Department of Economics
SOAS University of London

Working Paper Series

ISSN 1753 - 5816

Please cite this paper as:

van Huellen, Sophie. (2018), "Price Discovery in Commodity Futures and Cash Markets with Heterogenous Agents", SOAS Department of Economics Working Paper Series, No. 213, School of Oriental and African Studies, University of London.

No. 213

Price Discovery in Commodity Futures and Cash Markets with Heterogenous Agents

by

Sophie van Huellen

(November, 2018)

Department of Economics
School of Oriental and African Studies
University of London
WC1H 0XG
Phone: + 44 (0)20 7898 4730
Fax: 020 7898 4759
E-mail: economics@soas.ac.uk
http://www.soas.ac.uk/economics/

The **SOAS Department of Economics Working Paper Series** is published electronically by The School of Oriental and African Studies-University of London.

This and other papers can be downloaded without charge from:

SOAS Department of Economics Working Paper Series at
http://www.soas.ac.uk/economics/research/workingpapers/

Research Papers in Economics (RePEc) electronic library at
http://econpapers.repec.org/paper/

Design and layout: O.G. Dávila

Price Discovery in Commodity Futures and Cash Markets with Heterogenous Agents

Sophie van Huellen*

Abstract

The paper develops a price discovery model for commodity futures markets that accounts for two forms of limits to arbitrage caused by transaction costs and noise trader risk. Four market regimes are identified: (1) effective arbitrage, (2) transaction costs but no noise trader risk, (3) no transaction costs but noise trader risk and (4) both transaction costs and noise trader risk. It is shown that commodity prices are driven by both market fundamentals and speculative trader positions under the latter two regimes. Further, speculative effects spill over to the cash market under regime (3) but are confined to the futures market under regime (4). The model is empirically tested using data from six grain and soft commodity markets. While regime (4) is rare and short lived, regime (3) with some noise trader risk and varying elasticity of arbitrage prevails

Keywords: commodity futures, index investment, price discovery, speculation.

JEL classification: D84, G13, Q02, Q11

* SOAS University of London, Thornhaugh Street, Russell Square, London WC1H 0XG. Email: sv8@soas.ac.uk

I INTRODUCTION

The de-regulation of US American commodity futures markets through the Commodity Futures Modernization Act of 2000 had a lasting effect on the number and composition of traders in these markets. Commodity markets have seen an unprecedented inflow of liquidity, some of which has been linked to innovations in trading instruments, such as commodity indices. The Dodd-Frank Act of 2010 again altered the regulative environment of commodity futures markets. The liquidity inflow was visibly curbed with regards to over-the-counter transactions, but the general interest in commodity derivatives did not fade completely (BIS 2016). Both the volume and type of liquidity that entered commodity futures markets in the early 2000s triggered a debate around effects on price discovery mechanisms. Some welcomed the inflow as a trade facilitator while others feared potential price distortions.

With the intention to explore the effects of novel trading instruments on price discovery mechanisms in commodity markets, this paper amends existing price discovery models by the assumption of heterogeneous agents. Building on the work by Garbade and Silber (1983) and Figuerola-Ferretti and Gonzalo (2010), a pricing model with transaction cost induced limits to arbitrage between cash and futures markets is develop. The model is then extended by the assumption of heterogeneous agents in the form of informed arbitrage traders and uninformed noise traders, following the tradition of De Long et al. (1990). Under these assumptions, the relationship between cash and futures markets depends on the elasticity of arbitrage and the weight of uninformed traders in the market. Four different market regimes are identified: (1) effective arbitrage, (2) transaction cost induced limits to arbitrage and no noise trader risk, (3) effective arbitrage between cash and futures markets and noise trader risk and (4) limits to arbitrage due to both transaction cost and noise trader risk. It is shown that, firstly, price levels and changes are driven by both changes in market fundamentals and changes in noise trader positions under the latter two regimes. Noise trader effects spill over to the cash market if arbitrage between cash and futures markets is effective but is confined to the futures market if arbitrage between the two markets involves transaction costs. Secondly, if transaction costs hinder arbitrage, cointegration between spot and futures prices breaks down and the market basis follows a random walk process. If further, noise trader risk is present, the market basis reflects speculative price impulses.

The number of studies that theoretically explore the potential effects of heterogeneous traders on price discovery mechanisms in commodity futures markets is small yet growing.

Among notable contributions are studies by Basak und Pavlova (2016), Brunetti and Reiffen (2014), and Hamilton and Wu (2014, 2015). These authors build on the conjecture that the presence of noise traders, if acting systematic and not random, dilutes the market's information content, therefore feeding unwarranted speculative bubbles. Especially institutional investors that diversify into commodity indices were under suspicion for introducing systematic price pressure unrelated to market fundamentals.[1] This paper contributes to the growing literature, by development of a model that considers jointly the presence of institutional investors and limits to arbitrage. This contribution is particularly valuable as it formalises the link between index traders and an excessive market basis that was suggested for several US grain markets over recent years (Irwin, et al. 2011, Garcia, Irwin and Smith 2015, Van Huellen 2013).

The empirical validity of the model is tested for six grain and soft commodity markets traded at the Chicago Board of Trade (CBT): soft red winter wheat, number two yellow corn, number one yellow soybeans, cocoa, coffee 'c', and number two cotton. Arbitrage effectiveness is found to be low for corn and wheat markets, where non-convergence linked to bottlenecks in the delivery system occurred over recent years (Garcia et al. 2015). Trader position data is found to have a significant effect on the price discovery process in the short run and feedback effects to cash markets are strongest for markets for which the futures market is leading the price discovery process and arbitrage is effective. The coincidence of noise trader risk and perfectly inelastic arbitrage, corresponding to regime (4), is rare and short lived if it occurs. However, evidence for noise trader risk is found in all six markets with varying elasticities to arbitrage resulting in spill-over effects to cash markets.

The paper is structured in five parts. After this short introduction, the paper continues with the derivation of the price discovery model for commodity markets with heterogeneous agents and limits to arbitrage. The third part presents data and methodology for testing the model. The forth part reports empirical results and the fifth part concludes on the evidence.

II A HETEROGENOUS TRADER MODEL

The no-arbitrage condition between cash and futures prices is summarised in Eq. (1), with F_t being the futures price at time t with maturity date T, S_t being the spot price at t, r_t and w_t being continuously compounded risk free interest rate and storage costs over time $\tau = T - t$

[1] See Gilbert (2008), Nissanke (2012) and Irwin and Sanders (2012) for a summary of the debate and see Irwin (2013), Cheng and Xiong (2014) and Fattouh et al. (2013) for a summary of the empirical literature.

and y_t being the convenience yield; a utility based reward that accrues to the holder of inventories.

$$F_{t,T} = S_t e^{(r_t + w_t - y_t)\tau} \tag{1}$$

Taking logs of Eq. (1), with $\tau = 1$ for simplicity, yields Eq. (2):

$$f_{t,t+1} = s_t + r_t + w_t - y_t \tag{2}$$

Interest rate and storage costs less convenience yield can be summarised as net-carry costs c_t.

$$f_{t,t+1} = s_t + c_t \tag{3}$$

At maturity $\tau \to 0$ so that $f_{t,t} = s_t$ and market basis $b_t \equiv s_t - f_{t,t}$ is $b_t = 0$. Over a futures contract's life cycle, the size of the convenience yield in relation to interest and storage costs determines whether net-carry costs are positive or negative and hence whether the market is in contango or backwardation, that is, the futures exceeds the cash, or the cash exceeds the futures prices. While the extent of backwardation has not a limit, a contango has its maximum in the carry cost proper (Lautier 2005). Hence, a negative basis cannot exceed $r_t + w_t$, with $r_t, w_t \geq 0$, while a positive basis depends on the 'size' of the convenience yield y_t.

To understand market dynamics, it is useful to distinguish between two arbitrage mechanisms. *Spatial arbitrage* exploits deviations between spot and futures prices beyond the carry relationship. *Fundamental arbitrage* exploits deviations of price levels from the underlying fundamental value, but not necessarily relative prices. The former relies on arbitrage traders' willingness to trade the derivative and the physical commodity. The latter relies on the assumption that traders in both futures and cash markets make decisions based on the same information set so that a common (information) factor drives both markets. The common factor implies cointegration of the two price-series if $c_t \sim I(0)$; see Granger (1986) and Gonzalo and Granger (1995).

$$\begin{bmatrix} f_t \\ s_t \end{bmatrix} = \begin{bmatrix} C \\ 1 \end{bmatrix} \varphi_t + \begin{bmatrix} \tilde{f}_t \\ \tilde{s}_t \end{bmatrix} \tag{4}$$

With $(1, -C)$ being the cointegrating vector and φ_t the common factor, in our case the fundamental value. The reasonability of the assumption $c_t \sim I(0)$ is subject to debate. Figuerola-Ferretti and Gonzalo (2010) suggest that the convenience yield is non-stationary while Brenner and Kroner (1995) consider interest rates to be non-stationary. Both assumptions result in a rejection of $c_t \sim I(0)$. We will return to this this discussion in the following section.

The two arbitrage mechanisms, spatial and fundamental arbitrage, ensure that market dynamics obey Eq. (3) and Eq. (4) respectively. We, firstly, assume limits to spatial arbitrage in the form of transaction costs which result in a finite elasticity of arbitrage and hence deviations from Eq. (3). Secondly, we add the assumption of limits to fundamental arbitrage due to noise trader risk, so that traders do not evaluate a commodity and its derivative based solely on φ_t, which implies deviations from Eq. (4).

These two forms of limits to arbitrage are stepwise incorporated into a behavioural model for agents in the futures and spot market. Let there be N_s and N_f traders in the spot and futures market. $E_{i,t}$ and $E_{j,t}$ are the positions of the i[th] and j[th] trader in the spot and futures market respectively. The market clearing condition in the spot market is summarised in Eq. (5) with, H being elasticity of spatial arbitrage, A elasticity of demand, and $r_{i,t}$ the i[th] trader's reservation price with $A > 0, H > 0, i = 1, \dots, N_s$.[2]

$$\sum_{i=1}^{N_s} E_{i,t} = \sum_{i=1}^{N_s} \{E_{i,t} - A(s_t - r_{i,t})\} + H(f_t - s_t - c_t) \tag{5}$$

In correspondence to Eq. (5), market clearing condition in the futures markets are summarised in Eq. (6) with $r_{j,t}$ being the j[th] trader's reservation price and $j = 1, \dots, N_f$.

$$\sum_{j=1}^{N_f} E_{j,t} = \sum_{j=1}^{N_f} \{E_{j,t} - A(f_t - r_{j,t})\} - H(f_t - s_t - c_t) \tag{6}$$

[2] See Garbade and Silber (1983) for further details on the derivation of Eq. (5) and Eq. (6).

Eq. (5) and Eq. (6) can be solved for f_t and s_t with the reservation prices being expressed as the average reservation price of N_s and N_f traders in the respective markets so that $r_t^s = N_s^{-1}\sum_{i=1}^{N_s} r_{i,t}$ and $r_t^f = N_f^{-1}\sum_{j=1}^{N_f} r_{j,t}$.

$$s_t = \frac{\left[1+\left(\frac{H}{AN_f}\right)\right] r_t^s + \left(\frac{H}{AN_s}\right)(r_t^f - c_t)}{1+\left(\frac{H}{AN_f}\right)+\left(\frac{H}{AN_s}\right)}$$
$$f_t = \frac{\left[1+\left(\frac{H}{AN_s}\right)\right] r_t^f + \left(\frac{H}{AN_f}\right)(r_t^s + c_t)}{1+\left(\frac{H}{AN_f}\right)+\left(\frac{H}{AN_s}\right)} \quad (7)$$

Eq. (7) can be simplified into Eq. (8). According to Eq. (8) commodity futures and spot price are a function of the average reservation prices, net carry costs and the parameters a and b defined as: $a = \left(\frac{H}{AN_s}\right) / \left[1+\left(\frac{H}{AN_s}\right)+\left(\frac{H}{AN_f}\right)\right]$ and $b = \left(\frac{H}{AN_f}\right) / \left[1+\left(\frac{H}{AN_s}\right)+\left(\frac{H}{AN_f}\right)\right]$.

$$s_t = (1-a)r_t^s + ar_t^f - ac_t$$
$$f_t = br_t^s + (1-b)r_t^f + bc_t \quad (8)$$

Several implications under different degrees of spatial arbitrage can be derived from Eq. (8). These are summarised in Table 1.

Table 1. Spatial Arbitrage Regimes

	(1) Limits to spatial arbitrage	(2) Effective spatial arbitrage
H	$H \to 0$	$H \to \infty$
a, b	$\lim_{H\to 0} a = 0, \lim_{H\to 0} b = 0$	$\lim_{H\to \infty} a = \omega_f, \lim_{H\to \infty} b = \omega_s$
s_t	$s_t = r_t^s$	$s_t = \omega_s r_t^s + \omega_f(r_t^f - c_t)$
f_t	$f_t = r_t^f$	$f_t = \omega_f r_t^f + \omega_s(r_t^s + c_t)$
b_t	$b_t = r_t^s - r_t^f$	$b_t = -c_t$

Notes: Market basis $b_t \equiv s_t - f_t$. Weights $\omega_f = N_f/(N_s+N_f)$ and $\omega_s = N_s/(N_s+N_f)$ and hence $\omega_s = 1-\omega_f$.

Under limits to spatial arbitrage, the spot and futures price is equal to the respective average reservation price and the market basis is the difference of the two reservation prices. Under effective spatial arbitrage, cash and futures prices are the weighted averages of the reservation price of each market plus the net-carry costs and the market basis is equal to the net-carry costs in line with Eq. (3).

By proposing a data generating process for the average reservation prices, a dynamic relationship between cash and futures markets can be derived. Let two trader types exist in the market. Informed traders who base their positions on information about market fundamentals Φ and uninformed traders who follow an impulse Υ that is common among all uninformed traders but unrelated to market fundamentals. Only informed traders are present in the spot market, while both informed and uninformed traders are present in the futures market. This composition reflects the arrival of financial liquidity, channelled predominantly via indices, at commodity futures markets. Index positions are largely synchronised but unrelated to market fundamentals as their arrival is linked to global liquidity cycles and the roll-over of existing positions (Alam and Gilbert 2017). This justifies the assumption of a common impulse Υ.

With the division between informed and uninformed traders, the reservation prices of the i[th] and j[th] trader in the spot and futures market are $r^s_{i,t} = E_{i,t}[s_t|\Phi_i]$ and $r^f_{j,t} = (1-\beta)\{E_{j,t}[f_t|\Phi_j]\} + \beta\{E_{j,t}[f_t|\Upsilon_j]\}$, with $\Upsilon_j \in \Upsilon$ and $\Phi_j \in \Phi$ and with β being the probability that the j[th] trader is an uninformed trader. Hence, if $\sum E_{i,t}[s_t|\Phi_i] = \sum E_{j,t}[f_t|\Phi_i] = \varphi_t$ and $\sum E_{j,t}[f_t|\Upsilon_i] = \vartheta_t$, the evolution of the average reservation prices can be expressed as in Eq. (9a), with the white noise components $w_t \sim I(0)$ being the sum of idiosyncratic errors.[3]

$$\begin{aligned} r^s_t &= \varphi_t + w^s_t \\ r^f_t &= (1-\beta)\varphi_t + \beta\vartheta_t + w^f_t \end{aligned} \tag{9a}$$

By rewriting $\varphi_t = \varphi_{t-1} + \Delta\varphi_t$ and $\vartheta_t = \vartheta_{t-1} + \Delta\vartheta_t$ and assuming that each trader expects the last price to be the full information value so that s_{t-1} and f_{t-1} have to be the reservation prices immediately after the last clearing, it follows that $\varphi_{t-1} = \vartheta_{t-1} = f_{t-1}$ for

[3] Note that Eq. (9a) can be interpreted as factor decomposition, with common factor φ_t. Under limits to spatial arbitrage $H = 0$, Eq. (9a) is equivalent to Eq. (4) with $(1-\beta) = C$ and $s_t = r^s_t, f_t = r^f_t$ if assuming $\vartheta_t \sim I(0)$.

the futures market and $\varphi_{t-1} = s_{t-1}$ for the cash market. Therefore, Eq. (9a) can be transformed into Eq. (9b).

$$\begin{aligned} r_t^s &= s_{t-1} + \Delta\varphi_t + w_t^s \\ r_t^f &= f_{t-1} + (1-\beta)\Delta\varphi_t + \beta\Delta\vartheta_t + w_t^f \end{aligned} \tag{9b}$$

To proceed with the dynamics of the model, assumptions on the data generating process of net-carry costs are necessary. For simplicity, we assume that net-carry costs follow a white noise process with $\bar{c}$ being the mean net-carry costs.[4] This assumption will be eased in the following section to allow for variations in the data generating process of the different variables constituting net-carry costs.

$$c_t = \bar{c} + w_t^c \tag{9c}$$

Substituting Eq. (9) into Eq. (8) yields an error correction decomposition with the long run equilibrium error of Eq. (3) and short run shocks driven by traders' price impulses.

$$\begin{aligned} \Delta s_t &= a(f_{t-1} - s_{t-1} - \bar{c}) + (1-a\beta)\Delta\varphi_t + a\beta\Delta\vartheta_t + w_t^{s\prime} \\ \Delta f_t &= -b(f_{t-1} - s_{t-1} - \bar{c}) + (1-(1-b)\beta)\Delta\varphi_t + (1-b)\beta\Delta\vartheta_t + w_t^{f\prime} \end{aligned} \tag{10}$$

If uninformed traders induce systematic price impulses, so that $E[\Delta\vartheta_t] \neq 0$, the effect of these traders on the price discovery process depends on the elasticity of spatial arbitrage H and the weight of the futures market N_f relative to the cash market N_s. Note that these impulses by noise traders only enter prices in the short run but not the long run.

Table 2 distinguishes between four different arbitrage regimes: (1) effective spatial arbitrage and no noise trader risk, (2) limits to spatial arbitrage and no noise trader risk, (3) effective spatial arbitrage and noise trader risk and (4) limits to spatial arbitrage and noise trader risk.

[4] See for instance Low et al. (2002) who impose a similar assumption.

Table 2. Spatial Arbitrage Regimes with and without Noise Trader Risk

	(1) Effective spatial arbitrage, no noise trader risk	(2) Limits to spatial arbitrage, no noise trader risk	(3) Effective spatial arbitrage, noise trader risk	(4) Limits to spatial arbitrage, noise trader risk
H	$H \to \infty$	$H \to 0$	$H \to \infty$	$H \to 0$
β	$\beta = 0, \alpha = 1$	$\beta = 0, \alpha = 1$	$0 < \beta < 1$	$0 < \beta < 1$
a, b	$\lim_{H\to\infty} a = \omega_f,$ $\lim_{H\to\infty} b = \omega_s$	$\lim_{H\to 0} a = 0,$ $\lim_{H\to 0} b = 0$	$\lim_{H\to\infty} a = \omega_f,$ $\lim_{H\to\infty} b = \omega_s$	$\lim_{H\to 0} a = 0,$ $\lim_{H\to 0} b = 0$
s_t	$s_t = [\omega_s s_{t-1} + \omega_f(f_{t-1} - \bar{c})] + \Delta\varphi_t + w_t^s$	$s_t = s_{t-1} + \Delta\varphi_t + w_t^s$	$s_t = [\omega_s s_{t-1} + \omega_f(f_{t-1} - \bar{c})] + (1 - \omega_f\beta)\Delta\varphi_t + (\omega_f\beta)\Delta\vartheta_t + w_t^s$	$s_t = s_{t-1} + \Delta\varphi_t + w_t^s$
f_t	$f_t = [\omega_s(s_{t-1} + \bar{c}) + \omega_f f_{t-1}] + \Delta\varphi_t + w_t^f$	$f_t = f_{t-1} + \Delta\varphi_t + w_t^f$	$f_t = [\omega_s(s_{t-1} + \bar{c}) + \omega_f f_{t-1}] + (1 - \omega_f\beta)\Delta\varphi_t + (\omega_f\beta)\Delta\vartheta_t + w_t^f$	$f_t = f_{t-1} + (1 - \beta)\Delta\varphi_t + \beta\Delta\vartheta_t + w_t^f$
Δs_t	$\Delta s_t = \omega_f e_{t-1} + \Delta\varphi_t + w_t^s$	$\Delta s_t = \Delta\varphi_t + w_t^s$	$\Delta s_t = \omega_f e_{t-1} + (1 - \omega_f\beta)\Delta\varphi_t + (\omega_f\beta)\Delta\vartheta_t + w_t^s$	$\Delta s_t = \Delta\varphi_t + w_t^s$
Δf_t	$\Delta f_t = -\omega_s e_{t-1} + \Delta\varphi_t + w_t^f$	$\Delta f_t = \Delta\varphi_t + w_t^f$	$\Delta f_t = -\omega_s e_{t-1} + (1 - \omega_f\beta)\Delta\varphi_t + (\omega_f\beta)\Delta\vartheta_t + w_t^f$	$\Delta f_t = (1 - \beta)\Delta\varphi_t + \beta\Delta\vartheta_t + w_t^f$
b_t	$b_t = -\bar{c} + w_t^b$	$b_t = b_{t-1} + w_t^b$	$b_t = -\bar{c} + w_t^b$	$b_t = b_{t-1} + \beta(\Delta\varphi_t - \Delta\vartheta_t) + w_t^b$

Notes: Market basis $b_t \equiv s_t - f_t$. Weights $\omega_f = N_f/(N_s + N_f)$ and $\omega_s = N_s/(N_s + N_f)$ and hence $\omega_s = 1 - \omega_f$. Past error $e_{t-1} \equiv f_{t-1} - s_{t-1} - \bar{c}$.

If noise trader risk is present, as in (3) and (4), price levels and changes are driven by both changes in market fundamentals and changes in speculative demand. Speculative demand spills over to the cash market if spatial arbitrage is effective (3) but is restricted to the futures market if spatial arbitrage in limited (4). Further, if spatial arbitrage is limited, as in (2) and (4), the price levels are determined by the past settlement price of the respective markets resulting in the market basis being driven by its past values with a unit root process. If noise trader risk is present, as in (4), the market basis is further driven by speculative demand.

Several testable implications can be inferred from Table 2. If there are no limits to spatial arbitrage, the basis is driven solely by net-carry costs and futures and cash prices are cointegrated so that price dynamics can be expressed in an error correction decomposition. If limits to arbitrage are present, the basis follows a random walk and the error correction coefficient is insignificant due to the break in the cointegrating relationship between spot and futures markets. If, further, noise trader risk is present, changes in the futures price are linked to speculative demand while the spot price remains unaffected by this demand as long as the elasticity of arbitrage is low.

III DATA AND METHODOLOGY

A. Methodology

Before estimating Eq. (10), several considerations with regards to the variability of the weights a, b and β as well as the specification of the net-carry costs need to be addressed. Previous literature suggests that net-carry costs could be non-stationary due to a root component in the convenience yield or due to non-stationarity of interest rate or both (Figuerola-Ferretti and Gonzalo 2010, Brenner and Kroner 1995). If any component constituting the net-carry costs is non-stationary, the assumption $c_t \sim I(0)$ is invalid.

This question could be empirically settled if the elements of the net-carry costs were observed. Recall from Eq. (2) and Eq. (3), $c_t = r_t + w_t - y_t$. Of the three components, convenience yield is latent, while storage cost data is difficult to obtain. An elegant solution was proposed by Figuerola-Ferretti and Gonzalo (2010) who suggest modelling the convenience yield as a linear function of spot and futures prices. The model rests on the assumption of mean reversion of interest rate and storage costs, so that $r_t = \bar{r} + I(0)$, and $w_t = \bar{w} + I(0)$. With the data generating process for the convenience yield being $y_t = \gamma_1 s_t - \gamma_2 f_t$,

net-carry costs are specified as $c_t = \overline{rw} - \gamma_1 s_t + \gamma_2 f_t$. However, the mean reversion of interest rate and storage costs might not be supported empirically.

Another approach involves modelling storage costs and convenience yield as a function of level of inventory. The relationship between inventory and convenience yield is theoretically and empirically confirmed by Pindyck (2001), Bozic and Fortenbery (2011) and Pirrong (2011). Following this approach, $w_t(I_t) = w_0 + \gamma_2' I_t$ and $y_t(I_t) = y_0 - \gamma_2'' I_t$, so that $c_t = (w_0 + y_0) + \gamma_1 r_t + \gamma_2 I_t$, with convenience yield being inversely related to inventory and storage cost being positively related to inventory, which means c_t is increasing with I_t, with $\gamma_2 = \gamma_2' + \gamma_2''$. Both decompositions of net-carry costs can be incorporated into Eq. (10).

$$\begin{bmatrix}\Delta s_t \\ \Delta f_t\end{bmatrix} = \begin{bmatrix}a(1-\gamma_2) \\ -b(1-\gamma_2)\end{bmatrix}\begin{bmatrix}1 & -\frac{(1-\gamma_1)}{(1-\gamma_2)} & -\frac{\overline{rw}}{(1-\gamma_2)}\end{bmatrix}\begin{bmatrix}f_{t-1} \\ s_{t-1} \\ 1\end{bmatrix} + \begin{bmatrix}1-a\beta \\ 1-(1-b)\beta\end{bmatrix}\Delta\varphi_t + \begin{bmatrix}a\beta \\ (1-b)\beta\end{bmatrix}\Delta\vartheta_t + \begin{bmatrix}w_t^s \\ w_t^f\end{bmatrix} \tag{11a}$$

$$\begin{bmatrix}\Delta s_t \\ \Delta f_t\end{bmatrix} = \begin{bmatrix}a \\ -b\end{bmatrix}\begin{bmatrix}1 & -1 & -\gamma_1 & -\gamma_2 & (y_0 - w_0)\end{bmatrix}\begin{bmatrix}f_{t-1} \\ s_{t-1} \\ r_{t-1} \\ I_{t-1} \\ 1\end{bmatrix} + \begin{bmatrix}1-a\beta \\ 1-(1-b)\beta\end{bmatrix}\Delta\varphi_t + \begin{bmatrix}a\beta \\ (1-b)\beta\end{bmatrix}\Delta\vartheta_t + \begin{bmatrix}w_t^s \\ w_t^f\end{bmatrix} \tag{11b}$$

The ECMs in Eq. (11) reflect two different modelling choices for the net-carry costs. One problem remains in the estimation of Eq. (11a-b), which is the potential shift of the weights underlying a, b and β. Shifting weights imply parameter variance. A Markov regime switching model is suggested in order to account for the potential parameter invariance. Model specifications in Eq. (11) can be generalised into Eq. (12),

$$\Delta \boldsymbol{y_t} = \boldsymbol{\Pi y_{t-1}} + \sum_{j=0}^{r} \boldsymbol{\delta_j z_{t-j}} + \sum_{i=1}^{m} \boldsymbol{\pi_i} \Delta \boldsymbol{y_{t-i}} + \boldsymbol{w_t} \tag{12}$$

with $\boldsymbol{w_t} \sim IID(\mathbf{0}, \Omega)$, $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta'}$ and $\boldsymbol{\beta}$ being the cointegrating vector with $\boldsymbol{\beta' y_t} \sim \boldsymbol{I}(\mathbf{0})$, $\boldsymbol{y_t}$ being the vector of long run variables, that is spot and futures prices and net-carry costs and $\boldsymbol{z_t} \sim \boldsymbol{I}(\mathbf{0})$

being a set of additional explanatory variables such as changes in fundamental and speculative information.

Eq. (12) is estimated starting with modelling the long run cointegrating relationship then incorporating short run information shocks reflected in changing trader positions and finally accounting for regime changes reflected in changing coefficient estimates. Six estimation steps are conducted: 1) estimation of the cointegrating rank following Johansen (1988, 1991); 2) estimation of the reduced rank VECM; 3) test for long-run backwardation following Figuerola-Ferretti and Gonzalo (2010); 4) identification of the leading market in the price formation process by testing restrictions on $\alpha_{\perp}'$ following Gonzalo and Granger (1995); 5) test for the significance of information changes approximated by trader position indicators; and 6) considering weight changes by estimation of a Markov regime switching model on the market basis.

B. Data

Commodity prices are obtained from Thomson Reuters. Following Geman and Sarfo (2012), the futures price is constructed as a weighted average of all simultaneously traded contracts. Each contract is weighted by its share in total open interest. This way τ is held relatively constant over time which avoids cyclical contraction of carry variables due to maturity cycles.

The approximation of fundamental arbitrage and noise trader demand is challenging. We follow several empirical studies in using position data provided by the US Commodity Futures Trading Commission (CFTC), weekly Commitments of Traders Supplemental (CIT) Report—e.g. Irwin and Sanders (2010; 2012), Mayer (2012), Silvennoinen and Thorp (2013), and Singleton (2014). The report differentiates between long and short positions of commercial traders, index or portfolio insurance traders, non-commercial traders and small non-reporting traders. CFTC position data faces several shortcomings discussed elsewhere in the literature; e.g. Irwin and Sanders (2012). It is important to note here that the division of trader types is based on industry affiliation and not trading strategy. The information content of a trade cannot be observed, and we can only assume that commercial traders on average base their positioning on market fundamentals while institutional investors and portfolio insurance traders on average base their positioning on non-fundamental information. Against this conjecture, position change indicators are constructed for the commercial and index trader categories by dividing the respective net long positions (long minus short positions) by total open interest. The 3-

month LIBOR rate, obtained from Thomson Reuters, is used as an approximation for interest rate.

The paper focuses on grains and soft commodities which are storable. However, only for soft commodities data on inventories at exchange registered warehouses is available in weekly frequency. As for the price data, inventory data is obtained via Thomson Reuters. Weekly trader position data is available from the first week of 2006, while weekly inventory data for cocoa, coffee and cotton is available from 2010 week 32, 2011 week 2 and 2008 week 44 respectively. The samples used in this study end in 2016 week 29. Hence, Eq. (11b) is estimated, where at all possible, for a shorter time-period than Eq. (11a).

IV EMPIRICAL RESULTS

A. Model Estimation

Price variables and interest rate are found non-stationary and I(1), while inventory data is found stationary for cocoa and cotton, but not for coffee.

Table 3. Augmented Dickey Fuller Test to Establish Order of Integration

		x_t		Δx_t	
		c	c & t	c	c & t
Futures f_t	Wheat	-1.931 (0)	-1.776 (0)	-29.28** (0)	-29.30** (0)
	Corn	-1.527 (5)	-1.975 (5)	-12.87** (4)	-12.87** (4)
	Soy	-1.699 (2)	-2.490 (2)	-19.23** (1)	-19.22** (1)
	Cocoa	-2.039 (3)	-2.827 (3)	-19.36** (2)	-19.37** (2)
	Coffee	-1.267 (4)	-2.173 (4)	-16.43** (3)	-16.42** (3)
	Cotton	-2.033 (2)	-2.347 (2)	-19.25** (1)	-19.24** (1)
Cash s_t	Wheat	-2.263 (0)	-2.659 (0)	-30.81** (0)	-30.81** (0)
	Corn	-1.561 (5)	-1.999 (5)	-13.02** (4)	-13.02** (4)
	Soy	-1.582 (5)	-2.237 (5)	-13.99** (4)	-13.99** (4)
	Cocoa	-2.116 (3)	-2.670 (3)	-19.13** (2)	-19.15** (2)
	Coffee	-1.402 (0)	-2.791 (0)	-30.11** (0)	-30.10** (0)
	Cotton	-2.005 (0)	-2.331 (0)	-28.89** (0)	-28.87** (0)
Inventory I_t	Cocoa*	-4.634** (8)	-4.585** (8)	-6.168** (2)	-4.471** (5)
	Coffee†	-0.895 (4)	-0.329 (4)	-4.028** (3)	-4.589** (3)
	Cotton‡	-4.090** (2)	-4.291** (2)	-9.438** (1)	-9.438** (1)
Interest r_t	-	-1.431 (2)	-1.071 (2)	-8.104** (5)	-8.147** (5)

Notes: Augmented Dickey Fuller Test with critical values 5% = -2.87, 1% = -3.44 with constant and no trend included (c) and 5% = -3.42, 1% = -3.97 with constant and trend included (c & t). H_0: Time series has a unit root. * indicating 5% and ** indicating 1% significance level. Akaike Information Criteria (AIC) for choice of lag length (max 10 lags); lag length in (.) after the test statistic. * Available from 2010-32; † available from 2011-02; ‡ available from 2008-44. Remaining data spans from 2006-01 to 2016-29 in weekly frequency.

Having established the order of integration, the cointegrating rank is estimated with help of Johansen (1988) trace test. The test is conducted with correspondence to the cointegrating

vector $\boldsymbol{\beta}$ in Eq. (12). Interest rate is restricted to be exogeneous. Commodity futures and cash markets are found to be strongly cointegrated except for wheat where cointegration is weaker (Table 4a). If including inventory and interest rates, the finding of a single cointegrating vector does not change (Table 4b).

Table 4a. Trace Test s_t, f_t

	Wheat (10)	Corn (8)	Soy (7)	Cocoa (4)	Coffee (1)	Cotton (4)
H_0:	18.669	27.161	32.385	24.930	32.153	28.733
$r \leq 0$	[0.081]	[0.004] **	[0.000] **	[0.009] **	[0.000] **	[0.002] **
H_0:	7.3573	6.1318	5.8255	6.5158	4.6194	3.7936
$r \leq 1$	[0.111]	[0.187]	[0.212]	[0.159]	[0.339]	[0.456]

Table 4b. Trace Test s_t, f_t, I_t and r_t being treated as exogenous

	Cocoa (2)	Coffee (3)	Cotton (2)
$H_0: r \leq 0$	38.264	36.454	44.859
	[0.021]*	[0.035]*	[0.003]**
$H_0: r \leq 1$	14.417	9.7462	18.511
	[0.268]	[0.669]	[0.085]
$H_0: r \leq 2$	5.4704	1.2308	4.8546
	[0.244]	[0.905]	[0.310]

Notes: Lags in (.) selected by AIC. Constant included in the cointegration relationship as suggested by Eq. (11a-b). p-values in [.]. * indicating 5% and ** indicating 1% significance level. Interest rate r_t restricted to be exogenous in 4b.

The reduced rank model results for Eq. (11a) and Eq. (11b) are presented in Tables 5a-b. The system is restricted to one cointegrating relationship with interest rates being exogenous. In line with observations made for metal markets by Figuerola-Ferretti and Gonzalo (2010), cointegrating vectors deviate from the standard $(1, -1)$ with $\hat{\beta}_1 < -1$ for most markets, which indicates that backwardation prevails with a small unit root process in the net-carry variables. The deviation from unity is statistically significant for corn, soybeans, coffee and cotton. Interestingly, the unit root vanishes after addition of the remaining carry variables in confirmation of Eq. (11b). Dolatabadi et al (2015) account for the remaining unit root process by allowing for long memory in the equilibrium error by use of a fractional cointegration model. The findings here suggest that the long memory component is in fact due to model misspecification.

Table 5a. Cointegrating Vector $[s_t - \beta_0 - \beta_1 f_t]$

	$\hat{\beta}_0$	$\hat{\beta}_1$	$H_0: \beta_1 = 1^{\dagger}$
Wheat (10)	-0.037826	-0.97676	0.0079 (0.9292)
	[0.99616]	[0.15530]	
Corn (8)	1.6255	-1.2561	11.392 (0.0007)**
	[0.28783]	[0.046965]	
Soy (7)	0.37953	-1.0517	8.2781 (0.0040)**
	[0.092002]	[0.013186]	
Cocoa (4)	0.10118	-1.0280	0.4812 (0.4879)
	[0.25967]	[0.033137]	
Coffee (1)	0.71910	-1.1283	6.6131 (0.0101)*
	[0.20855]	[0.041531]	
Cotton (4)	0.64104	-1.1483	13.579 (0.0002)**
	[0.10474]	[0.024432]	

Table 5b. Cointegrating Vector $[s_t - \beta_0 - \beta_1 f_t - \beta_2 I_t - \beta_3 r_t]$

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$H_0: \beta_1 = 1^{\dagger}$
Cocoa (2)	0.63852	-1.1013	0.01476	0.00728	1.8827 (0.1700)
	[0.48585]	[0.059723]	[0.010561]	[0.02160]	
Coffee (3)	-0.031691	-1.0031	0.0084105	-0.05759	0.0213 (0.8839)
	[0.10056]	[0.017756]	[0.014929]	[0.01939]	
Cotton (2)	0.53517	-1.1139	0.21495	0.053922	2.5227 (0.1122)
	[0.19845]	[0.047992]	[0.049604]	[0.027963]	

Notes: Lag length in (.) determined by AIC (left column). Standard errors in [.]. † test statistic following a Chi-square distribution with one degree of freedom under the null; p-values in (.). * indicating 5% and ** indicating 1% significance level. Interest rate r_t restricted to be exogenous in 5b.

Table 6a reports tests on the $\alpha'_{\perp}$ derived from $\boldsymbol{\alpha}$ in Eq. (11a) to identify which market is leading the price formation process. Results suggest that the futures market is leading in most cases. The evidence is less conclusive for grains than for soft commodities and inconclusive wheat and corn markets. Tests on $\alpha'_{\perp}$ imply test for weak exogeneity. For Eq. (11a) the futures price is hence found weakly exogenous in most cases. For Eq. (11b) the lead-lag relationship cannot be clearly identified due to the potential endogeneity of the inventory market. Indeed, tests on $\boldsymbol{\alpha}$ suggest that both spot and futures price are weakly exogenous, except for the cocoa market, where futures prices and inventory are found weakly exogenous.

Table 6a. Test for leading market using $\alpha'_\perp$ for (11a)

	s_t	f_t	$\alpha'_\perp$	$H_0: \alpha'_\perp = (1,0)$ †	$H_0: \alpha'_\perp = (0,1)$ †
Wheat (10)	-0.03017	0.00469	(0.1344,	0.8014	0.0348
	[0.0203]	[0.0152]	0.8656)	(0.3707)	(0.8520)
Corn (8)	-0.01256	0.03304	(0.7245,	0.0988	0.9159
	[0.0343]	[0.0296]	0.2755)	(0.7533)	(0.3386)
Soy (7)	-0.18631	0.00700	(0.0362,	9.2955	0.0051
	[0.0845]	[0.0880]	0.9638)	(0.0023)**	(0.9431)
Cocoa (4)	-0.08025	0.05990	(0.4274,	2.1833	0.9770
	[0.0440]	[0.0492]	0.5726)	(0.1395)	(0.3230)
Coffee (1)	-0.08013	0.00537	(0.0629,	4.8895	0.0239
	[0.0331]	[0.0319]	0.9371)	(0.0270)*	(0.8772)
Cotton (4)	-0.12617	0.00737	(0.0552,	4.1316	0.0201
	[0.0575]	[0.0483]	0.9448)	(0.0421)*	(0.8874)

Notes: $\alpha'_\perp = (0,1)$ implies the futures market is leading the price discovery process and $\alpha'_\perp = (1,0)$ implies the cash market is leading the price discovery process. $\alpha'_\perp = (\alpha_2/(-\alpha_1 + \alpha_2), -\alpha_1/(-\alpha_1 + \alpha_2))$. Lag length in (.) determined by AIC (left column). Standard errors in [.]. † Test statistic following a Chi-squared distribution with two degrees of freedom under the null; p-values in (.). * indicating 5% and ** indicating 1% significance level.

Table 6b. Test for weak exogeneity using α for (11b)

	s_t	$H_0: \alpha_1 = 0$	f_t	$H_0: \alpha_2 = 0$	I_t	$H_0: \alpha_3 = 0$
Cocoa (2)	-0.12680	7.4133**	0.07724	1.8565	-0.21394	0.86845
	[0.0410]	(0.0065)	[0.0492]	(0.1730)	[0.1853]	(0.3514)
Coffee (3)	-0.39744	3.3839	-0.16917	0.67834	-0.17780	4.3135*
	[0.1854]	(0.0658)	[0.1809]	(0.4102)	[0.0775]	(0.0378)
Cotton (2)	-0.02673	0.36964	-0.02219	0.40164	-0.17040	12.134**
	[0.0367]	(0.5432)	[0.0315]	(0.5262)	[0.0341]	(0.0005)

Notes: Lag length in (.) determined by AIC (left column). Standard errors in [.]. Tests statistic on restrictions following Chi-squared distribution under the null hypothesis with one degree of freedom; p-values in (.). * indicating 5% and ** indicating 1% significance level.

Tables 7a-b report the coefficients on the information shocks. Estimations are based on single equation error correction models. Position data is found significant with regards to price impulses induced by commercial hedgers and impulses by index traders for most markets. This effect is more pronounced in the futures than the cash market as predicted under imperfect spatial arbitrage. For all but the wheat, corn and coffee market, position changes are significantly related to changes in the cash as well as futures prices. For soybeans and cotton, for which the futures market was found to be leading the price formation process, the effect of trader position changes on the cash market price are strongest.

Table 7a. Trader Position Indicators Eq. (11a)

	Cash market s_t		Futures market f_t	
	$\Delta\varphi_t$	$\Delta\vartheta_t$	$\Delta\varphi_t$	$\Delta\vartheta_t$
Wheat (4,4)	0.00683	-0.06675	-0.02499	0.04363
	[0.03917]	[0.06569]	[0.02854]	[0.04765]
Corn (7,7)	0.00781	0.00749	-0.03230	-0.01178
	[0.01765]	[0.02997]	[0.01583]**	[0.03152]
Soy (4,5)	0.05139	0.00414	-0.10055	-0.05267
	[0.02536]**	[0.03976]	[0.02488]***	[0.03279]
Cocoa (4,4)	-0.02984	0.02678	0.01354	-0.02122
	[0.01610]*	[0.04380]	[0.01617]	[0.04803]
Coffee (5,5)	-0.01496	-0.01607	-0.00503	0.00525
	[0.01562]	[0.02837]	[0.01626]	[0.02758]
Cotton (1,1)	0.01848	0.04216	-0.02377	-0.05767
	[0.01062]*	[0.03421]	[0.01015]**	[0.03045]*

Table 7b. Trader Position Indicators Eq. (11b)

	Cash market s_t		Futures market f_t	
	$\Delta\varphi_t$	$\Delta\vartheta_t$	$\Delta\varphi_t$	$\Delta\vartheta_t$
Cocoa (2,2)	-0.05302	0.14316	0.06598	-0.27159
	[0.02433]**	[0.07993]	[0.03029]**	[0.07590]***
Coffee (2,3)	0.01299	0.01306	-0.02455	0.00192
	[0.01610]	[0.02827]	[0.01720]	[0.02665]
Cotton (0,1)	0.03027	-0.01832	-0.03812	-0.04478
	[0.01425]**	[0.05073]	[0.01591]**	[0.04205]

Notes: Lag length in (.) chosen by general to specific modelling. Heteroscedasticity robust standard errors in [.]. * indicating 5% and ** indicating 1% significance level.

B. Test for Regime Shifts

A regime switching model is suggested for the estimation of the different regimes identified in Table 2 (Hamilton 2008). Under regimes (1–3) the first difference of the basis follows a white noise process, while under regime (4) the basis follows the difference of fundamental and speculative information weighted by the share of uninformed traders in the market:

$$\Delta b_t = \begin{cases} w_t^{b_1} & (1\text{–}3) \\ \beta\left(\Delta\varphi_t - \Delta\vartheta_t\right) + w_t^{b_2} & (4) \end{cases} \tag{13}$$

The regime-switching regression is specified in Eq. (14), with ζ_t being a random variable that assumes values $\zeta_t = 1$ or $\zeta_t = 2$ to differentiate between regimes (1–3) and regime (4). The probabilistic model of what causes the change from $\zeta_t = 1$ to $\zeta_t = 2$ is based on a two-state Markov chain with constant regime-switching probabilities.

$$\Delta b_t = \boldsymbol{\beta}'_{\zeta_t} \boldsymbol{z_t} + w_t^b \tag{14}$$

The set of explanatory variables, $\boldsymbol{z_t}$ is defined as in Eq. (12). Lagged values of Δb_t are added to control for autocorrelation in the residuals. Coefficients for lagged values are fixed while the remaining coefficient can vary across regimes. Regression results are summarized in Table 8.

Table 8. Markov Switching Regime Eq. (14)

	Wheat (4)	Corn (2)	Soy (4)	Cocoa (6)	Coffee (7)	Cotton (1)
			Regime 4			
c	-0.182***	0.032	0.088***	0.011	-0.057**	-0.121**
	[0.0013]	[0.3508]	[0.0028]	[0.2677]	[0.0102]	[0.0357]
$\Delta\varphi_t$	0.042	-0.367***	-0.426***	-0.145***	-0.618***	-0.566**
	[0.9108]	[0.0018]	[0.0046]	[0.0008]	[0.0000]	[0.0101]
$\Delta\vartheta_t$	0.479	-1.038***	-2.773***	-0.708***	-0.212	-0.155
	[0.2804]	[0.0004]	[0.0000]	[0.0000]	[0.3155]	[0.5518]
			Regime 1–3			
c	-0.008	0.005	0.003	-0.005	0.003	0.001
	[0.1989]	[0.4741]	[0.2062]	[0.2362]	[0.3489]	[0.9894]
$\Delta\varphi_t$	0.082**	0.018	0.011	0.016	-0.011	0.006
	[0.0261]	[0.3319]	[0.4042]	[0.4307]	[0.4424]	[0.8911]
$\Delta\vartheta_t$	0.121**	-0.009	-0.008	0.203***	-0.034	0.010
	[0.0271]	[0.8923]	[0.7187]	[0.0007]	[0.2363]	[0.9583]
			Transition Probabilities			
P11-C	1.075*	-0.108	-1.402	0.869	-1.166	-0.481
	[0.0953]	[0.7806]	[0.2162]	[0.1531]	[0.1965]	[0.6279]
P21-C	3.833***	-3.927***	-4.845***	1.184***	-3.171***	-4.252***
	[0.0000]	[0.0000]	[0.0000]	[0.0005]	[0.0000]	[0.0000]
i=1 j=1	0.254407	0.472982	0.197513	0.295400	0.237572	0.382117
i=1 j=2	0.745593	0.527018	0.802487	0.704600	0.762428	0.617883
i=2 j=1	0.021196	0.019330	0.007807	0.234298	0.040281	0.014038
i=2 j=2	0.978804	0.980670	0.992193	0.765702	0.959719	0.985962
Duration 1	1.341214	1.897470	1.246126	1.419245	1.311600	1.618430
Duration 2	47.17762	51.73336	128.0970	4.268066	24.82586	71.23351

Notes: Lag length in (.) decided by AIC. c is a constant; * for 10 %, ** for 5%, and *** for 1% significance level. Z-statistic with p-values in [.]; constant transition probabilities: P(i, k) = P(ζ(t) = k | ζ(t-1) = i); Duration is the constant expected duration for regime 1 and 2.

Regime (4), for which the coefficients for trader positions are found significant, is of much shorter duration, spanning one to two weeks only, than the prevailing regimes (1–3). This is expected as limits to arbitrage resulting in zero elasticity of arbitrage should be a rare incident if markets are functioning. Regime (4) shows non-zero and predominantly negative constants. This suggest that periods in which limits to arbitrage prevail are associated with a shrinking market basis whereby futures prices increase faster or more than spot prices. Coefficients for

trader positions are predominantly negative indicating that these are associated with the market becoming inverted. While these results support the model's predictions, they imply that the joint presence of limits to spatial arbitrage and noise trader risk, that is regime (4), is rare.

V CONCLUSION

The paper derives a model of price discovery in commodity markets that accounts for limits to arbitrage in two forms: limits to spatial arbitrage and limits to fundamental arbitrage. Limits to spatial arbitrage arise over transaction costs involved in the process of trading in commodity futures or cash markets. If transaction costs are significant, the elasticity of arbitrage is limited. Limits to fundamental arbitrage arise if systematic noise traders are present in the market which induce price deviations from market fundamentals in the short run.

It is shown that under these assumptions the relationship between cash and futures markets depends on the elasticity of arbitrage and the weight of uninformed traders in the market. Four different market regimes are identified: (1) effective arbitrage, (2) transaction cost induced limits to arbitrage and no noise trader risk, (3) effective arbitrage between cash and futures markets and noise trader risk and (4) limits to arbitrage due to both transaction cost and noise trader risk. It is shown that, firstly, price levels and changes are driven by both changes in market fundamentals and changes in speculative demand under the latter two regimes. Speculative demand spills over to the cash market if spatial arbitrage is effective but is restricted to the futures market if spatial arbitrage in limited. Secondly, if spatial arbitrage is limited the cointegrating relationship between cash and futures prices breaks and the market basis follows a random walk process. If further, noise trader risk is present, the market basis is also driven by speculative demand.

Empirical results largely confirm the model's predictions. Findings further suggest that regime (4) with zero elasticity of arbitrage is short lived and rare. A variation of regime (3) with imperfect elasticity of arbitrage prevails in most markets, which means that noise trader risk is more prevalent in the futures market than in the spot market.

The model and empirical evidence presented in this paper imply that the price discover process of commodity futures markets is affected by the presence of noise traders. The strength of the effect in both futures and cash markets depends on the relative market weight of noise traders in the futures markets and the elasticity of arbitrage. If the elasticity of arbitrage between cash and futures markets is high, price impulses by noise traders are likely to spill over to the

cash market. If the elasticity of arbitrage is low, noise trader risk is revealed in an excessive market basis.

BIBLIOGRAPHY

Alam, Md Rafayet, and Scott Gilbert. 2017. "Monetary policy shocks and the dynamics of agricultural commodity prices:evidence from structural and factor-augmented VAR analyses." *Agricultural Economics, 48 (1)* 15-27.

Basak, Suleyman, and Anna Pavlova. 2016. "A Model of Financialization of Commodities." *The Journal of Finance, 71(4)* 1511-1556.

BIS. 2016. *OTC Derivatives Statistics At End-June 2016.* Statistical Release, Basel, Switzerland: Bank for International Settlements, Monetary and Economic Department.

Bozic, Marin, and T. Randall Fortenbery. 2011. "Pricing Options on Commodity Futures: The Role of Weather and Storage." *Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011* .

Brenner, Robin J., and Kenneth F. Kroner. 1995. "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets." *The Journal of Financial and Quantitative Analysis, 30 (1)* 23-42.

Brunetti, Celso, and David Reiffen. 2014. "Commodity Index Trading and Hedging Costs." *Journal of Financial Markets, 21* (Federal Reserve Board) 153-180.

Cheng, Ing-Haw, and Wei Xiong. 2014. "The Financialization of Commodity Markets." *Annual Review of Financial Economics, 6* 419-41.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *The Journal of Political Economy, 98 (4)* 703-738.

Dolatabadi, Sepideh, Morten Ørregaard Nielsen, and Ke Xu. 2015. "A Fractionally Cointegrated VAR Analysis of Price Discovery in Commodity Futures Markets." *The Journal of Futures Markets, 35 (4)* 339–356.

Fattouh, Bassam, Lutz Kilian, and Lavan Mahadeva. 2013. "The Role of Speculation in Oil Markets: What Have We Learned So Far? ." *Energy Journal, 34* 7-33.

Figuerola-Ferretti, Isabel, and Jesus Gonzalo. 2010. "Modelling and measuring price discovery in commodity markets." *Journal of Econometrics, 158* 95–107.

Garbade, Kenneth D., and William L. Silber. 1983. "Price Movements and Price Discovery in Futures and Cash Markets." *The Review of Economics and Statistics, 65 (2)* 289-297.

Garcia, P., S. H. Irwin, and A. Smith. 2015. "Futures Market Failure?" *American Journal of Agricultural Economics, 97 (1)* 1-25.

Geman, Hélyette, and Seth Sarfo. 2012. "Seasonality in Cocoa Spot and Forward Markets: Empirical Evidence." *Journal of Agricultural Extension and Rural Development, Vol.4(8)* 164-180.

Gilbert, Christopher L. 2008. "Commodity Speculation and Commodity Investment." *Università degli Studi di Trento, Discussion Paper, No.20.*

Gonzalo, Jesus, and Clive Granger. 1995. "Estimation of Common Long-Memory Components in Cointegrated Systems." *Journal of Business & Economic Statistics, 13 (1)* 27-35.

Granger, Clive. 1986. "Developments in the Study of Cointegrated Economic Variables." *Oxford Bulletin of Economics and Statistics, 48 (3)* 213-228.

Hamilton, James D. 2008. "Regime Switching Models." In *The New Palgrave Dictionary of Economics, 2nd edition,* by N. Durlauf and L.E. Blume. Palgrave Macmillan.

Hamilton, James D., and Jing Cynthia Wu. 2015. "Effects Of Index-Fund Investing On Commodity Futures Prices." *International Economic Review, 56(1)* 187–205.

Hamilton, James D., and Jing Cynthia Wu. 2014. "Risk Premia In Crude Oil Futures Prices ." *Journal of International Money and Finance, 42* 9–37.

Irwin, S. H., P. Garcia, D. L. Good, and E. L. Kunda. 2011. "Spread and Non-Convergence in Chicago Board of Trade Corn, Soybeans, and Wheat Futures: Are Index Funds to Blame." *Applied Economics Perspective and Policy, 33 (1)* 116-142.

Irwin, Scott H. 2013. "Commodity Index Investment and Food Prices: Does the "Masters Hypothesis" Explain Recent Price Spikes?" *Agricultural Economics, 44 (supplement)* 29-41.

Irwin, Scott H., and Dwight R. Sanders. 2012. "Testing the Masters Hypothesis in commodity futures markets." *Energy Economics, 34* 256–269.

Johansen, Soren. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Autoregressive Models." *Econometrica, 59 (6)* 1551-1580.

Johansen, Soren. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control, 12 (2-3)* 231-254.

Lautier, Delphine. 2005. "Term Structure Models of Commodity Prices: A Review." *Journal of Alternative Investments, 8 (1)* 42-64.

Low, Aaron, Jayaram Muthuswamy, Sudipto Sakar, and Eric Terry. 2002. "Multiperiod hedging with futures contracts." *The Journal of Futures Markets, 22 (12)* 1179-1203.

Mayer, Jörg. 2012. "The Growing Financialisation of Commodity Markets: Divergences between Index Investors and Money Managers." *Journal of Development Studies, 48(6)* 751-767.

Nissanke, Machiko. 2012. "Commodity Market Linkages in the Global Financial Crisis: Excess Volatility and Development Impacts." *The Journal of Development Studies, 48 (6)* 732-750.

Pindyck, Robert S. 2001. "The Dynamics of Commodity Spot and Futures Markets: A Primer." *The Energy Journal, 22 (3)* 1-29.

Pirrong, Craig. 2011. "Stochastic Fundamental Volatility, Speculation, and Commodity Storage." In *Commodity Price Dynamics: A Structural Approach*, by Craig Pirrong, 109-130. Cambridge, UK: Cambridge University Press.

Silvennoinen, Annastiina , and Susan Thorp. 2013. "Financialization, Crisis and Commodity Correlation Dynamics." *Journal of International Financial Markets, Institutions and Money, 24* 42-65.

Singleton, Kenneth J. 2014. "Investor Flows and the 2008 Boom/Bust in Oil Prices." *Management Science, 60 (2)* 300-318.

van Huellen, Sophie (2013). 2013. "Price Non-Convergence in Commodities: A Case Study of the Wheat Conundrum ." *SOAS Department of Economics Working Paper Series, No. 185.*

ACKNOWLEDGMENT

An earlier version of this paper benefitted fom feedback from participants of the June 22, 2017 'Workshop Financialization of Commodity Markets' held in Bozen-Bolzano, Italy.