A note on the accuracy of commodity prices forecasts based on futures contracts

Marek Kwas[1], Michał Rubaszek[2]

Abstract

In this study we focus on the predictive power of futures for prices of six key global commodities. For that purpose, we use a comprehensive database of individual contracts to create continuous futures with weekly maturities ranging from one to fifty two weeks. We use this database to check how accurate futures-based forecasts in comparison to the random walk model are. We show that the futures curve does deliver accurate forecasts, which confirm the reliability of the common practice of financial institutions to use futures contracts in forecasting.

Keywords: *Commodity prices, futures contracts, forecasting*

JEL Classification: *G13, G17, Q02, Q47*

1. Introduction

Commodities play an important role in the world economy and commodity prices are important drivers of economic activity, inflation or trade balances. For that reason, understanding the dynamics and the ability to formulate reliable forecasts for commodity prices are important to many economic agents in decision making process. A question arises on whether it is possible to forecast commodity prices accurately. This question is the subject of a long-standing debate in the economic literature. On top of that, since commodity prices are more volatile than stock prices or exchange rates (Fratzscher et al., 2014), constructing a method that delivers accurate forecasts is considered to be a real challenge.

In this study we focus on the predictive power of futures for prices of six key global commodities: crude oil (West Texas Intermediate – WTI – as well as Brent), natural gas, gold, silver and copper. We analyse the accuracy of forecasts based on commodity futures, as they are regularly used in central banks and other policy institutions. This analysis supplements the earlier literature, which delivers ambiguous results. For instance, Alquist and Kilian (2010), Alquist et al. (2013) reported that futures-based forecasts of oil prices are no better than the no-change forecast from the random walk model. On the other hand, Chinn and Coibion (2014) showed that futures of energy commodities are essentially unbiased predictors of spot prices, providing good point and directional forecasts, whereas for industrial and precious metals this is not the case and futures perform relatively poorly.

We contribute to the literature by creating a comprehensive database of individual futures contracts, which are subsequently used to create continuous futures at weekly maturities ranging from one week to one year. We show that the futures curve does deliver accurate forecasts,

[1] SGH Warsaw School of Economics, Econometrics Institute, 162 Niepodległości, 02-554 Warsaw, Poland, marek.kwas@sgh.waw.pl.

[2] Corresponding author: SGH Warsaw School of Economics, Econometrics Institute, 162 Niepodległości, 02-554 Warsaw, Poland, michal.rubaszek@sgh.waw.pl.

which confirm the reliability of the common practice of financial institutions to use futures contracts in forecasting.

2. Data

We use weekly data (daily closing prices at the end of each week) for hundreds of individual futures contracts that were quoted between January 4th, 2009 and September 23rd, 2018. We consider prices of the following commodities: WTI and Brent crude oil, natural gas, gold, silver and copper. The data for individual contracts was downloaded from the CME Group, with an exception of Brent futures prices which were sourced from ICE. The time series of continuous weekly futures are derived as weighted averages of the prices of the two nearest active futures contracts, with the formula (Pindyck, 2001):

$$f_{t,h} = \left(\frac{h_2 - h}{h_2 - h_1}\right) f_{t,h_1} + \left(\frac{h - h_1}{h_2 - h_1}\right) f_{t,h_2}$$

where $f_{t,h}$ is the log price of a continuous futures expiring at horizon h observed in time t, and f_{t,h_1} and f_{t,h_2} denote the logs of prices of the two nearest active futures with $h_1 < h_2$. Following Pindyck (2001) we calculate the synthetic spot price as $s_t = f_{t,0}$. All futures analysed here are physically settled, with an exception of Brent futures which are settled financially. Thus, they all provide benchmarks for spot prices of respective commodities, derived in a similar way as described above.

The dynamics of time series for spot commodity prices are presented in Fig. 1. One can observe a qualitative difference of the natural gas market dynamics from the remaining five markets, the former characterised by rapidly changing trends and higher volatility. It is also possible to observe the post Great Financial Crisis decoupling of oil and natural gas markets, which is described in detail e.g. by Zhang and Ji (2018). Apart from the natural gas, the remaining markets exhibit relatively high degree of similarity in the analysed period.

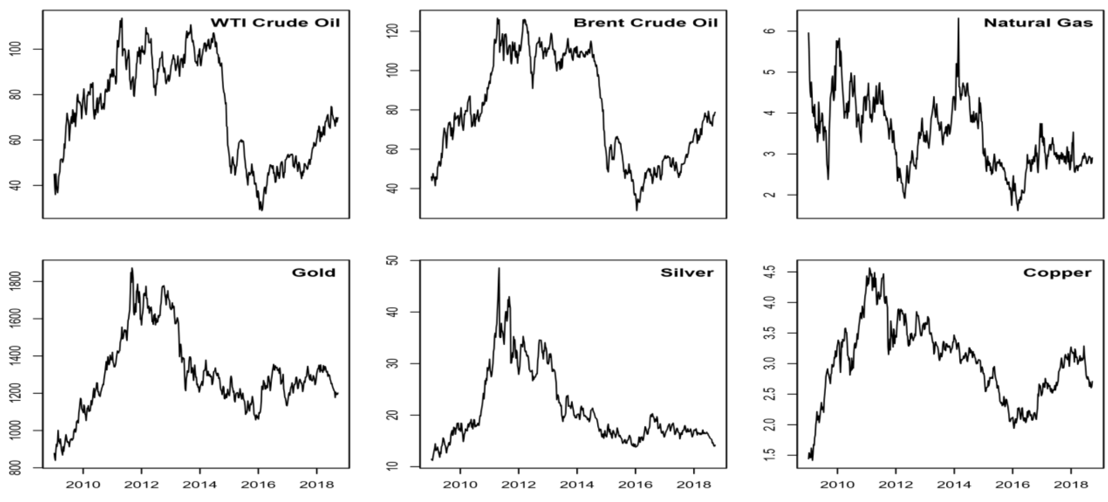


Fig. 1. The dynamics of spot commodity prices

The descriptive statistics for all commodities are presented in Table 1. The table shows gold and copper prices are the least volatile, whereas the variability of natural gas prices is the highest. In the case of skewness, the distributions of natural gas, WTI and copper prices turned out to be symmetric, whereas gold, Brent oil and silver prices were negatively skewed. Furthermore, the excess kurtosis statistics indicate heavy tails, which are the most pronounced for silver. The above is confirmed by the results of Jarque-Bera test, which show non-Gaussian distribution for all series. Let us point out that both WTI and Brent prices, although strongly correlated, show noticeable statistical differences. It illustrates the divergence of the US and European oil markets, with Brent oil gradually substituting WTI as the global benchmark (Manescu and Van Robays, 2014).

Table 1. Descriptive statistics

	Mean	**SD**	**Min.**	**Max.**	**Skew.**	**Kurt.**	**JB**
WTI	0.09	4.46	–15.91	23.01	–0.06	1.91	0.00
Brent	0.11	4.05	–15.26	13.97	–0.23	1.25	0.00
Nat. Gas	–0.14	6.44	–30.98	30.17	0.06	1.67	0.00
Gold	0.06	2.26	–10.13	6.78	–0.26	1.11	0.00
Silver	0.04	4.29	–31.98	13.37	–1.37	10.35	0.00
Copper	0.12	3.35	–18.05	13.94	0.07	2.18	0.00

Notes: The table presents the descriptive statistics for the weekly log-changes (×100) of analysed commodities. JB refers to the p value of the Jarque-Bera normality test.

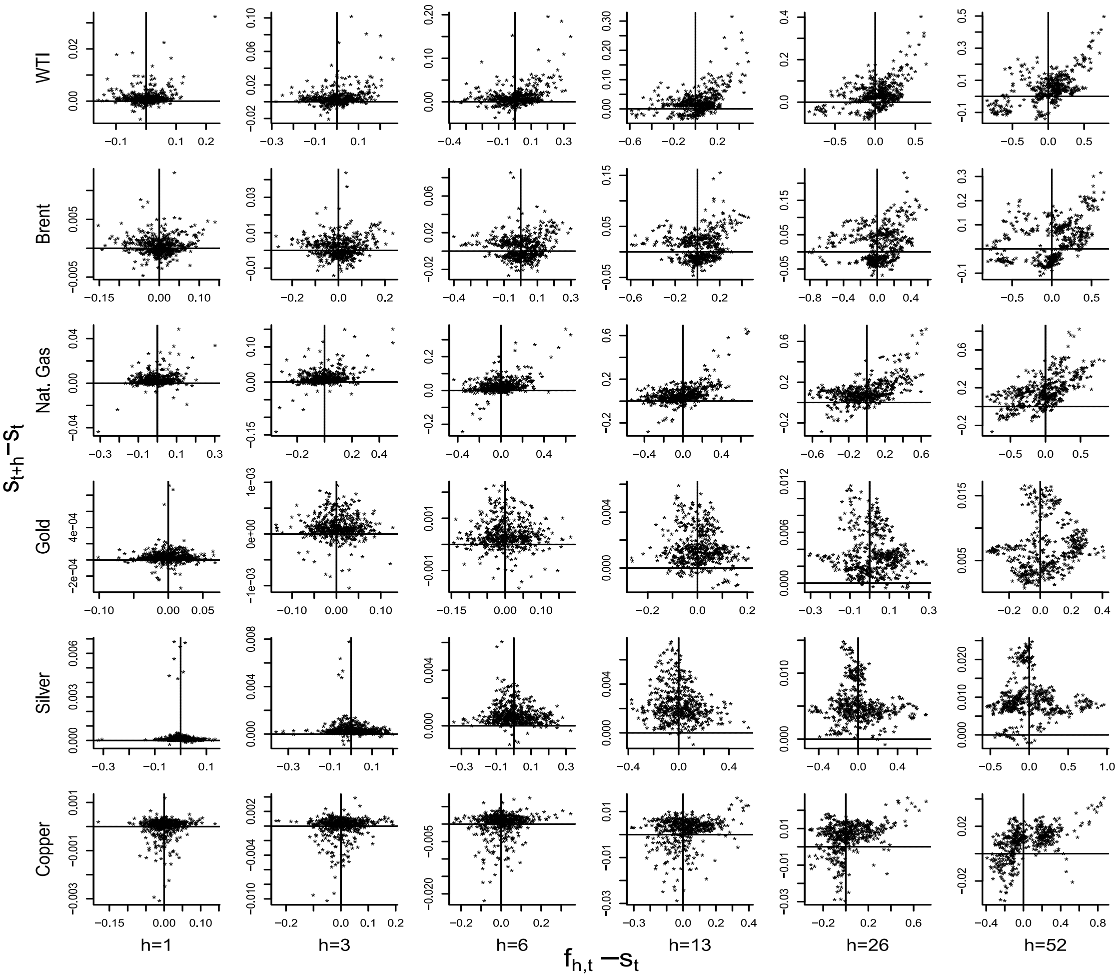
3. In-sample evidence

In this section we analyse how the spot price adjusts to the basis which is defined as the difference between futures and spot prices (Fama and French, 1987). Specifically, we estimate the parameters of regressions:

$$s_{t+\mathrm{h}} - s_t = \alpha_h + \beta_h(f_{h,t} - s_t) + \varepsilon_{t+h},$$

where s_t and $s_{t+\mathrm{h}}$ are the spot prices at times t and $t + h$ respectively, and $f_{h,t} - s_t$ is the basis at h weeks ahead horizon. Positive β_h means that the spot price adjusts to the basis. Moreover, if futures are effective predictors for the spot price the coefficients of the above regression should equal to $\alpha_h = 0$ and $\beta_h = 1$.

The relation between the basis and subsequent changes in spot prices over various horizons is illustrated by Fig. 2. It shows that the correlation is rather weak, which is confirmed by the estimates of the linear regression for six commodities and various horizons (see Table 2). Even if the estimates of are usually positive, the fit of the models is rather poor. The values of show that the model explains only a small fraction of spot price dynamics of investigated commodities. The poor fit of regression (1) might be due to two reasons: the predictive content of the basis is low or the relationship between the basis and the subsequent changes in the spot price is non-linear, which is suggested e.g. by Gulley and Tilton (2014) and Fernandez (2016).



Notes: The horizontal axes present the basis for weeks ahead horizon and the vertical axes describe the subsequent adjustment (log change) of spot prices over weeks horizon.

Fig. 2. The basis and commodity price adjustment

Table 2. The adjustment in the spot price to the basis

	WTI		**Brent**		**Nat. Gas**		**Gold**		**Silver**		**Copper**	
h	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2
1	2.65	0.03	0.78	0.00	2.91	0.07	2.77	0.00	-2.27	0.00	5.37	0.00
3	1.90	0.06	0.85	0.01	1.75	0.12	-2.45	0.00	-9.36	0.01	3.94	0.01
6	1.69	0.12	0.67	0.01	1.40	0.21	-2.55	0.00	-14.02	0.01	3.44	0.02
13	1.91	0.20	1.06	0.03	1.18	0.28	-6.55	0.01	-15.12	0.02	4.41	0.05
26	1.78	0.22	1.14	0.06	1.20	0.32	-3.47	0.01	-7.99	0.01	6.79	0.12
52	2.05	0.36	1.57	0.16	1.22	0.37	6.18	0.02	-1.42	0.00	11.24	0.30

Notes: The table presents the estimates of regression $s_{t+h} - s_t = \alpha_h + \beta_h(f_{h,t} - s_t) + \varepsilon_{t+h}$. The parameters were estimated with weekly data covering the period between January 4th, 2009 and September 23rd, 2018.

4. Forecasting contest results

We compare the accuracy of forecasts from two models. The first one is a widely used benchmark, i.e. the naive random walk (RW). From the perspective of a practitioner, there is nothing more conservative than assuming that the price will remain constant over the forecast horizon. The forecast for horizon h formulated in period t is:

$$s_{t+h,t}^{\mathrm{RW}} = s_t.$$

The second method, which we test in this study, is based on continuous futures. For this method the value of forecast formulated at period t for horizon h is:

$$s_{t+h,t}^{\mathrm{Fut}} = f_{h,t}.$$

where $f_{h,t}$ is the price of the continuous futures contract maturing at horizon h.

We evaluate the accuracy of forecasting for horizons ranging from one week to one year (h = 52 weeks). The first set of forecasts is formulated for the period between January 4th, 2009 and December 27th, 2009 and the last one for the period between September 16th, 2018 and September 15th, 2019. Given that our sample ends on September 16th, 2018, we assess the quality of forecasts using 507 forecast errors for one-week horizon, whereas for horizon the number of observations is equal to $508 - h$.

We begin our analysis by measuring the forecasting performance of the competing methods with the root mean squared and mean absolute forecast errors statistic (RMSFE and MAFE). Table 3 reports the values of RMSFE and MAFE for RW, whereas for the futures-based forecast it presents ratios relative to RW. Thus, the values below unity indicate the outperformance of the RW benchmark. We also test the null of equal forecast accuracy with the one-sided Coroneo and Iacone (2015) version of the Diebold and Mariano test, which offers relatively good finite sample size and power performance (Harvey et al., 2017).

Table 3. RMSFE and MAFE of RW and futures-based forecasts

	Random walk forecasts						**Futures-based forecasts**					
	1	**3**	**6**	**13**	**26**	**52**	**1**	**3**	**6**	**13**	**26**	**52**
							RMSFE					
WTI	0.05	0.08	0.11	0.17	0.23	0.32	0.991**	0.974**	0.949**	0.918*	0.908*	0.870*
Brent	0.04	0.07	0.10	0.16	0.23	0.32	0.999	0.997	0.996	0.981	0.966	0.931
Nat. Gas	0.07	0.11	0.14	0.19	0.25	0.32	0.982**	0.960**	0.926*	0.929	0.968	1.012
Gold	0.02	0.04	0.05	0.07	0.11	0.16	1.000	1.000	1.000	1.000	0.997	0.989
Silver	0.04	0.08	0.10	0.14	0.20	0.30	1.000	1.000	1.000	1.001	1.001	1.000
Copper	0.03	0.06	0.08	0.12	0.17	0.24	0.999**	0.998**	0.994**	0.988*	0.980*	0.970**

							MAFE					
WTI	0.03	0.06	0.08	0.13	0.17	0.24	0.991**	0.982**	0.943**	0.910*	0.893*	0.837*
Brent	0.03	0.06	0.08	0.12	0.17	0.25	0.999	1.001	0.999	0.963	0.941	0.908
Nat. Gas	0.05	0.08	0.11	0.15	0.20	0.25	0.992**	0.978**	0.956*	0.911	0.936	1.003
Gold	0.02	0.03	0.04	0.06	0.09	0.13	1.000	1.000	1.001	0.998	0.993	0.989
Silver	0.03	0.06	0.08	0.11	0.15	0.24	1.000	1.001	1.002	1.002	1.007	1.009
Copper	0.03	0.04	0.06	0.09	0.13	0.19	1.000**	0.998**	0.995**	0.990*	0.982*	0.979**

Notes: The figures describe the values of RMSFE and MAFE from futures-based forecasts in comparison to the RMSFE and MAFE from RW. Asterisks ***, ** and * denote the 1%, 5% and 10% significance levels of the one-sided Coroneo and Iacone (2015) version of the Diebold-Mariano test with the alternative that a given model performs better than RW.

A look at Table 3 leads to several immediate findings. The first one is that RMSFEs from RW for metal commodities are somewhat lower than those for energy commodities. The second finding is that for WTI, natural gas and copper, futures-based forecasts outperform RW consistently across horizons in both RMSFE and MAFE measures. The only exception is the longest horizon for natural gas prices. Moreover, for WTI and copper the gains are statistically significant, whereas for natural gas it holds for short to medium horizons. The third observation is that, surprisingly, one can observe a substantial difference for WTI and Brent oil prices. Even though for both commodities futures-based forecasts are better than those from the RW, only for WTI the gains are significant. Finally, in case of both precious metals – gold and silver – futures-based forecasts are of comparable accuracy to RW, with insignificant gains for gold for longer horizons. It can be added that our results are more optimistic than those of Alquist and Kilian (2010) and Alquist et al. (2013), who found that futures-based forecasts for oil prices are statistically indistinguishable from those of RW. On the other hand, our results confirm the findings of Chinn and Coibion (2014) or Fernandez (2017), who show that there is a predictive content in the futures prices.

We complement the forecast accuracy analysis with a visual illustration, by plotting the whole sequence of forecasts conducted at different points in time and comparing them to achieved values (see Fig. 3). A first inspection indicates that all models encounter problems to forecast sharp movements in prices, for instance the WTI price decline in 2015. It is mainly because forecasts are excessively conservative, in a sense that they do not deviate substantially from the last spot price.

We conclude the forecasting contest by calculating the percentage of correct sign forecasts together with the independence test of Pesaran and Timmermann (1992). Table 4 indicates that only for natural gas more than half of predictions are of correct sign consistently across horizons. In turn, for WTI, Brent and copper the share is above 50% only for longer horizons. For gold and silver the fractions of correct predictions are either below or insignificantly different from 50%.

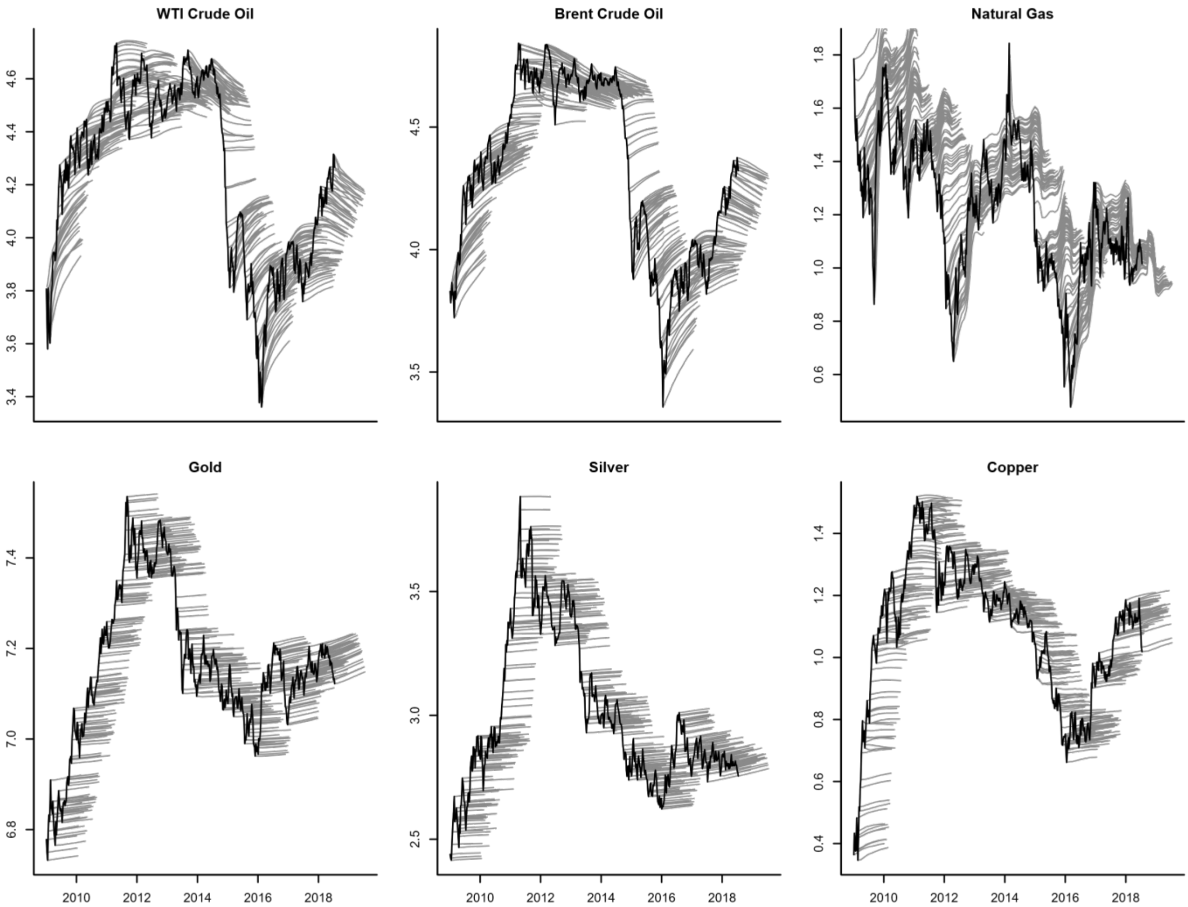


Fig. 3. Rolling forecasts for the log WTI price

Notes: Rolling futures-based forecasts for the logs of prices of six analysed commodities. The first set of forecasts is generated with data ending on December 28th, 2008 and the last set of forecasts is based on series ending on July 15th, 2018, which in total makes forecasts from each model.

Table 4. Fraction of correct sign predictions

	1	**3**	**6**	**13**	**26**	**52**
WTI	0.528	0.538	0.582***	0.604***	0.622***	0.765***
Brent	0.504	0.472	0.510	0.545**	0.600***	0.715***
Nat. Gas	0.516**	0.534***	0.534***	0.556***	0.588***	0.596***
Gold	0.512	0.480	0.496	0.545	0.577	0.564
Silver	0.478	0.480	0.454	0.501	0.459	0.456***
Copper	0.502	0.508	0.524	0.598***	0.631***	0.634***

Notes: The figures represent the fraction of futures-based forecasts that correctly predict the sign of the change. Asterisks ***, ** and * denote the rejection of the null of the independence test by Pesaran and Timmermann (1992) at the 1%, 5% and 10% significance levels.

Conclusions

A common practice of policy making institutions is to rely on futures prices while formulating predictions for commodity prices. It raises a question whether futures contracts contain information that could be extracted to precisely forecasts spot price movements. In this paper, we addressed this question by comparing the accuracy of futures-based forecasts for six main global commodities to those from the random walk model.

Using weekly data from the period 1999–2018, we found that futures-based forecasts are superior or comparable to the naive ones. It turned out that the gains are higher for longer rather than shorter horizons. Moreover, we showed that using futures is most advisable for energy commodities as well as copper, whereas the results for gold and silver are more mixed. A broad conclusion that emerges from our analysis is that it is justified to use forecasting methods that exploit information content of the futures contract prices rather than rely on no-change random walk forecast.

Acknowledgements

The project is financed by the Polish National Center of Science (Narodowe Centrum Nauki). Grant number 2017/25/B/HS4/00156.

References

Alquist, R., & Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* 25(4), 539–573.

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the price of oil. In: Elliott, G., Timmermann, A. (eds.), *Handbook of Economic Forecasting,* 2, 427–507.

Chinn, M.-D., & Coibion, O. (2014). The predictive content of commodity futures. *Journal of Futures Markets* 34(7), 607–636.

Coroneo, L., & Iacone, F. (2015). Comparing predictive accuracy in small samples. *Discussion Papers* 15/15, Department of Economics, University of York.

Fama, E.F., & French, K.R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *The Journal of Business* 60(1), 55–73.

Fernandez, V. (2016). Spot and futures markets linkages: Does contango differ from backwardation? *Journal of Futures Markets* 36(4), 375–396.

Fernandez, V. (2017). A historical perspective of the informational content of commodity futures. *Resources Policy* 51(C), 135–150.

Fratzscher, M., Schneider, D., & Van Robays, I. (2014). Oil prices, exchange rates and asset prices. *ECB Working Paper Series* 1689, European Central Bank.

Gulley, A., & Tilton, J.E. (2014). The relationship between spot and futures prices: An empirical analysis. *Resources Policy* 41(C), 109–112.

Harvey, D.I., Leybourne, S.J., & Whitehouse, E J. (2017). Forecast evaluation tests and negative long-run variance estimates in small samples. *International Journal of Forecasting* 33(4), 833–847.

Manescu, C., & Van Robays, I. (2014). Forecasting the brent oil price: Addressing time-variation in forecast performance. *ECB Working Paper Series* 1735, European Central Bank.

Pesaran, M.H., & Timmermann, A. (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics* 10(4), 561–565.

Pindyck, R.S. (2001). The dynamics of commodity spot and futures markets: A primer. *The Energy Journal* 22(3), 1–29.

Zhang, D., & Ji, Q. (2018). Further evidence on the debate of oil-gas price decoupling: A long memory approach. *Energy Policy* 113(C), 68–75.