Department of Economics

SOAS University of London

Working Paper Series

ISSN 1753 - 5816

Please cite this paper as:

Van Huellen, Sophie. (2018), “Too Much of a Good Thing? Speculative Effects on Commodity Futures Curves”, SOAS Department of Economics Working Paper Series, No. 211, School of Oriental and African Studies, University of London.

No. 211

Too Much of a Good Thing? Speculative Effects on Commodity Futures Curves

by

Sophie van Huellen

(July, 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/

Design and layout: O.G. Dávila

Too Much of a Good Thing?
Speculative Effects on Commodity Futures Curves

Sophie van Huellen*

Abstract

The increasing inflow of institutional investors replicating broad based indices into commodity futures markets has been linked to excessive calendar spreads and anomalies in futures curves. At the same time, these investors have been welcomed as liquidity providers. This paper hypothesises that this apparent dissent can be reconciled by considering the relative size of index positions to hedging positions, rather than the presence of index traders alone. The hypothesis is tested empirically for three soft commodity markets: cocoa, coffee, and cotton. By use of factor decomposition, the paper shows empirically that (a) index and hedging positions have inverse and offsetting effects on futures curves, and (b) index positions, net of hedging positions, are associated with upward sloping and peaked futures curves and occasionally wave-like shapes linked to roll-effects. The paper concludes that index traders are welcomed liquidity providers but can become 'too much of a good thing' if exceeding hedgers' demand for counterparty.

Keywords: financialisation; futures curve; speculation; soft commodities; term structure.

JEL classification: G13, G14, Q02, Q14

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

Introduction

Commodity markets have witnessed an increasing popularity among investors since the early 2000s, when deregulation of, foremost US American, commodity futures markets opened new opportunities for institutional investors and fund managers alike. Consequently, the number of commodity contracts traded at derivative markets quadrupled within less than a decade. A vivid debate arose around possible implications of 'financialisation' of commodity futures markets for price discovery and risk management. Financialisation in this context is understood as the increasing presence of non-traditional investors in commodity derivative markets using novel trading strategies and instruments such as commodity indices. Early on, critics of this development suspected adverse implications for the efficacy of risk management and price discovery; see for instance Masters (2008) and US Senate Subcommittee (2009).

This paper contributes to the debate, while focusing on implications for commodity markets' term structure. Institutional investors that seek exposure to commodity markets via broad based indices are hypothesised to contribute to excessive calendar spreads and convergence failure due to their low frequency and long-only investment behaviour; see for instance Irwin et al. (2011), and Mou (2011). Index traders invest into commodity futures long positions to replicate indices, thereby putting upward price pressure on the futures contract they are investing in. With reference to risk premium and hedging pressure theories, several studies have formalised these claims by linking the risk premium to hedging and index positions or *hedging pressure* and *index pressure*; see for instance Basak and Pavlova (2016), Brunetti and Reiffen (2014) and Hamilton and Wu (2014).

Building on the model proposed by Brunetti and Reiffen (2014), this paper shows that index pressure effects depend on the relative size of index positions to hedging positions, rather than the presence of index traders alone. This conclusion has implications for the measurement of index and hedging pressure and consequently the empirical testing of the financialisation hypothesis. A factor decomposition method developed by Nelson and Siegel (1987) and extended by Diebold and Li (2006) is used for an empirical analysis of this hypothesis. Thereby, non-linear shapes of commodity futures curves are captured, which are ignored by studies focusing on the simple spread between contracts; e.g. Irwin et al. (2011) and Brunetti and Reiffen (2014). Findings presented in this paper strongly support the index pressure hypothesis as well as more conventional theories of storage and convenience yield. The strength of the index pressure effect is found to vary with the market weight of index traders in the respective markets analysed.

We focus on three soft commodity markets that are cash crops for low and middle-income countries along the equatorial belt: cocoa, coffee and cotton.[1] These commodities are among the most volatile due to, among other factors, sensitivity to climate and weather conditions. At the same time, the industries around these commodities are highly centralised with few buyers and intermediaries executing substantial market power (Staritz, et al. 2018, Bargawi and Newman 2017). Additional risk exposure arising from financialisation or other developments are likely to be passed on upstream to commodity producers. Further, these

[1] Cotton is an exception here with China, India and the US leading global production.

soft commodities are traded at the Commodity Market Exchange (CME) with relatively long observed future curves. These markets have been studies less in the context of financialisation and are less liquid than the more commonly studied grain and oil markets. However, given their low liquidity compared to more prominent commodity markets, financialisation has potentially strong effects on price formation and risk management, whereby their long futures curve provides for robust empirical analysis.

The remainder of the paper is structured as following. Section II discusses theories of intertemporal pricing and the hypothesised factors driving the shape of the futures curve. Section III presents a method for extracting factors that parsimoniously capture the dynamics of the futures curve. Section IV outlines statistical methods and data used in the empirical analysis. Section V presents empirically results and Section VI concludes on the findings obtained.

Commodity Term Structure Models

Theories of storage and convenience yield and theories of risk premium and hedging pressure build the foundation of intertemporal pricing theories. Theories of storage and convenience yield are derived from a simple no-arbitrage condition between spot and futures prices as summarised in Eq. (1),

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

whereby F_t^T is the futures price at time t with maturity date T, S_t is the spot price at t, r_t and w_t are continuously compounded risk-free interest rate and storage costs over time $\tau = T - t$ and y_t is the convenience yield. The latter is a utility-based reward that accrues to the holder of inventories. At maturity $\tau \rightarrow 0$ so that $F_T^T = S_T$ and the market basis $B_t \equiv S_t - F_t^T$ is zero, implying convergence. A complementary pricing theory to convenience yield models is the risk premium approach, according to which the futures price is a biased predictor of the expected spot price so that:

$$F_t^T = E_t[S_T]e^{-\rho_t \tau} \tag{2}$$

with ρ_t being the risk premium. The bias arises due to an insurance premium demanded by speculators who provide a risk management service to hedgers; see Lautier (2005) for an overview of different interpretations of the risk premium.

Hypothetically, $F_t^{T_i}$ for each maturity date T_i with $T_i = t + i, i \in \mathbb{N}$. The set of prices of all futures contracts with maturity dates T_i is referred to as the term structure. By plotting the set of prices at time t, the futures curve is revealed (Borovkova 2010). However, the futures curve cannot be observed fully as at any point in time there is only a finite set of simultaneously traded contracts with a finite number of maturity dates. If the futures curve was linear, the curve could be approximated by the spot rate plus the gradient estimated as the difference between two consecutive futures contracts. Taking the logs of (1) and (2), the gradient according to the previously revised theories is:

$$\frac{f_t^{T_j} - f_t^{T_i}}{\tau_{ij}} = r_t + w_t - y_t \tag{3a}$$

$$\frac{f_t^{T_j} - f_t^{T_i}}{\tau_{ij}} = \frac{1}{\tau_{ij}} E_t[\Delta s_{\tau_{ij}}] - \rho_t \tag{3b}$$

whereby $\tau_{ij} = T_j - T_i$, with $i < j$. In other words, the gradient is the calendar spread ($s_{\tau,t}$) adjusted for differences in maturity dates. However, the imposition of linearity is too restrictive, as convenience yield and risk premium might not evolve linearly over the futures curve. Indeed, non-liner shapes have been identified as common features such as peaked slopes (Litterman and Scheinkman 1991). The simple spread of consecutive futures contracts hence falls short of capturing the dynamics of the futures curve.

While carry costs (interest and storage costs) are observed, expected prices, convenience yield and risk premium are latent constructs. The convenience yield is commonly conceptualised as an inverse and non-linear function of inventory (Pindyck 2001, Bozic und Fortenbery 2011, Pirrong 2011). The risk premium found more varied and potentially complementary interpretations in the literature. Broadly, two sets of theories have emerged: (1) theories of asset-pricing, which assign a risk premium to (systematic) risk (Kaldor 1939, Dusak 1973) and (2) theories of hedging pressure, which incorporate market imperfections into multiple-period pricing models (Hirshleifer 1988, 1990, Chang 1985, Bessembinder 1992).

Hedging pressure models have recently been complemented by a literature identifying a potentially offsetting price pressure effect emitted by institutional investors that invest in broad based commodity indices for portfolio diversification; see Basak and Pavlova (2016), Brunetti and Reiffen (2014) and Hamilton and Wu (2014). These authors show that the risk premium, under the assumption of market frictions, is driven by demand for hedging and index positions; described as *hedging pressure* and *index pressure*. Brunetti and Reiffen (2014) derive a two-period pricing model in which hedgers and speculators maximise their utility over consecutive trading cycles. Hedgers and speculators invest in futures contracts which mature in period one and two respectively. Their utility functions only differ in cash market positions, or inventory (ψ), that are realised by hedgers after the last trading cycle. Index traders' positions (I) are exogeneous in the model.[2]

The utility function of hedgers and speculators is given by:

$$U[W_0 + X_1^2 \Delta F_1^2 + X_2^2 \Delta F_2^2 + X_1^1 \Delta F_1^1 + F_2^2 \psi_k] \tag{4}$$

with W_0 being initial wealth, X_t^T being the trader's position at time $t = \{0,1,2\}$ in the futures contract that matures at time $T = \{1,2\}$ and $\Delta F_t^T = F_t^T - F_{t-1}^T$ being the respective price changes of the futures contracts. $\psi_k = \{0, \psi\}$ for speculators and hedgers respectively. Each trader maximises utility with respect to total wealth consumed in $t = 2$, W_2.

[2] An assumption that might not be fully justified and hence empirically problematic as discussed below.

Brunetti and Reiffen (2014) assume a standard exponential utility function Eq. (5) for traders and speculators alike and impose price changes to be normally distributed, so that traders' utility functions depend solely on the mean and variance-covariance of price changes.

$$U(W) = A - \exp(-\alpha W_2) \tag{5}$$

Market clearing conditions in both futures contracts $T = \{1,2\}$ with index trader positions are specified in Eq. (6-7):

$$X_1^1(N_H + N_s) = -\gamma I_1 \tag{6}$$

$$X_1^2(N_H + N_s) = N_H\psi - (1-\gamma)I_1 \tag{7}$$

whereby N_H, N_s is the number of hedgers and speculators respectively. I_1 is total index positions at $t = 0$ invested in both the nearby and deferred maturity contract and γ is the percentage share invested in the nearby contract $T = 1$. Further, $X_{1,H}^1 = X_{1,S}^1$ following from Eq. (4) as ψ_k becomes binding only in the deferred futures contract and therefore $X_{1,H}^2 + \psi = X_{1,S}^2$. Both hedgers and speculators act as counterparty for index traders in Eq. (6) and the demand for counterparty positions in Eq. (7) depends on net-hedging demand, net of index trader positions.

By use of backward induction Brunetti and Reiffen (2014) solve for F_0^1 and F_0^2. The spread between the two consecutive futures contracts can then be derived as $S_0 \equiv F_0^2 - F_0^1$, in line with Eq. (3b) with $\tau_{ij} = 1, t = 0$ and $T = \{2,1\}$, whereby the risk premium is defined as:

$$\rho_o = \frac{\alpha}{N_H+N_s}[I_1[(\sigma_1^{12} - (\sigma_1^1)^2)(\gamma) + ((\sigma_1^2)^2 - \sigma_1^{12})(1-\gamma)] + I_2R[\sigma_{12}^2 + \sigma_{12}^{12} - (\sigma_1^2)^2 - \sigma_1^{12}] + \psi[((\sigma_1^2)^2 - \sigma_1^{12})(N_S) - (\sigma_{12}^2 + \sigma_{12}^{12})(N_H R) - ((\sigma_1^2)^2 + \sigma_1^{12})(N_S R)]] \tag{8}$$

with $R \equiv [N_S^1/(N_H^1 + N_S^1)]/[N_S^0/(N_H^0 + N_S^0)]$.[3] It is easy to see that if $\psi = 0$ and $I = 0$, the spread is an unbiased estimator of future price changes, i.e. $S_0 = E_0[\Delta S_{\tau_{12}}]$ with $\rho_0 = 0$. Further, without hedging demand, that is $\psi = 0$, hedgers and speculators are identical (as in contract one) and the risk premium would only persist in the presence of index traders.

If $N_H\psi = (1-\gamma)I_1$ or $N_S\psi = -(1-\gamma)I_1$ and $\gamma = 0$ so that $N_H\psi = I_1$ and $N_S\psi = -I_1$, which implies no hedging or index pressure in the $T = 1$ contract and exactly offsetting positions by hedgers and index traders in the $T = 2$ contract:

$$\begin{aligned}\rho_o &= \frac{\alpha}{N_H + N_s}[(\sigma_{12}^2 + \sigma_{12}^{12})(I_2 - I_1)R - ((\sigma_2^1)^2 + \sigma_1^{12})(I_2 - I_1)R] \\ &= \frac{\alpha \Delta I_1 R}{N_H + N_s}[\sigma_{12}^2 + \sigma_{12}^{12} - (\sigma_2^1)^2 - \sigma_1^{12}]\end{aligned} \tag{9}$$

[3] See Brunetti and Reiffen (2014) for details.

Eq. (9) implies that if no new index traders enter at $t = 1$ into the $T = 2$ contract so that $I_2 = I_1$, then the risk premium is zero as hedging and index pressure cancel out. If additional index traders enter the market, their effect varies with the share of speculative traders R and the degree of risk aversion α. Three testable hypotheses regarding the risk premium can be derived from a comparative static analysis of Eq. (8):

$$\frac{\partial S_0}{\partial \gamma} = \frac{\alpha I_1}{N_s + N_H}[2\sigma_1^{12} - (\sigma_1^1)^2 - (\sigma_1^2)^2] \tag{10a}$$

$$\frac{\partial S_0}{\partial I_1} = \frac{\alpha}{N_s + N_H}[(2\gamma - 1)\sigma_1^{12} - \gamma(\sigma_1^1)^2 + (1-\gamma)(\sigma_1^2)^2] \tag{10b}$$

(1) If $\gamma \downarrow$, that is index traders roll from contract $T = 1$ to $T = 2$, the spread increases.

(2) If $I \uparrow$ and $\gamma = 1$, that is all index traders are in contract $T = 1$, the spread decreases.

(3) If $I \uparrow$ and $\gamma = 0$, that is all index traders are in contract $T = 2$, the spread increases.

These hypotheses are formulated with respect to the spread between two consecutively traded contracts but not the futures curve. It is immediately clear, that the shape of the futures curve depends on the location of index traders and hedgers in the market. Hedging pressure is seasonal, while index pressure is likely to cause non-linearity due to peaks in the contracts they are most active in. By use of non-publicly available data, Brunetti and Reiffen (2014) show that index positions are mostly located in the medium-term maturities. Following the model's predictions, index traders should hence contribute to a positive slope overall with an inverted U-shape, peaking at medium-term maturities. Hedging positions should have the reverse effect. A method that unveils the latent futures curve is needed to test these hypotheses.[4]

Factor Decomposition

Principal component analysis (PCA) has been used as a data reduction technique to decompose the variation of future curves, foremost yield curves; e.g. Barber and Copper (2012). However, PCA decomposition suffers from weight inconstancy and, resulting from its non-parametric nature, interpretability of the extracted components is difficult as no structure is imposed that could be linked to theory. Factor decomposition methods, such as suggested by Nelson and Siegel (1987), adjust for this shortcoming by presupposes structure. Therefore, factors are designed so that they satisfy certain properties which are regarded desirable for interpretability.

The Nelson and Siegel (1987) decomposition rests on a set of differential equations that capture dynamic components of the yield curve and thereby generate the typical shapes of the curve at any point in time. Diebold and Li (2006) show that by altering the original decomposition the three extracted factors can be interpreted as level, slope and curvature in a similar manner as it has been suggested by Litterman and Scheinman (1991), who were first to assign meaning to yield curve components extracted by PCA.

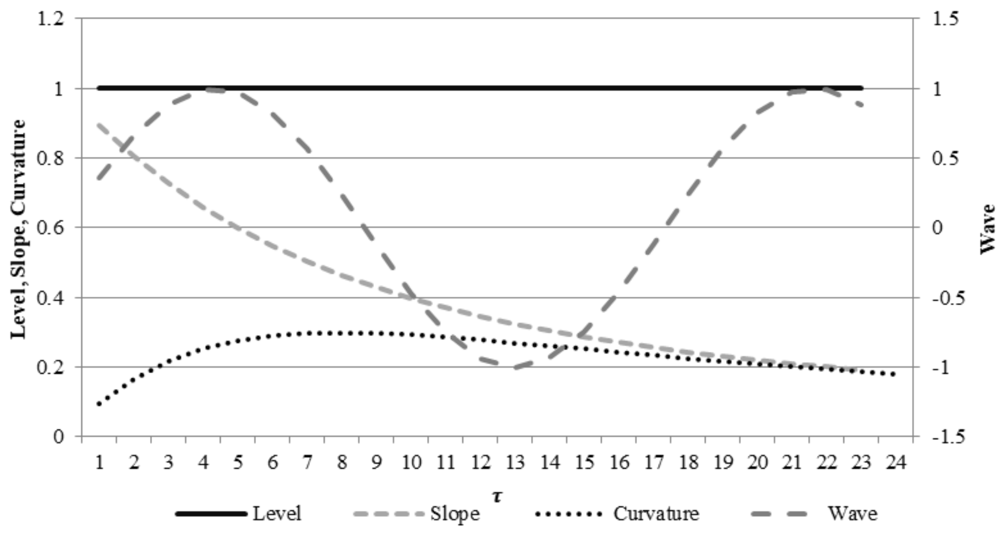
[4] Brunetti and Reiffen (2014) compensate for the shortcoming that spreads cannot capture non-linearities by use of non-publicly available data that provides information about the specific contracts in which traders are active.

The model by Diebold and Li (2006) takes the following form with level, slope and curvature $\{L, S, C\}$ jointly describing the futures curve.

$$f_t(\tau) = \beta_{t,L}L + \beta_{t,S}S(\tau) + \beta_{t,C}C(\tau) + \vartheta_t(\tau) \quad (11)$$

$f_t(\tau)$ is the price of the commodity futures at time t with time to maturity τ. The factor scores $\beta_{t,j}$, $j \in \{L, S, C\}$ can be extracted by firstly calculating the factor loadings for the slope $S(\tau) = \left(1 - e^{-\lambda\tau}/\lambda\tau\right)$ and the curvature $C(\tau) = S(\tau) - e^{-\lambda\tau}$ ($L = 1$ for all τ), for each contracts' maturity at each point in time and secondly using ordinary least square (OLS) to estimate $\beta_{t,j}$ at each t. It is easy to see that $f_t(\infty) = \beta_{t,L}$, $f_t(\infty) - f_t(0) = \beta_{t,S}$ and $\max C(\tau) = \lambda$. Therefore, the level factor has been termed the long run factor, while the slope factor is interpreted as the short run factor. The curvature is the medium run factor with λ governing at which month τ has a global maximum; see Figure 1.

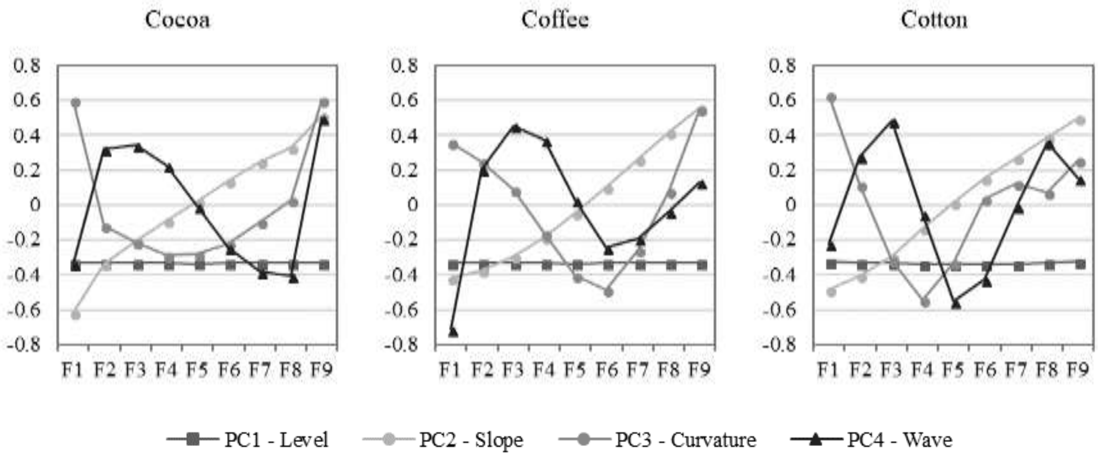
Figure 1. Nelson-Siegel Factor Loadings for L, S, C, W Eq. (12)



Note: λ is fixed at 0.22416, so that the curvature factor has its maximum at the 8th month, where the wave factor has its turning point.

Following Karstanje et al. (2017), we add a trigometric function. The addition is motivated by suspected seasonality in the data and the observation of a fourth wave-shaped component in the PCA decomposition for coffee, cocoa and cotton; see Figure 2. This finding supports the addition of a wave factor in the decomposition as in Eq. (12): $W(\tau) = \sin(\pi\lambda\tau)$. The functional form is simpler than in Karstanje et al. (2017) and offers a very good fit for the futures curves analysed.

Figure 2. First Four Principal Components.



Note: Contracts have been logged and standardised. Monthly data 2006M1-2017M7.

$$f_t(\tau) = \beta_{t,L}L + \beta_{t,S}S(\tau) + \beta_{t,C}C(\tau) + \beta_{t,W}W(\tau) + \vartheta_t(\tau) \quad (12)$$

While the shapes of the loadings are determined ex-ante, the rate of decay λ can be flexible. Hansen and Lunde (2013) find that allowing the decay factor to change over time does not improve the model fit, while Baruník and Malinská (2016) find that a time varying decay factor results in forecasting deterioration. Hence, we fix the decay factor, whereby the value of λ is found by grid search so that the best fit is reached. We further follow Karstanje et al. (2017) in re-centring[5] the loadings of the components for a clearer separation of the variation. The transformation is motivated by the observation that the slope and the curvature factor can become almost indistinguishable if the futures curve is flat.

Given the predetermined shapes of the factors, the obtained loadings are interpretable. The level reflects the overall price trend common across simultaneously traded contracts, while the slope indicates whether the market is normal or inverted. A positive factor value indicates a downward sloping futures curve, i.e., the contracts with longer maturities trade at a discount (inverted), and a negative value indicates an upward sloping futures curve, i.e., contracts with shorter maturity trade at a discount (normal). A positive value for the curvature coefficient indicates a convex and a negative a concave curve. A positive value for the wave component signals an N-shaped futures curve and a negative value signals an inverted N-shaped curve.

Recalling the model from Section II, we expect to find the following empirical relationships; hypotheses H(1) to H(3):

(1) Index pressure to be associated with
   a. a normal market, i.e. an upward sloping futures curve.
   b. a concave futures curve, i.e. a peaked curve
(2) Hedging pressure to be associated with
   a. an inverted market, i.e. a downward sloping futures curve.

[5] Re-centring is achieved by $S'(\tau) = S(\tau) - S(1)$, $C'(\tau) = C(\tau) - C(1)$ and $W'(\tau) = W(\tau) - W(1)$.

    b. a convex futures curve, i.e. a u-shaped futures curve.

(3) Inventory to be associated with
    a. an inverted market through the convenience yield.
    b. wave shapes, capturing seasonal cycles.

These hypotheses are corroborated by findings presented in the existing empirical literature. Etienne and Mattos (2016) apply the Nelson-Siegel factor decomposition to empirically test theories of storage and convenience yield. Karstanje et al. (2017) test the excessive comovement hypothesis, also linked to financialisation, on the extracted factors. Heidorn et al. (2015), like this study, link the extracted factors to trader positions in the oil market and find that index positions influence the slope and curvature factor but not the level. Their finding aligns with the argument made by Brunetti and Reiffen (2014) that price levels are dominantly driven by demand and supply fundamentals as well as noise. It is therefore empirically challenging to test for price pressure effect on the price level, unless all fundamentals are controlled for. However, since changes in fundamentals affect simultaneously traded contracts in a similar way, calendar spreads or futures curves provide more fertile ground to investigate price pressure effects as fundamental effects and noise driving the price level cancel out.

Method and Data

The empirical analysis is conducted in three stages: (1) hypothesis testing based on a simple spread analysis ignoring non-linearity, (2) factor decomposition and evaluation, (3) hypothesis testing based on factor decomposition. An unrestricted ECM specification is chosen for both hypothesis tests in (1) and (3). The model choice is motivated by the observation that extracted slope and curvature factors are trended. Unrestricted ECMs are robust in the presence of dominant trends and flexible as they incorporate both levels and differences, thereby incorporating potential long run equilibrium relations.

$$\Delta y_t = \alpha_0 + \alpha(L)\Delta y_{t-1} + \Omega(L)\Delta Z_t + \phi[y - \gamma' Z]_{t-1} + u_t \qquad (13)$$

$Z_t = [r_t, \psi_t, \psi_t^2, \rho_{H,t}, \rho_{I,t}]'$ is the set of explanatory variables informed by the two theoretical strands outlined in Eq. (3), whereby storage costs and convenience yield are modelled as functions of inventory ψ_t and the risk premium is modelled as hedging pressure $\rho_{H,t}$ and index pressure $\rho_{I,t}$ following Eq. (9). Ω is a coefficient matrix and γ' the cointegrating vector which can be recovered if ϕ is significant. The set of dependent variables is given by $y_t = \{s_t, \beta_{j,t}\}$, with $j = \{S, C, W\}$ from Eq. (3) and Eq. (12). The level factor is not considered for the analysis, as the derived hypotheses H(1) to H(3) are formulated with respect to extracted slope, curvature, and wave factors only. The appropriate lag length of Eq. (13) is found by general-to-specific modelling; see Campos et al. (2005).

As elaborated by Eq. (9), hedging and index pressure are offsetting forces. Index traders are liquidity providers if hedging pressure prevails and are liquidity consumers otherwise. Hence, indicators are created that capture the net effect of these price pressures. Trader position data is obtained from the Commodity Futures Trading Commission (CFTC) Commodity Index Supplement (CIT) which reports weekly open interest (OI) data disaggregated by long and short positions and trader type: non-commercial traders, commercial traders (com), index traders (ind) and small non-reporting traders. Hedging pressure is defined as the commercial net long positions (short hedging positions after internal netting) that are not covered by index net long positions (long positions after internal netting) as a share of total open interest. Hence, if index net positions exceed commercial net positions, we have index pressure and hedging pressure in the reverse case.

$$\rho_{H,t} = \begin{cases} \frac{|com_{nl}+ind_{nl}|}{OI_{total}} & if\ |com_{nl}| > ind_{nl} \\ 0 & if\ |com_{nl}| < ind_{nl} \end{cases} \tag{14a}$$

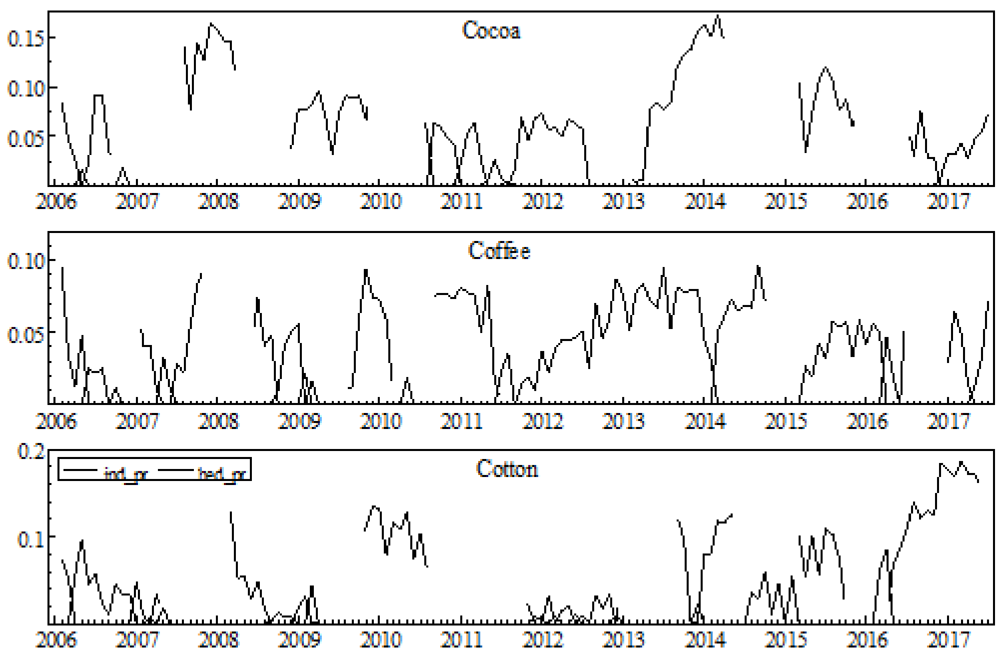
$$\rho_{I,t} = \begin{cases} \frac{|com_{nl}+ind_{nl}|}{OI_{total}} & if\ |com_{nl}| < ind_{nl} \\ 0 & if\ |com_{nl}| > ind_{nl} \end{cases} \tag{14b}$$

CIT data is available weekly from 2006, while inventory data is only available in monthly frequency. Hence, Eq. (13) is estimated using monthly data from January 2006M1 to 2017M6. The first twelve months 2006M1-2006M12 are used to evaluate appropriate lag structure. Price data for the different futures contracts and the 'risk free' interest rate, approximated by the 3-months Libor rate, is obtained from Thomson Reuters Datastream. Inventory data for cocoa and coffee is obtained from the CME Registry. Data for cotton inventory is obtained from the US Department of Agriculture (USDA).

Empirical Results

Index pressure is most pronounced in the coffee market and almost absent in the cotton market where short hedging positions outweigh long index positions at most times; see Figure 3. The average index pressure effect is hence expected to be smallest for cotton and largest for coffee.

Figure 3. Index $\rho_{I,t}$ and Hedging $\rho_{H,t}$ Pressure Eq. (14)



For estimation of Eq. (13) we use the calendar spread between the 8th nearest and 2nd nearest maturity date; 8-2 spread. At maturity of the nearest to maturity contract, the series are rolled over. We also experiment with the 5-2 and 8-5 spread, but results remain unchanged, which is unsurprising given the similarity of the different spreads; see Appendix Figure A1. Estimation results are reported in Table 1. We find both index and hedging pressure coefficients to be significant across markets, except for index pressure in cotton. As expected, index pressure is smallest for cotton and largest for coffee. Index pressure is associated with an increase in the spread while hedging pressure is associated with a decrease in the spread. We also find support for the convenience yield hypothesis for cocoa and cotton, while the coffee spread is dominated by the risk premium.

Table 1. Estimation of Eq. (13) with $y_t = s_t$.

	Constant	y_t	ψ_t	ψ_t^2	r_t	$\rho_{I,t}$	$\rho_{H,t}$
				Short-run			
	α_0	$(L)\alpha$	$(L)\omega_\psi$	$(L)\omega_{\psi^2}$	$(L)\omega_r$	$(L)\omega_{\rho_I}$	$(L)\omega_{\rho_H}$
Cocoa	0.0009	-0.2625** [1]	0.2154** [3]	-0.0738* [3]	0.1896* [2]	0.2396*	-0.2278**
	(0.002)	(0.078)	(0.073)	(0.029)	(0.088)	(0.108)	(0.042)
Coffee	0.0005	-0.3056** [1]	0.1198	-0.0264	0.1574	0.3314**	-0.3346**
	(0.002)	(0.079)	(0.089)	(0.059)	(0.174)	(0.102)	(0.062)
Cotton	-0.0006	0.2212* [1]	0.1413** [3]	-0.0434** [3]	-0.2904	0.1281	-0.5019**
	(0.004)	(0.085)	(0.052)	(0.015)	(0.5381)	(0.262)	(0.125)
				Long-run			
		ϕ	γ_ψ	γ_{ψ^2}	γ_r	γ_{ρ_I}	γ_{ρ_H}
Cocoa		-0.1850**	-	-	0.1358*	1.0363**	-
		(0.061)			[5.537][a]	[8.986][a]	

Note: Model estimated for 2007M1-2017M7 to allow enough room for appropriate lag selection. (.) Standard errors. [.] Lag length selected. * significant at the 5% level; ** significant at the 1% level. [a] Long-run coefficients are recovered as the ratio between $\hat{\gamma}$ and the absolute value of $\hat{\phi}$. [.] is the Wald test statistic of insignificance of the long run coefficient following a Chi Square distribution under the Null.

As mentioned previously, calendar spreads are unable to capture dynamics of futures curves fully which motivates the factor decomposition. Figure 4 depicts the extracted factor scores. The curvature factor is more volatile than the slope factor and volatility is highest for the wave factor. The autocorrelation functions of the three factors support this further with inertia for the curvature factor being less than for the slope factor. The wave factor exhibits seasonal patterns following harvest cycles; annual for coffee and cocoa and bi-annual for cotton due to the more geographically disperse production.

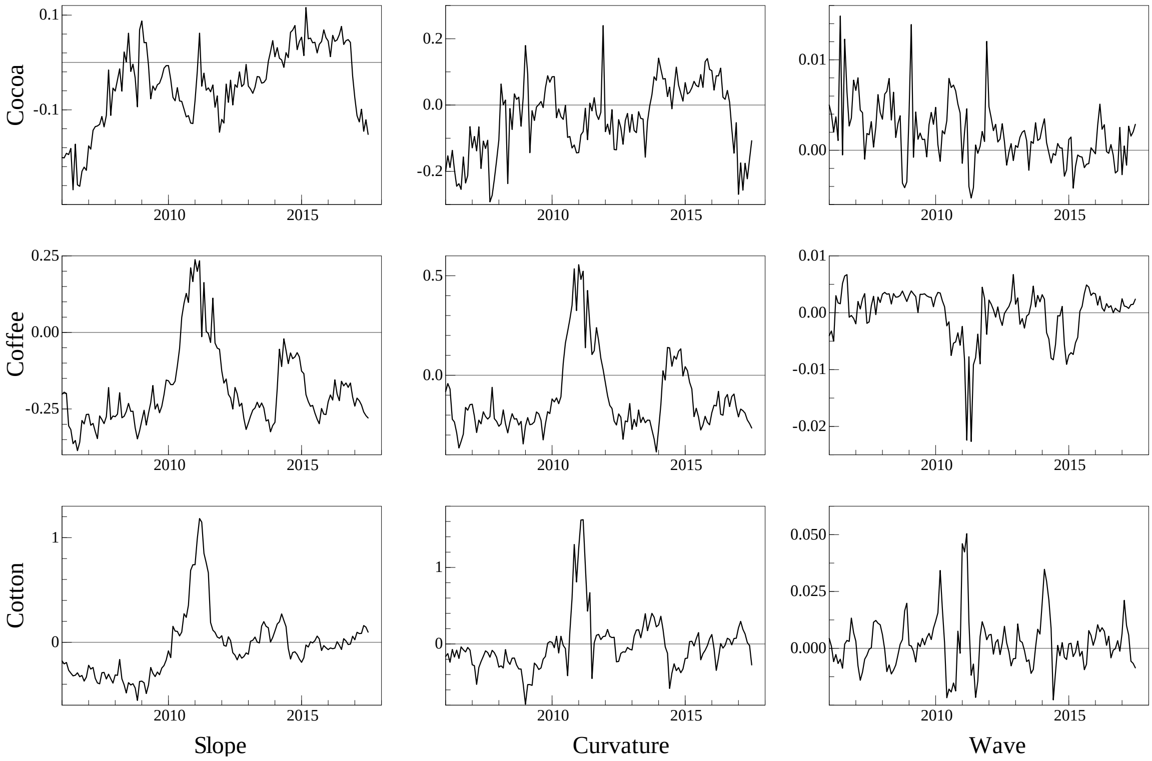
Table 2. Value of λ and Average R^2 for Eq. (12)

	τ	λ	R^2
Cocoa	8	0.22416	0.9673
Coffee	9	0.19925	0.9943
Cotton	7	0.25618	0.9170

Note: Values for τ and λ with $\max C(\tau) = \lambda$. R^2 is the average over 2006M1-2017M7.

On average, the factor decomposition of Eq. (12) captures more than 90 per cent of the variation of cotton futures curves and 99 per cent of the variation of coffee futures curves; see Table 2. The wave factor only marginally contributes to the average R-squares of Eq. (12) but remarkably improves the fit of the futures curves in instances of multiple extrema; a problem identified by Diebold and Li (2006) for their decomposition. The maximum for the curvature component, coinciding with the turning point of the wave factor, is found between the 7th and the 9th month for all three commodities. This is the end of the first third of the observed futures curve which spans over a maximum of 24 months for all three markets; see Figure 1.

Figure 4. Slope $\beta_{t,S}$, Curvature $\beta_{t,C}$, and Wave $\beta_{t,W}$ Factors Scores Eq. (12)



The extracted factors are then used to estimate Eq. (13). Results are reported in Table 3. As in the spread analysis, we find strong support for index and hedging pressure. Index pressure is associated with an upward sloping futures curve, while hedging pressure has the offsetting effect, in line with H(1a) and H(2a). The average index pressure effect is strongest for coffee and weakest (and again insignificant) for cotton. For coffee, index pressure is also associated with a concave futures curve, while hedging pressure is strongly associated with a convex futures curve for all three soft commodities as predicted in H(1b) and H(2b). Further, index pressure is found to be associated with N- shaped futures curves for all three commodities, potentially due to the roll-effects with the price of the maturing contract declining with the exit of index traders and the price of the deferred medium-term contract increasing with the entry of index traders. This observation is in line with Mou's (2011) finding of predictable and exploitable roll effects.

Table 3 also provides evidence for the convenience yield hypothesis for cocoa and cotton, including a non-linear relationship between convenience yield and level of storage. Inventory variables are leading the slope factor in line with H(3a). The signs of the coefficients are not easily interpretable because of their polynomial nature. A strong effect of levels of inventory is also found for the wave factor for all three commodity markets, supporting the argument that the wave component captures seasonal variations as suggested by H(3b).

Table 3. Estimation of Eq. (13) with $y_t = \beta_{j,t}, j = \{S, C, W\}$

		Constant	y_t	ψ_t	ψ_t^2	r_t	$\rho_{I,t}$	$\rho_{H,t}$
					Slope Factor			
		α_0	$(L)\alpha$	$(L)\omega_{\psi}$	$(L)\omega_{\psi^2}$	$(L)\omega_r$	$(L)\omega_{\rho_I}$	$(L)\omega_{\rho_H}$
SR	Cocoa	0.0005	-0.3538** [1]	-0.5571** [4]	0.2016** [4]	0.2227	-0.6369*	0.4882**
		(0.003)	(0.073)	(0.178)	(0.070)	(0.214)	(0.251)	(0.101)
	Coffee	-0.0014	-0.3455** [1]	0.1664 [1]	-0.2792* [1]	-0.2643	-0.7956**	0.9842**
		(0.004)	(0.069)	(0.213)	(0.143)	(0.429)	(0.248)	(0.156)
	Cotton	0.0050	0.1535 [1]	-0.2874* [3]	0.0915** [3]	2.5195* [2]	-0.2124	0.8740**
		(0.009)	(0.087)	(0.106)	(0.030)	(1.126)	(0.535)	(0.257)
			ϕ	γ_{ψ}	γ_{ψ^2}	γ_r	γ_{ρ_I}	γ_{ρ_H}
LR	Cocoa		-0.1631**	-	-	-	-2.6495**	-
			(0.048)				[7.868][b]	
					Curvature Factor			
		α_0	$(L)\alpha$	$(L)\omega_{\psi}$	$(L)\omega_{\psi^2}$	$(L)\omega_r$	$(L)\omega_{\rho_I}$	$(L)\omega_{\rho_H}$
SR	Cocoa	-0.0283*	-0.2622** [1]	0.6779	-0.2546	-1.1816*	-0.1109	0.4031
		(0.011)	(0.085)	(0.376)	(0.147)	(0.463)	(0.528)	(0.229)
	Coffee	-0.0028	-0.3408** [1]	-0.5034	0.1380	-0.6459	-1.1329*	1.0907**
		(0.007)	(0.081)	(0.396)	(0.266)	(0.790)	(0.455)	(0.282)
	Cotton[a]	0.0033	-	0.6716*	-0.1427*	2.1523* [1]	-0.5860	1.2125**
		(0.021)		(0.254)	(0.074)	(2.606)	(1.247)	(0.600)
			ϕ	γ_{ψ}	γ_{ψ^2}	γ_r	γ_{ρ_I}	γ_{ρ_H}
LR	Cocoa		-0.3997**	-	-	-0.6658**	-	1.1624**
			(0.089)			[18.22][b]		[11.96][b]
					Wave Factor			
		α_0	$(L)\alpha$	$(L)\omega_{\psi}$	$(L)\omega_{\psi^2}$	$(L)\omega_r$	$(L)\omega_{\rho_I}$	$(L)\omega_{\rho_H}$
SR	Cocoa[a]	0.0003	-	0.0383* [1]	-0.0140* [1]	-0.0450 [3]	0.0460	0.0177* [1]
		(0.000)		(0.015)	(0.006)	(0.018)	(0.020)	(0.008)
	Coffee[a]	0.0017	-0.2965* [1]	0.0208* [2]	-0.0157** [3]	0.0199	0.0387*	-0.0242*
		(0.001)	(0.082)	(0.009)	(0.006)	(0.035)	(0.019)	(0.011)
	Cotton[a]	0.0003	-	0.0318* [4]	-0.0089* [4]	0.1540** [3]	-0.0276 [3]	-0.0388 [1]
		(0.001)		(0.012)	(0.003)	(0.123)	(0.062)	(0.027)
			ϕ	γ_{ψ}	γ_{ψ^2}	γ_r	γ_{ρ_I}	γ_{ρ_H}
LR	Cocoa[a]		-0.6812**	-	-	0.0075**	0.0495**	-
			(0.079)			[6.662][b]	[8.735][b]	
	Coffee[a]		-0.2998**	-0.0306*	0.0198**	-	0.1185*	-
			(0.078)	[6.333][b]	[7.190][b]		[5.438][b]	

Note: Model estimated for 2007M1-2017M7 to allow enough room for appropriate lag selection. (.) Standard errors. [.] Lag length selected. * significant at the 5% level; ** significant at the 1% level. [a] Robust standard errors. [b] Long-run coefficients are recovered as the ratio between $\hat{\gamma}$ and the absolute value of $\hat{\phi}$. [.] is the Wald test statistic of insignificance of the long run coefficient following a Chi Square distribution under the Null. SR: short-run; LR: long-run.

In most instances, index and hedging pressure variables enter with contemporaneous terms. This is expected given the immediate pass through of changes in trader positions to changes in prices and hence futures curves. However, this implies that coefficient estimates presented in Table 3 potentially suffer from endogeneity bias. In this context, Gilbert and Pfuderer (2014) suggest an instrumental variable treatment for CIT positions with VIX based on S&P500 and Dow-Jones commodity index total returns as instruments. To test for the robustness of the results in Table 3, the instrumental variable (IV) estimation is replicated. Results are summarised in the Appendix Table A1, with only index and hedging pressure coefficients reported for brevity. As in Gilbert and Pfuderer (2014) instruments are found valid and index pressure coefficients are significantly larger (in absolute terms) than reported in Table 3. Hedging pressure coefficients are unchanged or insignificant. As hedging positions are unlikely to be correlated with the suggested financial instruments, the original set of instruments is extended by lagged inventory variables with hedging pressure coefficients being now closer to results in Table 3. Overall, results support the previous findings. Hedging pressure is a dominant driver of the slope and curvature, while index pressure is a dominant driver of the slope, curvature and wave-shapes of the futures curve. Slope and wave-shapes are further driven by level of inventory.

Conclusion

This paper reviews the argument that the presence of index traders has resulted in commodity futures curves to become uninformative or even misleading indicators of underlying demand and supply conditions. Index traders have been accused of inflating calendar spreads and causing exploitable anomalies due to their roll effect. On the other hand, index traders have been welcomed as liquidity providers, reducing hedging costs and thereby facilitating risk management and price discovery; see Sanders et al. (2010). Based on the heterogeneous agent model developed by Brunetti and Reiffen (2014), this paper shows that index traders indeed act as liquidity providers so long as their net-long positions do not exceed short hedgers' demand for counterparty. However, if index traders' net-long positions exceed hedgers' net-short positions, the presence of index traders is associated with the above-mentioned anomalies.

Following this hypothesis, the paper suggests indicators for hedging pressure and index pressure that capture the net price pressure effects of both traders. These indicators are used to investigate price pressure effects of these two trader types on commodity futures curves for cocoa, coffee and cotton futures markets. In contrast to previous studies relying on calendar spreads, this paper argues that calendar spreads are unable to capture the often complex and non-linear shapes of futures curves which have been linked to index positions in the literature. A factor decomposition method inspired by Nelson and Siegel (1987) is used instead.

Results presented in this paper show that index traders act as liquidity providers for the three soft commodity markets analysed and hence have a dampening effect on hedging pressure thereby reduce hedging costs as predicted by Brunetti and Reiffen (2014). However, when index long positions exceed short hedging demand, index traders' positions do not only offset hedging pressure effects but cause upward sloping and peaked futures curve with occasionally wave like shapes linked to roll-effects. These effects are clearly identified by use of the constructed indicators based on net-effects.

The findings presented in this paper lead to the conclusion that index traders' positions can turn into 'too much of a good thing'. While index traders are welcomed liquidity providers, their presence distorts price discover mechanisms if they become too dominant. This appears less of a threat for the cotton and cocoa market but is a relevant concern for the coffee market and other commodity markets in which index traders are more active, such as grains.

References

Barber, Joel R., and Mark L. Copper. 2012. "Principal Component Analysis of Yield Curve Movements." *Journal of Economics and Finance, 36* 750-765.

Bargawi, Hannah, and Susan Newman. 2017. “From Futures Markets to the Farm Gate: A Study of Price Formation along Tanzania’s Coffee Commodity Chain.” *Economic Geography 93 (2)* 162-184.

Baruník, Jozef, and Barbor Malinská. 2016. “Forecasting the term structure of crude oil futures prices with neural networks.” *Applied Energy, 164* 366-379.

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

Bessembinder, Hendrik. 1992. “Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets.” *The Review of Financial Studies, 5 (4)* 637-667.

Borovkova, Svetlana. 2010. “Detecting Market Transitions and Energy Futures Risk Management Using Principal Components.” *The European Journal of Finance, 12(6-7)* 495-512.

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* .

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

Campos, Julia, Neil R. Ericsson, and David F. Hendry. 2005. “General-to-specific Modeling: An Overview and Selected Bibliography.” *Board of Governors of the Federal Reserve System International Finance Discussion Papers, No. 838.*

Chang, Eric C. 1985. “Returns to Speculators and the Theory of Normal Backwardatio .” *The Journal of Finance, 40 (1)* 193-208.

Diebold, Francis X., and Canlin Li. 2006. “Forecasting the Term Structure of Government Bond Yields.” *Journal of Econometrics, 130* 337-364.

Dusak, Katherine. 1973. “Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums.” *Journal of Political Economy, 81 (6)* 1387-1406.

Etienne, Xiaoli L., and Fabio Mattos. 2016. “The Information Content in the Term Structure of Commodity Prices.” *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.*

Gilbert, Christopher L., and Simone Pfuderer. 2014. “The Role of Index Trading in Price Formation in the Grains and Oilseeds Markets.” *Journal of Agricultural Economics, 65 (2)* 303-322.

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

Hansen, Niels S., and Asger Lunde. 2013. "Analyzing Oil Futures with a Dynamic Nelson-Siegel Model." *CREATES Research Papers 2013-36, Department of Economics and Business Economics, Aarhus University.*

Heidorn, Thomas, Frieder Mokinski, Christoph Rühl, and Christian Schmaltz. 2015. "The impact of fundamental and financial traders on the term structure of oil." *Energy Economics, 48* 276-287.

Hirshleifer, David. 1990. "Hedging Pressure and Futures Price Movements in a General Equilibrium Model." *Econometrica, 58 (2)* 411-428.

Hirshleifer, David. 1988. "Residual Risk, Trading Costs, and Commodity Futures Risk Premia." *The Review of Financial Studies, 1 (2)* 173-193.

Irwin, Scott 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.

Kaldor, Nicholas. 1939. "Speculation and Economic Stability." *The Review of Economic Studies, 7 (1)* 1-27.

Karstanje, Dennis, Michel van der Wel, and Dick J. C. van Dijk. 2017. "Common Factors in Commodity Futures Curves." *http://dx.doi.org/10.2139/ssrn.2558014.*

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

Litterman, Robert, and José Scheinkman. 1991. "Common Factors Affecting Bond Returns." *The Journal of Fixed Income* 54-61.

Masters, Michael W. 2008. "Testimony of Michael W. Masters before the Committee on Homeland Security and Governmental Affairs United States Senate, May 20, 2008."

Mou, Yiqun. 2011. "Limits to Arbitrage and Commodity Index Investment: Front-Running the Goldman Roll." *Columbia Business School: Working Paper*.

Nelson, Charles R., and Andrew F. Siegel. 1987. "Parsimonious Modeling of Yield Curves." *The Journal of Business, 60(4)* 473-489.

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.

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

Staritz, Cornelia, Susan Newman, Bernhard Tröster, and Leonhard Plank. 2018. "Financialization and Global Commodity Chains: Distributional Implications for Cotton in Sub-Saharan Africa." *Development and Change* (forthcoming).

US Senate. 2009. *Excessive Speculation in the Wheat Market.* Washington D.C.: United States Senate Permanent Subcommittee on Investigations.

Acknowledgments

The research was supported financially by the German National Academic Foundation and Andrey Kuleshov and Elena Ivashentseva. This article greatly benefitted from comments by Prof Machiko Nissanke.