{**}$ | $1.908$ | $79.212^{***}$ | $32.977^{***}$ | $-28.945^{**}$ | $0.464$ | $-16.360^{***}$ | $7.913^{**}$ | $81.313^{***}$ | $79.126^{***}$ | $50.597^{***}$ |
| | (7.158) | (7.670) | (8.795) | (7.258) | (9.967) | (11.583) | (3.158) | (3.096) | (3.568) | (5.334) | (4.570) | (7.523) |

*Note:* Standard errors in parentheses. Statistical significance is reported as $^{*}$ for $p < 0.1$, $^{**}$ for $p < 0.05$ and $^{***}$ for $p < 0.01$.

Table 5: Results of Regressions with Excessive Speculation ($W_t$)

| | Corn | | | Soybeans | | | Cocoa | | | Feeder Cattle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS |
| $W_t$ | $27.751^{*}$ | $68.690^{***}$ | $101.306^{***}$ | $42.402^{***}$ | $165.849^{***}$ | $204.366^{***}$ | $35.705^{**}$ | $110.087^{***}$ | $145.279^{***}$ | $-3.077$ | $20.473^{***}$ | $67.815^{***}$ |
| | (15.197) | (16.530) | (18.500) | (15.107) | (20.184) | (23.848) | (16.623) | (16.460) | (18.542) | (5.327) | (4.530) | (7.274) |
| $VOL$ | $-0.005$ | $-0.011$ | $-0.022^{***}$ | $-0.022^{***}$ | $0.011$ | $0.044^{***}$ | $0.732^{***}$ | $0.941^{***}$ | $0.789^{***}$ | $-0.806^{***}$ | $0.304^{***}$ | $2.518^{***}$ |
| | (0.006) | (0.007) | (0.008) | (0.008) | (0.010) | (0.012) | (0.063) | (0.063) | (0.071) | (0.131) | (0.111) | (0.179) |
| $BAS$ | $5.728$ | $15.422^{***}$ | $18.848^{***}$ | $10.409^{***}$ | $17.814^{***}$ | $11.203^{***}$ | $-0.086^{***}$ | $-0.095^{***}$ | $-0.039^{***}$ | $-0.555^{***}$ | $-0.408^{***}$ | $-0.402^{**}$ |
| | (4.870) | (5.297) | (5.928) | (2.625) | (3.506) | (4.143) | (0.008) | (0.008) | (0.009) | (0.125) | (0.106) | (0.171) |
| $Cons.$ | $15.473$ | $-29.752$ | $-60.951^{***}$ | $9.323$ | $-134.314^{***}$ | $-187.751^{***}$ | $-22.113$ | $-109.627^{***}$ | $-125.125^{***}$ | $84.983^{***}$ | $53.581^{***}$ | $-39.532^{***}$ |
| | (17.881) | (19.450) | (21.769) | (17.694) | (23.640) | (27.932) | (17.752) | (17.579) | (19.802) | (7.781) | (6.616) | (10.624) |

*Note:* Standard errors in parentheses. Statistical significance is reported as $^{*}$ for $p < 0.1$, $^{**}$ for $p < 0.05$ and $^{***}$ for $p < 0.01$.

Table 6: Results of Regressions with Total Hedging ($H_t$)

| | Corn | | | Soybeans | | | Cocoa | | | Feeder Cattle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS |
| $H_t$ | $0.399^{*}$ | $0.248$ | $-0.347$ | $0.996^{***}$ | $0.210$ | $-1.256^{***}$ | $-1.624^{***}$ | $-2.088^{***}$ | $-1.523^{***}$ | $0.485^{***}$ | $-0.612^{***}$ | $-2.493^{***}$ |
| | $(0.209)$ | $(0.229)$ | $(0.257)$ | $(0.210)$ | $(0.291)$ | $(0.343)$ | $(0.109)$ | $(0.104)$ | $(0.127)$ | $(0.106)$ | $(0.090)$ | $(0.134)$ |
| $VOL$ | $0.007$ | $0.001$ | $-0.020^{**}$ | $-0.009$ | $0.015$ | $0.031^{**}$ | $0.173^{***}$ | $0.338^{***}$ | $0.464^{***}$ | $-0.470^{***}$ | $0.024$ | $1.244^{***}$ |
| | $(0.008)$ | $(0.008)$ | $(0.010)$ | $(0.008)$ | $(0.011)$ | $(0.013)$ | $(0.067)$ | $(0.063)$ | $(0.078)$ | $(0.145)$ | $(0.123)$ | $(0.183)$ |
| $BAS$ | $2.543$ | $9.917^{*}$ | $13.020^{**}$ | $6.636^{***}$ | $9.932^{***}$ | $4.319$ | $-0.031^{***}$ | $-0.029^{***}$ | $0.005$ | $-0.501^{***}$ | $-0.459^{***}$ | $-0.625^{***}$ |
| | $(4.803)$ | $(5.265)$ | $(5.925)$ | $(2.548)$ | $(3.533)$ | $(4.160)$ | $(0.008)$ | $(0.008)$ | $(0.009)$ | $(0.124)$ | $(0.106)$ | $(0.157)$ |
| $Cons.$ | $21.859$ | $34.453^{**}$ | $80.367^{***}$ | $-2.411$ | $46.459^{**}$ | $128.198^{***}$ | $156.590^{***}$ | $188.109^{***}$ | $160.865^{***}$ | $57.691^{***}$ | $112.131^{***}$ | $176.276^{***}$ |
| | $(13.756)$ | $(15.077)$ | $(16.969)$ | $(12.985)$ | $(18.008)$ | $(21.204)$ | $(9.699)$ | $(9.216)$ | $(11.338)$ | $(5.064)$ | $(4.298)$ | $(6.391)$ |

*Note:* Standard errors in parentheses. Statistical significance is reported as $^{*}$ for $p < 0.1$, $^{**}$ for $p < 0.05$ and $^{***}$ for $p < 0.01$.

Table 7: Results of Robustness Exercise

| | Corn | | | Soybeans | | | Cocoa | | | Feeder Cattle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS | CS | IS | ILS |
| $S_t(200)$ | $1.900^{***}$ | $2.871^{***}$ | $1.615^{***}$ | $0.287$ | $2.115^{***}$ | $3.141^{***}$ | $0.778^{***}$ | $1.499^{***}$ | $1.755^{***}$ | $-0.138$ | $0.048$ | $0.194$ |
| | $(0.240)$ | $(0.286)$ | $(0.326)$ | $(0.264)$ | $(0.366)$ | $(0.430)$ | $(0.155)$ | $(0.156)$ | $(0.184)$ | $(0.118)$ | $(0.104)$ | $(0.165)$ |
| $S_t(375)$ | $2.297^{***}$ | $3.206^{***}$ | $1.319^{***}$ | $-0.267$ | $1.118^{***}$ | $2.559^{***}$ | $0.924^{***}$ | $1.540^{***}$ | $1.420^{***}$ | $0.310^{***}$ | $0.126$ | $-0.256$ |
| | $(0.247)$ | $(0.254)$ | $(0.318)$ | $(0.274)$ | $(0.328)$ | $(0.424)$ | $(0.150)$ | $(0.156)$ | $(0.179)$ | $(0.108)$ | $(0.092)$ | $(0.156)$ |
| $S_t(500)$ | $1.466^{***}$ | $2.106^{***}$ | $1.361^{***}$ | $-0.659^{**}$ | $0.028$ | $1.332^{***}$ | $0.873^{***}$ | $1.568^{***}$ | $1.696^{***}$ | $0.420^{***}$ | $0.067$ | $-0.362^{**}$ |
| | $(0.224)$ | $(0.237)$ | $(0.308)$ | $(0.277)$ | $(0.300)$ | $(0.429)$ | $(0.152)$ | $(0.155)$ | $(0.186)$ | $(0.103)$ | $(0.089)$ | $(0.152)$ |
| $W_t(200)$ | $44.666^{***}$ | $75.374^{***}$ | $58.247^{***}$ | $11.885$ | $164.111^{***}$ | $265.082^{***}$ | $5.000$ | $79.105^{***}$ | $142.803^{***}$ | $-10.272^{*}$ | $22.021^{***}$ | $73.259^{***}$ |
| | $(14.104)$ | $(17.040)$ | $(18.873)$ | $(15.167)$ | $(20.728)$ | $(23.903)$ | $(16.467)$ | $(16.822)$ | $(19.679)$ | $(5.529)$ | $(4.829)$ | $(7.484)$ |
| $W_t(375)$ | $72.632^{***}$ | $143.382^{***}$ | $121.009^{***}$ | $87.774^{***}$ | $174.527^{***}$ | $150.683^{***}$ | $63.926^{***}$ | $141.433^{***}$ | $142.046^{***}$ | $8.716^{*}$ | $18.345^{***}$ | $41.521^{***}$ |
| | $(14.616)$ | $(15.014)$ | $(18.052)$ | $(15.463)$ | $(18.118)$ | $(24.310)$ | $(15.887)$ | $(16.539)$ | $(18.954)$ | $(5.137)$ | $(4.346)$ | $(7.340)$ |
| $W_t(500)$ | $44.836^{***}$ | $85.886^{***}$ | $76.733^{***}$ | $30.178^{*}$ | $111.318^{***}$ | $134.642^{***}$ | $95.296^{***}$ | $173.392^{***}$ | $148.137^{***}$ | $14.516^{***}$ | $15.511^{***}$ | $36.857^{***}$ |
| | $(13.057)$ | $(13.848)$ | $(17.697)$ | $(15.878)$ | $(16.846)$ | $(24.387)$ | $(16.016)$ | $(16.295)$ | $(19.770)$ | $(4.902)$ | $(4.208)$ | $(7.136)$ |
| $H_t(200)$ | $-0.009$ | $0.114$ | $0.245$ | $1.078^{***}$ | $0.151$ | $-1.368^{***}$ | $-1.546^{***}$ | $-2.078^{***}$ | $-1.632^{***}$ | $0.627^{***}$ | $-0.697^{***}$ | $-2.816^{***}$ |
| | $(0.195)$ | $(0.236)$ | $(0.260)$ | $(0.210)$ | $(0.299)$ | $(0.351)$ | $(0.108)$ | $(0.106)$ | $(0.135)$ | $(0.110)$ | $(0.096)$ | $(0.136)$ |
| $H_t(375)$ | $0.231$ | $-0.540^{**}$ | $-1.524^{***}$ | $-0.077$ | $-0.614^{**}$ | $-0.731^{**}$ | $-1.515^{***}$ | $-2.036^{***}$ | $-1.610^{***}$ | $0.262^{**}$ | $-0.577^{***}$ | $-1.980^{***}$ |
| | $(0.203)$ | $(0.214)$ | $(0.249)$ | $(0.220)$ | $(0.264)$ | $(0.346)$ | $(0.105)$ | $(0.106)$ | $(0.129)$ | $(0.104)$ | $(0.087)$ | $(0.139)$ |
| $H_t(500)$ | $0.115$ | $-0.474^{**}$ | $-1.125^{***}$ | $-0.155$ | $-0.758^{***}$ | $-0.942^{***}$ | $-1.479^{***}$ | $-2.102^{***}$ | $-1.924^{***}$ | $0.089$ | $-0.554^{***}$ | $-1.930^{***}$ |
| | $(0.180)$ | $(0.193)$ | $(0.243)$ | $(0.223)$ | $(0.240)$ | $(0.345)$ | $(0.107)$ | $(0.105)$ | $(0.133)$ | $(0.101)$ | $(0.085)$ | $(0.137)$ |

*Note*: The displayed regressions include all the variables of the baseline regression of equation (16), however, only the coefficient estimates of the speculation-related variables are reported. Standard errors in parentheses. Statistical significance is reported as $^{*}$ for $p < 0.1$, $^{**}$ for $p < 0.05$ and $^{***}$ for $p < 0.01$.