AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

A fractionally cointegrated VAR analysis of price discovery in commodity futures markets

Sepideh Dolatabadim, Morten Ørregaard Nielsen and Ke Xu

CREATES Research Paper 2014-24

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

A fractionally cointegrated VAR analysis of price discovery in commodity futures markets*

Sepideh Dolatabadi
Queen's University

Morten Ørregaard Nielsen†
Queen's University and CREATES

Ke Xu
Queen's University

August 14, 2014

Abstract

In this paper we apply the recently developed fractionally cointegrated vector autoregressive (FCVAR) model to analyze price discovery in the spot and futures markets for five non-ferrous metals (aluminium, copper, lead, nickel, and zinc). The FCVAR model allows for long memory (fractional integration) in the equilibrium errors, and, following Figuerola-Ferretti and Gonzalo (2010), we allow for the existence of long-run backwardation or contango in the equilibrium as well, i.e. a non-unit cointegration coefficient. Price discovery can be analyzed in the FCVAR model by a relatively straightforward examination of the adjustment coefficients. In our empirical analysis we use the data from Figuerola-Ferretti and Gonzalo (2010), who conduct a similar analysis using the usual (non-fractional) CVAR model. Our first finding is that, for all markets except copper, the fractional integration parameter is highly significant, showing that the usual, non-fractional model is not appropriate. Next, when allowing for fractional integration in the long-run equilibrium relations, fewer lags are needed in the autoregressive formulation, further stressing the usefulness of the fractional model. Compared to the results from the non-fractional model, we find slightly more evidence of price discovery in the spot market. Specifically, using standard likelihood ratio tests, we do not reject the hypothesis that price discovery takes place exclusively in the spot (futures) market for copper, lead, and zinc (aluminium and nickel).

JEL Codes: C32, G13.

Keywords: fractional cointegration, futures markets, price discovery, vector error correction model.

1 Introduction

Futures markets have traditionally been given (at least) two important roles in financial economics. First, they offer a method of hedging risk, and second, they contribute to the price discovery process (Working, 1948). Price discovery generally refers to the process of revealing an asset's permanent or fundamental value. This can be different from the actual observed price, which can be decomposed into the fundamental value as well as a transitory effect consisting of price changes due to bid-ask bounce, order book imbalances, etc.

*This article was prepared for a keynote address at the Conference on Performance of Financial Markets and Credit Derivatives, Centre for Economics and Financial Econometrics, Deakin University, Australia, which was given by the second author in May, 2014. We are grateful to the participants at the conference as well as Federico Carlini, Søren Johansen, Maggie Jones, and Michał Popiel for useful comments, suggestions and insights, and to the Canada Research Chairs program, the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation, DNRF78) for financial support.

†Corresponding author. E-mail: mon@econ.queensu.ca

A long literature has attempted to determine whether price discovery takes place primarily in the spot or futures markets; seminal contributions to this literature include, e.g., Stein (1961), Garbade and Silber (1983), Hasbrouck (1995), and Gonzalo and Granger (1995). Cointegration methods have played an important role in this development, both to asses futures market efficiency (e.g., Chowdhury, 1991; Kellard, Newbold, Rayner, and Ennew, 1999; Krehbiel and Adkins, 1993) and to analyze price discovery (e.g., Schroeder and Goodwin, 1991; Quan, 1992; Schwarz and Szakmary, 1994) for applications to commodity futures markets. More recently, Figuerola-Ferretti and Gonzalo (2010) (FG hereafter) develop an equilibrium model with finite elasticity of supply of arbitrage services and endogenously modeled convenience yields, and show how this model implies cointegration between log spot prices, s_t, and log futures prices, f_t. They also show how the cointegrated vector autoregressive (CVAR) model of Johansen (1995) is then a natural empirical model within which to analyze the pair (s_t, f_t) and, in addition, how it provides a simple way of determining which of the two prices is long-run dominant in the price discovery process.

Recently, much research has focused on the potential presence of fractional integration (or long memory) in the equilibrium relation between spot and futures prices. Some definitions of fractional differences and fractional integration are given in the Appendix. Specifically, with respect to futures markets, recent studies have found evidence of fractional integration in the forward premium, i.e. in the difference between the spot and futures prices; see, e.g., Baillie and Bollerslev (1994), Lien and Tse (1999), Maynard and Phillips (2001), Kellard and Sarantis (2008), and Coakley, Dollery, and Kellard (2011). Fractional integration in the forward premium thus implies fractional cointegration between futures and spot prices in the sense that, although the prices themselves are I(1), the spread between the two prices is fractionally integrated of a lower order. This generalizes the usual notion of cointegration where the spread would be I(0).

The main contribution of this paper is the empirical analysis of price discovery using the recently developed fractionally cointegrated vector autoregressive (FCVAR) model of Johansen (2008) and Johansen and Nielsen (2012), which provides a natural methodology for the analysis of multiple fractional time series variables by generalizing the well-known CVAR model of Johansen (1995) to fractional time series processes. The empirical analysis using the FCVAR model is justified from economic theory based on a variation of the equilibrium model developed by FG in which spot and futures prices are fractionally cointegrated.

This fractionally cointegrated equilibrium model is suggested by Dolatabadi, Nielsen, and Xu (2014) (henceforth DNX), and is briefly reviewed in Section 2.3 below. DNX apply this model to analyze the fractional cointegration relationship empirically using the FG data set within the FCVAR model. The main finding from that paper is that, when allowing for the possibility that the linear combination of s_t and f_t is fractionally integrated, there is more support in the data for a $(1, -1)$ cointegration vector. That is, there is more support for stationarity of the spread, $s_t - f_t$, and less support for long-run backwardation or contango compared to the analysis in FG based on the CVAR model. The present paper then adds to this empirical work by analyzing price discovery using the newly developed FCVAR model.

The FCVAR model has many advantages when estimating a system of possibly cointegrated fractional time series variables. The flexibility of the model permits one to determine the cointegrating rank, or number of equilibrium relations, via statistical tests and to jointly estimate the adjustment coefficients and the cointegrating relations, while accounting for the short-run dynamics. These features each bear some relevance to our research question. For example, the cointegrating rank is the number of long-run equilibria that exist between the spot and futures prices, and the cointegrating relations themselves are the linear combinations of these variables that form a stationary equilibrium. The most important parameters for our particular analysis are the adjustment coefficients, which tell us how the variables adjust to changes in the equilibrium, and these are

hence informative about price discovery.

The asymptotic theory for estimation and inference in the FCVAR model was developed recently in a series of papers by Johansen and Nielsen (2010, 2012, 2014). They prove asymptotic distribution results for the maximum likelihood estimators and for the likelihood ratio tests for cointegration rank. Furthermore, Nielsen and Morin (2014) provide an accompanying Matlab package for calculation of estimators and test statistics, and MacKinnon and Nielsen (2014) provide accompanying computer programs for calculation of P values and critical values for the cointegration rank tests. Taken together, these contributions imply that the FCVAR framework is now ready to be fully applied empirically. Finally, the possibility of deterministic trends in the observed variables is analyzed by DNX, who provide representation theory for this extension of the basic FCVAR model, and we will also entertain the possibility of deterministic trends in our empirical models.

In our empirical analysis, we apply the FCVAR model to the dataset from FG which consists of daily observations on spot and forward prices in five commodity markets for non-ferrous metals (aluminium, copper, lead, nickel, and zinc) from January 1989 to October 2006. In all markets the spot and futures prices are cointegrated, and for all markets except copper the fractional integration parameter is highly significant, showing that the usual non-fractional model is not appropriate. When allowing for fractional integration in the long-run equilibrium relations, fewer lags are needed in the autoregressive formulation compared to the non-fractional model, further stressing the usefulness of the fractional model.

With respect to (long-run) price discovery, we find that the spot market is more dominant in the discovery process compared to the results from the non-fractional model. Specifically, using standard likelihood ratio tests, we do not reject the hypothesis that price discovery takes place exclusively in the spot (futures) market for copper, lead, and zinc (aluminium and nickel). In the non-fractional CVAR model, for comparison, FG find that price discovery takes place exclusively in the spot (futures) market for lead (aluminium, nickel, and zinc), while their results for copper are inconclusive.

The remainder of the paper is organized as follows. Section 2 describes the equilibrium model in FG and the extension that results in fractional cointegration. Section 3 presents the empirical methodology based on the FCVAR model and links it to the analysis of price discovery. In Section 4 we describe the data and empirical results and Section 5 concludes. Some definitions related to fractional integration are given in the appendix.

2 Theoretical framework

The theoretical framework for the dynamics of spot and futures prices is the equilibrium model developed by FG which in turn builds on Garbade and Silber (1983). In this section we first briefly review the FG model by presenting the two cases of their model separately: (i) infinite elasticity of supply of arbitrage services and (ii) finite elasticity of supply of arbitrage services. The third subsection then describes an extension of their model that will establish a link to the FCVAR model described in Section 3.

2.1 FG equilibrium model of spot and futures prices with infinite elasticity of supply of arbitrage services

We begin with the following set of rather standard market conditions, which we collectively refer to as Assumption A.

A.1 No taxes or transaction costs.

A.2 No limitations on borrowing.

A.3 No costs other than financing a futures position (short or long) and storage costs.

A.4 No limitations on short sale in the spot market.

Let s_t and f_t denote the log-spot price of a commodity in period t and the contemporaneous log-futures price for a one-period-ahead futures contract, and let r_t and c_t denote the continuously compounded interest rate and storage cost applicable to that period. The following conditions, collectively referred to as Assumption B, are made in reference to the time-series behavior of these variables.

B.1 $r_t = \bar{r} + u_{rt}$, where $\bar{r}$ denotes the mean of r_t and u_{rt} denotes an I(0) process with mean zero and finite positive variance.

B.2 $c_t = \bar{c} + u_{ct}$, where $\bar{c}$ denotes the mean of c_t and u_{ct} denotes an I(0) process with mean zero and finite positive variance.

B.3 Δs_t is an I(0) process with mean zero and finite positive variance.

Under Assumption A, no-arbitrage equilibrium conditions imply

$$f_t = s_t + r_t + c_t, \tag{1}$$

so that, imposing also Assumption B,

$$f_t - s_t = \bar{r} + \bar{c} + u_{rt} + u_{ct}, \tag{2}$$

which implies that s_t and f_t are both clearly I(1) and cointegrate to I(0) with cointegration vector $(1, -1)$.

2.2 FG equilibrium model of spot and futures prices with finite elasticity of supply of arbitrage services

Finite elasticity of arbitrage services reflects the existence of factors such as basis risk, convenience yields, constraints on storage space and other factors that make arbitrage transactions risky. FG focus on convenience yield, i.e. the benefit associated with storing the commodity instead of holding the futures contract (Kaldor, 1939). More generally, by the definition of Brennan and Schwartz (1985), convenience yield is "the flow of services that accrues to an owner of the physical commodity but not to an owner of a contract for future delivery of the commodity". Accordingly, backwardation is defined by FG as "the present value of the marginal convenience yield of the commodity inventory". When this is negative, the market is said to be in contango.

Denoting the convenience yield by y_t, the no-arbitrage condition (1) is modified to

$$f_t + y_t = s_t + r_t + c_t. \tag{3}$$

FG approximate y_t by a linear combination of s_t and f_t, i.e. $y_t = \gamma_1 s_t - \gamma_2 f_t$ with $\gamma_i \in (0, 1)$ for $i = 1, 2$. Imposing Assumption B then implies the equilibrium condition

$$s_t = \beta_2 f_t + \beta_3 + u_{rt} + u_{ct}, \tag{4}$$

where β_2 and β_3 are simple functions of the model parameters. In particular, β_2 can take three different values:

(i) $\beta_2 > 1$: there is long-run backwardation ($s_t > f_t$).
(ii) $\beta_2 < 1$: there is long-run contango ($s_t < f_t$).
(iii) $\beta_2 = 1$: there is neither backwardation nor contango in the long run.

That is, the FG model admits the (empirically warranted) theoretical possibility of having a cointegration coefficient β_2 different from unity. We next describe a simple modification which will link this theoretical framework to the FCVAR model.

2.3 Fractionally cointegrated equilibrium model

We propose a simple variation of the FG model described in the previous subsection. Specifically, we replace Assumption B by the following conditions, collectively referred to as Assumption C.

C.1 $r_t = \bar{r} + v_{rt}$, where $\bar{r}$ denotes the mean of r_t and v_{rt} denotes an I$(1-b)$ process with $b > 1/2$, mean zero, and finite positive variance.

C.2 $c_t = \bar{c} + v_{ct}$, where $\bar{c}$ denotes the mean of c_t and v_{ct} denotes an I$(1-b)$ process with $b > 1/2$, mean zero, and finite positive variance.

C.3 Δs_t is an I(0) process with mean μ and finite positive variance.

Conditions C.1 and C.2 generalize B.1 and B.2 to fractionally integrated interest rates and storage costs. While storage costs are basically unobserved, interest rates are observed and are typically not found to be I(0). The assumption that both interest rates and storage costs have the same order of fractional integration, i.e. that both v_{rt} and v_{ct} are I$(1-b)$, is made only to simplify notation. If, instead, the two b parameters were different, then the b parameter in the following analysis would be replaced simply by the minimum of the two. The assumption that $b > 1/2$ ensures that the processes v_{rt} and v_{ct} are stationary, since then $1 - b < 1/2$. Finally, condition C.3 allows a possible drift in spot prices (when $\mu \neq 0$), which appears reasonable from an empirical point of view. For example, it is not at all clear from Figure 1, which presents the data series analyzed in the empirical section below, whether the series considered in our empirical analysis are drift-less, and hence it seems prudent to allow a possible drift rather than to rule it out a priori.

Imposing Assumption C instead of Assumption B, the equilibrium condition (4) becomes

$$s_t = \beta_2 f_t + \beta_3 + v_{rt} + v_{ct}. \tag{5}$$

Thus, the modification of the FG model with Assumption C instead of Assumption B implies the same cointegration vector, but differs from the FG model in that the long-run equilibrium errors are fractionally integrated of order $1 - b$ rather than I(0). That is, s_t and f_t are fractionally cointegrated.

3 Econometric model

Based on the above economic model with fractional cointegration, our empirical analysis uses the FCVAR model, see Johansen (2008) and Johansen and Nielsen (2010, 2012, 2014). The model is a generalization of Johansen's (1995) CVAR model to allow for fractional processes of order d that cointegrate to order $d - b$. In Sections 3.1 and 3.2 we review the FCVAR model and the extension from DNX to accommodate a linear time trend (drift) in the data. Section 3.3 then moves on to discuss the analysis of price discovery within the FCVAR.

3.1 The FCVAR model and interpretation of the parameters

To introduce the FCVAR model, we begin with the well-known, non-fractional, CVAR model. Let $Y_t, t = 1, \ldots, T$, be a p-dimensional I(1) time series. Then the CVAR model is

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta Y_{t-i} + \varepsilon_t = \alpha\beta' L Y_t + \sum_{i=1}^{k} \Gamma_i \Delta L^i Y_t + \varepsilon_t, \tag{6}$$

where, as usual, ε_t is p-dimensional independent and identically distributed with mean zero and covariance matrix Ω. The simplest way to derive the FCVAR model is to replace the difference and lag operators, Δ and L, in (6) by their fractional counterparts, Δ^b and $L_b = 1 - \Delta^b$, respectively. We then obtain

$$\Delta^b Y_t = \alpha\beta' L_b Y_t + \sum_{i=1}^{k} \Gamma_i \Delta^b L_b^i Y_t + \varepsilon_t, \tag{7}$$

which we apply to $Y_t = \Delta^{d-b} X_t$ to obtain the FCVAR model,

$$\Delta^d X_t = \alpha\beta' L_b \Delta^{d-b} X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t. \tag{8}$$

The well-known parameters from the CVAR model have the usual interpretations also in the FCVAR model. In particular, α and β are $p \times r$ matrices with $0 \leq r \leq p$. The columns of β are the cointegrating vectors such that $\beta' X_t$ are the stationary combinations of the variables in the system, i.e. the long-run equilibrium relations. The coefficients in α are the adjustment or loading coefficients which represent the speed of adjustment towards equilibrium for each of the variables. The parameters Γ_i govern the short-run dynamics of the variables.

There are two additional parameters in the FCVAR model compared with the CVAR model. The parameter d denotes the fractional integration order of the observable time series. As would presumably be the case for most — if not all — financial assets, we assume in our study that $d = 1$ in accordance with Assumptions B.3 and C.3. That is, we consider $d = 1$ to be fixed and known, and therefore not estimated. The parameter b is estimated and determines the degree of fractional cointegration, i.e. the reduction in fractional integration order of $\beta' X_t$ compared to X_t itself. The relevant ranges for b are $(0, 1/2)$, in which case the equilibrium errors are fractional of order greater than $1/2$, and are thus non-stationary although mean reverting, and $(1/2, 1]$, in which case the equilibrium errors are fractional of order less than $1/2$ and are thus stationary. Note that, for $b = 1$, the FCVAR model reduces to the CVAR model, which is therefore nested in the FCVAR model as a special case.

Thus, the FCVAR model enables simultaneous modelling of the long-run equilibria (with tests to determine how many such equilibria exist), the adjustment responses to deviations from the equilibria, and the short-run dynamics of the system. In addition, the FCVAR model makes it possible to evaluate model fit, i.e. whether the assumptions underlying the asymptotic distribution theory are likely satisfied, by examining the model residuals using, for instance, tests for serial correlation.

In our empirical analysis, following Assumption C.3 in the theoretical model, we allow for a drift in the variables. That is, we want to model X_t by

$$X_t = \tau_1 \pi_t(1) + \tau_2 \pi_t(2) + X_t^0, \tag{9}$$

where X_t^0 is the FCVAR in (8) (with $d = 1$) and $\pi_t(\cdot)$ is defined in (25) in the appendix. Specifically, with $1_{\{A\}}$ denoting the indicator function of the event A, using $\pi_t(1) = 1_{\{t \geq 0\}}$ and $\pi_t(2) = (t + 1)1_{\{t \geq 0\}}$ is convenient in the mathematical derivations rather than using $(1, t)$ because $\Delta^b \pi_t(a) = \pi_t(a - b)$. In any case, these deterministic terms are such that the parameters τ_1 and τ_2 in (9) allow for a linear deterministic trend in X_t. However, the trend is not empirically warranted in the equilibrium relation (5), and consequently we impose the restriction $\beta' \tau_2 = 0$. The latter restriction implies, in particular, that $\beta' X_t = \beta' \tau_1 \pi_t(1) + \beta' X_t^0$, which is $\mathrm{I}(1 - b)$ under Assumption C and (8).

The representation theory for this extension of the basic FCVAR model is derived by DNX, where it is shown that it implies the error correction equation

$$\Delta X_t = \alpha L_b \Delta^{1-b} (\beta' X_t - \rho' \pi_t(1)) + \sum_{i=1}^{k} \Gamma_i \Delta L_b^i X_t + \xi \pi_t(1) + \varepsilon_t. \tag{10}$$

Here, ρ is interpreted as the mean of the stationary linear combinations, $\beta' X_t$, and ξ gives rise to a linear deterministic trend in the levels of the variables. In the terminology of Johansen (1995), (10) contains both a restricted constant, $\rho \pi_t(1)$, and an unrestricted constant, $\xi \pi_t(1)$.

3.2 Estimation and inference in the FCVAR model

Estimation and inference for the model is as discussed in Johansen and Nielsen (2012) and Nielsen and Morin (2014), with the latter providing Matlab computer programs for the calculation of estimators and test statistics. The programs can be applied also to our model with deterministic trends with only minor modifications. Specifically, for fixed b, (10) is estimated by reduced rank regression of ΔX_t on $L_b\Delta^{1-b}(X_t', -\pi_t(1))'$ corrected for $\{\Delta L_b^i X_t\}_{i=1}^k$ and $\pi_t(1)$. The resulting profile likelihood is then a function only of b which is maximized numerically.

It is important to note that the fractional difference operator, see (24), is defined in terms of an infinite series, but that any observed sample will include only a finite number of observations, thus prohibiting calculation of the fractional differences as defined. An assumption that would allow calculation of the fractional differences is that X_t were zero before the start of the sample. However, in our case, as would often be the case, we cannot reasonably make such an assumption. The bias introduced by making such an assumption to allow calculation of the fractional differences is analyzed by Johansen and Nielsen (2014) using higher-order expansions in a simpler model. Their analysis reveals several ways to alleviate this bias. In particular, they show, albeit in a simpler model, that this bias can be alleviated by splitting the observed sample into initial values to be conditioned upon and observations to include in the likelihood—similar to the way a sample is divided into k initial values and $T-k$ observations in estimation of AR(k) models to reduce conditional maximum likelihood estimation to least squares regression. In our empirical analysis, we follow Johansen and Nielsen (2012, 2014) and apply maximum likelihood inference conditional on initial values.

The asymptotic analysis in Johansen and Nielsen (2012) shows that the maximum likelihood estimator of $(b, \alpha, \Gamma_1, \ldots, \Gamma_k)$ is asymptotically normal, while the maximum likelihood estimator of (β, ρ) is asymptotically mixed normal when $b_0 > 1/2$ and asymptotically normal when $b_0 < 1/2$. The important implication is that asymptotic χ^2-inference can be conducted on the parameters $(b, \rho, \alpha, \beta, \Gamma_1, \ldots, \Gamma_k)$ using likelihood ratio (LR) tests.

We will test a number of interesting hypotheses on the model parameters in our empirical analysis. The general theory of hypothesis testing for the CVAR model (Johansen, 1995) carries over almost unchanged to the FCVAR model. In particular, the degrees of freedom is equal to the number of overidentifying restrictions under the null. Although counting the degrees of freedom is non-standard because of the normalization required to separately identify α and β, this is done in the same way for the FCVAR as for the CVAR model.

Specifically, in the empirical analysis, the main hypotheses of interest are hypotheses on α. These can be formulated as

$$\alpha = A\psi, \tag{11}$$

where the known $p \times m$ matrix A specifies the restriction(s) and ψ is an $m \times r$ matrix of free parameters with $m \geq r$. In (11) the same restriction is imposed on each column of α, as is relevant in our empirical analysis. The degrees of freedom of the test is $\text{df} = (p-m)r$.

We next show how price discovery can be analyzed in the FCVAR model and how hypotheses relevant to price discovery can be formulated in terms of (11).

3.3 Price discovery in the FCVAR model

We now move beyond describing the empirical model of Johansen and Nielsen (2012) and DNX, and discuss how to analyze price discovery within the FCVAR model based on the permanent-transitory (PT) decomposition of Gonzalo and Granger (1995) applied to the FCVAR model. An alternative to the PT decomposition is the information shares metric of Hasbrouck (1995), but as described in detail in FG, there is "a perfect link between an extended Garbade and Silber (1983) theoretical

model and the PT decomposition".

According to PT decomposition, the common permanent component of $X_t = (s_t, f_t)'$ is $W_t = \alpha_\perp' X_t$, where $\alpha_\perp$ is such that $\alpha_\perp'\alpha = \alpha'\alpha_\perp = 0$. This common permanent component constitutes the dominant price or the long-run market price, in the sense that information that does not affect W_t will not have a permanent effect on X_t. Thus, the key parameter $\alpha_\perp$ is a direct measure of the contribution of each market (spot and futures) to the price discovery process. Because linear hypotheses on $\alpha_\perp$ can be tested either directly on $\alpha_\perp$ or alternatively on α itself by considering the corresponding mirror hypothesis, it follows from the discussion in the previous subsection that LR tests of such hypotheses are simple and critical values can be taken from the χ^2-distribution.

For example, to test the hypothesis that price discovery is exclusively in the spot market, i.e. $\alpha_\perp = (a, 0)'$, we can equivalently test the mirror hypothesis $\mathscr{H}_\alpha^1 : \alpha = (0, \psi)'$. Similarly, to test the hypothesis that price discovery is exclusively in the futures market, i.e. $\alpha_\perp = (0, a)'$, we test the mirror hypothesis $\mathscr{H}_\alpha^2 : \alpha = (\psi, 0)'$. The tests are implemented as in (11) with, in the case of $\mathscr{H}_\alpha^1$,

$$\underset{p\times m}{A} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where $(p, m, r) = (2, 1, 1)$ and ψ is a freely varying scalar. The degrees of freedom for the test of $\mathscr{H}_\alpha^i, i = 1, 2$, is $\text{df} = (p - m)r = (2 - 1)1 = 1$.

An alternative, yet strongly related, interpretation of the coefficient α is that of an adjustment coefficient that measures how the previous periods disequilibrium error feeds into today's (fractional) changes in X_t. Under this interpretation, the natural question to ask about the adjustment coefficients is whether some coefficients in α are zeros as in $\mathscr{H}_\alpha^i, i = 1, 2$, in which case the variable in question is weakly (or long-run) exogenous for the parameters α and β. For example, under the hypothesis $\mathscr{H}_\alpha^1$, the parameter $\alpha_1 = 0$ such that spot prices do not react to the disequilibrium error, i.e. the transitory component, implying that spot prices are the main contributors to price discovery.

4 Data and empirical results

4.1 Data description

The data set used in the empirical analysis is the same as that in FG to facilitate comparison. The data set includes daily observations (business days only) from the London Metal Exchange on spot and 15-month forward prices for aluminium, copper, lead, nickel, and zinc for the period from January 1989 to October 2006. The sample period has approximately 4484 observations. The London Metal Exchange data has the advantage that there are simultaneous spot and forward prices, for fixed forward maturities, on every business day. The (logarithmic) data is shown in Figure 1. For details, see FG.

4.2 Model selection

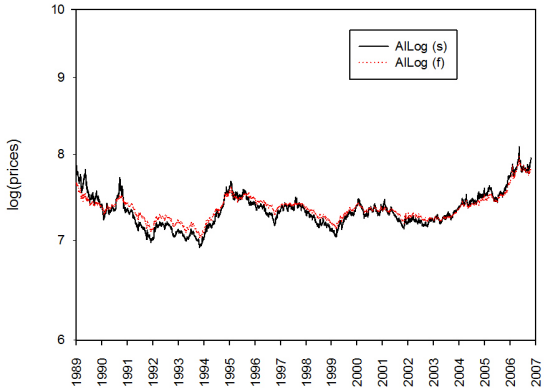
Before estimating the FCVAR model and testing the hypotheses of interest, $\mathscr{H}_\alpha^i, i = 1, 2$, there are several choices to be made. First of all, throughout we apply estimation conditional on $N = 21$ initial values.[1] Next, there are three additional elements in the specification of the FCVAR model: the lag length (k), the deterministic components, and the cointegration rank (r).

To select the lag length we carefully apply several sources of information, namely the Bayesian Information Criterion (BIC), the LR test statistic for significance of Γ_k, and univariate Ljung-Box Q tests (with $h = 10$ lags) for each of the two residual series. In each case these are based on

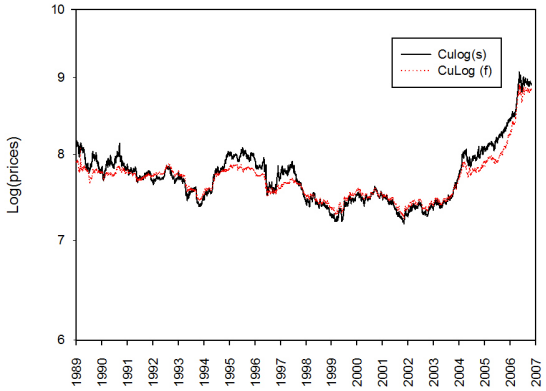
[1] Recall that there are 21 business days in a month. For robustness, we also computed the results with $N = 0$, $N = 5$, and $N = 63$ and these were very similar to those reported.

Figure 1: Daily spot and futures log-prices

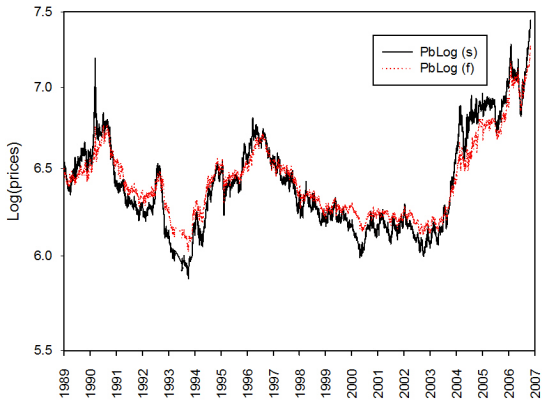
(a) aluminium



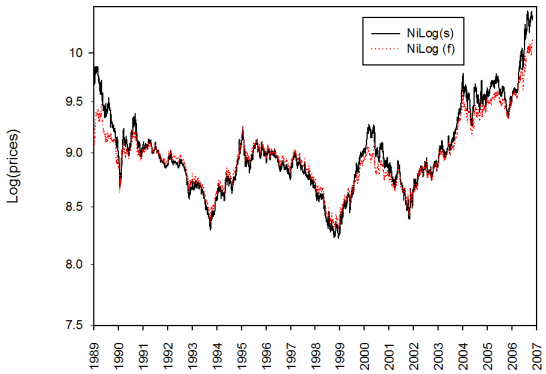
(b) copper



(c) lead



(d) nickel



(e) zinc

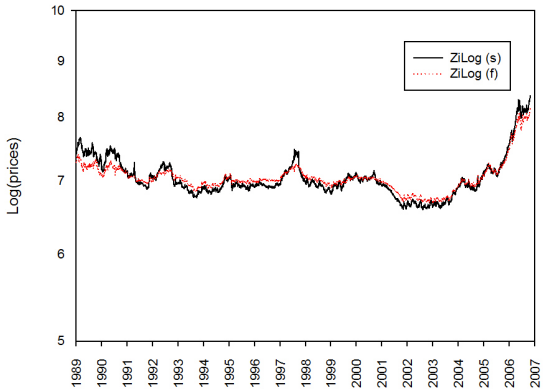


Table 1: Model specification results

	Aluminium	Copper	Lead	Nickel	Zinc (a)	Zinc (b)
CVAR lag length (FG)	17	14	15	18	16	16
FCVAR lag length, k	5	4	4	4	1	1
Cointegration rank, r	1	1	1	1	1	1
Restricted constant	yes	yes	yes	yes	yes	yes
Unrestricted constant	no	no	no	no	no	yes
$\mathscr{H}_b : b = 1$	no	yes	no	no	no	no
$\mathscr{H}_\beta : \beta = (1, -1)'$	yes	yes	no	no	yes	yes

Notes: The table summarizes the FCVAR model selection and hypothesis testing results from DNX for each of the five commodity time series. For comparison, the CVAR lag length from FG is included as well. For the hypothesis tests, a 'no' indicates that the hypothesis is rejected, while a 'yes' indicates non-rejection.

the model that includes all the deterministic components considered and has full rank $r = p$. We also examined the unrestricted estimates of b and β_2 which, when the lag length is misspecified, will sometimes be very far from what should be expected. In particular, there is an identification problem in the FCVAR model when the lag length is misspecified, which can result in, e.g., $\hat{b} = 0.05$ or similar, see Johansen and Nielsen (2010, Section 2.3) and Carlini and Santucci de Magistris (2014) for a theoretical discussion of this phenomenon. Thus, for each commodity, we first use the BIC as a starting point for the lag length, and from there we find the nearest lag length which satisfies the criteria (i) Γ_k is significant based on the LR test, (ii) the unrestricted estimates of b and β_2 are reasonable (very widely defined), and (iii) the Ljung-Box Q tests for serial correlation in the two residual series do not show signs of misspecification.

After deciding on the lag length, we need to select the deterministic components and the cointegrating rank (r). For the former, we will work under the maintained hypothesis that the restricted constant, $\rho\pi_t(1)$, is present based on the theoretical framework in Section 2. The selection of deterministic components thus comes down to the absence or presence of the unrestricted constant, $\xi\pi_t(1)$, i.e. the trend component. Because the limit distribution of the test of cointegrating rank depends on the actual cointegrating rank and on the presence or absence of the trend, we have to simultaneously decide the cointegration rank and whether or not to include the trend. A detailed discussion of how to test both hypothesis together can be found in Johansen (1995, pp. 170–174).

DNX analyze the same data set as we do, using also the FCVAR model, with the purpose of hypothesis testing on β. Their model specification results for each of the five metals are summarized in Table 1 and we will apply these as well. The first row shows the lag length from the CVAR analysis in FG for comparison. We find that, allowing for the possibility of fractional cointegration, fewer lags are needed to adequately model the data. In FG, they select $k = 17, 14, 15, 18, 16$ lags for aluminium, copper, lead, nickel, and zinc, respectively. As shown in the second row, we select $k = 5, 4, 4, 4, 1$ lags for the five metals. Thus, allowing for b to be fractional we can select a much smaller number of lags while maintaining white noise residuals, although part of the difference can likely be attributed to FG's use of the AIC for lag length selection. Next, the third row shows the chosen cointegration rank, where all five metals have $r = 1$ as expected from theory, and also in accordance with the CVAR analysis in FG.

For the selection of deterministics, shown in the next two rows, we choose the model with only a restricted constant for aluminium, copper, nickel, and lead. For zinc, however, the tests are inconclusive; specifically the model with only a restricted constant term is rejected at 15% level and its estimate of α_1 has the wrong sign such that spot prices do not move towards the equilibrium. Hence, for the remainder of the analysis we present two sets of results for zinc: with a restricted

Table 2: FCVAR results for aluminium

Unrestricted $\hat{\alpha} = \begin{bmatrix} -0.036 & 0.015 \\ (0.017) & (0.013) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 0.294 & 0.706 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$	$\mathscr{H}_\beta \cap \mathscr{H}_\alpha^2$
df	1	1	2
LR	6.434	1.812	5.564
P value	0.011	**0.178**	**0.062**

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} -0.041 \\ 0.000 \end{bmatrix} z_t + \sum_{i=1}^{5} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + \hat{\varepsilon}_t \tag{12}$$

$$\hat{b} = 0.754, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.641)}{7.875}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.729)}{6.959}, \quad \log(\mathscr{L}) = 30430.601$$

Equilibrium relation:

$$s_t = -0.038 + f_t + z_t \tag{13}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4487$.

constant only (denoted zinc (a)) and with both a restricted and an unrestricted constant term (denoted zinc (b)).

The final two rows of Table 1 shows the results of LR tests of the hypotheses $\mathscr{H}_b : b = 1$ and $\mathscr{H}_\beta : \beta = (1, -1)'$, respectively, with a 'no' indicating that the hypothesis is rejected and a 'yes' indicating non-rejection. With the exception of copper, the test of $\mathscr{H}_b : b = 1$ rejects (very strongly) for all metals. This finding shows, again with the exception of copper, that the FCVAR model is more appropriate for these data sets than the CVAR model (which has exactly $b = 1$). The last hypothesis in the table, $\mathscr{H}_\beta : \beta = (1, -1)'$, is that there is no long-run backwardation or contango, and this hypothesis is rejected only for lead and nickel.

4.3 Empirical results

We now move to the main empirical analysis, namely that of price discovery for the spot and futures commodity markets using the FCVAR model. The empirical results are presented in Tables 2-7 with one table for each of the five metals (two for zinc). Each table is laid out in the same way with four panels. The first panel presents the unrestricted estimate of α with standard errors in parentheses along with the unrestricted estimate of $\alpha_\perp$. The latter is normalized such that the two elements add to one, and can therefore be interpreted as proportions of spot and futures price contributions to the price discovery process. The second panel is a subtable with the results of the tests of the hypotheses $\mathscr{H}_\alpha^i, i = 1, 2$, and also joint tests of these along with non-rejected tests from Table 1. In the third and fourth panels of each table, the estimation results for the FCVAR model are presented with all non-rejected hypotheses imposed. In addition, Ljung-Box Q tests for serial correlation up to lag $h = 10$ are also presented for each of the two residual series.

The results for aluminium, presented in Table 2, show that the α parameter for spot prices appears significant while that for the futures prices does not. That is, futures prices may be weakly exogenous. In terms of $\alpha_\perp$, this suggests that the first element of $\alpha_\perp$ may be zero such that spot prices do not contribute to the long-run market price, i.e. the futures price is dominant in the price

Table 3: FCVAR results for copper

Unrestricted $\hat{\alpha} = \begin{bmatrix} -0.001 & 0.005 \\ (0.003) & (0.003) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 0.833 & 0.167 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$	$\mathscr{H}_b \cap \mathscr{H}_\beta \cap \mathscr{H}_\alpha^1$
df	1	1	3
LR	0.214	3.400	0.412
P value	**0.643**	0.065	**0.937**

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} 0.000 \\ -0.004 \end{bmatrix} z_t + \sum_{i=1}^{4} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + \hat{\varepsilon}_t \tag{14}$$

$$\hat{b} = 1, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.547)}{8.842}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.925)}{4.440}, \quad \log(\mathscr{L}) = 27554.114$$

Equilibrium relation:

$$s_t = 0.008 + f_t + z_t \tag{15}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4496$.

discovery process. This finding is partially confirmed by the reported estimate of $\alpha_\perp$, which shows that futures prices constitute 71% of the price discovery process, and is a slightly lower proportion than in FG, where it was estimated at 91%.

This finding is also supported by the hypothesis tests, where $\mathscr{H}_\alpha^2$ is not rejected (P value of 0.178) and the joint hypothesis $\mathscr{H}_\beta \cap \mathscr{H}_\alpha^2$ is not rejected.[2] This implies that the long-run equilibrium is $\beta = (1, -1)'$ such that there is neither backwardation nor contango in the long run, and, for our purposes more importantly, that price discovery is exclusively in the futures market. The latter finding is also in line with the results in FG and, e.g., Figuerola-Ferretti and Gilbert (2005).

The last two panels of the table show the restricted final model for aluminium, including serial correlation tests for the residuals that show no signs of serial correlation, thus indicating that the model is well specified. We note, in particular, that the fractional parameter is estimated at $\hat{b} = 0.754$ which implies that $\beta' X_t$ is a stationary process with long memory, i.e. with integration order in the range $[0, 1/2)$.

We next move to Table 3 which shows the results for copper. These are rather different from the results for the other metals, in the sense that copper is the only metal for which $\mathscr{H}_b : b = 1$ is not rejected (see Table 1). It therefore appears that a CVAR is in fact adequate to model copper spot and futures prices. The unrestricted estimate $\hat{\alpha}$ shows that both coefficients are, strictly speaking, insignificant (at the 5% level), although the coefficient on spot prices more so. In terms of price discovery proportions, these are estimated at 83% and 17% for spot and futures prices, respectively. These findings are reflected in the hypothesis tests, which also show that the joint hypothesis $\mathscr{H}_b \cap \mathscr{H}_\beta \cap \mathscr{H}_\alpha^1$, i.e. the hypothesis that $b = 1, \beta = (1, -1)'$, and $\alpha_1 = 0$, is strongly supported by the data with a P value of 0.937.

Thus, for copper we find that the spot market is dominant in the price discovery process.

[2]Note that, as in e.g. Johansen (1995), we test hypotheses on α after imposing any non-rejected hypotheses on β, since β is estimated super-consistently.

Table 4: FCVAR results for lead

Unrestricted $\hat{\alpha} = \begin{bmatrix} -0.007 & 0.045 \\ (0.015) & (0.015) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 0.865 & 0.135 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$
df	1	1
LR	0.224	14.812
P value	**0.635**	0.000

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} 0.000 \\ 0.047 \end{bmatrix} z_t + \sum_{i=1}^{4} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + \hat{\varepsilon}_t \tag{16}$$

$$\hat{b} = 0.743, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.319)}{11.510}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.061)}{17.650}, \quad \log(\mathscr{L}) = 25662.120$$

Equilibrium relation:

$$s_t = -1.066 + 1.160 f_t + z_t \tag{17}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4482$.

These results are comparable to, although slightly more clear-cut than, those in FG. In their analysis the price discovery proportions are estimated at 58% and 42%, respectively, and both hypotheses $\mathscr{H}_\alpha^i, i = 1, 2$, have high P values (0.384 and 0.123, respectively), and hence their results are somewhat inconclusive. However, $\mathscr{H}_\alpha^1$ does have the higher P value, and hence they also find some support for the notion that the spot price is the more dominant price.

Next, in the cases of lead and nickel in Tables 4 and 5, we find from both the unrestricted $\hat{\alpha}$ and $\hat{\alpha}_\perp$ estimates and from the hypothesis tests that price discovery takes place in the spot market for lead and in the futures market for nickel. This is in accordance with FG, where strong support is also found for price discovery in the spot market for lead and in the futures market for nickel. In the estimation of the final restricted models for lead and nickel, we find again that there are no signs of model misspecification based on the Ljung-Box serial correlation tests on the residuals, and that $\hat{b}$ is different from one (and significantly so, see Table 1).

In Tables 6 and 7 the results for the two different model specifications for zinc are presented. Table 6 shows results for the model specification with a restricted constant only, and Table 7 shows results for the model specification that includes both the restricted and the unrestricted constant terms. The results are very similar, but differ in one important aspect. In Table 6 the unrestricted estimate $\hat{\alpha}_1$ is such that spot prices do not in fact adjust towards the equilibrium (because it is positive rather than negative), even though $\hat{\alpha}_2$ is such that the pair (s_t, f_t) still jointly moves towards equilibrium. For this reason, we prefer the specification (b) with both constant terms. In the latter model in Table 7, the unrestricted $\hat{\alpha}$ and $\hat{\alpha}_\perp$ estimates show that spot prices dominate with 97.1% of the price discovery process.

In terms of the tests of $\mathscr{H}_\alpha^i, i = 1, 2$ and the restricted models, the results in Tables 6 and 7 are very similar. In both tables the hypothesis $\mathscr{H}_\alpha^1$ and the joint hypothesis $\mathscr{H}_\beta \cap \mathscr{H}_\alpha^1$ are not rejected, and therefore spot prices appear to be dominant in the price discovery process for zinc regardless of which of the two model specifications are applied. For this particular metal it therefore appears that the results using the new FCVAR framework are in contrast to those obtained by FG using

Table 5: FCVAR results for nickel

Unrestricted $\hat{\alpha} = \begin{bmatrix} -0.059 & 0.035 \\ (0.038) & (0.025) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 0.372 & 0.628 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$
df	1	1
LR	4.982	1.996
P value	0.025	**0.157**

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} -0.107 \\ 0.000 \end{bmatrix} z_t + \sum_{i=1}^{4} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + \hat{\varepsilon}_t \tag{18}$$

$$\hat{b} = 0.607, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.635)}{7.931}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.884)}{5.092}, \quad \log(\mathscr{L}) = 25440.643$$

Equilibrium relation:

$$s_t = -2.222 + 1.244 f_t + z_t \tag{19}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4484$.

the CVAR (where it was found that futures prices were dominant in the price discovery for zinc).

In summary, using the new FCVAR framework, the results from the estimation of α and hypothesis testing on α show that for aluminium and nickel, futures prices are dominant in the price discovery process, whereas for copper, lead, and zinc, spot prices dominate the price discovery process. For aluminium, lead, and nickel, these results correspond to the findings in FG using the CVAR. However, for copper our findings are slightly more clear-cut with P values of 0.643 and 0.065 for $\mathscr{H}_\alpha^1$ and $\mathscr{H}_\alpha^2$, respectively, compared to those in FG, where the two hypotheses have P values of 0.384 and 0.123, respectively. Finally, for zinc our conclusion that the spot price is dominant is reversed compared with FG who find that the futures price is dominant. Overall, it therefore seems that our analysis using the new FCVAR model gives slightly more support to price discovery in the spot market compared with the analysis using the CVAR model in FG.

5 Concluding remarks

The present paper applies the recently developed fractionally cointegrated vector autoregressive (FCVAR) model to five different commodities (aluminium, copper, lead, nickel, and zinc) to analyze price discovery in spot and futures markets. The empirical model has economic foundation using a variation of the economic equilibrium model of FG to capture the existence of backwardation and contango in the long-run equilibrium relationship of spot and futures prices.

In the empirical analysis, the data set from FG is used to facilitate comparison with their empirical analysis using the non-fractional cointegrated VAR model. The results show that spot and futures prices for all metals are cointegrated, and—with the exception of copper—the cointegration is of the fractional type, where the long-run equilibrium errors are stationary and fractionally integrated, i.e., have long memory. When taking this into account with the FCVAR model rather than the CVAR model, two important implications arise: (i) fewer lags are needed in the autoregressive augmentation of the model and (ii) there appears to be slightly more support for price discovery in the spot market when using the FCVAR model, with the latter finding mainly for zinc. An

Table 6: FCVAR results for zinc (a)

Unrestricted $\hat{\alpha} = \begin{bmatrix} 0.015 & 0.173 \\ (0.062) & (0.085) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 1.095 & -0.095 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$	$\mathscr{H}_\beta \cap \mathscr{H}_\alpha^1$
df	1	1	2
LR	0.048	6.852	0.068
P value	**0.826**	0.008	**0.966**

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} 0.000 \\ 0.156 \end{bmatrix} z_t + \sum_{i=1}^{1} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + \hat{\varepsilon}_t \tag{20}$$

$$\hat{b} = 0.349, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.046)}{18.539}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.210)}{13.247}, \quad \log(\mathscr{L}) = 27810.365$$

Equilibrium relation:

$$s_t = -0.240 + f_t + z_t \tag{21}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4484$.

interesting topic for future research would be to reconcile the finding of more price discovery in the spot market with economic fundamentals for the particular metals involved.

Appendix: fractional differencing and fractional integration

The fractional (or fractionally integrated) time series models are based on the fractional difference operator,

$$\Delta^d X_t = \sum_{n=0}^{\infty} \pi_n(-d) X_{t-n}, \tag{24}$$

where the fractional coefficients $\pi_n(u)$ are defined in terms of the binomial expansion $(1-z)^{-u} = \sum_{n=0}^{\infty} \pi_n(u) z^n$, i.e.,

$$\pi_n(u) = \frac{u(u+1)\cdots(u+n-1)}{n!}. \tag{25}$$

For details and many intermediate results regarding this expansion and the fractional coefficients, see, e.g., Johansen and Nielsen (2014, Appendix A). Efficient calculation of fractional differences, which we apply in our estimation, is discussed in Jensen and Nielsen (2014).

With the definition of the fractional difference operator in (24), a time series X_t is said to be fractional of order d, denoted $X_t \in \mathrm{I}(d)$, if $\Delta^d X_t$ is fractional of order zero, i.e. if $\Delta^d X_t \in \mathrm{I}(0)$. The I(0) property can be defined in the frequency domain as having spectral density that is finite and non-zero near the origin or in terms of the linear representation coefficients if the sum of these is non-zero and finite, see, e.g., Johansen and Nielsen (2012). An example of a process that is I(0) is the stationary and invertible ARMA model.

References

Baillie, R. T. and T. Bollerslev (1994). The long memory of the forward premium. *Journal of International Money and Finance 13*, 565–571.

Table 7: FCVAR results for zinc (b)

Unrestricted $\hat{\alpha} = \begin{bmatrix} -0.003 & 0.102 \\ (0.045) & (0.050) \end{bmatrix}'$ and $\hat{\alpha}_\perp = \begin{bmatrix} 0.971 & 0.029 \end{bmatrix}'$

Hypothesis tests:

	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$	$\mathscr{H}_\beta \cap \mathscr{H}_\alpha^1$
df	1	1	2
LR	0.004	5.916	0.466
P value	**0.949**	0.015	**0.792**

Restricted model:

$$\Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} = \Delta^{1-\hat{b}} \begin{bmatrix} 0.000 \\ 0.131 \end{bmatrix} z_t + \sum_{i=1}^{1} \Gamma_i L_{\hat{b}}^i \Delta \begin{bmatrix} s_t \\ f_t \end{bmatrix} + 10^{-3} \begin{bmatrix} 0.204 \\ 0.308 \end{bmatrix} + \hat{\varepsilon}_t \tag{22}$$

$$\hat{b} = 0.382, \quad Q_{\hat{\varepsilon}_1}(10) = \underset{(0.043)}{18.737}, \quad Q_{\hat{\varepsilon}_2}(10) = \underset{(0.184)}{13.754}, \quad \log(\mathscr{L}) = 27811.208$$

Equilibrium relation:

$$s_t = -0.035 + f_t + z_t \tag{23}$$

Notes: The table shows unrestricted α and $\alpha_\perp$ estimates from the FCVAR as well as hypothesis tests and estimation results with non-rejected hypotheses imposed. In the hypothesis tests subtable, P values in bold denote hypotheses that are imposed in the restricted model. Standard errors are in parentheses below $\hat{\alpha}$ and P values are in parentheses below $Q_{\hat{\varepsilon}_i}$, which is the Ljung-Box Q test for serial correlation in the i'th residual. The sample size is $T = 4484$.

Brennan, M. J. and E. S. Schwartz (1985). Evaluating natural resource investments. *Journal of Business 58*, 135–157.

Carlini, F. and P. Santucci de Magistris (2014). On the identification of fractionally cointegrated VAR models with the $\mathcal{F}(d)$ condition. CREATES research paper 2013-44, Aarhus University.

Chowdhury, A. R. (1991). Futures market efficiency: evidence from cointegration tests. *Journal of Futures Markets 11*, 577–589.

Coakley, J., J. Dollery, and N. Kellard (2011). Long memory and structural breaks in commodity futures markets. *Journal of Futures Markets 31*, 1076–1113.

Dolatabadi, S., M. Ø. Nielsen, and K. Xu (2014). A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. QED working paper 1327, Queen's University.

Figuerola-Ferretti, I. and C. L. Gilbert (2005). Price discovery in the aluminium market. *Journal of Futures Markets 25*, 967–988.

Figuerola-Ferretti, I. and J. Gonzalo (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics 158*, 95–107.

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

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

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

Jensen, A. N. and M. Ø. Nielsen (2014). A fast fractional difference algorithm. *Journal of Time Series Analysis*, forthcoming.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. New York: Oxford University Press.

Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory 24*, 651–676.

Johansen, S. and M. Ø. Nielsen (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics 158*, 51–66.

Johansen, S. and M. Ø. Nielsen (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica 80*, 2667–2732.

Johansen, S. and M. Ø. Nielsen (2014). The role of initial values in nonstationary fractional time series models. QED working paper 1300, Queen's University.

Kaldor, N. (1939). Speculation and economic stability. *Review of Economic Studies 7*, 1–27.

Kellard, N., P. Newbold, T. Rayner, and C. Ennew (1999). The relative efficiency of commodity futures markets. *Journal of Futures Markets 19*, 413–432.

Kellard, N. and N. Sarantis (2008). Can exchange rate volatility explain persistence in the forward premium? *Journal of Empirical Finance 15*, 714–728.

Krehbiel, T. and L. C. Adkins (1993). Cointegration tests of the unbiased expectations hypothesis in metal markets. *Journal of Futures Markets 13*, 753–763.

Lien, D. and Y. K. Tse (1999). Fractional cointegration and futures hedging. *Journal of Futures Markets 19*, 457–474.

MacKinnon, J. G. and M. Ø. Nielsen (2014). Numerical distribution functions of fractional unit root and cointegration tests. *Journal of Applied Econometrics 29*, 161–171.

Maynard, A. and P. C. B. Phillips (2001). Rethinking an old empirical puzzle: econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics 16*, 671–708.

Nielsen, M. Ø. and L. Morin (2014). FCVARmodel.m: a Matlab software package for estimation and testing in the fractionally cointegrated VAR model. QED working paper 1273, Queen's University.

Quan, J. (1992). Two-step testing procedure for price discovery role of futures prices. *Journal of Futures Markets 12*, 139–149.

Schroeder, T. C. and B. K. Goodwin (1991). Price discovery and cointegration for live hogs. *Journal of Futures Markets 11*, 685–696.

Schwarz, T. V. and A. C. Szakmary (1994). Price discovery in petroleum markets: arbitrage, cointegration, and the time interval of analysis. *Journal of Futures Markets 14*, 147–167.

Stein, J. L. (1961). The simultaneous determination of spot and futures prices. *American Economic Review 51*, 1012–1025.

Working, H. (1948). Theory of the inverse carrying charge in futures markets. *Journal of Farm Economics 30*, 1–28.

Research Papers 2013

2014-07: Markku Lanne and Jani Luoto: Noncausal Bayesian Vector Autoregression

2014-08: Timo Teräsvirta and Yukai Yang: Specification, Estimation and Evaluation of Vector Smooth Transition Autoregressive Models with Applications

2014-09: A.S. Hurn, Annastiina Silvennoinen and Timo Teräsvirta: A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market

2014-10: Marcelo Fernandes and Cristina M. Scherrer: Price discovery in dual-class shares across multiple markets

2014-11: Yukai Yang: Testing Constancy of the Error Covariance Matrix in Vector Models against Parametric Alternatives using a Spectral Decomposition

2014-12: Stefano Grassi, Nima Nonejad and Paolo Santucci de Magistris: Forecasting with the Standardized Self-Perturbed Kalman Filter

2014-13: Hossein Asgharian, Charlotte Christiansen and Ai Jun Hou: Macro-Finance Determinants of the Long-Run Stock-Bond Correlation: The DCC-MIDAS Specification

2014-14: Mikko S. Pakkanen and Anthony Réveillac: Functional limit theorems for generalized variations of the fractional Brownian sheet

2014-15: Federico Carlini and Katarzyna Łasak: On an Estimation Method for an Alternative Fractionally Cointegrated Model

2014-16: Mogens Bladt, Samuel Finch and Michael Sørensen: Simulation of multivariate diffusion bridges

2014-17: Markku Lanne and Henri Nyberg: Generalized Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models

2014-18: Dragan Tevdovski: Extreme negative coexceedances in South Eastern European stock markets

2014-19: Niels Haldrup and Robinson Kruse: Discriminating between fractional integration and spurious long memory

2014-20: Martyna Marczak and Tommaso Proietti: Outlier Detection in Structural Time Series Models: the Indicator Saturation Approach

2014-21: Mikkel Bennedsen, Asger Lunde and Mikko S. Pakkanen: Discretization of Lévy semistationary processes with application to estimation

2014-22: Giuseppe Cavaliere, Morten Ørregaard Nielsen and A.M. Robert Taylor: Bootstrap Score Tests for Fractional Integration in Heteroskedastic ARFIMA Models, with an Application to Price Dynamics in Commodity Spot and Futures Markets

2014-23: Maggie E. C. Jones, Morten Ørregaard Nielsen and Michael Ksawery Popiel: A fractionally cointegrated VAR analysis of economic voting and political support

2014-24: Sepideh Dolatabadim, Morten Ørregaard Nielsen and Ke Xu: A fractionally cointegrated VAR analysis of price discovery in commodity futures markets