AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Dolatabadi, S., Narayan, P. K., Nielsen, M. Ø., & Xu, K. (2018). Economic significance of commodity return forecasts from the fractionally cointegrated VAR model. *Journal of Futures Markets*, 38(2), 219-242.
<https://doi.org/10.1002/fut.21866>

Publication metadata

Title:	Economic significance of commodity return forecasts from the fractionally cointegrated VAR model
Author(s):	Dolatabadi, S., Narayan, P. K., Nielsen, M. Ø., & Xu, K.
Journal:	<i>Journal of Futures Markets</i>
DOI/Link:	10.1002/fut.21866
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

Economic significance of commodity return forecasts from the fractionally cointegrated VAR model*

Sepideh Dolatabadi

Queen's University, Kingston, Ontario, Canada

Paresh Kumar Narayan

Deakin University, Burwood, Victoria, Australia

Morten Ørregaard Nielsen[†]

Queen's University, Kingston, Ontario, Canada, and CREATES, Aarhus, Denmark

Ke Xu

University of Victoria, Victoria, British Columbia, Canada

July 30, 2017

Abstract

We model and forecast commodity spot and futures prices using fractionally cointegrated vector autoregressive (FCVAR) models generalizing the well-known (non-fractional) CVAR model to accommodate fractional integration. In our empirical analysis to daily data on 17 commodity markets, the fractional model is statistically superior in terms of in-sample fit and out-of-sample forecasting. We analyze economic significance of the forecasts through dynamic (mean-variance) trading strategies, leading to statistically significant and economically meaningful profits in most markets. We generally find that the fractional model generates higher profits on average, especially in the futures markets.

Keywords: commodity markets, economic significance, forecasting, fractional cointegration, futures markets, price discovery, trading rule, vector error correction model.

JEL Classification: C32, G11.

1 Introduction

The forecastability of commodity market returns is a very active area of research in financial economics. In particular, recent research has shown that commodity spot and futures prices are fractionally cointegrated; see, *inter alia*, Baillie and Bollerslev (1994), Lien and Tse (1999), Maynard and Phillips (2001), Coakley, Dollery, and Kellard (2011), and Dolatabadi, Nielsen, and Xu (2016). An implication is that a fractionally cointegrated model may provide a better statistical fit for

*We are grateful to an anonymous referee, Peter Extercate, Michał Popiel, and participants at the Conference on Recent Developments in Financial Econometrics and Applications at Deakin University, Australia (December, 2014) for many useful comments. Financial support from 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) is gratefully acknowledged.

[†]Corresponding author. Mailing address: Department of Economics, Dunning Hall Room 307, Queen's University, 94 University Avenue, Kingston, Ontario, K7L 3N6, Canada. Phone: (+1) 613-533-2262. Email: mon@econ.queensu.ca

commodity prices and returns. Relatedly, the understanding of how commodity market return forecasts can be used to devise trading strategies appears still rather limited.

In this paper, we make two contributions to this literature. Our first contribution is methodological. We propose to model and forecast commodity spot and futures prices using the recently developed fractionally cointegrated vector autoregressive (FCVAR) model of Johansen (2008) and Johansen and Nielsen (2012). Specifically, we derive the best linear predictor for the FCVAR model and show that it takes a relatively simple form due to the autoregressive structure. We thus demonstrate how to forecast commodity spot and futures prices and returns based on the FCVAR model, and we evaluate these using statistical measures of forecast performance. Our second contribution is to investigate the economic significance of the FCVAR model forecasts through a dynamic trading strategy based on a portfolio of two assets with portfolio weights derived from a mean-variance utility function and from return forecasts. Throughout, we compare with forecasts from the more standard (non-fractional) cointegrated vector autoregressive (CVAR) model of Johansen (1995).

In our empirical analysis, we consider spot and futures prices of 17 commodities, and we demonstrate that the FCVAR model provides superior statistical in-sample fit compared with the more standard CVAR model. We also estimate price discovery from both models (e.g., Dolatabadi, Nielsen, and Xu, 2015, and references therein), which may be important from a forecasting point of view since historical information from the dominant market could be useful in forecasting prices and returns in the non-dominant market. In any case, both the FCVAR and CVAR models are joint models of the two prices series, and as such they automatically take into account the price discovery information in modeling and forecasting. With both the CVAR and FCVAR models we find that there is significant price discovery in both the spot and futures markets for many commodities, although the general tendency is that the futures market has a larger share of the price discovery process, as much theory predicts (e.g., Hasbrouck, 1995).

For the statistical forecast comparison we consider several forecast horizons and a variety of out-of-sample forecast evaluation metrics. The general finding is that the FCVAR model outperforms the CVAR model. Specifically, in terms of statistical tests of forecast superiority at the short horizon ($h = 1$), these favor the FCVAR model in almost all cases and are statistically significant at standard levels in most cases. At longer horizons, most statistical tests continue to favor the FCVAR model, although fewer are now statistically significant. Among those that are statistically significant for longer horizon forecasting ($h = 5$ or $h = 21$), 22 out of 23 favor the FCVAR model.

Finally, as an additional metric of forecast performance and comparison, we also examine the economic—as opposed to purely statistical—significance of return forecasts. We do this by investigating whether the return forecasts can generate significant excess returns when implemented in a dynamic portfolio trading strategy. For our main empirical analysis we find that using return forecasts from both FCVAR and CVAR models in simple mean-variance trading strategies leads to statistically significant and economically meaningful profits in most commodity markets, although there is much heterogeneity in profits across different markets. Furthermore, we find that profits are higher on average when based on forecasts from the FCVAR model rather than the CVAR model, especially in the futures markets where trading is also more practical.

Our finding that profits from commodity markets are statistically significant and economically meaningful is consistent with a broad range of studies which show, using different approaches, that commodity markets are profitable. For example, Miffre and Rallis (2007) and Narayan, Ahmed, and Narayan (2014) show profitability using technical and momentum trading strategies. However, given the profitability of these approaches, only limited focus has been on using a model-based forecasting approach to estimate profits (an exception is Narayan, Narayan, and Sharma, 2013).

In spite of the limited attention to model-based forecasting approaches, there is a clear acceptance of the fact that a forecasting based trading model that draws its profitability analysis from

a utility function, such as a mean-variance utility function, has theoretical appeal, see e.g. Marquering and Verbeek (2004) and Campbell and Thompson (2008). On the basis of this evidence, commodity markets are treated as an investment class. As the focus on theoretically motivated profitability analysis gains momentum, following, for example, the works mentioned above, the emphasis on and hence demand for appropriate forecasting models will increase.

We note from the outset that, although trading strategies based on commodity spot prices are not really feasible, because it would be too expensive to take possession of the commodity, we nonetheless consider simultaneous modeling of commodity spot and futures prices. In terms of applying these as forecasting models for futures returns, it has no relevance whether spot prices can be traded on or not, and hence this point is irrelevant for all our results regarding futures markets, futures price and return forecasting, and trading strategies involving commodity futures. For trading strategies involving commodity spot markets, these can still be considered a useful metric for comparison of forecast performance in terms of economic significance, even if the trading strategies are infeasible; a related point was also made in, e.g., Graham-Higgs, Rambaldi, and Davidson (1999), Wang (2000), and Narayan, Narayan, and Sharma (2013). Thus, even if portfolios involving commodity spot positions are infeasible, we consider such “artificial portfolios” as a means of forecast evaluation and comparison.

To demonstrate the robustness of our empirical results, our analysis is conducted with several different variations. First, we forecast returns over both short and long horizons. Second, we use several statistical out-of-sample forecast evaluation metrics. Third, when estimating profits using the mean-variance investor utility function, where the choice of the investor’s risk-aversion coefficient influences portfolio weights, we consider low, medium, and high risk-aversion investors. Fourth, we calculate Sharpe ratios. Finally, when estimating profits we consider several alternative restrictions on short-selling and leverage/borrowing. In general, all these results confirm (i) that portfolio excess returns are statistically different from zero and economically meaningful in many commodity markets, and (ii) that portfolio returns derived from CVAR and FCVAR model forecasts are similar, although the latter are slightly higher on average.

The remainder of the paper is organized as follows. The econometric model is explained in the next section, where, in particular, the best linear predictor is derived and forecasting is discussed in Section 2.3. The following section discusses the economic equilibrium model and shows how fractional cointegration can arise from an economic model which thus provides a link between economic theory and econometric modeling. In Section 4 we discuss the commodity data and conduct some preliminary data analysis. Section 5 contains the empirical results, and is divided into subsections on estimation, statistical forecast comparison, economic significance of forecasts, and robustness of economic significance. Finally, in Section 6 we provide some concluding remarks.

2 Econometric methodology: fractionally cointegrated VAR model

Our empirical analysis applies the FCVAR model, see Johansen (2008) and Johansen and Nielsen (2012), which is a generalization of Johansen’s (1995) CVAR model to allow for fractionally integrated (or just fractional) time series.

2.1 Fractional integration and cointegration

Fractional time series models are based on the fractional difference operator

$$\Delta^d X_t = \sum_{n=0}^{\infty} \pi_n(-d) X_{t-n}, \quad (1)$$

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.,

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

For details and for many intermediate results regarding this expansion and the fractional coefficients, see, for example, Johansen and Nielsen (2016, 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 (1), a time series X_t is said to be fractional of order d , denoted $X_t \in I(d)$, if $\Delta^d X_t$ is fractional of order zero, i.e., if $\Delta^d X_t \in I(0)$. The latter 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, for example, Johansen and Nielsen (2012, p. 2672). An example of a process that is fractional of order zero is the stationary and invertible ARMA model. Finally, then, a p -dimensional time series $X_t \in I(d)$ for which one or more linear combinations are fractional of a lower order, i.e., for which there exists a $p \times r$ matrix β such that $\beta' X_t \in I(d-b)$ with $b > 0$, is said to be (fractionally) cointegrated.

2.2 The FCVAR model and interpretation

For a p -dimensional time series, Y_t , the well-known CVAR model has the error correction form

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

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 from the CVAR is to replace the difference and lag operators, Δ and $L = 1 - \Delta$, in (2) by their fractional counterparts, Δ^b and $L_b = 1 - \Delta^b$, respectively, and apply the resulting model to $Y_t = \Delta^{d-b} X_t$. We then obtain

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

where Δ^d is the fractional difference operator and $L_b = 1 - \Delta^b$ is the fractional lag operator.¹ When the so-called “restricted constant term”, denoted ρ , is included, the CVAR and FCVAR models are

$$\Delta Y_t = \alpha (\beta' Y_{t-1} + \rho') + \sum_{i=1}^k \Gamma_i \Delta Y_{t-i} + \varepsilon_t \quad (2)$$

and

$$\Delta^d X_t = \alpha \Delta^{d-b} L_b (\beta' X_t + \rho') + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \quad (3)$$

respectively, see Johansen and Nielsen (2012) or Dolatabadi, Nielsen, and Xu (2016).

Model (3) nests Johansen’s (1995) CVAR model in (2) as the special case $d = b = 1$. Some parameters are well-known from the CVAR model and have the usual interpretations also in the FCVAR model. Most importantly the long-run parameters α and β , which are $p \times r$ matrices

¹Both the fractional difference and fractional lag operators are defined in terms of their binomial expansion in the lag operator, L , as in (1). Note that the expansion of L_b has no term in L^0 and thus only lagged disequilibrium errors appear on the right-hand side of the error correction model.

with $0 \leq r \leq p$. The rank r is termed the cointegration, or cofractional, rank. The columns of β constitute the r cointegration (cofractional) vectors such that $\beta'X_t$ are the cointegrating combinations of the variables in the system, i.e. the long-run equilibrium relations. The parameters in α are the adjustment or loading coefficients which represent the speed of adjustment towards equilibrium for each of the variables. The restricted constant ρ is interpreted as the mean level of the long-run equilibria when these are stationary, i.e. $E\beta'X_t + \rho' = 0$. Finally, short-run dynamics is governed by the autoregressive parameters $(\Gamma_1, \dots, \Gamma_k)$.

The FCVAR model has two additional parameters compared with the CVAR model, namely the fractional parameters d and b . Here, d denotes the fractional integration order of the observable time series. As would presumably be the case for most—if not all—financial asset prices, we assume that these are integrated of order $d = 1$ (Assumption B.1 below). That is, we consider $d = 1$ to be fixed and known, and therefore not estimated. On the other hand, the parameter b is estimated jointly with the remaining parameters, 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 FCVAR model (3) thus has the same main structure as the standard CVAR model (2), in that it allows for modeling of both cointegration and adjustment towards equilibrium, but is more general since it accommodates fractional integration and fractional cointegration.

We note that the fractional difference defined in (1) is an infinite series, but any observed sample will include only a finite number of observations. In practice, therefore, the summation in (1) is truncated at $n = t - 1$. This truncation is analyzed by Johansen and Nielsen (2012, 2016), who argue that the effects of the truncation can be alleviated by conditioning the statistical (maximum likelihood) analysis on a number of initial values, denoted N . Conditional inference is standard in autoregressive models; for example, conditional maximum likelihood estimation of an $AR(k)$ model leads to least squares estimation, which is commonly applied. Furthermore, making the estimation conditional on a number of initial values also alleviates the effect of a non-zero starting point for the first observation, i.e., for X_1 . In our empirical work we follow this suggestion (setting $N = 10$) and apply the version of the FCVAR model given in (3) and the CVAR model in (2).

The asymptotic analysis of the FCVAR model is provided in Johansen and Nielsen (2010, 2012), where it is shown that the maximum likelihood estimator of $(b, \alpha, \Gamma_1, \dots, \Gamma_k)$ is asymptotically normal, while the maximum likelihood estimator of (β, ρ) is asymptotically mixed normal when $b > 1/2$ and asymptotically normal when $b < 1/2$. The important implication is that standard asymptotic inference can be applied to all these parameters.

Likelihood ratio (trace-type) tests for cointegration rank can be calculated as well, and hypotheses on the cointegration rank can be tested as in the CVAR model. In the FCVAR model, the asymptotic distribution of the tests for cointegration rank depends on the unknown (true value of the) parameter b , which complicates empirical analysis compared to the CVAR model. However, the distribution can be simulated on a case-by-case basis. Calculation of maximum likelihood estimators and test statistics is discussed in detail in Johansen and Nielsen (2012) and Nielsen and Popiel (2014), with the latter providing Matlab programs that we apply in our empirical work.

2.3 Forecasting from the FCVAR model

We now discuss how to forecast log-prices, that is X_t , as well as returns, $r_t = \Delta X_t$, from the FCVAR model. Because the model is autoregressive, the best linear predictor is relatively straightforward to calculate. We rearrange (3) with $d = 1$ as

$$X_{t+1} = X_t + \alpha \Delta^{1-b} L_b(\beta'X_{t+1} + \rho') + \sum_{i=1}^k \Gamma_i \Delta L_b^i X_{t+1} + \varepsilon_{t+1}. \quad (4)$$

Note that $L_b = 1 - \Delta^b$ is a lag operator, so that $L_b^i X_{t+1}, i \geq 1$, is known at time t .

We let conditional expectation given the information set at time t be denoted $E_t(\cdot)$, and the best linear predictor of any variable Z_{t+1} given information available at time t be denoted $\hat{Z}_{t+1|t} = E_t(Z_{t+1})$. Clearly, $\hat{\varepsilon}_{t+1|t} = E_t(\varepsilon_{t+1}) = 0$, and $\hat{X}_{t+1|t}$ is then easily found from (4). Inserting also coefficient estimates based on data available up to time t , denoted² $(\hat{b}, \hat{\alpha}, \hat{\beta}, \hat{\rho}, \hat{\Gamma}_1, \dots, \hat{\Gamma}_k)$, we have that

$$\hat{X}_{t+1|t} = X_t + \hat{\alpha} \Delta^{1-\hat{b}} L_{\hat{b}} (\hat{\beta}' X_{t+1} + \hat{\rho}') + \sum_{i=1}^k \hat{\Gamma}_i \Delta L_{\hat{b}}^i X_{t+1}. \quad (5)$$

This defines the forecast of log-prices for period $t+1$ given information at time t . From (5) we can derive the forecast of returns as

$$\hat{r}_{t+1|t} = \hat{X}_{t+1|t} - X_t. \quad (6)$$

We note that, after constructing a series of one-step ahead log-price forecasts, $\hat{X}_{t+1|t}$ for a range of t , the return forecast (6) is different from $\Delta \hat{X}_{t+1|t}$, which may seem the obvious forecast of returns based on forecasts of log-prices, given the definition of returns as the first difference of log-prices. However, since X_t is known at time t , clearly (6) is the appropriate forecast of returns.

Multi-period ahead forecasts can be generated recursively. That is, first generalize (5) as

$$\hat{X}_{t+j|t} = \hat{X}_{t+j-1|t} + \hat{\alpha} \Delta^{1-\hat{b}} L_{\hat{b}} (\hat{\beta}' \hat{X}_{t+j|t} + \hat{\rho}') + \sum_{i=1}^k \hat{\Gamma}_i \Delta L_{\hat{b}}^i \hat{X}_{t+j|t}, \quad (7)$$

where $\hat{X}_{s|t} = X_s$ for $s \leq t$. Then forecasts are calculated recursively from (7) for $j = 1, 2, \dots, h$ to generate h -step ahead forecasts of log-prices, $\hat{X}_{t+h|t}$. Given these, h -step ahead forecasts of returns are calculated as in (6) using the recursively generated log-price forecasts on the right-hand side. We will apply the return forecast (6) in our empirical analysis below for several forecast horizons, h , and we will compare with the similarly obtained return forecast based on the CVAR model (2).

3 Economic model and price discovery

The economic model for the dynamics of spot and futures prices, that will provide the theoretical foundation for our empirical analysis, was developed by Dolatabadi, Nielsen, and Xu (2015, 2016) building on Figuerola-Feretti and Gonzalo (2010). We briefly outline the model here to establish a natural connection to the FCVAR model described in Section 2 above, and refer the reader to Dolatabadi, Nielsen, and Xu (2015, 2016) for details.

3.1 Fractionally cointegrated equilibrium model

We begin with the following set of standard market conditions, which are collectively referred 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.

We denote the log-spot price of a commodity in period t by s_t and the contemporaneous log-futures price for a one-period-ahead futures contract by f_t , while r_{ft} and c_t denote the continuously

²To emphasize that these estimates are based on data available at time t , they could be denoted by a subscript t . However, to avoid cluttering the notation we omit this subscript and let it be understood in the sequel.

compounded interest rate and storage cost, respectively, for period t . The time series behavior of these variables is described in the following conditions, collectively referred to as Assumption B.

- B.1 Δs_t is an $I(0)$ process with mean zero and finite positive variance.
- B.2 $r_{ft} = \bar{r}_f + v_{rt}$, where $\bar{r}_f$ denotes the mean of r_{ft} and v_{rt} denotes an $I(1-b)$ process with $b > 1/2$, mean zero, and finite positive variance.
- B.3 $c_t = \bar{c} + v_{ct}$, 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.

In Assumption B.1 log-prices are $I(1)$, and hence their first difference given by the continuously compounded returns are $I(0)$. Of course, standard asset pricing theory assumes that s_t is a martingale to rule out arbitrage, which is more restrictive than — and implies — our Assumption B.1.³

While storage costs are basically unobserved, interest rates are observed and are typically not found to be $I(0)$, and thus Assumption B allows both to be fractionally integrated. To simplify notation we assume that interest rates and storage costs have the same order of fractional integration, i.e. both v_{rt} and v_{ct} are $I(1-b)$. Also, the assumption that $b > 1/2$ ensures that v_{rt} and v_{ct} are stationary, since then $1-b < 1/2$. Neither of these assumptions are critical, nor even necessary for the economic equilibrium model, but they facilitate interpretation of the parameters in the FCVAR model. For example, with two different fractional integration orders of v_{rt} and v_{ct} , the sum $v_{rt} + v_{ct}$ would be fractional of the highest of the two orders. Furthermore, if $b < 1/2$, then v_{rt} , for example, would not be stationary and in that case we would define $\bar{r}_f$ simply as a constant, rather than interpret it as the mean of r_{ft} , and v_{rt} would denote an $I(1-b)$ process initialized at zero.

Next, a central concept in the theory of commodity pricing is the convenience yield, which is the benefit associated with storing the commodity instead of holding the futures contract (Kaldor, 1939). With the convenience yield denoted by y_t , and imposing Assumption A, no-arbitrage equilibrium conditions imply that

$$f_t + y_t = s_t + r_{ft} + c_t. \quad (8)$$

It is commonplace to characterize convenience yield as 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

$$s_t + \beta_2 f_t + \rho = v_{rt} + v_{ct}, \quad (9)$$

where β_2 and ρ are simple functions of the model parameters. In particular, β_2 can take three different values (with the interpretations assuming a small enough value of ρ):

- (i) $-\beta_2 > 1$: long-run backwardation is when $s_t > f_t$ (backwardation is defined as the present value of the marginal convenience yield of the commodity inventory).
- (ii) $-\beta_2 < 1$: long-run contango is when $s_t < f_t$ (defined as negative present value of the marginal convenience yield of the commodity inventory).
- (iii) $-\beta_2 = 1$: neither backwardation nor contango in the long run.

Note that the equilibrium condition (9) is often stated with only s_t on the left-hand side, which is the reason why we interpret the cointegration coefficient in terms of $-\beta_2$. However, the econometric model specifies the equilibrium in terms of $\beta'X_t + \rho$ and to unify our notation, we have specified the equilibrium (9) in the same way. Thus, the equilibrium model (9) admits the (empirically warranted) theoretical possibility of having a cointegration coefficient $-\beta_2$ different from unity. Moreover, under (9), s_t and f_t are fractionally cointegrated such that the FCVAR model of Section 2 is directly applicable to this economic model.

³Nevertheless, empirical evidence to support Assumption B.1 is presented in Section 4.

3.2 Price discovery

We now briefly review how to analyze price discovery within the FCVAR model based on the discussion in Figuerola-Feretti and Gonzalo (2010) and Dolatabadi, Nielsen, and Xu (2015). The analysis applies the permanent-transitory (PT) decomposition of Gonzalo and Granger (1985) to the FCVAR model. In the notation of the previous subsections, we let $X_t = (s_t, f_t)'$, where s_t and f_t denote the log-spot and log-futures prices at time t , respectively. According to the PT decomposition, X_t may be decomposed into a transitory (stationary) part, $\beta'X_t$, and a permanent part, $W_t = \alpha'_\perp X_t$, using the identity

$$X_t = (\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp + \alpha(\beta'\alpha)^{-1}\beta')X_t = A_1W_t + A_2\beta'X_t,$$

where $\alpha_\perp$ is such that $\alpha'_\perp\alpha = \alpha'\alpha_\perp = 0$.

Here, W_t is the common permanent component of X_t . In the case of spot and futures log-prices, W_t is interpreted as the long-run fundamental (or efficient) market price, in the sense that information that does not affect W_t has no permanent effect on X_t . Thus, the proportions of price discovery attributable to each market may be inferred from the elements of the parameter $\alpha_\perp$, after being normalized so that the elements sum to unity. For details, we refer to Gonzalo and Granger (1995), Figuerola-Feretti and Gonzalo (2010), and Dolatabadi, Nielsen, and Xu (2015).

An alternative, yet strongly related, interpretation of the coefficient α is that of an adjustment coefficient that measures how the disequilibrium errors in previous periods feed into today's changes in X_t . Under this interpretation, the natural question is whether some coefficients in α are zeros, in which case the variable in question is weakly (or long-run) exogenous for the parameters α and β . For example, if $\alpha_2 = 0$, futures prices do not react to the disequilibrium error, i.e. the transitory component, implying that futures prices are the main contributors to price discovery.

4 Data description and preliminary model estimation

4.1 Data description

In our empirical analysis we have data on 17 commodity spot and futures markets. These are canola, cocoa, coffee, copper, corn, crude oil, gasoline, gold, heating oil, palladium, platinum, silver, soybean, soybean meal, soybean oil, sugar, and wheat. The spot and futures price series are both closing prices and are obtained from the Commodity Research Bureau (CRB) database. The same data set (except gasoline and heating oil) was used by Narayan, Ahmed and Narayan (2014) to analyze momentum-based trading strategies in commodity futures markets.⁴

Some facts about the data sources and commodity markets are collected in Table 1. The second column of Table 1 lists the CRB commodity identifier symbol. The spot and futures price series come from different sources/exchanges, which are listed in the third and fourth columns of the table, respectively. The next column shows the start date for our sample period, which varies by commodity although for most commodities it is March 30, 1983. For all commodities, the end date is October 12, 2012. This gives a total of 7708 observations, except for gasoline (7270 observations) and platinum (7100 observations). Our choice of commodities is determined by the availability of time-series data for both spot and futures prices in the CRB database. There are several commodities for which either spot or futures prices are unavailable or only available for a short time period (natural gas, for example, only has futures price data starting in 1990 and spot prices in 1993). For other commodities (cotton and orange juice, for example) there are many months or even years of missing data.

⁴Futures price series are obtained by taking a long position in the nearest-to-expiration futures contract, and 10 days prior to its expiry rolling over to the next nearest contract (see Narayan, Ahmed and Narayan, 2014, for details). This approach follows the literature; see, for example, Gorton, Hayashi, and Rouwenhorst (2013) and Yang (2013).

Table 1: Data description for commodity spot and futures prices

Commodity	CRB	Spot	Futures	Start	Volume	Open int.	Vol. 2012	O.I. 2012
Canola	WC	Vancouver	WCE	3/30/83	6255	53,266	18,649	209,368
Cocoa	CC	New York	ICE	3/30/83	5441	53,443	23,763	177,720
Coffee	KC	New York	ICE	3/30/83	8337	47,609	24,387	139,664
Copper	HG	New York	NYMEX	3/30/83	12,627	68,929	68,716	147,796
Corn	C-	Chicago	CBOT	3/30/83	61,239	320,828	322,417	1,227,418
Crude oil	CL	DJES	NYMEX	3/30/83	202,147	573,095	582,596	1,477,764
Gasoline	RB	DOE	NYMEX	12/3/84	43,308	116,930	148,492	313,488
Gold	GC	Composite	NYMEX	3/30/83	52,332	198,512	179,622	421,913
Heating oil	HO	DJES	NYMEX	3/30/83	38,857	129,577	141,131	303,555
Palladium	PA	New York	NYMEX	3/30/83	899	7531	4353	21,528
Platinum	PL	Engelhard	NYMEX	7/29/85	2697	14,869	9040	46,519
Silver	SI	Composite	NYMEX	3/30/83	20,017	101,307	53,797	115,476
Soybean	S-	Central IL	CBOT	3/30/83	50,882	165,613	211,884	692,521
Soy meal	SM	Decatur, IL	CBOT	3/30/83	18,142	75,344	75,682	236,697
Soy oil	BO	Decatur, IL	CBOT	3/30/83	20,198	91,680	108,576	344,871
Sugar	SB	New York	ICE	3/30/83	23,694	166,924	109,728	699,786
Wheat	W-	St. Louis	CBOT	3/30/83	21,605	102,263	117,191	445,275

Notes: This table provides information on the data sources and commodity markets for the spot and futures prices. The first four columns show the CRB identifier symbol, spot data source, futures data source, and sample start date for each commodity. The end date is 10/12/12 for all data series. The next two columns show the daily volume (number of contracts) and open interest averaged over the entire sample period, while the last two columns show the daily volume and open interest averaged over the 2012 part of the sample.

Some further characteristics of the commodity futures markets by way of volume and open interest are presented in the last four columns of Table 1. These statistics are daily time-series averages for each commodity, where average volume reflects the number of futures contracts traded each day and average open interest reflects the total number of outstanding futures contracts held by market participants each day. In other words, open interest can be used to gauge the liquidity situation of the futures market. The higher the number of open interest contracts the larger the market activity and hence liquidity. The volume of contracts is largest for crude oil at an average of over 200,000 contracts per day over the full sample period and over 580,000 contracts per day in 2012. Based on trading volume, crude oil makes up over one-third of all commodity contracts traded in the market. Corn, gold, and soybean contracts each make up 8–10% of the market, while heating oil and gasoline make up another approximately 7% of the market each. The rest of the commodities each contribute less than 5% to total volume of trade in the commodity futures markets. Similarly, open interest data suggests that roughly a quarter of all outstanding contracts belong to the crude oil market, followed by corn (14%), soybean (9%), and gold (8%).

The spot and futures log-price series are plotted in Figure 1. Three tendencies emerge from the figure. First, the log-price series do not appear to have obvious time trends. This finding was supported by statistical tests (not reported) in the estimation of our models below, where any trend included was statistically insignificant and hence removed. Second, the spot and futures log-price series appear to move together in the long-run, supporting the notion that they are cointegrated. Third, there is clearly heterogeneity among the commodities.

Table 2 presents some descriptive statistics for each continuously compounded percentage return series (first difference of log-prices, multiplied by 100). These confirm the tendencies observed in

Figure 1: Daily commodity log-prices

Notes: Each plot shows daily commodity log-prices. The blue lines are spot prices and the red lines are futures prices. The sample start dates vary by commodity, see Table 1, and the end date is 10/12/12 for all commodities.

Table 2: Selected descriptive statistics of daily commodity spot and futures returns

Commodity	Spot market returns					Futures market returns				
	Mean	S.d.	Skew.	Kurt.	LWE	Mean	S.d.	Skew.	Kurt.	LWE
Canola	0.009	1.437	-0.282	13.968	0.004	0.009	1.295	-1.169	21.957	-0.004
Cocoa	0.005	1.784	-0.018	6.625	0.006	0.004	1.929	-0.007	5.962	0.000
Coffee	0.004	1.931	-0.042	17.243	-0.006	0.003	2.279	0.012	11.817	-0.001
Copper	0.020	1.656	-0.294	8.725	0.011	0.021	1.722	-0.426	8.035	0.010
Corn	0.011	1.664	-0.139	6.546	-0.004	0.012	1.659	-1.003	21.212	-0.007
Crude oil	0.006	1.052	-0.792	19.979	-0.006	0.006	0.939	-0.877	20.787	-0.006
Gasoline	0.020	2.647	-0.350	10.526	-0.042	0.018	2.255	-0.457	11.274	0.029
Gold	0.019	0.995	-0.131	10.097	0.006	0.019	1.003	-0.103	10.153	0.006
Heating oil	0.008	1.047	-0.812	22.937	-0.084**	0.008	0.997	-1.364	21.283	0.003
Palladium	0.022	1.958	-0.106	12.200	0.025	0.024	1.961	-0.234	8.802	0.015
Platinum	0.025	1.502	-0.331	26.744	0.015	0.025	1.405	-0.584	9.091	0.015
Silver	0.015	1.794	-1.063	15.324	0.014	0.015	1.802	-0.697	10.896	0.014
Soybean	0.012	1.503	-0.568	7.950	-0.008	0.011	1.539	-0.934	20.509	-0.008
Soy meal	0.012	1.699	-0.200	6.556	0.004	0.012	1.678	-0.904	12.272	0.000
Soy oil	0.013	1.570	0.020	5.137	-0.013	0.013	1.509	-0.047	5.661	-0.014
Sugar	0.015	2.231	-0.133	9.567	0.007	0.014	2.514	0.218	15.249	0.063*
Wheat	0.011	2.049	-0.273	9.052	-0.000	0.011	1.857	-1.217	29.765	-0.004

Notes: This table reports selected descriptive statistics for the 17 commodity spot and futures return series based on percentage returns. Specifically, the table reports the sample mean, standard deviation, skewness, kurtosis, and the exact local Whittle estimate (LWE) of the fractional integration order, d . For the latter, we let one, two, and three asterisks denote significance at the 10%, 5%, and 1% level, respectively.

Figure 1. For example, the sample mean returns in the spot (futures) market vary from 0.004% to 0.031% (0.003% to 0.030%) per day. A similar variation in sample standard deviation, skewness, and kurtosis is found, which suggests that these commodities can offer investors quite different risk-return trade-offs when considered from an investment portfolio point of view. The last column in the table reports the exact local Whittle estimate (LWE) of the fractional integration order, d , for each series with bandwidth choice $m = T^{0.6}$ (the asymptotic standard error is $1/\sqrt{4m}$, see Shimotsu and Phillips, 2005). These estimates are all very small, ranging from -0.084 to 0.063 , and only one of the 34 series showed $\hat{d}$ significantly different from zero at the 5% level (heating oil spot returns). This strongly suggests that all return series are stationary $I(0)$ processes, which in turn supports our Assumption B.1 and confirms our modeling choice of fixing $d = 1$ in the FCVAR analysis. More generally, it suggests that past returns alone will likely not be very good predictors of future returns.

4.2 Estimation results

Before we can estimate the FCVAR model and apply it in forecasting, we have to make some model selection choices. First, as discussed in Section 2 above, we apply estimation conditional on $N = 10$ initial values for all our FCVAR results, corresponding to conditioning on the first two weeks of observations. Experimentation with different values of N showed little effect. For the CVAR model we applied estimation conditional on $k + 1$ initial values, such that maximum likelihood estimation is reduced rank regression (Johansen, 1995). Second, we have to specify the lag length, k , in the vector error correction models (2) and (3). For the CVAR model we select the lag length to minimize

the Bayesian Information Criterion (BIC) based on the model that has full rank, $r = p$, where p is the dimension of the system.⁵ For the FCVAR model, we apply several different statistics to select the lag length, namely the BIC, the LR test statistic for significance of Γ_k , and univariate Ljung-Box Q tests (with $m = 10$ lags) for each of the two residual series, in each case based on the model that has full rank, $r = p$. In addition, we examined the unrestricted estimates of b which, when the lag length is over-specified, will sometimes be very far from what should be expected, e.g. $\hat{b} = 0.05$ or similar, due to a non-identification issue in the FCVAR model with over-specified lag length, see Johansen and Nielsen (2010, Section 2.3) or Carlini and Santucci de Magistris (2017) for a theoretical discussion of this phenomenon.⁶ Thus, for each commodity, we use the BIC as an initial guide to choosing the lag length, and starting 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 estimate of b is reasonable (very widely interpreted), and (iii) the Ljung-Box Q tests for serial correlation in the two residual series do not show signs of misspecification. Third, after choosing the number of lags, we select the cointegrating rank, r , by sequentially testing (using the LR trace statistic) the hypotheses $r = 0, 1, 2$ until rejection, choosing the last non-rejected hypothesis to be the cointegration rank. The critical values for the rank tests are simulated case-by-case for the FCVAR model, and for the CVAR model we used Johansen (1995, Table 15.2).

Table 3 reports results⁷ from estimation of CVAR and FCVAR models for commodity spot and futures log-prices, i.e. with $X_t = (s_t, f_t)'$ in the notation of Section 2. For estimation we use the first 75% of the sample, and reserve the remainder for out-of-sample forecasting. This leaves $T = 5781$ observations in the estimation sample, except for gasoline ($T = 5452$) and platinum ($T = 5325$).

The second column of Table 3 shows the chosen lag-order (k) for each commodity. It is clear that fewer lags are usually needed in the lag-augmentation for the FCVAR model compared with the CVAR model (only crude oil has more lags in the FCVAR model specification). This is expected since the FCVAR model includes the additional parameter b to accommodate serial dependence. Although the FCVAR model in this way includes one additional parameter (b), each additional lag included in the CVAR model requires four additional parameters (Γ_i) to be estimated, and hence the CVAR model in most cases includes a larger number of parameters than the FCVAR model.

The third column of Table 3 shows the estimated fractional parameter, $\hat{b}$. Point estimates range from 0.194 to 0.955, showing a wide variety of fractional cointegration properties across the different commodities. Relating the estimates of b to the theoretical model in Section 3.1, we note that the heterogeneity in these estimates derives from heterogeneity in the storage cost in Assumption B.3, since the interest rate in Assumption B.2 presumably will be the same for different commodities. The CVAR model is nested within the FCVAR model by imposing the hypothesis $b = 1$, so that we can test the CVAR model against the more general FCVAR model by testing the restriction $b = 1$. In Table 3, we use one, two, and three asterisks on the estimates $\hat{b}$ to denote when the fractional parameter is significantly different from unity at the 10%, 5%, and 1% level, respectively. From these tests we note that the CVAR model is rejected against the FCVAR model for 15 of 17 commodities at the 1% level, and for one more commodity at the 10% level. Only palladium appears to be well-described by a CVAR model in terms of in-sample fit when judged by this statistical test. Thus, the FCVAR model provides a better statistical in-sample fit in most cases.⁸

⁵When calculating the BIC for different values of k , we use $N = 10$ initial values for all k to have the same effective sample size—and hence facilitate comparison—across different values of k .

⁶In fact, as shown in Carlini and Santucci de Magistris (2017, Lemma 1), this identification problem is not present when d is fixed as in our model.

⁷Full estimation results are available from the authors upon request.

⁸We note that the log-likelihood—and hence the BIC—for the CVAR and FCVAR models are based on different effective sample sizes, even for the same commodity, because the number of initial values are different. Therefore,

Table 3: Estimation results for CVAR and FCVAR models of commodity prices

Commodity	k	$\hat{b}$	$-\hat{\beta}_2$	$-\hat{\rho}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_{\perp,1}$	$\hat{\alpha}_{\perp,2}$
Panel A: CVAR model								
Canola	3	1	1.036	-0.199	-0.019	0.000	-0.002	1.002***
Cocoa	3	1	0.916	0.777	0.000	0.012	0.982**	0.018
Coffee	2	1	0.956	0.303	-0.018	0.004	0.171	0.829***
Copper	5	1	0.993	0.060	-0.025	0.004	0.143	0.857***
Corn	1	1	1.057	-0.298	-0.034	0.015	0.305***	0.695***
Crude oil	1	1	1.006	-0.007	-0.247	0.035	0.124**	0.876***
Gasoline	1	1	0.993	-0.015	-0.046	-0.011	-0.329**	1.329***
Gold	2	1	1.000	-0.003	-0.506	0.148	0.226***	0.774***
Heating oil	6	1	1.005	0.004	-0.075	-0.022	-0.414**	1.414***
Palladium	2	1	1.002	-0.002	-0.057	0.037	0.392***	0.608***
Platinum	7	1	0.999	0.006	-0.076	0.107	0.586***	0.414***
Silver	3	1	0.993	0.042	-0.352	0.106	0.232***	0.768***
Soybean	1	1	1.029	-0.201	-0.036	0.061	0.628***	0.372***
Soy meal	3	1	1.042	-0.211	-0.019	0.007	0.255	0.745***
Soy oil	4	1	1.179	-0.552	0.001	0.012	1.062**	-0.062
Sugar	3	1	1.119	-0.219	-0.002	0.011	0.862***	0.138
Wheat	1	1	1.110	-0.652	-0.012	0.010	0.471***	0.529***
Panel B: FCVAR model								
Canola	2	0.776***	1.031	-0.177	-0.063	-0.004	-0.066	1.066***
Cocoa	2	0.752***	0.966	0.407	0.014	0.043	1.502**	-0.502
Coffee	1	0.528***	1.034	-0.131	-0.116	0.091	0.441**	0.559***
Copper	0	0.341***	1.053	-0.194	-0.920	0.056	0.057	0.943***
Corn	1	0.923*	1.057	-0.297	-0.044	0.021	0.325***	0.675***
Crude oil	3	0.194***	1.007	-0.008	-44.874	2.691	0.057	0.943***
Gasoline	1	0.735***	1.027	0.005	-0.125	-0.027	-0.270*	1.270***
Gold	1	0.795***	1.002	-0.015	-0.878	0.181	0.171***	0.829***
Heating oil	4	0.694***	1.014	0.006	-0.206	-0.055	-0.364	1.364***
Palladium	2	0.955	1.002	0.000	-0.066	0.043	0.392***	0.608***
Platinum	1	0.664***	0.996	0.022	-0.507	0.213	0.296***	0.704***
Silver	1	0.672***	0.998	0.011	-1.174	-0.027	-0.023	1.023***
Soybean	1	0.872***	1.028	-0.198	-0.050	0.098	0.661***	0.339***
Soy meal	1	0.518***	1.138	-0.700	-0.115	0.103	0.472**	0.528***
Soy oil	2	0.488***	1.294	-0.891	-0.041	0.046	0.529	0.471
Sugar	3	0.606***	1.090	-0.188	-0.003	0.086	0.969**	0.031
Wheat	0	0.717***	1.140	-0.820	-0.027	0.048	0.640***	0.360***

Notes: This table reports estimation results for CVAR (Panel A) and FCVAR (Panel B) models applied to the first 75% of the sample of commodity spot and futures log-prices. The columns include lag-order, k , estimates of the fractional parameter, $\hat{b}$, the cointegration coefficient, $-\hat{\beta}_2$, the restricted constant term, $-\hat{\rho}$, the adjustment coefficients, $\hat{\alpha}_1$ and $\hat{\alpha}_2$, and the price discovery coefficients, $\hat{\alpha}_{\perp,1}$ and $\hat{\alpha}_{\perp,2}$. The latter are normalized to add to unity. For the fractional parameter we let one, two, and three asterisks denote significant difference from unity at the 10%, 5%, and 1% level, respectively, and for the price discovery coefficients, asterisks denote significant difference from zero.

The next four columns of Table 3 report estimates of the cointegration coefficient, $-\hat{\beta}_2$, the restricted constant term, $-\hat{\rho}$, and the adjustment coefficients, $\hat{\alpha}_1$ and $\hat{\alpha}_2$. The estimates $-\hat{\beta}_2$ of the cointegration coefficients are close to unity, as expected from an efficient markets hypothesis point of view, although they may deviate from unity still. The latter reflects a market in long-run backwardation (when $-\beta_2 > 1$) or contango (when $-\beta_2 < 1$), see Section 3 above. Generally, the estimates $-\hat{\beta}_2$ suggest that backwardation is more common across markets. This is especially the case for estimates from the FCVAR model, which indicate backwardation in 14 of 17 commodity markets, with the remaining three having estimates $-\hat{\beta}_2$ very close to unity. As discussed in the economic equilibrium model in Section 3, the extent of backwardation or contango is related to the convenience yield, which is expected to differ across different commodities.

As discussed briefly in Sections 2.2 and 3.2, the adjustment coefficients α_1 and α_2 determine the speed of adjustment towards equilibrium for the two variables when the system is in disequilibrium. In our model, $\alpha_1 < 0$ and $\alpha_2 > 0$ imply adjustment towards equilibrium, with the other sign indicating adjustment away from equilibrium. For example, we note that the estimates for gasoline suggest that the futures prices do not adjust towards equilibrium (and in fact adjust away from equilibrium, although this is only mildly significant, see below), whereas spot prices do adjust towards equilibrium. This behavior is found only for a few commodities. Note also that the system can still exhibit overall adjustment towards the equilibrium, even if one price series does not adjust towards equilibrium. This happens, of course, if the price series that does adjust towards equilibrium does so at a faster rate than the other price series moves away from equilibrium. For most commodities, however, both prices adjust towards equilibrium but at different speeds.

The final two columns show price discovery coefficients, $\hat{\alpha}_{\perp,1}$ and $\hat{\alpha}_{\perp,2}$, normalized to add to unity, so that $\hat{\alpha}_{\perp,1}$ ($\hat{\alpha}_{\perp,2}$) is the proportion of price discovery in the spot (futures) market. We test the hypothesis that the price discovery coefficients are zero and let one, two, and three asterisks denote significance at the 10%, 5%, and 1% level, respectively.⁹ Thus, to test whether, for example, the futures market is dominant in the sense that price discovery takes place exclusively in the futures market, we test the hypothesis that $\alpha_{\perp,1} = 0$. We note that, corresponding to α coefficients that do not indicate adjustment towards equilibrium, there are a few negative $\hat{\alpha}_{\perp,i}$ coefficients. However, only two are significantly negative in the CVAR results and only one is significantly negative (and only at the 10% level) in the FCVAR results.

According to point estimates of the price discovery coefficients from both the CVAR and FCVAR models, the futures market dominates price discovery for most commodities, as expected from theory (e.g., Hasbrouck, 1995), with average futures market price discovery coefficients of 0.665 and 0.659 from the CVAR and FCVAR models, respectively. The CVAR and FCVAR models agree that there is statistical evidence that the spot market is dominant in the price discovery process for cocoa and sugar, while the futures market is dominant for canola, copper, gasoline, and heating oil. For the remaining commodities, there is evidence of price discovery taking place in both the spot and futures markets, although with some differences between the CVAR and FCVAR models. One possible reason for the empirically observed spot market dominance for a few commodities could be the decline in futures trading volume during and immediately after the recent crisis.

In the literature on price discovery in commodity markets, several studies show that price discovery is dominated by the futures market (e.g., Figuerola-Ferretti and Gonzalo, 2010, Dolatabadi, Nielsen, and Xu, 2015, and references cited therein). Our results broadly confirm this view, but at the same time point to a few commodities where price discovery is not dominated by the futures

these measures cannot be used directly as a means of statistical comparison of the CVAR and FCVAR in-sample fit.

⁹Hypotheses on $\alpha_{\perp}$ are in fact tested by testing the mirror hypotheses on α . For example, as is obvious from the definition of $\alpha_{\perp}$, the hypothesis $\alpha_{\perp,1} = 0$ is equivalent to the mirror hypothesis $\alpha_2 = 0$, which is straightforward to test within the CVAR and FCVAR models; see also the discussion in Dolatabadi, Nielsen, and Xu (2015).

Figure 2: Daily crude oil forecasts

Notes: The four plots show forecasts as well as subsequently realized values for (a) log-spot prices, (b) log-futures prices, (c) spot returns, and (d) futures returns. Each plot shows the last 20 observations in the estimation sample together with the first 50 out-of-sample one-step ahead forecasts and the subsequently realized values. In each subplot there are three lines: The blue denotes the data observations, the red line are the recursive FCVAR forecasts, and the green line the recursive CVAR forecasts.

market, so that futures market dominance in price discovery is not completely general—a finding consistent with Dolatabadi, Nielsen, and Xu (2015).

These price discovery results are not trivial outcomes because the dominance of one market over another indicates the market which has the highest information content. This has implications for investors because the market which has most information can then be used to forecast the market which has less information. In univariate regression-style forecasting models, one would then consider using past information from the dominant market to forecast prices or returns in the non-dominant market. However, our FCVAR (and CVAR) models are joint models for spot and futures price series, and will therefore forecast both series simultaneously and hence automatically take the price discovery information in both markets into account.

5 Out-of-sample forecasting and economic significance

5.1 Statistical out-of-sample forecast comparison

We now consider out-of-sample forecasting. Specifically, starting from the estimation results in the previous section, we recursively generate one-step ahead (daily) return forecasts, re-estimating the model each period. We generate a total of 1927 out-of-sample return forecasts in this manner (only 1818 for gasoline and 1775 for platinum), to match the remaining 25% of our observations. This allows us to compare our forecasts with the actually observed out-of-sample returns series.

In Figure 2 we show forecasts as well as subsequently realized values for daily forecasting

Figure 3: Daily crude oil equilibrium error forecasts

Notes: The two plots show the last observation of the model equilibrium error, i.e. $\hat{\beta}'X_t$, together with the recursive h -period ahead forecasts, $\hat{\beta}'\hat{X}_{t+h|t}$ for $h = 1, \dots, 100$, generated from (a) the CVAR model and (b) the FCVAR model. In each panel, the mean is indicated by a horizontal line given by $-\hat{\rho}'$.

of crude oil (a) log-spot prices, (b) log-futures prices, (c) spot returns, and (d) futures returns. We choose crude oil for the illustration because it is the most heavily traded commodity in our dataset, see Table 1. Each plot shows the last 20 observations in the estimation sample (i.e. X_t or $r_t = \Delta X_t$ for $t = 5762, \dots, 5781$) together with the first 50 out-of-sample one-step ahead forecasts ($\hat{X}_{t+1|t}$ or $\hat{r}_{t+1|t}$ for $t = 5781, \dots, 5830$) and the corresponding realized values (X_{t+1} or r_{t+1} for $t = 5781, \dots, 5830$). In each plot there are three lines: The blue line denotes the data observations, while the red and green lines show the recursive FCVAR and CVAR forecasts, respectively. We note that log-price forecasts track the subsequently realized observations quite well, whereas returns are clearly predicted much less accurately. This is, of course, expected from no-arbitrage theory of efficient markets. However, we do note a slightly better forecasting performance of spot returns compared with futures returns, at least for this particular part of the sample, which is also in accordance with our finding of price discovery in the futures market for crude oil, see Table 3.

The difference between FCVAR and CVAR forecasts shows most clearly in the forecasts of the equilibrium error series, $\beta'X_t$, which are depicted in Figure 3. Here we show the last (daily) observation on the equilibrium error (i.e. $\hat{\beta}'X_t$ for $t = 5781$), together with recursive h -period ahead forecasts of these, $\hat{\beta}'\hat{X}_{t+h|t}$ for $t = 5781$ and $h = 1, \dots, 100$, generated from (a) the CVAR model and (b) the FCVAR model. In each panel, the horizontal line indicates the mean of the equilibrium relation given by $-\hat{\rho}$. Again, forecasts are depicted for crude oil as an illustration. Figure 3 clearly shows that the CVAR model equilibrium error forecasts return to their mean value much more quickly than the FCVAR model equilibrium error forecasts. This reflects the $I(0)$ nature of $\beta'X_t$ in the CVAR model and the fractional integration nature of $\beta'X_t$ in the FCVAR model, where $\beta'X_t$ is estimated to be $I(0.806)$, see Table 3, which is nonstationary but mean-reverting. This feature of the multi-step ahead forecasts may impact the relative forecasting performance at longer horizons.

In Table 4 we report some out-of-sample forecast comparison statistics for the one-step ahead (daily, $h = 1$) forecasts calculated from either the FCVAR model or the CVAR model. In particular, we first report the Clark and West (2007, Section 2) test statistic for equal predictive ability, which is a modification of the Diebold and Mariano (1995) test statistic to account for the fact that the CVAR model is nested within the FCVAR model class¹⁰ (see also Giacomini and White, 2006, for the point about nested model classes). The CW statistic is asymptotically standard normally distributed and favors the FCVAR model forecasts when it is positive. The null hypothesis of

¹⁰We are grateful to Peter Extercate for bringing this point to our attention.

Table 4: Statistical out-of-sample forecast comparison at daily ($h = 1$) horizon

Commodity	CW statistic		Relative RMSE		R^2_{OOS} CVAR		R^2_{OOS} FCVAR	
	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures
Canola	2.3240**	2.0699**	-0.1932	-0.1516	0.4010	-0.0008	0.4033	0.0023
Cocoa	4.9334***	2.4235***	-0.6563	-0.1389	0.0037	-0.0030	0.0167	-0.0002
Coffee	10.1160***	0.4626	-2.3397	0.0455	0.4386	0.0003	0.4646	-0.0006
Copper	2.4182***	1.1569	-0.2294	-0.0591	-0.0013	-0.0014	0.0033	-0.0003
Corn	0.7305	0.8105	-0.0182	-0.0181	0.0009	0.0030	0.0013	0.0033
Crude oil	2.0993**	2.1385**	-0.3290	-0.2452	-0.0461	0.0026	-0.0393	0.0075
Gasoline	2.1831**	-0.3623	-0.1532	0.0202	0.0177	0.0021	0.0207	0.0017
Gold	1.3540*	0.5399	-0.0672	-0.0079	-0.0399	0.0426	-0.0385	0.0428
Heating oil	2.2040**	1.5268*	-0.1608	-0.1013	-0.0129	-0.0059	-0.0096	-0.0039
Palladium	3.4676***	2.0645**	-0.0962	-0.0239	0.1165	0.0091	0.1182	0.0096
Platinum	4.2108***	2.4841***	-0.8827	-0.3388	0.1241	-0.0032	0.1395	0.0036
Silver	1.6530**	2.0524**	-0.1440	-0.1923	-0.0740	0.0204	-0.0709	0.0241
Soybean	-0.2842	1.7922**	0.0553	-0.3238	0.0103	-0.0080	0.0092	-0.0015
Soy meal	0.6216	0.5917	0.0220	0.0229	0.0055	0.0055	0.0051	0.0051
Soy oil	1.3304*	1.4648*	-0.0726	-0.0888	0.0048	-0.0064	0.0062	-0.0046
Sugar	-0.8063	1.6087*	0.2434	-0.0854	0.0511	-0.0015	0.0465	0.0003
Wheat	1.5979*	1.4354*	-0.1292	-0.0668	-0.0030	-0.0035	-0.0005	-0.0021

Notes: This table reports out-of-sample forecast comparison statistics for one-step ahead ($h = 1$) return forecasts. The statistics reported are the Clark and West (2007) test statistic, the relative RMSE, and the out-of-sample R^2 . The CW statistic is asymptotically standard normally distributed and positive values favors the FCVAR model. Statistical significance (one-sided) of the CW statistic at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively. The relative RMSE is calculated such that it favors FCVAR model when it is negative.

the nested CW test is that the CVAR model forecasts are at least as good as the FCVAR model forecasts, and the alternative hypothesis is that the FCVAR forecasts are superior. Note, therefore, that this is a one-sided test, and we report one-sided significance using asterisks in the table. The next statistic is the relative root mean squared error (RMSE) of the two forecasts (from the FCVAR and CVAR models, respectively), and this is calculated such that negative values favor the FCVAR model forecasts. Finally, we report the out-of-sample R^2 for both sets of forecasts.

The results in Table 4 clearly favor the FCVAR model forecasts. Specifically, the CW statistic favors the FCVAR model forecasts in 15/17 spot markets and 16/17 futures markets, and is significant at the 10% level or better in 13 and 11 of those cases, respectively. None of the three negative CW statistics are significant even at the 10% level. The relative RMSE prefers the FCVAR model forecasts for 14/17 commodities in both the spot and futures markets.

The final columns in Table 4 report the out-of-sample R^2 for both sets of return forecasts. We note that the forecastability of returns vary greatly across commodities, and also between spot and futures markets for the same commodity. Comparing out-of-sample R^2 values for FCVAR and CVAR forecasts, supports the previous conclusions with R^2 being higher on average for the FCVAR model. Also, in most cases the out-of-sample R^2 is higher for the spot market than for the futures market, confirming earlier results on the relative forecastability of returns from the two markets.

We investigated the robustness of the Table 4 results by considering forecasting at longer horizons, specifically the weekly and monthly horizons based on daily data, i.e., $h = 5$ and $h = 21$ periods ahead. The motivation is that these horizons could correspond to an investor that rebal-

ances the portfolio weekly or monthly and hence needs only forecasts at those horizons. With this motivation we consider only non-overlapping forecasts, i.e. forecasts calculated every h periods (days). These results (omitted for brevity, but available from the authors upon request) are generally similar to those presented in Table 4, although not quite as favorable towards the FCVAR model and with fewer statistically significant CW statistics. Of course, part of the explanation is that the number of non-overlapping monthly forecasts is much smaller than the number of daily forecasts. Among the statistically significant CW statistics for longer horizon forecasting ($h = 5$ or $h = 21$), 22 out of 23 favor the FCVAR model forecasts over the CVAR model forecasts.

Generally, our findings suggest that it is more difficult to predict returns at longer horizons in both the spot and futures commodity markets, at least for the horizons we considered. We expect this to have implications for the profitability and economic significance of longer horizon forecasts, and we return to this point below.

5.2 Mean-variance utility function based profits

Another metric of comparison and evaluation of forecasting performance is economic—rather than purely statistical—significance. That is, whether the forecasts can generate statistically significant profits when incorporated into a dynamic trading strategy. In this comparison, we also investigate whether the improved statistical in-sample fit and forecast performance of the FCVAR model relative to the CVAR model translate into differences in profits, and whether the relatively strong forecastability in some markets compared with others translate into different profits across markets.

Consider a dynamic trading strategy where the investor rebalances the portfolio every h days; in particular $h = 1$, $h = 5$, and $h = 21$, corresponding to daily, weekly, and monthly rebalancing, respectively. There are two assets: a risk-free asset with (continuously compounded) return from period t to period $t + h$ given by $r_{f,t+h}$ and a risky asset with (continuously compounded) return $r_{t+h} = s_{t+h} - s_t$ when the risky asset is a spot position or $r_{t+h} = f_{t+h} - f_t$ when the risky asset is a futures position. With weight w_{t+h} on the risky asset, the return on the investor's portfolio is

$$\begin{aligned} r_{p,t+h} &= w_{t+h}r_{t+h} + (1 - w_{t+h})r_{f,t+h} - \theta|w_{t+h} - w_t| \\ &= w_{t+h}(r_{t+h} - r_{f,t+h}) + r_{f,t+h} - \theta|w_{t+h} - w_t|, \end{aligned} \quad (10)$$

where θ denotes a transaction cost, which is applied to the value of the fraction of the portfolio that is being traded and hence multiplied by $|w_{t+h} - w_t|$. The size of transaction costs in commodity markets is subject to some discussion in the literature. Locke and Venkatesh (1997, p. 239) write “Overall, transaction costs appear to be relatively low for futures trading, as these numbers translate to between 0.0004% and 0.033% of notional value, much less than the 1% or so often cited for equities.” We take the average of their suggested range and use $\theta = 0.000167$ in (10) as the transaction cost in both futures and spot markets, i.e. 0.0167% of nominal value. We note from the outset that, as also discussed in the introduction, futures market trading is much more practical than trading in the spot market. Indeed, the latter may not even be feasible for some commodities. Nonetheless, for futures market trading in particular, these transaction costs appear reasonable.

In the equity return forecasting literature, a mean-variance utility function is often used to derive the trading strategy, e.g., Marquering and Verbeek (2004) and Campbell and Thompson (2008). Following this literature, the weight on the risky asset maximizes the investor's mean-variance utility function,

$$U(r_{p,t+h}) = E_t(r_{p,t+h}) - \frac{1}{2}\gamma Var_t(r_{p,t+h}), \quad (11)$$

where $E_t(\cdot)$ and $Var_t(\cdot)$ denote conditional mean and variance given information at time t and γ is the investor's coefficient of relative risk aversion. Maximizing $U(r_{p,t+h})$ with respect to w_{t+h} yields

the optimal weight

$$w_{t+h}^* = \frac{E_t(r_{t+h}) - r_{f,t+h}}{\gamma \text{Var}_t(r_{t+h})}, \quad (12)$$

noting that the risk-free rate carries no risk and hence does not contribute to the variance of the portfolio. Following the literature, we further constrain the optimal weight and impose $w_{t+h}^* \geq -0.5$ (at most 50% short-selling) and $w_{t+h}^* \leq 1.5$ (at most 50% borrowing/leverage). In the next subsection we consider robustness to alternative restrictions on short-selling and leverage.

Thus, to calculate the trading strategy profits, or economic significance of the forecasts, three steps are performed. First, we forecast commodity (spot or futures) returns at each time period using (6) in Section 2.3. Second, we calculate portfolio weights from (12) given return forecasts $E_t(r_{t+h}) = \hat{r}_{t+h|t}$. For the risk-free return $r_{f,t+h}$ we use the return on the US three-month Treasury bill, which is assumed known at time t (since it is risk-free). The risk-aversion coefficient is set at $\gamma = 6$, corresponding to an investor that takes a medium level of risk, and for robustness we also consider $\gamma = 3$ and $\gamma = 12$ in the next subsection. Following standard practice, we estimate the time-varying variance of the risky asset by a GARCH(1,1) model using all observations available at time t . Third, given the portfolio weights, portfolio returns are computed from (10) for each period. These are aggregated across time and reported as an annualized average portfolio excess return to facilitate comparison across different values of the rebalancing horizon, h .

Because futures market trading is more practical than spot market trading, we begin with the former. In Table 5 we report annualized average excess portfolio returns (multiplied by 100 to yield a continuously compounded percentage return) for commodity futures markets. The returns are reported as excess returns above and beyond the risk-free rate (which had an average return of 1.753% per annum over the out-of-sample forecasting period). We report results for daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 21$) rebalancing and for forecasts from both the CVAR and FCVAR models. Standard errors are reported in parentheses. For each commodity and rebalancing horizon, we conduct a two-sided statistical test of the null hypothesis that the excess portfolio return is zero, and we interpret this as a (statistical) test of economic significance of the return forecasts.

The results in Table 5 show that excess portfolio returns average about 9.0% and 10.8% per annum with daily rebalancing, based on CVAR and FCVAR forecasts, respectively, 0.2% and 0.9% with weekly rebalancing, and 1.9% and 2.5% with monthly rebalancing. Compared with the average annual return of just under 1.8% on the risk-free asset and the very small average returns on the commodities themselves, as reported in Table 2, we find that the excess returns in Table 5 with daily rebalancing are impressive for most commodities. Furthermore, with daily rebalancing, the excess returns are significantly positive in 14 of 34 cases, whereas only one of the negative returns is significant (and only at the 10% level). For longer horizons, there are fewer significant returns because the returns are smaller and their standard errors are higher. Finally, by far the highest returns on average are found with daily rebalancing, so it would appear that more opportunities to rebalance clearly outweighs the additional transaction costs imposed.

Comparing portfolio returns using weights calculated from CVAR and FCVAR based forecasts show that the FCVAR model produces slightly higher returns on average, although there are sometimes substantial differences for individual commodities. Specifically, the FCVAR forecasts generate higher returns for 13 out of 17 commodities with daily rebalancing, but only 10/17 commodities with weekly rebalancing and 9/17 commodities with monthly rebalancing. Quite surprisingly, given our expectations of the FCVAR model as a superior long-horizon forecasting model, the CVAR and FCVAR models basically perform equally well with longer horizon rebalancing. Of course, as noted earlier, the accuracy of the forecasts deteriorates at longer horizons, which is clearly part of the reason why profits are lower with longer horizon rebalancing.

Table 5: Annualized excess portfolio returns for commodity futures markets

Commodity	Daily		Weekly		Monthly	
	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR
Canola	14.042*** (5.430)	16.732*** (5.202)	2.778 (4.799)	0.426 (3.760)	-3.707 (6.370)	-4.566 (3.713)
Cocoa	-6.157* (3.659)	-0.547 (3.747)	0.666 (4.320)	-3.621 (6.469)	6.824 (8.222)	0.568 (9.337)
Coffee	2.565 (4.240)	3.910 (4.655)	-6.106 (5.908)	-0.328 (7.818)	-10.748 (9.563)	2.572 (9.584)
Copper	-3.810 (3.858)	-0.574 (0.952)	-4.338 (3.222)	-1.852 (2.077)	7.636 (5.827)	2.386 (3.848)
Corn	14.163** (6.066)	15.545** (6.330)	-12.351* (7.053)	-9.570 (7.548)	-3.461 (14.955)	-3.397 (15.914)
Crude oil	-1.040 (1.778)	4.362 (3.592)	0.612*** (0.054)	-0.101 (0.068)	0.684*** (0.101)	0.496*** (0.093)
Gasoline	6.705 (4.833)	7.133 (4.688)	8.915 (6.844)	9.493 (6.679)	29.981* (17.561)	25.199 (16.170)
Gold	36.877*** (4.765)	36.726*** (4.680)	0.816 (5.546)	0.600 (5.131)	8.067 (6.811)	7.425 (7.173)
Heating oil	-1.289 (4.047)	0.696 (3.566)	-0.653*** (0.097)	-0.354*** (0.088)	0.108 (0.117)	0.323*** (0.105)
Palladium	26.677*** (6.483)	27.852*** (6.594)	-1.351 (5.878)	-1.879 (5.572)	7.067 (7.730)	6.935 (6.992)
Platinum	1.322 (3.629)	6.199** (2.784)	-3.099 (4.062)	-1.673 (3.651)	9.072 (6.819)	15.287** (6.269)
Silver	38.723*** (6.306)	32.990*** (4.788)	1.455 (10.523)	-3.333 (3.987)	4.737 (11.122)	8.816** (4.234)
Soybean	-3.796 (4.362)	-1.052 (5.055)	-2.159 (6.880)	2.166 (7.496)	-11.097 (10.621)	-8.732 (10.996)
Soy meal	13.839** (6.513)	13.379** (6.078)	-2.907 (5.796)	6.155 (6.615)	-11.337** (5.325)	-10.342 (8.705)
Soy oil	2.272 (6.447)	4.173 (6.052)	8.841 (5.375)	0.582 (5.185)	-12.082 (8.978)	-4.102 (5.258)
Sugar	8.066 (7.912)	14.399* (8.467)	22.104*** (8.359)	25.621** (10.245)	-10.072 (15.294)	-6.563 (16.114)
Wheat	3.496 (5.229)	1.162 (5.398)	-10.276 (7.501)	-7.763 (9.294)	21.121 (12.938)	9.477 (13.676)
Average	8.980	10.770	0.174	0.857	1.929	2.458

Notes: This table reports annualized average excess portfolio percentage returns for commodity futures markets for a medium risk-aversion investor ($\gamma = 6$) with weights restricted to the interval $[-0.5, 1.5]$. The results are reported for daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 21$) rebalancing and for forecasts based on both the CVAR and FCVAR models. Standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively. The final row is the average return across all commodities.

Table 6: Annualized excess portfolio returns for commodity spot markets

Commodity	Daily		Weekly		Monthly	
	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR
Canola	213.317*** (9.436)	209.812*** (9.237)	-11.277 (17.420)	-15.609 (16.861)	-31.636 (24.099)	-21.274 (21.505)
Cocoa	38.237*** (8.354)	44.565*** (8.347)	6.299 (13.653)	3.763 (13.765)	-7.236 (14.43)	3.089 (10.201)
Coffee	189.528*** (7.534)	203.214*** (7.551)	-20.420** (9.647)	-26.835** (10.751)	-5.803 (9.791)	-7.660 (9.290)
Copper	2.171 (5.528)	9.339* (5.375)	6.801 (5.012)	8.163 (8.418)	-7.620** (3.841)	-11.765* (6.090)
Corn	20.802** (8.550)	18.228** (7.640)	13.662 (16.969)	1.606 (14.547)	9.211 (32.283)	4.107 (23.750)
Crude oil	-2.398 (6.141)	0.152 (6.004)	-1.206*** (0.130)	-0.816*** (0.124)	-0.118 (0.168)	0.004 (0.142)
Gasoline	35.462*** (9.450)	37.033*** (10.290)	26.439* (15.947)	31.775* (18.325)	71.912*** (27.360)	51.143** (25.841)
Gold	1.481 (6.876)	3.081 (7.072)	-4.146 (8.562)	-3.547 (9.151)	31.451** (15.472)	29.186* (15.603)
Heating oil	-3.760 (4.540)	-0.739 (4.328)	-0.913*** (0.117)	-0.554*** (0.098)	-0.452** (0.206)	-0.128 (0.157)
Palladium	119.516*** (11.495)	119.772*** (11.397)	-2.883 (11.939)	-0.186 (11.542)	5.831 (8.862)	5.460 (7.974)
Platinum	29.971*** (4.358)	31.787*** (4.209)	-7.316 (6.641)	-2.421 (6.694)	-6.068 (11.043)	3.809 (9.081)
Silver	-1.480 (10.523)	-9.107 (11.099)	-10.603 (9.604)	-15.683 (10.832)	37.507 (27.947)	26.621* (15.009)
Soybean	20.292*** (6.158)	14.973*** (4.879)	11.153 (10.179)	6.761 (7.167)	2.326 (10.262)	-1.766 (4.424)
Soy meal	16.737** (8.015)	14.153* (8.222)	16.537* (10.005)	4.633 (8.946)	-8.751 (23.589)	-3.818 (15.889)
Soy oil	17.370** (7.245)	14.659** (6.356)	11.073 (7.919)	3.613 (5.923)	-23.826 (15.467)	-13.026 (10.303)
Sugar	94.437*** (9.944)	90.336*** (10.335)	-13.574 (8.579)	-17.619 (11.959)	0.312 (6.516)	-7.855 (12.984)
Wheat	2.486 (6.197)	2.259 (3.203)	28.017* (16.775)	6.269 (8.844)	-49.919 (38.537)	-10.141 (12.648)
Average	46.716	47.266	2.802	-0.982	1.007	2.705

Notes: This table reports annualized average excess portfolio percentage returns for commodity spot markets for a medium risk-aversion investor ($\gamma = 6$) with weights restricted to the interval $[-0.5, 1.5]$. The results are reported for daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 21$) rebalancing and for forecasts based on both the CVAR and FCVAR models. Standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively. The final row is the average return across all commodities.

Table 7: Sharpe ratios of excess portfolio returns for commodity futures markets

Commodity	CVAR	FCVAR	Commodity	CVAR	FCVAR
Canola	0.950*** (0.368)	1.182*** (0.368)	Palladium	1.512*** (0.368)	1.551*** (0.368)
Cocoa	-0.618* (0.367)	-0.054 (0.367)	Platinum	0.139 (0.383)	0.852** (0.383)
Coffee	0.222 (0.367)	0.309 (0.367)	Silver	2.256*** (0.369)	2.531*** (0.370)
Copper	-0.363 (0.367)	-0.221 (0.367)	Soybean	-0.320 (0.367)	-0.076 (0.367)
Corn	0.858** (0.368)	0.902** (0.368)	Soy meal	0.780** (0.368)	0.809** (0.368)
Crude oil	-0.215 (0.367)	0.446 (0.367)	Soy oil	0.129 (0.367)	0.253 (0.367)
Gasoline	0.525 (0.378)	0.575 (0.378)	Sugar	0.374 (0.367)	0.625* (0.367)
Gold	2.843*** (0.370)	2.882*** (0.370)	Wheat	0.246 (0.367)	0.079 (0.367)
Heating oil	-0.117 (0.367)	0.072 (0.367)			

Notes: This table reports estimated annualized Sharpe ratios of excess portfolio returns for commodity futures markets with daily ($h = 1$) rebalancing corresponding to the portfolio returns in Table 5. Standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively.

The results for the spot markets are presented in Table 6 (laid out as in Table 5) and show even larger excess returns with daily rebalancing than for the futures markets. We notice from Table 2 that there is no substantial difference between average returns in the spot and futures markets, so different profits in the two markets cannot be attributed simply to differences in unconditional average returns. The differences in profits reflect earlier findings from Figure 2 and Table 4 that futures returns are more difficult to forecast than spot returns, and from Table 3 that price discovery is primarily in the futures market for most commodities. There are four obvious exceptions, namely crude oil, gold, heating oil, and silver, where excess portfolio returns in the spot markets are lower than in the futures markets, which also agrees with the out-of-sample R^2 statistics in Table 4.

Since spot market positions are not as easily tradable as futures market positions, and since our transaction costs more closely match those found in futures markets, we do not focus too much on spot market profits results, but use them only as a metric of forecast comparison between the CVAR and FCVAR models. Generally, the results in Table 6 are similar to those in Table 5. Excess returns are highest with daily rebalancing, and this is also where the most significant returns are found. For a few commodities, these are in the hundreds of percent per annum. Comparing average returns of CVAR and FCVAR based forecasts (last row of the table), shows that the two models produce very similar returns on average, with a slight advantage to the FCVAR model in the daily and monthly rebalancing cases and a slight advantage to the CVAR model with weekly rebalancing.

As an additional metric of performance and comparison between the CVAR and FCVAR models, we report in Table 7 the estimated Sharpe ratios of excess portfolio returns for the commodity futures markets with daily rebalancing, i.e., corresponding to the first two columns of Table 5. Of course, the Sharpe ratio is defined simply as the average excess return divided by the standard

deviation of the excess return, and the reported Sharpe ratios are annualized. The asymptotic standard error of the Sharpe ratio, say $\hat{S}$, can be estimated by $s.e.(\hat{S}) = (1 + \hat{S}^2/2)^{1/2}/T^{1/2}$ (Jobson and Korkie, 1981, p. 893), which we report in parentheses. Since Sharpe ratios are fairly trivially calculated from the excess returns already presented, and to conserve space, we only report Sharpe ratios for our benchmark case. The Sharpe ratio measures the performance of a portfolio adjusted for its risk, but of course the usual disadvantage is that the Sharpe ratio relies on the notions that risk equals variance and that all variance is bad, such that it penalizes strategies with upside potential the same as strategies with similar magnitude downside risk. Thus, any conclusions drawn from the Sharpe ratios should be considered in this light.

The Sharpe ratios reported in Table 7 vary from -0.618 to 2.882 and reflect the heterogeneity of excess returns across commodities in Table 5. Only one negative Sharpe ratio is significant (and only at the 10% level, corresponding to the significantly negative return for cocoa in Table 5), whereas 14 of the positive Sharpe ratios are significant. In terms of Sharpe ratios, portfolios based on FCVAR model forecasts have an advantage over those based on CVAR model forecasts. In Table 7, Sharpe ratios for the FCVAR based portfolio are higher than those for the CVAR based portfolio in 16 of 17 futures markets. This finding is confirmed by the average Sharpe ratios across commodities, which are 0.541 and 0.748 for the CVAR and FCVAR models, respectively.¹¹

In general, both portfolio profits and Sharpe ratios are quite heterogeneous across commodities. This is not surprising since all studies that we are aware of in the literature on commodity futures markets (see examples in the introduction) find heterogeneous profits across commodities. While the goal of our paper is not to explain heterogeneity in profits, our analysis does offer some insights into why profits are heterogeneous. Table 1 shows that trading volume and liquidity of commodity futures are very heterogeneous and these are fundamental to the nature and magnitude of profits. Further, Table 2 shows that commodities differ on standard statistical features, such as mean, standard deviation, and higher moments of returns. Taken together, these differences in commodity fundamentals and statistical features suggest potentially quite different risk-return relationships, and it would appear that this translates into heterogeneous portfolio profits across commodities.

5.3 Robustness of economic significance

In this section, we investigate the robustness of our findings in two directions: (i) the choice of risk-aversion coefficient, γ , and (ii) the restrictions on the portfolio weights.¹² So far we considered an investor who takes a medium level of risk, i.e. risk-aversion parameter $\gamma = 6$, and trading regulations allowing short-selling and borrowing/leverage up to 50%. The choice of γ has obvious implications for portfolio returns via the calculation of weights in (12), from which we note that a lower (higher) risk-aversion investor places higher (lower) weight on the risky asset, although of course the investor remains risk averse for any $\gamma > 0$. Similarly, closer analysis of the earlier results reveals that optimal weights are quite often restricted by the 50% bound on short-selling and borrowing/leverage. Thus, robustness to these two model assumptions is relevant.

In our first set of robustness results, presented in Table 8, we report annualized excess portfolio returns for commodity spot and futures markets with daily rebalancing ($h = 1$) for both a lower risk-aversion investor ($\gamma = 3$) and a higher risk-aversion investor ($\gamma = 12$) with weights restricted to the interval $[-0.5, 1.5]$. Overall, these results support our earlier findings. First, excess portfolio

¹¹Results for the spot markets corresponding to Table 7 are available from the authors upon request. These Sharpe ratios range from -0.304 to 9.886 and are significant in 23 of 34 cases. The FCVAR based Sharpe ratio is higher than that for the CVAR in 10 of 17 commodities and their average Sharpe ratios are 2.100 (FCVAR) and 2.012 (CVAR).

¹²We also compared our results with returns obtained from a simple technical trading rule given by a moving-average (MA) crossover strategy. In spite of the fact that the MA trading rule assumes risk neutrality of the investor, while our mean-variance trading strategy assumes risk-aversion, the MA trading rule returns were less statistically significant and their Sharpe ratios substantially smaller. These results are available from the authors upon request.

Table 8: Annualized daily excess portfolio returns for alternative risk coefficients

Commodity	Lower risk-aversion ($\gamma = 3$)				Higher risk-aversion ($\gamma = 12$)			
	Spot markets		Futures markets		Spot markets		Futures markets	
	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR
Canola	217.200*** (9.538)	213.876*** (9.513)	16.683** (6.951)	21.239*** (6.802)	199.847*** (8.770)	196.809*** (8.602)	8.978** (3.504)	10.069*** (3.347)
Cocoa	2.459 (8.455)	12.209 (7.708)	-6.183 (6.477)	-1.118 (1.740)	2.928 (3.462)	7.346** (3.397)	-1.990 (2.049)	-0.386 (0.527)
Coffee	43.667*** (9.917)	49.239*** (9.969)	-10.053* (5.905)	-2.935 (5.805)	35.504*** (6.690)	40.463*** (6.734)	-2.681 (2.016)	0.056 (2.159)
Copper	194.244*** (7.708)	209.350*** (7.759)	1.377 (6.761)	3.402 (6.409)	175.775*** (7.184)	187.932*** (7.219)	1.346 (2.159)	2.229 (2.703)
Corn	27.831** (12.035)	23.984** (11.027)	21.781** (8.812)	23.310** (9.233)	11.881** (4.922)	9.703** (4.420)	8.543** (3.983)	9.240** (4.091)
Crude oil	-3.731 (6.657)	0.060 (6.424)	-1.814 (2.486)	7.555 (4.962)	-0.695 (5.223)	0.244 (5.267)	-0.236 (1.312)	2.290 (2.149)
Gasoline	38.279*** (12.484)	42.372*** (13.542)	12.029 (7.409)	13.034* (7.465)	26.995*** (6.466)	31.328*** (7.075)	4.252 (2.912)	4.248 (2.759)
Gold	1.863 (7.815)	2.451 (7.988)	49.086*** (5.970)	48.040*** (5.825)	0.355 (5.657)	2.451 (5.787)	25.540*** (3.599)	25.890*** (3.567)
Heating oil	-4.639 (5.436)	-1.006 (5.283)	-0.843 (5.039)	-0.008 (4.698)	-2.347 (3.338)	-0.231 (2.948)	-1.122 (2.771)	0.220 (2.218)
Palladium	123.770*** (12.786)	124.800*** (12.687)	38.578*** (9.055)	40.022*** (9.218)	102.359*** (9.416)	102.123*** (9.336)	16.267*** (4.173)	16.942*** (4.200)
Platinum	30.840*** (4.486)	32.429*** (4.357)	0.200 (4.178)	6.633* (3.476)	28.774*** (4.064)	30.297*** (3.902)	1.638 (2.714)	4.954** (1.961)
Silver	-1.256 (12.331)	-11.248 (13.206)	59.202*** (9.245)	48.506*** (6.595)	-2.913 (8.215)	-8.000 (8.447)	22.530*** (3.589)	21.024*** (3.136)
Soybean	22.161*** (8.370)	14.621** (6.795)	-4.968 (6.079)	0.507 (6.781)	15.260*** (3.973)	11.833*** (3.315)	-3.720 (3.019)	-0.098 (3.413)
Soy meal	22.883** (10.720)	20.831** (10.877)	15.138* (8.785)	17.995** (8.039)	13.319** (5.234)	10.967** (5.300)	10.698** (4.478)	9.791** (4.148)
Soy oil	21.473** (9.937)	18.789** (8.727)	5.366 (8.412)	6.511 (7.980)	13.557*** (4.740)	12.271*** (4.239)	-0.010 (3.940)	0.972 (3.688)
Sugar	102.521*** (11.420)	96.460*** (11.851)	10.986 (11.209)	18.920 (11.723)	77.431*** (8.010)	74.422*** (8.285)	4.783 (4.859)	7.054 (5.083)
Wheat	5.086 (11.317)	3.973 (5.922)	8.418 (7.477)	3.620 (7.890)	1.408 (3.140)	1.151 (1.615)	1.440 (3.338)	0.302 (3.277)
Average	49.685	50.188	12.646	15.014	41.143	41.830	5.662	6.753

Notes: This table reports annualized average excess portfolio percentage returns for commodity spot and futures markets for daily ($h = 1$) rebalancing with weights restricted to the interval $[-0.5, 1.5]$. The results are reported for a lower risk-aversion investor ($\gamma = 3$) and a higher risk-aversion investor ($\gamma = 12$) and for forecasts based on both the CVAR and FCVAR models. Standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively. The final row is the average return across all commodities.

returns are large and positive in almost all cases, and many are statistically significant. Second, and not surprisingly, average portfolio returns are higher for the lower than for the higher risk-aversion investor, with those for the medium risk-aversion investor in Tables 5 and 6 falling in between. Third, spot markets profits are once more found to be higher than futures markets profits, with the exception of crude oil, gold, heating oil, and silver as above. In fact, the lower risk-aversion investor is able to obtain about a 50% excess return per annum in the gold and silver futures markets.

Comparing CVAR and FCVAR profits in Table 8 using the average across all 17 commodities shows that the FCVAR model produces slightly higher excess returns on average for both the lower risk-aversion and the higher risk-aversion investor and for both the spot and futures markets. More specifically, focusing on the futures markets, which is the more difficult to predict but also where trading is more practical, the FCVAR model outperforms the CVAR model for 14 of 17 commodities when $\gamma = 3$ and for 13 of 17 commodities when $\gamma = 12$. Thus, for the majority of commodity futures markets, the FCVAR model forecasts produce higher portfolio returns.

In our second set of robustness results, presented in Table 9, we report annualized excess portfolio returns for commodity spot and futures markets with daily rebalancing ($h = 1$) for a medium risk-aversion investor ($\gamma = 6$) with weights restricted either to the interval $[0, 1]$, where no short-selling or borrowing/leverage is allowed or to the interval $[-1, 2]$, where up to 100% short-selling and borrowing/leverage is allowed. These results again support our earlier findings, and also illustrate the importance of the restrictions placed on the optimal portfolio weights.

When optimal portfolio weights are restricted to $[0, 1]$, returns are somewhat smaller than in the benchmark case in Tables 5 and 6 (where the interval is $[-0.5, 1.5]$). On the other hand, when optimal weights are restricted to $[-1, 2]$, which is less restrictive than the benchmark case in Tables 5 and 6, portfolio returns are much higher on average than in Tables 5 and 6. This shows the importance of the restrictions on the optimal weights. As with the earlier results, a few excess returns are negative but these are not statistically significant. The empirical result from Table 9 that eliminating (allowing more) short-selling and borrowing reduces (increases) profits is due to the fact that negative price movements (i.e., negative returns) are often predicted quite accurately by the models, thus generating a short position that is as big as allowed within the weight restrictions. Some of these negative returns are very large, and hence the profits made from the associated short positions can be very large as well. Indeed, with up to 100% short-selling and borrowing, the investor is able to obtain a 40–45% excess return per annum in the gold and silver futures markets.

Comparing the CVAR and FCVAR models in Table 9 shows a very similar picture to that in Table 8. On average across all 17 commodities, the FCVAR model produces slightly higher excess returns in both the spot and futures markets and for both the more and the less restrictive weights. Focusing on the futures markets, which is the more difficult to predict but also where trading is more practical, the FCVAR model outperforms the CVAR model for 14 of 17 commodities whether weights are restricted to $[0, 1]$ or are restricted to $[-1, 2]$. Thus, for the majority of commodities, the FCVAR model forecasts once more produce higher portfolio returns in the futures markets.

The overall implication is that the statistical superiority of the FCVAR, as documented in Sections 4 and 5.1, mostly extends to its economic importance, in particular for the futures markets where returns are more difficult to predict and trading is more practical. Therefore, certainly statistically and to some extent also economically, the FCVAR model offers investors a better guide to undertaking investment portfolio decisions.

6 Concluding remarks

This paper has analyzed the link between statistical models of forecasting for commodity prices and returns and their implications for investors. Identifying suitable forecasting models for asset returns is at the forefront of research in asset pricing. This is so because the accuracy of forecasts

Table 9: Annualized daily excess portfolio returns for alternative weight restrictions

Commodity	Weights in $[0, 1]$				Weights in $[-1, 2]$			
	Spot markets		Futures markets		Spot markets		Futures markets	
	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR	CVAR	FCVAR
Canola	110.428*** (6.308)	108.366*** (6.234)	9.976*** (3.862)	11.073*** (3.707)	311.796*** (12.859)	306.644*** (12.599)	17.238** (6.782)	19.856*** (6.511)
Cocoa	4.166 (4.293)	6.325 (4.205)	-1.729 (2.855)	0.112 (0.195)	3.391 (6.494)	12.348** (6.300)	-3.980 (4.099)	-0.771 (1.053)
Coffee	22.362*** (5.799)	25.209*** (5.816)	-3.856 (2.789)	-1.600 (2.561)	54.396*** (10.933)	64.245*** (10.993)	-5.451 (4.010)	0.063 (4.309)
Copper	96.182*** (5.051)	103.297*** (5.138)	2.101 (3.482)	2.930 (3.188)	279.279*** (10.469)	298.721*** (10.420)	2.709 (4.317)	4.105 (5.209)
Corn	15.988** (6.837)	15.208** (6.156)	11.340*** (4.349)	12.519*** (4.602)	22.870** (9.464)	19.265** (8.515)	16.966** (7.888)	18.197** (8.089)
Crude oil	-0.260 (4.137)	1.373 (3.913)	-0.672 (0.870)	3.303 (2.774)	-2.792 (8.062)	-0.829 (8.119)	-0.472 (2.624)	4.669 (4.259)
Gasoline	18.063*** (6.457)	17.586** (7.415)	4.452 (3.246)	5.293 (3.436)	49.838*** (12.301)	53.914*** (12.920)	8.688 (5.755)	8.713 (5.450)
Gold	5.062 (4.708)	5.641 (4.844)	24.049*** (3.292)	22.846*** (3.160)	-1.791 (9.186)	1.313 (9.365)	46.537*** (6.214)	46.650*** (6.119)
Heating oil	-1.599 (3.159)	0.459 (3.014)	0.221 (2.911)	1.235 (2.577)	-5.482 (5.768)	-1.498 (5.482)	-1.505 (5.110)	0.760 (4.339)
Palladium	65.158*** (7.832)	65.956*** (7.749)	17.816*** (4.435)	18.741*** (4.564)	166.335*** (15.381)	165.970*** (15.290)	32.535*** (8.345)	33.902*** (8.400)
Platinum	15.974*** (2.781)	16.750*** (2.685)	0.790 (2.482)	3.334* (1.883)	43.791*** (6.111)	46.825*** (5.945)	2.345 (4.720)	8.647** (3.662)
Silver	4.048 (7.060)	-1.404 (7.718)	25.397*** (4.820)	19.547*** (3.151)	-5.612 (14.193)	-16.779 (14.393)	44.150*** (7.072)	41.638*** (6.221)
Soybean	12.789*** (4.824)	8.566** (3.733)	0.666 (2.659)	1.926 (3.064)	26.803*** (7.123)	20.756*** (5.811)	-7.042 (5.994)	-0.517 (6.696)
Soy meal	10.317* (6.027)	11.377* (6.196)	5.935 (4.629)	9.500** (4.031)	23.614** (9.788)	19.179* (9.977)	20.559** (8.487)	18.485** (7.993)
Soy oil	10.708* (5.731)	8.789* (4.854)	2.571 (4.773)	3.434 (4.473)	23.080*** (8.447)	19.463*** (7.536)	1.174 (7.700)	2.364 (7.282)
Sugar	51.447*** (6.818)	48.738*** (7.052)	5.817 (5.872)	10.866* (6.355)	131.696*** (13.138)	125.969*** (13.582)	9.908 (9.564)	14.788 (9.977)
Wheat	3.279 (5.604)	1.912 (3.020)	5.330 (3.405)	3.501 (3.709)	2.827 (6.279)	2.303 (3.230)	3.085 (6.626)	0.577 (6.547)
Average	26.124	26.126	6.482	7.562	66.120	66.930	11.026	13.066

Notes: This table reports annualized average excess portfolio percentage returns for commodity spot and futures markets for a medium risk-aversion investor ($\gamma = 6$) with daily ($h = 1$) rebalancing. The results are reported for weights restricted to the intervals $[0, 1]$ and $[-1, 2]$ and for forecasts based on both the CVAR and FCVAR models. Standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% level is denoted by one, two, and three asterisks, respectively. The final row is the average return across all commodities.

has direct implications for investors' decision making, particularly with regard to portfolio choice. In this paper we take a step in this direction by proposing an FCVAR model for forecasting commodity spot and futures returns, based on recent empirical evidence of fractional cointegration in commodity spot and futures markets. We derive the best linear predictor for the FCVAR model and perform an out-of-sample forecast comparison with forecasts from the more standard CVAR model. In our empirical analysis of 17 commodity spot and futures markets, the fractional model is found to be statistically superior in terms of both in-sample fit and out-of-sample forecasting.

In terms of economic significance of the forecasts, we analyze this through a dynamic trading strategy based on a portfolio with weights derived from a mean-variance utility function. This analysis leads to statistically significant and economically meaningful profits in most commodity markets, and also that the statistical superiority of the FCVAR model extends to an economic advantage, at least in the futures markets where the fractional model has an advantage on average in terms of both portfolio excess returns and their Sharpe ratios. Our results are robust on several fronts. First, our out-of-sample forecasting evaluation exercise applies a number of statistical metrics. Second, we analyze several forecasting horizons. Third, we show that our results on profitability are robust to an investor's level of risk aversion as measured by the coefficient of relative risk aversion that enters in the utility function and hence in the calculation of portfolio weights. Finally, our results are robust to alternative restrictions on the optimal weights, reflecting alternative restrictions on short-selling and on borrowing/leverage.

References

1. Baillie, R.T., and Bollerslev, T. (1994) The long memory of the forward premium, *Journal of International Money and Finance*, 13, 565–571.
2. Campbell, J.Y., and Thompson, S.B., (2008) Predicting the equity premium out of sample: can anything beat the historical average?, *Review of Financial Studies*, 21, 1509–1531.
3. Carlini, F., and Santucci de Magistris, P., (2017) On the identification of fractionally cointegrated VAR models with the F(d) condition, *Journal of Business and Economic Statistics*, forthcoming.
4. Clark, T.E., and West, K.D., (2007) Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics*, 138, 291–311.
5. Coakley, J., Dollery, J., and Kellard, N., (2011) Long memory and structural breaks in commodity futures markets, *Journal of Futures Markets*, 31, 1076–1113.
6. Diebold, F.X., and Mariano, R.S., (1995) Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253–263.
7. Dolatabadi, S., Nielsen, M.Ø., and Xu, K., (2015) A fractionally cointegrated VAR analysis of price discovery in commodity futures markets, *Journal of Futures Markets*, 35, 339–356.
8. Dolatabadi, S., Nielsen, M.Ø., and Xu, K., (2016) A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets, *Journal of Empirical Finance* 38, 623–639.
9. Figuerola-Ferretti, I., and Gonzalo, J., (2010) Modelling and measuring price discovery in commodity markets, *Journal of Econometrics*, 158, 95–107.
10. Giacomini, R., and White, H., (2006) Tests of conditional predictive ability, *Econometrica*, 74, 1545–1578.
11. Gonzalo, J., and Granger, C.W.J., (1995) Estimation of common long-memory components in cointegrated systems, *Journal of Business and Economic Statistics*, 13, 27–36.
12. Gorton, G.B., Hayashi, F., and Rouwenhorst, K.G., (2013) The fundamentals of commodity futures returns, *Review of Finance*, 17, 35–105.
13. Graham-Higgs, J., Rambaldi, A., and Davidson, B., (1999) Is the Australian wool futures

- market efficient as a predictor of spot prices, *Journal of Futures Markets*, 19, 565–582
14. Hasbrouck, J., (1995) One security, many markets: determining the contribution to price discovery, *Journal of Finance*, 50, 1175–1199.
 15. Jensen, A.N., and Nielsen, M.Ø., (2014) A fast fractional difference algorithm, *Journal of Time Series Analysis*, 35, 428–436.
 16. Jobson, J.D., and Korkie, B.M., (1981) Performance hypothesis testing with the Sharpe and Treynor measures, *Journal of Finance*, 36, 889–908.
 17. Johansen, S., (1995) *Likelihood-based inference in cointegrated vector autoregressive models*. New York: Oxford University Press.
 18. Johansen, S., (2008) A representation theory for a class of vector autoregressive models for fractional processes, *Econometric Theory*, 24, 651–676.
 19. Johansen, S., and Nielsen, M.Ø., (2010) Likelihood inference for a nonstationary fractional autoregressive model, *Journal of Econometrics*, 158, 51–66.
 20. Johansen, S., and Nielsen, M.Ø., (2012) Likelihood inference for a fractionally cointegrated vector autoregressive model, *Econometrica*, 80, 2667–2732.
 21. Johansen, S., and Nielsen, M.Ø., (2016) The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models, *Econometric Theory* 32, 1095–1139.
 22. Kaldor, N. (1939) Speculation and economic stability, *Review of Economic Studies*, 7, 1–27.
 23. Lien, D., and Tse, Y.K., (1999) Fractional cointegration and futures hedging, *Journal of Futures Markets*, 19, 457–474.
 24. Locke, P.R., and Venkatesh, P.C. (1997) Futures market transaction costs, *Journal of Futures Markets*, 17, 229–245.
 25. Marquering, W., and Verbeek, M., (2004) The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis*, 39, 407–429.
 26. Maynard, A., and Phillips, P.C.B., (2001) Rethinking an old empirical puzzle: econometric evidence on the forward discount anomaly, *Journal of Applied Econometrics*, 16, 671–708.
 27. Miffre, J., and Rallis, G., (2007) Momentum strategies in commodity futures markets, *Journal of Banking and Finance*, 31, 1863–1886.
 28. Narayan, P.K., Narayan, S., and Sharma, S.S., (2013) An analysis of commodity markets: what gain for investors?, *Journal of Banking and Finance*, 37, 3878–3889.
 29. Narayan, P.K., Ahmed, H.A., and Narayan, S., (2014) Do momentum-based trading strategies work in the commodity futures markets?, *Journal of Futures Markets*, 35, 868–891.
 30. Nielsen, M.Ø., and Popiel, M.K., (2014) A Matlab program and user’s guide for the fractionally cointegrated VAR model, QED working paper 1330, Queen’s University.
 31. Shimotsu, K., and Phillips, P.C.B., (2005) Exact local Whittle estimation of fractional integration, *Annals of Statistics*, 33, 1890–1933.
 32. Wang, J., (2000) Trading and hedging in S&P500 spot and futures markets using genetic programming, *Journal of Futures Markets*, 20, 911–942.
 33. Yang, F., (2013) Investment shocks and the commodity basis spread, *Journal of Financial Economics*, 110, 164–184.