The price discovery function of agricultural commodities - Evidence from smooth transition regressions

This version: April 2012

Abstract
The movement of food prices remains a controversial issue due to the engagement of speculative capital which arguably introduces volatility and price movements unrelated to changes in traditional demand and supply factors. The relationship between spot and futures prices is an important topic in this context as the current period’s price of a futures contract should be an unbiased estimator of next period’s spot price under the joint assumption of risk neutrality and rationality. This study contributes to the literature by allowing for a smooth switching regression in the relationship between spot and futures of agricultural commodities after testing for cointegration between them. Our findings show that a linear framework, which has been adopted by previous studies, neglects important dynamics as spot prices only adjust to futures prices under specific circumstances. In some cases, future spot changes even tend to be negatively related to the forward premium.

*JEL Classification:* G13, G14, Q14, Q18

*Keywords:* agriculture, cointegration, commodities, spot and futures markets, smooth transition regression
1. Introduction

Food prices have drawn considerable attention after the increase over 2007 and 2008. It has been argued that financial rather than real factors determined the dynamics of agricultural commodities during this period which was characterized by volatility and spikes, unrelated to long-term trends (Modena, 2011). In particular, the possible engagement of speculative capital which arguably introduces volatility and price movements unrelated to demand and supply effects such as changes in the world population, economic growth or agricultural production is an important issue. In this context, the role of futures markets is also controversially discussed (Piesse and Thirile, 2009; Wright, 2011). On the one hand, volatility evokes risk for producers as well as consumers and futures markets are needed to reduce that risk. They allow for the transfer of risk from commercial traders who are exposed to futures price movements to non-commercial traders which are frequently labeled as speculators and take short (long) futures positions in the hope of yielding a capital gain as a result of the fall (rise) in prices. Classically, a limited number of traders which supported investment and stabilized futures prices were engaged in agricultural futures markets (Pace et al., 2008). However, this has changed recently as another group of investors have also entered the food market. Focusing on portfolio optimization, they regard food futures as an asset class and are labeled as futures investors instead of futures speculators (Gilbert, 2010). Having no intention of selling in the real market, their purchasing possibly introduces volatility as well as upward or downward pressure on prices (Pace et al., 2008). In this vein, Borin and Di Nino (2012) distinguish between traditional speculators and a new type of commodity investors which enter agricultural markets through the intermediation of 'swap dealers'. Masters (2008) as well as Masters and White (2008) argue that extensive buy-side pressure from index funds recently created a speculative bubble in commodity prices, with the consequence that prices heavily exceeded their fundamental values at the highest level. Applying GARCH models and Granger causality tests, Borin and Di Nino (2012) find that traditional futures speculators react to futures price changes rather than cause them while the engagement of 'swap dealers' seems to amplify price volatility in some markets. Their findings are supported by Von Braun and Tadesse (2012) who find that speculation effects are stronger then demand- and supply-side shocks for short-term price spikes. On the other hand, the US Commodity Futures Trading Commission argues that the level of speculation in agricultural commodities has remained relatively constant in percentage as prices have risen (CFTC, 2008).

The relationship between spot and futures prices is an important topic when analyzing the price discovery role of futures markets which might help to reduce uncertainty (Hernandez and Torero, 2010). The reason is that although futures markets can offer the possibility to gain arbitrage revenues and thus, exhibit speculation they may also form the mechanism by which new information are incorporated in prices if markets are efficient. Under the joint assumption of risk neutrality and rationality the current forward price is an unbiased estimator of the expected future spot price (Aulton et al., 1997; Kellard, 2002; Lin and Liang, 2010; Gilbert, 2010). On the other hand, an implication of standard
theory is that futures prices should follow a random walk with the price innovations introducing new information and mostly uninformed speculators trying to follow informed market participants. In this vein, a distinction between informed and uninformed traders has been provided in the sense that informed traders will trigger a return to a fundamental value through trading if uninformed traders previously have moved a market away from its fundamental value (Gilbert, 2010).

Early studies regarding the relationship between spot and futures prices for commodities conducted Granger (1969) causality tests and standard cointegration techniques such as the Engle and Granger (1987) methodology or the Johansen (1988) framework (MacDonald and Taylor, 1988; Oellemann et al., 1989; Koontz et al., 1990; Serletis and Banack, 1990; Schroeder and Goodwin, 1991; Quan, 1992; Crowder and Hamed, 1993; Schwartz and Szakmary, 1994; Foster, 1996; Aulton et al., 1997; Peroni and McNown, 1998; McKenzie and Holt, 2002; McAleer and Sequeira, 2004; Wang and Ke, 2005). However, due to financial market frictions such as high transaction costs, the role of noise traders, different types of risk-averse agents, and the microstructure effects of commodity markets, spot and futures markets could be characterized by a nonlinear price adjustment (Silvapulle and Moosa, 1999; Chen and Lin, 2004; Bekiros and Diks, 2008; Huang et al., 2009; Lin and Liang, 2010; Hernandez and Torero, 2010). This nonlinear structure of the relationship between spot and futures prices could depend on the current value of the basis or rather the forward premium which is the difference between the current period’s price of a futures contract for delivery in the next period and the current spot price.\(^1\) If the spot price is above the futures price, what is known as backwardation, the basis is negative, and if the magnitude of the basis exceeds a certain level, investors would start selling commodities at the spot market and buying futures contracts. Conversely, if the futures price is above the spot price, what is called contango, the basis is positive, and if again the magnitude of the basis exceeds a certain threshold, investors in this case would start selling futures contracts and buying commodities at the spot market.\(^2\) In between these two scenarios, investors may show just a slight reaction within a certain range (Huang et al., 2009). This argument is related to the existence of limits to speculation which states that investors only follow an investment strategy if the expected yield is higher than the one implied by other strategies (Sarno et al., 2006). Irwin et al. (2009), Sanders and Irwin (2010, 2011) as well as Irwin and Sanders (2011, 2012) cannot confirm the above mentioned Masters (2008)

---

\(^1\)In the literature the basis is often referred to as spot price minus futures price as well. The behavior of the basis determines the hedging usefulness of futures markets. From a theoretical point of view spot and futures prices should converge as maturity approaches. In general, the basis literature can be sub-divided in two categories: one deals with the determinants of the basis and the other intends to forecast the future basis level. Furthermore, the basis behavior is different for storable and non-storable commodities since the latter exhibit no direct functional relationship between spot and futures prices. Nevertheless, modeling the basis relationship for non-storable commodities is also important since spot and futures prices are expected to converge at maturity as well. For details regarding the basis relationship, the hedging function of futures markets and market efficiency of agricultural commodities see Garcia and Leuthold (2004) as well as Pennings and Leuthold (2000, 2001).

\(^2\)The threshold could be characterized by 'carrying costs' which makes an investor indifferent of buying a spot commodity or a futures contract.
hypothesis of a speculative price bubble in commodity and agricultural prices empirically by using returns and volatility in 19 commodity futures markets and concluded that the price discovery role in commodity futures markets is not harmed by speculators. However, they also only apply techniques that presume linearity such as cross-sectional regression and Granger (1969) causality tests.

In the recent literature nonlinearity has been accounted for by using methods such as the threshold autoregressive (TAR) model developed by Tong (1983), the momentum TAR (MTAR) model suggested by Enders and Granger (1998) as well as Enders and Siklos (2001), the multivariate TAR (MTAR) model proposed by Tsay (1998) and the threshold vector error correction (TVECM) model developed by Hansen and Seo (2002) (Ewing et al., 2006; Huang et al., 2009; Lin and Liang, 2010; Mamatzakis and Remoundos, 2011). However, all those studies allow for a discrete switching from one scenario to the other. Such a pattern seems inadequate in cases where investors with different expectations are involved. Market participants may not all act promptly and uniformly as they are confronted with different information and opportunity costs which implies different bands of inaction. In addition, their reaction to new information might also exhibit different delays (Teräsvirta, 1998). A more appropriate modeling strategy which corresponds to a smooth switching between two extreme regimes has been inspired by the work of Teräsvirta (1994).

Hence, this paper contributes to the literature by applying the corresponding smooth transition regressive (STR) models for the spot and futures relationship. Based on different economic conditions reflected by a transition variable, our approach is well suited to distinguish periods where futures markets are able to perform a price discovery function from times where other market factors, such as speculative pressure, outshine such a function. More precisely, the family of STR models offers the main advantage of allowing for different dynamics which are determined by the choice of the transition function. The latter is either of exponential form to account for a symmetric but size increasing adjustment in both extreme regimes or logistic form to allow for a more flexible asymmetric adjustment above and below a threshold. It is worth mentioning that we do not focus on the question whether the data generating process of the individual series for futures and spot prices follow a nonlinear pattern. Instead, we analyze whether the adjustment pattern of spot prices to a linear relationship between spot and futures is nonlinear. Teräsvirta et al. (2010) provide a detailed discussion of the relation between nonstationarity and nonlinearity. They state that a nonlinear data generating process may still be approximated as integrated of order one, e.g. $I(1)$, which implies that the concept of cointegration may still be applied. However, we also consider nonlinear unit root tests to exclude the possibility that we approximate a nonlinear stationary process as $I(1)$.

The remainder of the paper is organized as follows. In the next section we present the underlying theoretical considerations, which draw the connection between spot and futures

---

3Furthermore, Cunado and Perez De Gracia (2003) as well as Maslyuk and Smyth (2009) used the Gregory and Hansen (1996) framework to account for the presence of structural breaks in the relationship between spot and futures prices. Recently, Lee and Zeng (2011) used quantile cointegration regression to describe the connection between spot and futures oil prices.
prices. The empirical part of the study is presented in section 3. After a brief discussion of the data used we analyze the time series properties and test for cointegration between spot and futures prices in a linear framework. This seems necessary as many studies assume a long-run relationship between spot and futures prices without explicitly testing. We then proceed by allowing for a nonlinear adjustment of the spot return with respect to the forward premium. More precisely, we test for linearity against nonlinearity, choose the appropriate transition function and finally, estimate the specified equations. Section 4 concludes and gives the limitations of our study as well as an outlook on further research.

2. The model

Under the joint assumption of risk neutrality and rationality the current period’s price of a futures contract for delivery in the next period is an unbiased estimator of the expected next period’s spot price (Kellard, 2002; Huang et al., 2009; Lin and Liang, 2010). Thus, the corresponding hypothesis is given below

\[ E_t(s_{t+k}) = f_{t,k}, \]  

where \( s_{t+k} \) denotes the logarithm of the spot price at time \( t+k \), \( f_{t,k} \) represents the logarithm of the price of a futures contract observed at time \( t \) for delivery at time \( t + k \), and \( E_t(\cdot) \) gives the expectations operator conditional on information available at time \( t \). Equation (1) could be transformed as follows

\[ s_{t+k} = f_{t,k} + u_{t+k}, \]  

where \( u_{t+k} \) indicates an uncorrelated random error term with zero mean and constant variance. Moreover, subtracting \( s_t \), the logarithm of the spot price at time \( t \), on both sides of equation (2) yields the commonly known linear regression provided by Fama (1984)

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k}, \]  

where \( \Delta s_{t+k} \equiv s_{t+k} - s_t \) denominates the spot return and the unbiasedness hypothesis given above presumes \( \alpha = 0 \) and \( \beta = 1 \). Due to our attempt of analyzing the spot and futures markets of different agricultural products we rely on this framework. A similar approach which is based on the theory of uncovered interest rate parity (UITP) has been used by Sarno et al. (2006), Baillie and Kilic (2006), Hochradl and Wagner (2010), Olmo and Pilbeam (2011) as well as Pilbeam and Olmo (2011) to analyze the relationship between the spot return and the forward premium for different exchange rates. As outlined in the introduction, equation (3) neglects the possibility of nonlinearity in the relationship between the spot return and the forward premium represented by the difference of the current futures and spot price. To allow for this kind of dynamics, we augment equation

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k}, \]  

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k} \] 

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k} \] 

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k} \] 

\[ \Delta s_{t+k} = \alpha + \beta (f_{t,k} - s_t) + u_{t+k} \]
(3) as follows

\[ \Delta s_{t+k} = [\alpha_1 + \beta_1(f_{t,k} - s_t)] + [\alpha_2 + \beta_2(f_{t,k} - s_t)]F(z_t, \gamma, c) + u_{t+k}, \quad (4) \]

where \( F(z_t, \gamma, c) \) is a transition function which ascertains the speed of adjustment and could either be a logistic or an exponential function. Hence, equation (4) can be interpreted as a nonlinear error correction framework for the spot return with respect to a proportional long-run relationship between \( f_{t,k} \) and \( s_t \). The terms \( \alpha_1 \) and \( \beta_1 \) correspond to the lower regime, while \( (\alpha_1 + \alpha_2) \) and \( (\beta_1 + \beta_2) \) belong to the upper regime of the adjustment process (Van Dijk et al., 2002). Thus, \( \beta_1 \) and \( (\beta_1 + \beta_2) \) can be interpreted as the error correction coefficients implying that values between 0 and 1 indicate that the spot return adjusts to the premium spread.\(^5\) Strictly speaking, any deviation from 1 implies that speculative behavior in terms of simultaneously selling or buying at the spot and futures market may gain excess revenues. In terms of hedging, investors who omit adjustment to the cointegrating relationship between spot and futures prices as given by equation (3) will adopt a suboptimal portfolio. As long as \( \beta \geq 0 \), the optimal futures price tends to be too small. The opposite holds if \( \beta < 0 \). Both cases result in a relative poor hedging performance (Lien, 1996).\(^6\)

While the strategy to select the adequate transition function will be discussed and applied in the next section, it seems reasonable to point out the main differences between a logistic and an exponential formulation at this stage of the analysis. Although both configurations are close substitutes, they refer to different patterns of nonlinearities. In a nutshell, a logistic transition function allows for different adjustment above and below a threshold while the exponential transition function allows for a distinction between small and large deviations from the threshold. To explain the underlying dynamics in greater detail, we first consider the case where \( F(z_t, \gamma, c) \) is a bounded continuous logistic transition function which lies between 0 and 1. Thus, it has the following form:

\[ F(z_t, \gamma, c) = \left[1 + \exp(-\gamma(z_t - c)/\sigma_{z_t})\right]^{-1} \quad \text{with} \quad \gamma > 0, \quad (5) \]

where \( z_t \) indicates the transition variable, \( \sigma_{z_t} \) represents its standard deviation, \( \gamma \) denotes a slope parameter and \( c \) is a location parameter. In order to create a scale-free smoothness parameter, \( \gamma \) is normalized by the standard deviation of the transitional variable \( z_t \), as suggested by Teräsvirta (1998). A natural choice for the transition variable is the forward premium \( f_{t,k} - s_t \) with several lags \( j \) up to one week. An interesting alternative is the average absolute difference between spot and futures during the last week which is labeled

---

\(^5\)Values above unity indicate an overshooting of the error correction and values below zero report that the dynamic pattern is explosive.

\(^6\)See Pennings and Leuthold (2000) for an excellent overview on motivations for hedging.
as \( v_t \) in the following and can be written as

\[
v_t = \frac{1}{j} \sum_{i=1}^{j} |f_{t-i,k} - s_{t-i}|
\]

with \( j = 5 \). Such a measure is, for instance, adopted by Franses and Van Dijk (2000) and has the advantage of quantifying the overall volatility instead of focusing on prices of a certain day.

Thus, this implies that the lower (upper) regime is associated with negative (positive) values of the forward premium \( f_{t,k} - s_t \) or the overall volatility relative to the location parameter \( c \). The logistic function increases monotonically from 0 to 1 as the transition variable increases, so that \( F(z_t, \gamma, c) \rightarrow 0 \) as \( z_t \rightarrow -\infty \) and \( F(z_t, \gamma, c) \rightarrow 1 \) as \( z_t \rightarrow +\infty \) while it takes the value 0.5 if \( z_t = c \). Hence, the location parameter can be interpreted as a threshold value dividing equation (4) into three different extreme regimes corresponding to \( \lim_{z_t \rightarrow -\infty} F(z_t, \gamma, c) \), \( \lim_{z_t \rightarrow +\infty} F(z_t, \gamma, c) \) and \( z_t = c \). In case of \( z_t = c \) equation (4) reduces to the linear model given by equation (3) with \( \alpha = \alpha_1 + 0.5\alpha_2 \) and \( \beta = \beta_1 + 0.5\beta_2 \).

The slope parameter \( \gamma \) determines the speed of the transition between the extreme regimes (Baillie and Kilic, 2006).

Turning to the second possibility, where \( F(z_t, \gamma, c) \) is modeled by a bounded continuous exponential transition function which lies between 0 and 1 and thus, has the following form:

\[
F(z_t, \gamma, c) = 1 - \exp(-\gamma(z_t - c)^2/\sigma_z) \quad \text{with} \quad \gamma > 0.
\]

The exponential transition function (7) is symmetrically inverse-bell-shaped (\( F(z_t, \gamma, c) \rightarrow 1 \) for \( z_t \rightarrow \pm \infty \)), so that an adjustment for deviations of the basis above and below the threshold \( c \) is symmetric as opposed to the logistic function. The parameter \( \gamma \) again determines the smoothness of the transition, with lower absolute values implying slower transition (Taylor et al., 2001).

In the following the use of equation (4) in combination with equation (5) and (7) is referred to as the logistic STR or LSTR and the exponential STR or ESTR, respectively. In terms of interpretation, it is worthwhile mentioning that the threshold \( c \) is unrestricted in the present study. If a strictly proportional long-run relationship between spot and futures exist, the latter might also be restricted to zero in the fashion of a smooth transition error correction model suggested by Van Dijk et al. (2002). However, although this would facilitate the interpretation of the results, an unrestricted threshold seems more appropriate as a proportional long-run relationship cannot be verified for all commodities according to the results in the next section. Owing to the fact that we also consider a volatility measure, this does not alter the interpretability of the results.
3. Data, methodology and empirical results

3.1. Data

Our analysis is based on data from the Dow Jones UBS Commodity Index (DJ-UBSCI) which is composed of commodities traded on U.S. exchanges and provided by Dow Jones Indexes (http://www.djindexes.com/commodity/). The DJ-UBSCI is weighted by the relative amount of trading activity of a particular commodity and has been known as the Dow Jones AIG Commodity Index until 2009. Beside the S&P Goldman Sachs Commodity Index (GSCI) the DJ-UBSCI is one of the two largest indices by market share. More precisely, we apply the spot prices and the three month nearby futures contracts prices of the commodity subindex for agriculture and also analyze the corresponding individual commodities, namely coffee, corn, cotton, soybeans, soybean oil, sugar, and wheat on a daily basis. Our sample period covers each working day from January 2, 1991 to October 19, 2011 (5210 observations) and thus, exhibits the largest available sample size which contains the low volatility period until the early 2000’s as well as the high volatility period thereafter. In general prices of agricultural products mostly rose considerably in the mid-90s due to the crop price shock as well as between 2007 and 2008 owing to the recent food crisis. As it is common practice, each series is taken as log.

To estimate equation (4) for each spot and futures market in the context of cointegration we first have to assure that each of the current spot and futures prices is integrated of order one, e.g. $I(1)$, and both are cointegrated as well as each spot return is $I(0)$. Thus, we use the augmented Dickey-Fuller (ADF) test to check the null of a unit root in each series (Dickey and Fuller, 1979). We apply an auxiliary regression with an intercept, but without a trend regressor since a graphical inspection shows that neither series exhibits a time dependent mean. Therefore, we test the null of a random walk process without drift against the alternative of a stationary process with non-zero mean for the level of each series. The results are displayed in Table 1 and indicate that the null cannot be rejected for each series. These results continue to hold if unit root tests which account for nonlinearity are considered. To test the null of a unit root against a nonlinear but globally stationary exponential smooth transition autoregressive process we conduct the $t_{NL}$ test proposed by Kapetanios et al. (2003). These statistics are reported in Table 1 as well. Thus, each spot and futures price can be regarded as $I(1)$ since the same null can clearly be rejected for the first difference of each series denoted by $\Delta$. Moreover, the spot return is constructed by $\Delta s_{t+k} \equiv s_{t+k} - s_t$ with $k = 66$ since we use three month futures to estimate equation (4) and 66 equals the number of working days during three months. As can be seen in Table 1, each spot return is stationary, e.g. $I(0)$.\footnote{See Tang and Xiong (2010) and Gilbert (2010) for details regarding the DJ-UBSCI and its subindices. Following Tang and Xiong (2010) the correlation between the GS and the DJ-UBS commodity indices is over 0.9. As a result, using GSCI would not change our findings qualitatively. Among others, Irwin and Sanders (2012) applied the DJ-UBSCI data set for a similar purpose.}

\footnote{We have also applied the more powerful Ng and Perron (2001) $MZ_\alpha$ test to ascertain the}
3.2. Methodology and empirical results

In the following we have to ascertain that each spot and futures price pair is cointegrated such that the forward premium is $I(0)$.$^9$ After checking for cointegration we test the null of a linear specification given by equation (3) against a nonlinear specification of the form shown in equation (4) and we apply an empirical procedure to choose the appropriate transition function.

3.2.1. Testing for cointegration

We start our analysis by applying the multivariate cointegration test by Johansen (1988, 1991), which draws upon the following vector autoregression representation (VAR):

$$\Delta Y_t = \Pi Y_{t-1} + \Gamma(L)\Delta Y_{t-1} + \Phi D_t + \varepsilon_t, \quad t = 1, \ldots, T,$$

where $Y_t = [f_t, s_t]'$. The non-stationary behavior of the series is accounted for by a reduced rank ($r < p$) restriction of the long-run level matrix $\Pi$, which can be fragmented into two $r \times p$ matrices $\alpha$ and $\beta'$ ($\Pi = \alpha\beta'$). $\beta'$ gives the coefficients of the variables for the $r$ long-run relation, while $\alpha$ contains the adjustment coefficients describing the reaction of each variable to disequilibria from the $r$ long-run relations given by the $r \times 1$ vector $\beta'Y_{t-1}$. The deterministic components are given by the $(p \times 1)$ vector $\Phi D_t$, while $\varepsilon_t$ describes an independent and identically distributed error term. The term $\Gamma(L)\Delta Y_{t-1}$ describes the short-run dynamics of the model using $p$ equations between current variables, $L$-lagged variables and equilibrium errors (Juselius, 2006).

To identify the rank, that is, the number of cointegrating relations $r$, we rely on the trace test developed by Johansen (1988). The idea of the test is to separate the eigenvalues $\lambda_i$, $i = 1, \ldots, r$, which correspond to stationary relations, from those eigenvalues $\lambda_i$, $i = r + 1, \ldots, p$ which belong to non-stationary eigenvectors. The test statistic of the corresponding likelihood test, the so-called trace test, is given by $trace(r) = -T \sum_{i=r+1}^{p} \log(1-\tilde{\lambda}_i)$. Starting with the hypothesis of full rank, the rank $r$ is determined by using a top-bottom procedure until the null cannot be rejected (Juselius, 2006). For the specification of all models, the choice of lag length is based on tests for autocorrelation. The results of the trace tests are shown in Table 2.

---

$^9$Recently, a similar approach is conducted by Asche et al. (2012) to test sorghum market integration for different regions in Tanzania.
According to the test statistic, the hypothesis of a rank of one cannot be rejected at the 5% level for most spot and futures market relation. However, spot and futures of the indices coffee, cotton, and soybeans do not share a common stochastic trend according to the results and are consequently dropped from the analysis in the following. The findings in these cases imply that the forward spread cannot be applied as the transition variable. The reason is that the latter should be a stationary linear combination of the $I(1)$ variables under observation. After the determination of the rank, we followed the Johansen (1988, 1991) approach and computed the maximum likelihood estimates of the unrestricted cointegrating relations $\beta'Y_{t-1}$ for the remaining configurations.

As a next step, we analyze whether a strictly proportional long-run relationship between spot and futures holds by testing the restriction $\beta' = (1, -1)$ based on a likelihood ratio procedure suggested by Juselius (2006). The results, which are also given in Table 2, show that this restriction is rejected for some commodities (sugar and wheat). Nevertheless, each pair of coefficients is theory-consistent in terms of sign and magnitude with spot and futures prices always being positively related. However, the failure to find a strictly proportional long-run relationship between them might be attributed to nonlinear adjustment which may bias downward the speed of adjustment if a linear specification is used (Taylor, 2006). Thus, in the following we assume that each spot and futures price pair is cointegrated with a coefficient of unity when allowing for nonlinearities in the adjustment of spot returns to the forward spread between both. This line of reasoning facilitates the interpretation of the findings in the next step. However, the results remain qualitatively unchanged if the estimated long-run coefficients are applied for sugar and wheat.

### 3.2.2. Testing for linearity against nonlinearity

Next, it is necessary to formally test for nonlinearity, though it is also important to choose an adequate transition variable, which in the present study means the choice of a lag order for the basis. Both issues can be tackled by applying a lagrange multiplier (LM) test introduced by Luukkonen et al. (1988) which is based on the following third order Taylor approximation of the transition function (Teräsvirta, 1998; Franses and Van Dijk, 2000):

$$
\Delta s_{t+k} = \varphi_0 + \varphi_1(f_{t+k} - s_t)z_t + \varphi_3(f_{t+k} - s_t)z_t^2 + \varphi_4(f_{t+k} - s_t)z_t^3 + \varepsilon_{t+k}.
$$

(9)

The null hypothesis which refers to the linear model being adequate is tested as $H_0: \varphi_i = 0$ with $i = 2, 3, 4$ against the alternative $H_1$ that at least one $\varphi_i \neq 0$, implying that the higher

---

**Table 2 about here**

**According to the test statistic, the hypothesis of a rank of one cannot be rejected at the 5% level for most spot and futures market relation. However, spot and futures of the indices coffee, cotton, and soybeans do not share a common stochastic trend according to the results and are consequently dropped from the analysis in the following. The findings in these cases imply that the forward spread cannot be applied as the transition variable. The reason is that the latter should be a stationary linear combination of the $I(1)$ variables under observation. After the determination of the rank, we followed the Johansen (1988, 1991) approach and computed the maximum likelihood estimates of the unrestricted cointegrating relations $\beta'Y_{t-1}$ for the remaining configurations.**

As a next step, we analyze whether a strictly proportional long-run relationship between spot and futures holds by testing the restriction $\beta' = (1, -1)$ based on a likelihood ratio procedure suggested by Juselius (2006). The results, which are also given in Table 2, show that this restriction is rejected for some commodities (sugar and wheat). Nevertheless, each pair of coefficients is theory-consistent in terms of sign and magnitude with spot and futures prices always being positively related. However, the failure to find a strictly proportional long-run relationship between them might be attributed to nonlinear adjustment which may bias downward the speed of adjustment if a linear specification is used (Taylor, 2006). Thus, in the following we assume that each spot and futures price pair is cointegrated with a coefficient of unity when allowing for nonlinearities in the adjustment of spot returns to the forward spread between both. This line of reasoning facilitates the interpretation of the findings in the next step. However, the results remain qualitatively unchanged if the estimated long-run coefficients are applied for sugar and wheat.

### 3.2.2. Testing for linearity against nonlinearity

Next, it is necessary to formally test for nonlinearity, though it is also important to choose an adequate transition variable, which in the present study means the choice of a lag order for the basis. Both issues can be tackled by applying a lagrange multiplier (LM) test introduced by Luukkonen et al. (1988) which is based on the following third order Taylor approximation of the transition function (Teräsvirta, 1998; Franses and Van Dijk, 2000):

$$
\Delta s_{t+k} = \varphi_0 + \varphi_1(f_{t+k} - s_t)z_t + \varphi_3(f_{t+k} - s_t)z_t^2 + \varphi_4(f_{t+k} - s_t)z_t^3 + \varepsilon_{t+k}.
$$

(9)

The null hypothesis which refers to the linear model being adequate is tested as $H_0: \varphi_i = 0$ with $i = 2, 3, 4$ against the alternative $H_1$ that at least one $\varphi_i \neq 0$, implying that the higher

---

To save space we do not report the coefficients, but these are available upon request. Among others Crowder and Hamed (1993), Aulton et al. (1997), Peroni and McNoon (1998), McKenzie and Holt (2002) as well as Wang and Ke (2005) support our finding of a cointegration relationship between spot and futures prices for commodities.

In the case of small samples in combination with a large number of explanatory variables, $F$ versions of the LM test statistics are preferable, as they have better size properties (Granger and Teräsvirta, 1993; Teräsvirta, 1998; Van Dijk et al., 2002).
order terms are significant (Teräsvirta, 1998). The test statistic has a $\chi^2$ distribution with three degrees of freedom.\textsuperscript{12} In the case of the hypothesis of linearity being rejected, a method for choosing the transition variable lies in computing the test statistic for several transition functions, i.e. different values of the lag order $j$, and selecting the configuration for which its value is maximized (Taylor et al., 2001; Van Dijk et al., 2002). Teräsvirta (1994, 1998) has shown that this approach works well in most cases.

In the present study, delays from one to five days are considered.\textsuperscript{13} The results of the LM tests presented in Table 3 show that the hypothesis of linearity is rejected for all commodities and lag orders $j$.

Table 3 about here

Hence, the overall conclusion is that a nonlinear framework is adequate. An inspection of the tests statistics shows that the optimal transition variable differs with different lag orders for the forward spread considered to be the most adequate choice. Although, the forward premium is preferred compared to the volatility measure in each case, in order to achieve comparable and robust results, two different models will be estimated for each commodity in the following: One configuration based on the volatility measure and another on the lagged spread associated with the highest test statistic. This proceeding always includes the optimal transition variable for each commodity and enables us to draw clear conclusions as the thresholds for the lagged basis and the volatility hold different implications.

3.2.3. Testing for the appropriate specification

The above shown LM testing procedure can also be applied to distinguish between a logistic and an exponential transition function and thus, to choose the appropriate specification (Granger and Teräsvirta, 1993; Teräsvirta, 1994, 1998; Van Dijk et al., 2002). If the linearity null has been rejected, equation (9) is used to test the following hypotheses successively

$$
H_{01} : \varphi_4 = 0, \\
H_{02} : \varphi_2 = 0 | \varphi_3 = 0, \\
H_{03} : \varphi_3 = 0 | \varphi_4 = 0.
$$

(10)

The decision rule to select the most adequate transition function introduced by Teräsvirta (1994) is as follows. If the rejection of $H_{03}$ is the strongest one in terms of largest test statistic, the ESTR model should be chosen, otherwise the LSTR model should be preferred.\textsuperscript{14} Table 4 displays the empirical realizations of the test statistics and corresponding $p$-values for the two configurations of each commodity.

\textsuperscript{12}The number of degrees of freedom $3p$ refers to the number of regressors $p$ which in our case is one. Furthermore, the test assumes that all regressors as well as the transition variable $z_t$ should be stationary and uncorrelated with the error in equation (4) $u_{t+k}$ (Teräsvirta, 1998). As shown in section 3.2.1 our only regressor and transition variable $(f_{t,k} - s_t)$ is a stationary linear combination of non-stationary $I(1)$ variables.

\textsuperscript{13}Longer delays have turned out to be less suitable in previous estimations carried out by the authors. The results are available upon request.

\textsuperscript{14}See Granger and Teräsvirta (1993) or Teräsvirta (1994) for details.
As can been seen, the ESTR specification turns out to be most adequate for the agriculture index and for the wheat subindex while the LSTR specification is selected for corn, soybean oil and sugar.

3.2.4. Estimation results

The results of the estimation by nonlinear least squares are presented in Table 5.

It may be worthwhile mentioning first of all that in some cases observed insignificance of the transition parameter $\gamma$ cannot be interpreted as evidence against a smooth transition model, since the $t$-statistics need to be interpreted with caution (Taylor et al., 2001; Van Dijk et al., 2002). As Figure 1 shows, each transition function is identified, where we have plotted all estimated transition functions against the lagged forward premium as the corresponding transition variable. Each of the small blue circles represents an actual observation. It becomes evident that the conducted smooth switching approach is more appropriate than a discrete switching framework applied in previous studies since for each spot and futures relation we have enough observations in each regime and particularly in the region of transition between the extreme regimes. If a discrete switching process would be appropriate the transition function would only take the values of zero and unity and no values in between, but the graphs clearly show that a very huge amount of values between zero and unity have been observed. Thus, the plots display nearly continuous functions in each case. In case of a discontinuous function, one could see huge gaps from one circle to the other and this is unambiguously not the case. Hence, a discrete function is not able to specify the dynamics adequately.

The first main outcome regarding the estimation results given in Table 5 is that the adjustment of the spot return frequently differs between the two regimes, regardless if an exponential or a logistic function is chosen. This strongly supports our presumption of a nonlinear adjustment process in spot and futures markets for agricultural commodities.

Starting with exponential transition frameworks, the findings for both the whole agriculture index and the individual wheat index display a positive coefficient $\beta_1$ while the sum of $\beta_1$ and $\beta_2$ turns out to be negative for wheat and close to zero for agriculture. Hence, spot returns positively adjust to the forward spread in case of a small spread or low volatility with the high $\beta_1$ coefficients for agriculture even suggesting an overshooting behavior of future spot prices. The former scenario might be labeled as normal in the sense that spot and futures prices rarely depart from each other. However, if the previous basis and volatility, respectively, has been large a positive spread may even result in a drop of future spot prices and vice versa for wheat while the influence becomes very low for agriculture in the extreme regime. As large forward spreads seem to be rather exception than rule, such a scenario might be referred to as a time of turbulences. However, both patterns obviously imply arbitrage opportunities which may be traced back to the engagement of speculators.
An overshooting of spot returns might result from upward pressure of prices if speculators believe that prices will continue to rise as, for instance, observed in 2007 and 2008 when it was argued that such a positive feedback loop generated an economic bubble (Pace et al., 2008). A weak or inverse relationship might arise if market participants do not consider futures prices to obtain relevant information or even expect market prices to turnaround in the future.

Turning to the case of a logistic function, a positive sign of the coefficient $\beta_1$ can be observed for both transition variables in case of corn and the volatility configuration of sugar. Applying the lagged spread as the transition variable for sugar gives insignificant coefficients. However, the link between the forward spread and the spot return weakens in the other three cases in the second regime with $\beta_2$ turning out to be negative although the sum of $\beta_1$ and $\beta_2$ remains positive but draws near zero for corn and sugar. These findings imply that future spot changes tend to be less related to the forward premium if the latter is above its threshold in previous periods while a positive coefficient can be observed if the lagged forward spread turns out to be below the threshold. Considering that the threshold is positive for the three configurations, $\beta_1$ corresponds to a falling basis which might become negative while the sum of $\beta_1$ and $\beta_2$ refers to an increasing positive basis. In terms of volatility, a value above the threshold corresponds to a large average absolute basis during the previous week while a value below the threshold corresponds to low volatility. Hence, a similar pattern as argued in case of the exponential transition function arises as the forward spread exhibits significantly more predictive power for spot returns if the volatility and previous period’s spread, respectively, has been low.

A reversed causality can be observed for the price of soybean oil where $\beta_1$ is negative and $\beta_2$ positive while the sum of $\beta_1$ and $\beta_2$ remains negative. Therefore, forward spread and future spot returns seem to be inversely related in both regimes. Similar to the arguments outlined for wheat and the whole agriculture index above, the patterns might also point to arbitrage gains from speculation which may be traced back to different expectations over future market movements and the information embedded in futures prices.

Several modified estimations have been carried out to test for the robustness of the overall results. In particular, we have modified our estimations with respect to the choice of the transition variable by introducing different lags for the lagged forward premium. The results remain qualitatively unchanged, for instance, if the lag order is chosen to be one for each commodity. Estimating all models with exponential or logistic transition functions also suggests that the established outcomes continue to hold. Overall, the results show that our findings are robust with respect to different configurations.\footnote{To save space, the corresponding results are not presented here, but are available upon request.}

Our results are also compatible with the findings of the cointegration analysis in section 3.2.1 where a long-run relationship between spot and futures has been found while a proportional relationship between both has partly been rejected by the data. The different adjustment patterns to such a proportional relationship which partly contradict error correction due to incorrect signs or overshooting offer a possible explanation for the rejection

of proportionality. As outlined above, these findings might be explained by the engagement and expectations of different investors, in particular speculators. However, it is worthwhile mentioning that they do not necessarily imply that uninformed traders are responsible for the market movements as systematic forecast errors can also occur under rational expectations (Lewis, 1989). Hence, our findings also do not contradict the hypothesis of rational expectations. Altogether, the results have confirmed that nonlinearities need to be accounted for when the relationship between spot and futures prices for food commodities is analyzed. This is also important since the different adjustment states may correspond to varying degrees of hedging effectiveness as outlined in section 2. Predictive power observed in cases of low volatility behavior corresponds to a smaller than optimal futures position while the opposite holds in case of an inverse relationship.

4. Conclusion

The study has focused on the relationship between spot and futures prices of agriculture products. Testing for a long-run relationship between both and allowing for different patterns of nonlinear adjustment of future spot changes to the forward premium, we have shown that previous studies suffer from at least two shortcomings. Firstly, cointegration between spot and futures is often assumed without explicitly testing which turns out to be problematic in some cases as a long-run relationship is not supported by the data for three food commodities. More important, studies which adopt a linear framework neglect important dynamics as spot prices only proportionally adjust to futures prices under specific circumstances. In some cases, future spot changes even tend to be negatively related to the forward spread.

Our findings also suggest that a smooth switching approach should be adopted instead of a discrete threshold model. As far as the predictive power of current forward premium for spot returns is concerned, our results show that spot returns mostly only adjust to the forward spread in case of a small spread or low volatility. However, if volatility is large, future spot returns are not only increasingly detached from the current forward premium, a positive forward spread may even result in a drop of future spot prices and vice versa. Overall, the observed patterns hint to arbitrage opportunities which may be traced back to the engagement of speculators. The overshooting of spot returns might result from upward pressure while a weak or inverse relationship possibly arises if market participants do not consider futures prices to obtain relevant information or expect market prices to turnaround in the future. The first pattern is also compatible with the findings by Borin and Di Nino (2012) who identify a trend-following behavior of speculators. Overall, the engagement of speculative investors and the hedging effectiveness seems to be reversely related.

Our outcomes are also important in terms of policy implications. Von Braun and Torero (2008, 2009) have suggested the specification of a price band which would be a signal (threat) to speculators on food markets in the sense that a market assessment based on
virtual reserve is likely to occur when futures prices exceed the upper limit of this band. Although, the concrete implementation of such a mechanism is beyond the scope of this paper, our results indeed display that spot prices in many cases only show positive adjustment to futures prices if the latter do not depart extensively from the current spot prices. Furthermore, speculators seem to become increasingly engaged once futures prices depart from spot prices to a certain degree.

An interesting question for further research is whether the results remain unchanged if a different sample is analyzed. One would, for instance, expect that the extremes of overshooting behavior and the inverse relationship would dominate markets during specific time periods. In this vein, another major task is to dissect the role of speculation with respect to long-term price trends, volatility and spikes.

Acknowledgement

Thanks are due to Matin Qaim for his useful comments and helpful suggestions.

References


CFTC (2008): “Written Testimony of Jeffrey Harris, Chief Economist and John Fenton, Director of Market Surveillance Before the Subcommittee on


## A. Tables

### Table 1: Unit root tests

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Series</th>
<th>Level</th>
<th>Δ</th>
<th>Level</th>
<th>Δ</th>
<th>Level</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>t-stat</td>
<td></td>
<td>t-stat</td>
<td></td>
<td>t-stat</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lags</td>
<td></td>
<td>Lags</td>
<td></td>
<td>Lags</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>$f_t$</td>
<td>-0.99</td>
<td>[0]</td>
<td>-68.57***</td>
<td>[0]</td>
<td>-3.49</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$s_t$</td>
<td>-1.43</td>
<td>[0]</td>
<td>-68.47***</td>
<td>[0]</td>
<td>-2.33</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-5.99***</td>
<td>[0]</td>
<td>-53.40***</td>
<td>[0]</td>
<td>-4.26***</td>
<td>[0]</td>
</tr>
<tr>
<td>Coffee</td>
<td>$f_t$</td>
<td>-1.68</td>
<td>[0]</td>
<td>-70.68***</td>
<td>[0]</td>
<td>-1.79</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-6.72***</td>
<td>[0]</td>
<td>-71.71***</td>
<td>[0]</td>
<td>-3.73***</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-5.15***</td>
<td>[0]</td>
<td>-39.98***</td>
<td>[0]</td>
<td>-5.31***</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$s_t$</td>
<td>-0.67</td>
<td>[0]</td>
<td>-67.19***</td>
<td>[0]</td>
<td>0.10</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-5.44***</td>
<td>[0]</td>
<td>-30.87***</td>
<td>[0]</td>
<td>-4.60***</td>
<td>[0]</td>
</tr>
<tr>
<td>Soybean oil</td>
<td>$f_t$</td>
<td>-1.57</td>
<td>[0]</td>
<td>-69.03***</td>
<td>[0]</td>
<td>-4.4</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$s_t$</td>
<td>-2.01</td>
<td>[0]</td>
<td>-69.21***</td>
<td>[0]</td>
<td>-2.46</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-5.54***</td>
<td>[0]</td>
<td>-58.62***</td>
<td>[0]</td>
<td>-3.71***</td>
<td>[0]</td>
</tr>
<tr>
<td>Soybeans</td>
<td>$f_t$</td>
<td>-0.96</td>
<td>[1]</td>
<td>-70.35***</td>
<td>[0]</td>
<td>-2.46</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$s_t$</td>
<td>-1.44</td>
<td>[0]</td>
<td>-70.52***</td>
<td>[0]</td>
<td>-3.51</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-6.12***</td>
<td>[0]</td>
<td>-64.63***</td>
<td>[0]</td>
<td>-4.80***</td>
<td>[0]</td>
</tr>
<tr>
<td>Sugar</td>
<td>$f_t$</td>
<td>-0.81</td>
<td>[0]</td>
<td>-72.01***</td>
<td>[0]</td>
<td>1.07</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-6.59***</td>
<td>[0]</td>
<td>-59.50***</td>
<td>[0]</td>
<td>-4.46***</td>
<td>[0]</td>
</tr>
<tr>
<td>Wheat</td>
<td>$f_t$</td>
<td>-1.18</td>
<td>[0]</td>
<td>-69.16***</td>
<td>[0]</td>
<td>-4.1</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$s_t$</td>
<td>-0.14</td>
<td>[0]</td>
<td>-69.03***</td>
<td>[0]</td>
<td>1.00</td>
<td>[0]</td>
</tr>
<tr>
<td></td>
<td>$\Delta s_{t+66}$</td>
<td>-5.79***</td>
<td>[0]</td>
<td>-59.61***</td>
<td>[0]</td>
<td>-5.38***</td>
<td>[0]</td>
</tr>
</tbody>
</table>

**Note:** * Statistical significance at the 10% level, ** at the 5% level, *** at the 1% level. The ADF test equation is estimated including an intercept (c) for the levels and without deterministic regressors (n) for the first differences. For the ADF test critical values are taken from MacKinnon (1996): (c) 10% -2.57, 5% -2.86, 1% -3.43 and (n) 10% -1.62, 5% -1.94, 1% -2.57, respectively. For the $M_z$ test critical values are taken from Ng and Perron (2001): 10% -5.7, 5% -8.1, 1% -13.8. For the $t_{NL}$ test critical values are taken from Kapetanios et al. (2003): (c) 10% -2.66, 5% -2.93, 1% -3.48 and (n) 10% -1.92, 5% -2.22, 1% -2.82, respectively. The lag length is chosen by minimizing the Schwarz information criterion. Maximum lag length has been set to 32.
Table 2: Bi-variate Johansen cointegration rank test

<table>
<thead>
<tr>
<th>Spot and futures relation</th>
<th>Lags</th>
<th>Trace stat.</th>
<th>( (1,1) ) ( \chi^2 )-stat.</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>2</td>
<td>24.925*</td>
<td></td>
<td>0.072</td>
</tr>
<tr>
<td>( H_0: r \leq 1 \ vs. \ H_1: r \geq 2 )</td>
<td>2</td>
<td>2.870</td>
<td></td>
<td>0.788</td>
</tr>
<tr>
<td>Coffee</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>4</td>
<td>8.814</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>3</td>
<td>20.784*</td>
<td></td>
<td>0.212</td>
</tr>
<tr>
<td>( H_0: r \leq 1 \ vs. \ H_1: r \geq 2 )</td>
<td>3</td>
<td>3.177</td>
<td></td>
<td>0.645</td>
</tr>
<tr>
<td>Cotton</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>3</td>
<td>13.236</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybean oil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>3</td>
<td>21.780*</td>
<td></td>
<td>2.449</td>
</tr>
<tr>
<td>( H_0: r \leq 1 \ vs. \ H_1: r \geq 2 )</td>
<td>3</td>
<td>2.731</td>
<td></td>
<td>0.118</td>
</tr>
<tr>
<td>Soybeans</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>3</td>
<td>14.543</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>4</td>
<td>30.160*</td>
<td></td>
<td>10.608</td>
</tr>
<tr>
<td>( H_0: r \leq 1 \ vs. \ H_1: r \geq 2 )</td>
<td>4</td>
<td>3.137</td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Wheat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 \ vs. \ H_1: r \geq 1 )</td>
<td>4</td>
<td>51.570*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r \leq 1 \ vs. \ H_1: r \geq 2 )</td>
<td>4</td>
<td>4.059</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * Statistical significance at the 5% level. The constant is restricted to the cointegration relation, allowing for no linear trend neither in the data nor in the cointegrating equation. 5% critical values for testing \( H_0: r = 0 \) and \( H_0: r \leq 1 \) are taken from MacKinnon et al. (1999): 20.164 and 9.142, respectively. The lag length is chosen based on tests for autocorrelation.

Table 3: Teräsvirta test for nonlinearity and choice of the delay parameter

<table>
<thead>
<tr>
<th>( j )</th>
<th>( t - 1 )</th>
<th>( t - 2 )</th>
<th>( t - 3 )</th>
<th>( t - 4 )</th>
<th>( t - 5 )</th>
<th>Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>133.487***</td>
<td>134.140***</td>
<td>130.845***</td>
<td>125.440***</td>
<td>127.668***</td>
<td>118.811***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Corn</td>
<td>70.809***</td>
<td>72.242***</td>
<td>71.618***</td>
<td>68.988***</td>
<td>67.589***</td>
<td>54.327***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Soybean oil</td>
<td>562.434***</td>
<td>565.290***</td>
<td>571.224***</td>
<td>582.472***</td>
<td>570.635***</td>
<td>313.023***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Sugar</td>
<td>352.648***</td>
<td>356.280***</td>
<td>370.026***</td>
<td>385.722***</td>
<td>393.652***</td>
<td>273.881***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Wheat</td>
<td>399.623***</td>
<td>398.131***</td>
<td>397.822***</td>
<td>401.340***</td>
<td>398.084***</td>
<td>372.250***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Note: The table displays the test statistic of the LM test for nonlinearity as described in Section 3.2.2 for different lag orders \( j \) and the volatility measure, with \( p \)-values in parentheses. The test is distributed as \( \chi^2 \) with three degrees of freedom. For details, see Teräsvirta (1998). */**/*** implies rejection of the null hypothesis at the 10/5/1% significance level.
Table 4: Teräsvirta test for LSTR vs. ESTR

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Optimal lag ($j$)</th>
<th>Volatility</th>
<th>$H_{01}$</th>
<th>$H_{02}$</th>
<th>$H_{03}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>2</td>
<td>60.572***</td>
<td>75.634***</td>
<td>0.033</td>
<td>53.765***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.856)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Corn</td>
<td>2</td>
<td>38.250***</td>
<td>2.336</td>
<td>46.071***</td>
<td>27.364***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.126)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.318)</td>
</tr>
<tr>
<td>Soybean oil</td>
<td>4</td>
<td>66.221***</td>
<td>45.153***</td>
<td>3.390*</td>
<td>56.310***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.066)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Sugar</td>
<td>5</td>
<td>52.209***</td>
<td>64.883***</td>
<td>158.524***</td>
<td>13.456***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Wheat</td>
<td>4</td>
<td>1.922</td>
<td>298.349***</td>
<td>26.809***</td>
<td>0.526</td>
</tr>
<tr>
<td></td>
<td>(0.166)</td>
<td>(0.000)</td>
<td>(0.468)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Note: The table displays the test statistic for the selection of the appropriate specification as described in Section 3.2.3 for the optimal lag orders $j$ and the volatility measure, with $p$-values in parentheses. The test is distributed as $\chi^2$ with one degree of freedom. For details, see Teräsvirta (1998). */**/*** implies rejection of the null hypothesis at the 10/5/1% significance level.

Table 5: Estimation results

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Trans.func.</th>
<th>Trans.var.</th>
<th>$\alpha_1$</th>
<th>$\beta_1$</th>
<th>$\alpha_2$</th>
<th>$\beta_2$</th>
<th>$\gamma$</th>
<th>$c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>exponential</td>
<td>Basis ($t-2$)</td>
<td>-2.047***</td>
<td>7.808***</td>
<td>2.038***</td>
<td>-7.762***</td>
<td>48.024***</td>
<td>0.237***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.339)</td>
<td>(1.362)</td>
<td>(0.339)</td>
<td>(1.366)</td>
<td>(11.755)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volatility</td>
<td>-2.289***</td>
<td>8.723***</td>
<td>2.280***</td>
<td>-8.671***</td>
<td>51.207***</td>
<td>0.235***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.655)</td>
<td>(2.621)</td>
<td>(0.655)</td>
<td>(2.630)</td>
<td>(19.427)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Corn</td>
<td>logistic</td>
<td>Basis ($t-2$)</td>
<td>-0.065***</td>
<td>0.530***</td>
<td>-0.046***</td>
<td>-0.381***</td>
<td>158.561</td>
<td>0.212***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.083)</td>
<td>(0.017)</td>
<td>(0.087)</td>
<td>(99.559)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volatility</td>
<td>-0.065***</td>
<td>0.500***</td>
<td>-0.044***</td>
<td>-0.358***</td>
<td>329.988*</td>
<td>0.211***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.089)</td>
<td>(0.018)</td>
<td>(0.094)</td>
<td>(171.968)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Soybean oil</td>
<td>logistic</td>
<td>Basis ($t-4$)</td>
<td>0.014*</td>
<td>-0.411***</td>
<td>0.064</td>
<td>0.263*</td>
<td>21.858***</td>
<td>0.221***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.009)</td>
<td>(0.103)</td>
<td>(0.045)</td>
<td>(0.145)</td>
<td>(9.167)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volatility</td>
<td>0.010</td>
<td>-0.357***</td>
<td>0.073</td>
<td>0.196</td>
<td>25.056*</td>
<td>0.224***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.009)</td>
<td>(0.102)</td>
<td>(0.049)</td>
<td>(0.150)</td>
<td>(13.841)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Sugar</td>
<td>logistic</td>
<td>Basis ($t-5$)</td>
<td>0.003</td>
<td>-0.155</td>
<td>-0.009</td>
<td>0.272</td>
<td>34.939*</td>
<td>-0.125***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.045)</td>
<td>(0.216)</td>
<td>(0.046)</td>
<td>(0.220)</td>
<td>(978.065)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volatility</td>
<td>-0.004</td>
<td>3.212***</td>
<td>0.025</td>
<td>-3.162**</td>
<td>25.411</td>
<td>0.026***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
<td>(1.508)</td>
<td>(0.019)</td>
<td>(1.599)</td>
<td>(18.334)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>exponential</td>
<td>Basis ($t-4$)</td>
<td>-0.222***</td>
<td>0.284***</td>
<td>0.306***</td>
<td>-0.418***</td>
<td>0.878*</td>
<td>0.684***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.036)</td>
<td>(0.052)</td>
<td>(0.038)</td>
<td>(0.070)</td>
<td>(0.525)</td>
<td>(0.100)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volatility</td>
<td>-0.308***</td>
<td>0.466</td>
<td>0.348**</td>
<td>-0.522</td>
<td>6.503*</td>
<td>0.438***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.163)</td>
<td>(0.392)</td>
<td>(0.163)</td>
<td>(0.366)</td>
<td>(2.606)</td>
<td>(0.057)</td>
<td></td>
</tr>
</tbody>
</table>

Note: * Statistical significance at the 10% level, ** at the 5% level, *** at the 1% level. The coefficients are estimated by nonlinear least squares. Newey-West standard errors are given in parentheses.
B. Figures

Figure 1: Estimated transition functions