Visnyk of the National Bank of Ukraine, 2024, No. 254, pp. 15–28

SHORT-TERM FORECASTING OF GLOBAL ENERGY AND METAL PRICES: VAR AND VECM APPROACHES

DIANA BALIOZ[a]

[a]National Bank of Ukraine
E-mail: Diana.Balioz@bank.gov.ua

Abstract This study introduces a set of multivariate models with the aim of forecasting global prices of 1) crude oil, 2) natural gas, 3) iron ore, and 4) steel. Various versions of vector autoregression and error-correction models are applied to monthly data for the short-term prediction of nominal commodity prices six months ahead. The fundamentals for metal and energy price predictions include inter alia, stock changes, changes in commodity production volumes, export volumes by the largest players, changes in the manufacturing sector of the largest consumers, the state of global real economic activity, freight rates, and a recession indicator. Kilian's (2009) index of global real economic activity is found to be a useful proxy for global demand and a reliable input in forecasting both energy and metal prices. The findings suggest that models with smaller lag orders tend to outperform those with a higher number of lags. At the same time, selected individual models, while showing a standalone high performance, have varying forecast precision during different periods, and no individual model outperforms others consistently throughout the forecast horizon. Note that the price projections obtained from the models could be used further for the longer-term forecasting of commodity prices. Our short-term hands-on framework could be a useful forecasting tool for central bank policymakers and researchers.

JEL Codes C32, C53, Q02

Keywords forecasting, commodities, forecast evaluation, VAR models, VECMs

1. INTRODUCTION

Commodity prices play an increasingly important role in influencing global inflation and the macroeconomic environment. For many developing economies, primary commodities remain the main drivers of the balance of payments, while price fluctuations affect their macroeconomic performance. Energy transition, the COVID-19 pandemic, and russia's war against Ukraine have led to sharp price changes, highlighting the high volatility of the commodity markets and the vulnerability of commodity-dependent countries to price shocks (Baffes and Nagle (eds.), 2022). Therefore, a deeper understanding of commodity price movements and the factors behind them are crucial to policymakers, international institutions, and think tanks.

The approaches used to forecast the prices of energy and metals differ in many ways depending on the purpose of studies, the benchmarks chosen, the frequency of data, and forecasting techniques. There are papers that employ univariate techniques (Tularam and Saeed, 2016; Nademi and Nademi, 2018; Hosseinipoor et al., 2016), multivariate econometric models (Nick and Thoenes, 2014; Berrisch and Ziel, 2022), and machine learning approaches (Kriechbaumer et al., 2014; Li et al., 2020). Various research studies focus either on short- or long-term forecasting tools: they aim to predict the spot (nominal or real) prices or futures prices of commodities, and find the interrelation between commodity prices and their potential impact on one another (West and Wong, 2014).

This paper introduces the hands-on approach of multivariate models for the short-term forecasting of global prices for crude oil, natural gas, steel, and iron, and analyzes the forecasting performance of these techniques. More specifically, this study focuses on predicting the spot nominal monthly prices of commodities six months ahead, while the majority of papers develop models to predict either futures prices (Bowman and Husain, 2004; Reichsfeld and Roache, 2011; Ambya et al., 2020), spot real quarterly prices (Baumeister and Kilian, 2013, 2014; Wårell, 2018) or price indices (Chou et al., 2012). Our short-term hands-on framework could be a useful tool for central banks and analysts, while the price projections obtained from these models could be used further as assumptions for the longer-term forecasting of commodity prices.

The paper is organized as follows. Section 2 provides a review of relevant research literature on commodity price forecasting, and examines the modern econometric approaches used to predict oil, gas, iron ore and steel prices. Section 3 describes the general methodology of VAR/VECM models, and is divided into subsections to analyze in detail the specifications of the models and data for each of the four commodities discussed. In Section 4, we look at the results of our short-term models and assess their forecasting properties. Finally, Section 5 offers conclusions and recommendations on how this forecasting approach can be improved.

2. LITERATURE REVIEW

Forecasting commodity prices is generally considered a challenging task, and rightfully so, given their volatility, dependence on many economic and financial factors, trend changes over time, and the huge variety of methods and approaches used in forecasting. The literature on commodity price forecasting differs significantly, depending on the purposes of the studies, the techniques used and the features of each commodity market. There are papers that employ econometric approaches, namely univariate and multivariate forecasting models, and those that use machine learning and non-parametric techniques. Different studies also focus on short- or long-term forecasting tools; they aim to predict spot or futures prices of commodities, nominal or real (spot) prices; and seek to find the interrelation between different commodity prices and their potential impact on one another. In this section, we review the literature by the commodities of interest.

The literature on predicting global crude oil prices is probably the most extensive, compared to other commodity groups, due to the impact oil prices have on inflation and macroeconomic development. For example, Tularam and Saeed (2016) focus on univariate time-series models to predict oil price movements, and find that the ARIMA model is a better fit for daily WTI oil prices than the exponential smoothing and Holt-Winters models. Conversely, Nademi and Nademi (2018) find that the semiparametric Markov-switching AR-ARCH model outperforms other simple approaches, including ARIMA and GARCH, when it comes to forecasting OPEC, WTI and Brent oil prices. However, univariate forecasting models rely only on one input – in this case, the price of crude oil itself and its past patterns – and do not take into account other factors that might impact the price. Whereas multivariate models are more sophisticated and include the economic determinants of price movements.

It is worth noting that a number of research papers use oil *futures* prices to predict movements in *spot* prices. The intuition behind this is that oil is both a physical commodity and a financial asset. Thus, it is often argued that there is a theoretical link between its spot and futures prices, and the slope of oil futures prices may help predict the movements in spot prices. The empirical evidence, however, is mixed. Chernenko et al. (2004), for instance, argue that oil futures prices (just as natural gas futures prices) show little evidence for risk premiums and can be used to forecast spot prices. Some central banks tend to use futures curves for the short-term forecasting of oil spot prices as they are simple and easy to communicate. Reichsfeld and Roache (2011) prove empirically that futures-based forecasts outperform random walk models over a three-month horizon, but not over longer forecast periods. In contrast, Alquist et al. (2011) conclude that futures prices are not good predictors of nominal oil prices and do not outperform no-change forecasts. There is also an arbitrage relationship between oil futures and spot prices, which, inter alia, means that the slope of oil futures prices is rather flat relative to the changes in oil spot prices (Nixon and Smith, 2012). Moreover, due to oil being a physical and storable good with limited inventories, its futures price curve is downward sloping most of the time, except for some occasions of contango (i.e. an upward sloping curve) when there is ample supply and a high level of oil stocks. In general, futures-based models alone do not prove accurate in predicting spot oil prices, Therefore, other approaches or even combinations of different models should be used (ECB, 2015).

A growing number of recent research papers focus on vector autoregression (VAR) models to predict nominal and real *spot* prices on the commodity markets as these models take into account the economic determinants of price movements and market fundamentals. Such models consider each variable as a function of its own past values and past values of other variables in the model. They also provide estimates of the impact of supply and demand shocks on commodity prices, which makes such models a useful analytical tool. VAR and structural VAR models have smaller forecast errors and prove to be more accurate in forecasting oil price movements than other time-series techniques, especially in the short run (as discussed in Baumeister and Kilian (2013, 2014), Kilian and Murphy (2014) etc.). For example, Baumeister and Kilian (2014), in their seminal work, study real-time forecasting techniques, including forecast combinations, to predict the quarterly real price of oil over short-term horizons. The authors use market fundamental variables, such as a change in crude oil production, Kilian’s (2009) index of global real economic activity, a change in oil inventories and so on to conclude that VAR models based on monthly data are the most accurate tools for predicting real oil prices on a quarterly basis. At the same time, one may argue that the accuracy and stability of individual forecasts are time varying, and different models might be suitable for different periods. Thus, the combination of individual forecasts should improve the accuracy of forecasts and help overcome the potential misspecifications of individual models. Baumeister and Kilian (2014) developed a number of forecasting models to test an equal-weighted combination of a monthly VAR model (as the best-performing one among the individual approaches) and the futures-based approach, which provides some MSPE improvement of the forecast of the U.S. real refiners’ acquisition cost (RAC) and WTI price, while deteriorating directional accuracy. However, this combination method does not improve forecast accuracy for Brent oil real prices at all, thus the results are mixed. In their later study, Baumeister and Kilian (2013) demonstrate that the combination of four models (namely a VAR model, a model based on non-oil commodity prices, a method based on futures spreads, and a time-varying product spread approach) with inverse MSPE weights actually provides better forecast accuracy. The results hold for the U.S. refiners’ acquisition cost for crude oil imports and WTI oil prices over January 1992 through September 2012, but there are no results for Brent oil prices, due to the lack of suitable data.

In contrast, Manescu and Van Robays (2014) focus on the current international benchmark price and prove empirically that for Brent oil prices for the period from Q1 1995 through Q4 2012, a four-model combination, which consists of futures, risk-adjusted futures approaches, Bayesian VAR, and a DGSE model, is the best forecasting technique. This equal-weighted model combination produces robust forecasts of oil prices over the studied period, reduces the forecast bias, and outperforms simple models in out-of-sample exercises. At the same time, this combination approach is found to improve the forecasts of the futures-based model and the random walk model (on average, up to 11 quarters ahead), but there is no evidence that it outperforms other forecasting approaches. When compared to benchmarks other than predictions by futures prices, it may produce worse results. Moreover, given the latest patterns on the global oil market, it is unclear if this particular combination of models can be robust over a more recent period than that discussed in a paper by Manescu and Van Robays (2014), i.e. after 2012.

The research papers mentioned above describe the methods of forecasting *real* oil prices or *real* RAC (refiner acquisition cost) based on global supply and demand variables, according to economic theory. These real price forecasts could then be used by analysts and policy makers. Nominal price forecasts, which are usually of interest, could be derived from them, using separate forecasts of the CPI. Thus, the models proposed in the aforementioned papers cannot directly predict the *nominal* price of oil and require other models or external forecasts for that. Meanwhile, Beckers and Beidas-Strom (2015) introduce CPI inflation into the VAR model to fill in this gap in the literature, and they find that this VAR model outperforms the random walk and futures-based forecasts. The authors also conclude that there is merit in combining forecasts of futures and VAR models, although only for horizons beyond 18 months.

Just as in the case of forecasting oil prices, the literature on predicting natural gas prices differs in terms of the purpose of forecasts, chosen benchmarks or markets, the use of additional price determinants, and forecasting methodology. For example, Hosseinipoor et al. (2016) apply the ARIMA/GARCH combined approach to predict Henry Hub (U.S. market) monthly spot prices in the long run. In contrast, Jin and Kim (2015) suggest using a combination of wavelet decomposition and the ARIMA model, for more precise forecasting of Henry Hub weekly spot prices.

A growing body of literature, however, studies the impact of additional factors on natural gas prices and suggests that multivariate models are more precise for forecasting. For instance, Nick and Thoenes (2014) develop a structural VAR model for the German (NCG) gas hub over the period of 2008-2011 and find that in the short-term gas prices depend on the temperature, storage and supply shortfalls, while in the long-term crude oil and coal prices have an impact on gas price developments. Moreover, the authors argue that while supply interruptions have an impact on NCG gas prices, their effects might be overestimated, since some of the supply shortfalls overlapped with extraordinary demand-side conditions. Thus, it is important to not only focus on the supply-side aspects of the gas market in order to improve its security, but also to address the flexibility side of market demand.

Hamie (2020) tests an extensive set of methodologies to model natural gas prices, including game theory, information theory, records theory, non-parametric approaches, and the multivariate regression analysis. As for multivariate models, the author employs the VECM (vector error-correction model) to account for the effects of fundamental variables on gas price formation. Hamie (2020) argues that natural gas prices in the German hub (NCG benchmark) are affected by the weather conditions measured by heating degree days, the storage utilization of gas, coal and crude oil prices, the euro-dollar exchange rate, as well as by the lags of their own prices. At the same time, it is argued that many other factors might determine natural gas prices, including oil storage inventories, extreme weather events, political factors, and financial market conditions. Similarly, Berrisch and Ziel (2022) use state-space models to forecast daily and monthly gas prices based on various factors, such as seasonality, air temperature, risk premiums, storage levels, the price of European Emission Allowances, and the prices of oil, coal and electricity. As can be seen even from the examples above, apart from the main supply and demand factors, it is common to use weather-related variables to model natural gas prices. Temperature has an impact on gas consumption as the primary usage of gas is for heating purposes. Moreover, gas is also used in hot weather to cool buildings. Heating degree days (HDD) and cooling degree days (CDD) are the measures that quantify respective energy needs depending on the temperature (Sharma et al., 2021).

Gao et al. (2021) develop a class of hybrid time-varying parameter models (i.e. combinations of TVPSV and Markov switching classes of models) for three gas markets, namely the U.S., and the European and Japanese markets. The authors find evidence that time-varying models are better for forecasting European and Japanese monthly gas prices than static models, while for the U.S. market a simple AR model outperforms other studied approaches.

In recent literature, more and more papers focus on the Dutch TTF price as a benchmark, because the TTF is currently Europe's main gas hub, and it is becoming widely internationalized. Hulshof et al. (2016), for example, prove that daily TTF gas prices predominantly depend on market fundamentals, such as weather and storage availability, while the linkage between crude oil and natural gas prices is not strong over the period of 2011–2014, and coal prices are insignificant for the day-ahead forecasting of gas prices. Obadi and Korcek (2020) examine month-ahead TTF contracts over 2016–2019 and find evidence that monthly gas prices are driven by demand and supply fundamentals, like the price of German power and the price of coal, changes in total demand for gas, storage capacity, and in LNG variables.

The literature on predicting iron ore and steel prices is interrelated, given the direct links between these two commodity markets. Iron ore is the primary raw material that is used in the production of steel and steel products. Almost all iron ore that is mined is used in steelmaking, thus the demand for iron ore is primarily defined by the demand for steel. Therefore, the factors impacting global iron ore and steel prices are related.

The Asian market, more specifically Chinese consumers, play a great role in shaping the global iron ore market. China is a dominant consumer of metals in general and iron ore in particular, as it is the world's largest producer of steel. The growing importance of the Chinese market is often seen as one of the main reasons behind the transition of the iron ore pricing mechanism from an annual negotiation system to spot market pricing in late 2008–2010 (Wårell, 2014; Wårell, 2018). The author also finds GDP growth in China to have the strongest impact on iron ore prices. Ma and Zhen (2020) analyze spot prices for iron ore in 2014-2018 and find evidence that China's steel production affects the volatility of iron ore prices, while the mean and volatility of prices are also influenced by changes in port stocks.

Mei and Chen (2018) study the factors influencing steel overcapacity on the Chinese market and find that they include the steel export rate, investment in fixed assets, the growth rate of real estate construction areas, concentration levels in the iron ore and steel industry, iron ore prices, and local government investment. The capacity utilization rate has an impact on market competition and prices. For example, moderately excess capacity can improve competition and boost technological innovation in the industry, whereas severe overcapacity might provoke vicious competition, weak prices and a deterioration in the business environment.

There is a growing bulk of literature suggesting that there are links between the prices of different commodity groups. For example, Campiche et al. (2007) find cointegrating relationships between crude oil, corn and soybean prices over 2006–2007. Nazlioglu and Soytas (2012) prove the presence of dynamic cointegration links between the global oil price and prices of twenty-four agricultural commodities over an extended period of 1980–2010. Meanwhile, West and Wong (2014) employ factor models to predict the monthly prices of energy, metals and agricultural commodities using a sample of 1996 to 2012. Ding and Zhang (2020) use cross-market information from long-run equilibrium models to predict commodity prices, such as oil, copper, cattle, corn, and gold.

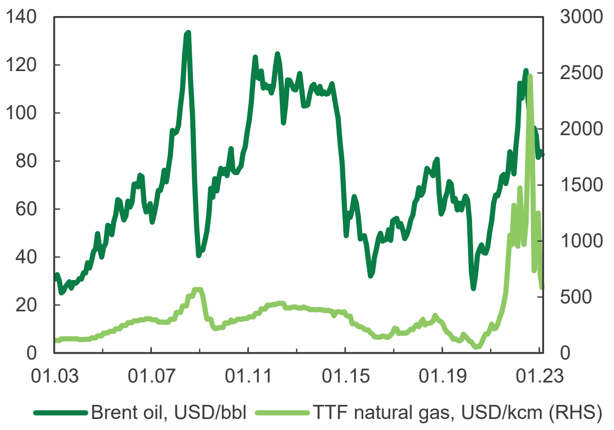
A number of papers argue that energy and crude oil prices can determine the prices of other commodities, including metals. The intuition behind this argument is that oil constitutes an important operational expense and a power source for the shipping industry, and commodities, such as metals, are often transported by sea. Therefore, the literature provides empirical evidence of crude oil prices having long-term cointegration relations with other commodity prices. For instance, Chou et al. (2012) prove, using a VARMA model, that global steel prices measured by the CRU steel price index were cointegrated with, and affected by, crude oil prices over the period of 2000 to 2010. Similarly, Asmoro (2017) has evidence that hot rolled coil (HRC) and billet steel prices are impacted by crude oil prices over the period from May 1996–December 2016. Moreover, studies by Chou et al. (2012) and Asmoro (2017) suggest that there is a unidirectional causal relation between crude oil and steel, i.e. the steel price is impacted by the crude oil price, whereas changes in oil prices are not influenced by steel prices.

Therefore, the literature on modeling the prices of energy and metals is quite extensive. However there are still some shortcomings, which could be improved. One of them is the primary usage of quarterly frequency data to predict commodity prices, which makes forecasts less detailed and ignores some important price reactions to changes in fundamental factors. Another drawback is that many authors focus on predicting real prices, commodity indices or futures prices, while changes in nominal spot prices are often of higher interest to central banks, researchers and think tanks. Moreover, the models used in the literature do not always incorporate a sufficient number of explanatory variables, focusing rather on the impact of a limited number of factors on commodity price developments. This study adds to the literature by 1) focusing on monthly rather than quarterly data to predict commodity prices in the short-run, 2) predicting the nominal spot prices of the commodities of interest, 3) using up-to-date global benchmarks for commodity prices, and 4) accounting for the comprehensive set of supply and demand factors that determine price movements. Moreover, the models applied in this study do not completely repeat the specifications used previously in the literature, but represent a hands-on approach to predicting commodity prices, taking into account the perspectives of central banks.

3. DATA AND METHODOLOGY

3.1. General Methodology

We use monthly data for the periods from 2003, 2004 or 2008 (depending on the availability of the data required by model specifications) up to February 2023 to examine the determinants of global commodity prices and construct 6-month-ahead forecasts. For instance, the models that include global manufacturing PMI as a proxy for global demand have their estimation samples starting in 2008 due to the limited availability of this data. Overall, the chosen time span is sufficient to analyze the dynamics of commodity prices and allows for some adjustment of the models to previous episodes of price volatility, relatively similar to the current ones. Although macro-forecasting processes in central banks normally focus on quarterly data, monthly variables better capture price developments and market changes, while also allowing one to make a more precise prediction of commodity prices than with quarterly data. Moreover, the quarterly projections of prices can easily be derived from our monthly forecasts by averaging, and can be used further for macroeconomic forecasting (Figure 1).



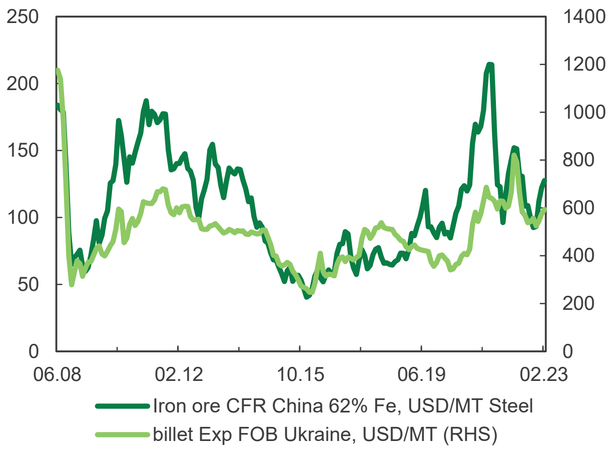


Figure 1. Nominal Global Prices of Energy Commodities and Metals

Source: World Bank, Thomson Reuters, Delphica.

In the models that are described below, the variables chosen have inter-links that are explained by economic theory. Moreover, they demonstrate Granger-causality relative to one another, which justifies the use of VAR models. All variables are tested for unit roots, and nonstationary series are transformed into stationary ones by simple differencing or log-differencing (see Table 5). In the case of cointegration relationships between the variables, error-correction models are used.

In order to account for the main demand and supply factors that affect price changes and for the impact of past observations on commodity prices, we employ standard vector autoregressive (VAR) and vector error-correction models (VECM) to regress the world prices of crude oil, natural gas, steel and iron ore, and to make projections. Although the choice of variables in the models and of forecasting approaches is based on the literature and economic theory, they do not completely repeat the models used previously in other studies. The methodology of this study represents a hands-on approach to predicting commodity prices, taking into account the perspectives of central banks, and shows the impact of various factor combinations on price changes. We compare the forecast accuracy of the models (measured by root mean square errors) to that of a random walk forecast and perform an out-of-sample forecasting exercise.

The VAR model

The general representation of a standard reduced-form VAR model with p lags can be written as follows:

$$y_t = c + B_1 y_{t-1} + B_2 y_{t-2} + \ldots + B_p y_{t-p} + u_t, \quad u_t \sim (0, \Omega), \quad (1)$$

where y_t and c are $K \times 1$ vectors of K monthly variables and constants, respectively, and $B_i, i = 1, \ldots, p$ are $K \times K$ matrices of coefficients. The residuals u_t are assumed to be i.i.d. N(0, Ω), where Ω is the variance-covariance matrix of innovations.[1]

Equation (1) indicates that any series in the model depends on the past values of all the K series through their lags. For example, if the number of variables K in the system equals 2, and the number of lags p = 2, the VAR(2) process can be rewritten as follows:

$$\begin{bmatrix} y_{t,1} \\ y_{t,2} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} B^1_{1,1} & B^1_{1,2} \\ B^1_{2,1} & B^1_{2,2} \end{bmatrix} \begin{bmatrix} y_{t-1,1} \\ y_{t-1,2} \end{bmatrix} + \begin{bmatrix} B^2_{1,1} & B^2_{1,2} \\ B^2_{2,1} & B^2_{2,2} \end{bmatrix} \begin{bmatrix} y_{t-2,1} \\ y_{t-2,2} \end{bmatrix} + \begin{bmatrix} u_{t,1} \\ u_{t,2} \end{bmatrix}, \quad (2)$$

where the subscripts indicate equation and variable numbers, and the superscripts refer to the lag number. Thus, the VAR(p) is an example of a seemingly unrelated regression (SUR) model with lagged variables and deterministic terms as common regressors (Zivot and Wang, 2003). In addition, some other deterministic terms (seasonal dummies, a linear trend, and a set of exogenous variables) can be included in the VAR system.

The VECM Model

When variables in such a system are cointegrated, the vector error correction model (VECM) should be used rather than a standard VAR model. Variables are said to be cointegrated if each of them are non-stationary with a unit root (I(1)), while there is some linear combination $a'y_t$ of these series which are I(0), i.e. stationary. Here a is a non-zero $K \times 1$ vector.

Let us consider again a VAR(p) process as in equation (1). In lag operator notation, this equation can be written as follows:

$$B(L) Y_t = c + u_t, \quad (3)$$

where $B(L) = I_K - B_1 L - \ldots - B_p L^p$. This VAR system is stable if the roots of the polynomial

$$\det\left(I_K - B_1 z - B_2 z^2 - \ldots - B_p z^p\right) = 0$$

lie outside the complex unit circle, or have a modulus greater than 1. If at least one series among $y_{K,t}$ is I(1), the VAR(p) is unstable, since $\Pi = -(I_K - B_1 - B_2 - \ldots - B_p)$ is singular, $\det(\Pi) = 0$, and the roots lie on the unit circle.

In general, cointegration means that there is some long-term relationship between the individual elements of y_t, which is represented by the linear combination $a'y_t$. A VECM is a special type of VAR model, which introduces error-correction terms into the system. A VECM focuses on differences to account for short-run relationships between variables (as represented in a VAR), while its error-correction terms and cointegrating equations account for short-run adjustments and long-run cointegrating relationships. For example, the VECM(2) system for two variables $y_{t,1}$ and $y_{t,2}$ can be specified as:

$$\begin{aligned} \Delta y_{t,1} &= c + B^1_{1,1}\Delta y_{t-1,1} + B^2_{1,1}\Delta y_{t-2,1} + B^1_{1,2}\Delta y_{t-1,2} + \\ &+ B^2_{1,2}\Delta y_{t-2,2} - \lambda_1\left(y_{t-1,1} - \alpha_0 - \alpha_1 y_{t-1,2}\right) + u_{t,1} \end{aligned};$$

$$\begin{aligned} \Delta y_{t,2} &= c + B^1_{2,1}\Delta y_{t-1,1} + B^2_{2,1}\Delta y_{t-2,1} + B^1_{2,2}\Delta y_{t-1,2} + \\ &+ B^2_{2,2}\Delta y_{t-2,2} - \lambda_2\left(y_{t-1,1} - \alpha_0 - \alpha_1 y_{t-1,2}\right) + u_{t,2} \end{aligned}, \quad (4)$$

where $y_{t,1} = \alpha_0 + \alpha_1 y_{t,2}$ is the long-run cointegrating relationship between the variables $y_{t,1}$ and $y_{t,2}$, and λ_1 and λ_2 are the error-correction terms. The error-correction terms measure the response of the variables $y_{t,1}$ and $y_{t,2}$ to deviations from long-run equilibrium. As in (2), the subscripts in the system indicate equation and variable numbers, and the superscripts refer to the lag number. If a VEC model has more than two variables, it means that there can be more than one cointegrating relationship. The number of these relationships can be determined using cointegration tests. Also one should note, that the VECM(2) in example (4) is derived from the VAR(3) model, since the VECM focuses on differences, and for a VAR(p) model the corresponding VECM would be with (p – 1) lags.[2]

3.2. Crude Oil Price Forecasting

To model global crude oil prices and to produce short-term forecasts, we use general VAR methodology introduced by Baumeister and Kilian (2013, 2014), with some adjustments. We use four standard VAR models with slightly different specifications to forecast monthly crude oil prices in the short run. The specifications of the models, which are based on the literature and economic theory, were adjusted in order to incorporate additional factors of interest that are relevant to the current period.

[1] The number of lags (p) is obtained on the basis of some theoretical models, by using a rule of thumb, or by statistical selection criteria, such as the Akaike information criterion (AIC), the Schwarz-Bayesian criterion (BIC) and Hannan-Quinn (HQ) criterion.

[2] For a more detailed description of VAR and VECM estimations, please refer to Hamilton (1994), Zivot and Wang (2003), Ouliaris et al. (2018).

They were also adjusted for analytical purposes to test the impact of various factor combinations on prices. For the first model, the vector of endogenous variables consists of 1) the real price of Brent oil (the nominal price deflated by the U.S. CPI) as an international benchmark, 2) the percentage change in crude oil production, 3) the percentage change in OECD petroleum inventories (as a proxy for changes in global inventories), and 4) the index of global real economic activity also known as the Kilian (2009) shipping index (as a proxy for global demand). One of the important factors affecting real commodity prices is the shift in demand for commodities which, in turn, is caused by unexpected fluctuations in real global economic activity. Kilian's index is a business-cycle indicator, which is derived from global bulk dry cargo shipping rates, and is expressed as a percentage deviation from the trend. Kilian's index proves to be a good monthly indicator of the state of, and changes in, the global economy. We also find that it is a more convenient indicator than the index of monthly GDP of OECD+Major 6 NMEs calculated by the OECD, as the latter is available with significant lags, and does not capture major global economic fluctuations. For the discussion of the advantages of Kilian's index, see Kilian (2009), and Kilian and Zhou (2018). Since our aim is to forecast *nominal* rather than *real* oil prices, we follow Beckers and Beidas-Strom (2015) and introduce a fifth variable, the U.S. CPI (index 1982 – 1984 = 100, seasonally adjusted, from the FRED database of the St. Louis Fed), into the vector of endogenous variables in our VAR models. That makes it possible to produce forecasts for both the real price of oil and the consumer price index, while also deriving a forecast of the nominal oil price. The vector of exogenous variables includes constants and eleven seasonal dummies, as in a paper by Beckers and Beidas-Strom (2015). The standard methodologies used in the aforementioned papers suggest including 12 lags as a rule of thumb for the models based on monthly data, or four lags for those based on quarterly data, respectively. Notwithstanding that, our standard VAR model has six lags, which is explained by the Akaike and Schwarz-Bayesian information criteria, and the model's estimation sample starts in August 2003. The real price variable and the CPI are log-differenced to make them stationary, Kilian's index is taken in the first difference, while two supply-side determinants expressed as percentage changes are already stationary.

The second model's specification is slightly different for analytical purposes and it shows the impact of a different combination of explanatory factors on price changes. Here we use the J. P. Morgan Global Manufacturing PMI as our global demand proxy, instead of Kilian's index to test if a model with a different demand-side variable would prove more accurate in terms of forecasting. We also change the representation of the oil production variable by expressing it as the log-difference of production levels rather than calculating the percentage change in production. Due to the limited availability of the PMI time series, the sample for this model is shorter and starts in April 2008. We also employ two more models with the same specifications and sample length as the second one, but with a different U.S. inflation variable, which is a non-seasonally adjusted index, 2010 = 100 (as a result, the real price of crude oil differs too). The model number three has six lags, like the first two models, which is based on the AIC and BIC. The fourth model has three lags in order to better capture the most recent movements in oil prices. Moreover, lag exclusion tests also show that a higher number of lags might be unnecessary for this model.

3.3. Natural Gas Price Forecasting

In order to model and forecast TTF gas prices, we apply VAR and VECM approaches. The choice of the explanatory variables is based on the research literature and the fundamentals for the European gas market (Nick and Thoenes, 2014; Hamie, 2020; Berrisch and Ziel, 2022). The first model is a VAR(3) that uses the price of Dutch TTF gas, the Brent oil price, Kilian's index of global real economic activity, the global manufacturing PMI, changes in natural gas reserves in the Netherlands, gas stock changes, and natural gas supply variables in first difference, as well as a vector of constants in exogenous variables. The price of oil is included into the model since it is a close substitute for natural gas, and the prices of these two energy resources normally tend to move in similar directions. The Kilian and PMI indices are used as proxies for global demand factors, whereas stock changes and gas supply and stocks represent the supply-side determinants of gas prices. The second model basically has the same specifications, except that natural gas stocks are an exogenous variable.

With a view to conducting an in-depth analysis as to whether or not gas prices have similar determinants as crude oil prices, we also apply two of the oil forecasting models to predict natural gas prices. Thus, model number three uses real rather than nominal prices of TTF gas, the change in oil production, the change in petroleum inventories, Kilian's index, and the U.S. CPI. It also incorporates the vectors of constants and seasonal dummies into the set of exogenous variables, as in the oil forecasting models, but has 12 lags, as suggested in the literature and is confirmed by the information criteria. The fourth model, VAR(6), incorporates the real price of gas, an oil production variable (the first difference of natural logarithms rather than a percentage change), the PMI, the change in petroleum inventories, and the CPI, which are all used as endogenous variables. Similarly, there are seasonal dummies and constants used in the vectors of exogenous variables. The number of lags for this model's specification is explained by the Akaike criterion and lag exclusion tests. These models are meant to test whether or not the factors influencing oil prices can be reliably used to model and predict natural gas prices, without including gas-specific market determinants.

Finally, the fifth model includes gas and oil nominal prices, Kilian's index and the PMI, gas supply, and gas stock changes, as in the case of our second model. However, it has a different set of exogenous variables, namely seasonal dummies (as in the third and fourth models) and a weather conditions proxy measured in degree days, or DDs. As described earlier in the literature review section, temperature conditions play a crucial role in shaping natural gas consumption and, consequently, prices. When air temperatures are abnormally low, there is greater demand for heating and natural gas, which leads to higher gas prices. Likewise, very high temperatures increase the need to cool buildings, and natural gas is also widely used for these purposes. Heating and cooling degree days (HDDs and CDDs, respectively) are weather-based technical indicators that measure the energy requirements of buildings in terms of heating and cooling. For example, if one compares energy needs in 1979 and

2022 in the EU, HDD values declined by 19% during this period indicating that heating needs in 2022 were roughly two-tenths lower on average than those in 1979. At the same time, CDD values in the EU were almost four times higher in 2022 compared to 1979, showing the increased need for air conditioning and higher energy consumption over decades (Eurostat, 2023).

Although we are interested in modeling gas prices for the Dutch TTF hub, which is a benchmark for the European market, degree days in the model are those related to weather conditions in Germany, rather than in the Netherlands. The intuition behind this is that Germany is the largest natural gas consumer in the EU. According to Eurostat's final energy consumption indicator, which measures the energy consumption of end-users (industry, transport, households, agriculture, and services), Germany accounted for almost 27% of the total consumption of natural gas in the European Union in 2021. For comparison, Italy takes second place, but its end-users consume only about 16% of the total natural gas quantities consumed in the EU. Moreover, according to annual data for 2021, Germany is the dominant leader in the final consumption of heat with a share of over 21% of the EU's total, and the top electricity consumer with a share that almost equals 20%. And these shares have been stable or even growing over the years. Therefore, we include degree days data for Germany in our model number five. Since DD is an exogenous variable, we need readymade forecast values of it for the whole forecast horizon, but getting these weather forecasts on a country level is too complicated. Thus, we focus on regional data and gather DDs for the most populous cities, such as Berlin and Munich, and important industrial towns, including Ludwigshafen am Rheine, Wittenberg, and Hamburg. According to the Federal Statistical Office of Germany, the latter three cities are leaders in natural gas consumption in German industry. We collect historical degree days from Eurostat, where these indicators are calculated as follows:

$$\textit{If } T_m \leq 15°\text{C} \quad \textit{Then } [\text{HDD} = \textstyle\sum_i (18°\text{C} - T_m^i)] \quad \textit{Else } [\text{HDD} = 0]$$

$$\textit{If } T_m \geq 24°\text{C} \quad \textit{Then } [\text{CDD} = \textstyle\sum_i (T_m^i - 21°\text{C})] \quad \textit{Else } [\text{CDD} = 0], \quad (5)$$

where T_m^i is the mean air temperature of day i. The base temperatures for HDDs and CDDs are set to 15°C and 24°C, respectively, in accordance with the general climatological approach. These calculations are made on a daily basis and then added up to monthly figures, which we use. We then calculate total monthly degree days (DDs) as the sum of HDDs and CDDs of all chosen cities as a proxy for German energy needs. As Eurostat updates degree days once per year for the full year that has passed, we gather daily air temperature data for the aforementioned cities from the website https://www.weather25.com/ and use the formulae to produce the missing actual values of monthly DDs. We obtain the projected HDD and CDD values that are needed for the model from https://cds.climate.copernicus.eu/. These values are modelled on the basis of historical averages over a 30-year period taking into account climatological inputs. This model is a VECM since there are cointegration relationships between the variables in this specification, and the inclusion of two lags is explained by tests and information criteria.

The estimation samples of all of the models that predict natural gas prices start in either April or May 2008 and run through February 2023, except for the third model, which has its estimation sample starting in February 2004, as data for that period was more readily available.

3.4. Iron Ore Price Forecasting

We apply standard VAR methodology to model iron ore prices. Again, we use four VAR models with slightly different specifications to better capture the impact that various combinations of factors might have on the price. After that, we construct a baseline forecast as an average of four approaches, which helps to combine the benefits of individual forecasting models and performs equally well during different periods. The models applied to use the monthly spot prices of iron ore fines, CFR China 62% Fe, from the World Bank database, as it is the most commonly used global benchmark. We deflate the nominal price by the U.S. CPI (index, 1982 – 1984 = 100) to obtain the real price. Similarly, the CPI is also included in the vector of endogenous variables, and the forecasts of the two variables are then used to obtain nominal price predictions.

The iron ore market is significantly influenced by the steel market since iron ore is primarily used in steelmaking. Moreover, China's large steel market makes it a big player, so it has a great impact on iron ore prices, especially from the demand side. China is the dominant consumer of metals in general and iron ore in particular, as it is the world's largest producer of steel. Moreover, the Chinese construction sector and infrastructure projects require substantial amounts of materials, such as steel. With that in mind, we use China's crude steel production (monthly growth rate expressed as a percentage) as a proxy for global demand for iron ore in all four models. Models one and two also use changes in the Baltic Dry Index, i.e. the Baltic Exchange's main sea freight index. We expect an increase in freight rates to push up commodity prices as well. The Baltic Dry Index is available on a daily basis, but we transform it into a monthly series and then calculate monthly percentage changes. From the supply side, we add a change in Brazil's exports of iron ore as this country, together with Australia, are traditionally major exporters of iron ore. As Australia's detailed export data was not made available when we were collecting data, we focused on Brazil's exports as the second largest iron ore exporter to China. The weather conditions in Brazil, other disruptions to its economy and the mining sector, as well as the policies of the mining giant Vale are known to influence Brazil's iron ore production and export volumes and, through these, world prices. Two out of four VAR models have these variables and a vector of constants to model iron ore prices, however, they have different numbers of lags – two and one, respectively. This is attributable to different lag suggestions by statistical information criteria, and to ambiguous test results for the optimal number of lags that should be used.

The third model incorporates China's Manufacturing Purchasing Managers' Index (PMI), which is published by the National Bureau of Statistics of China on a monthly basis. This variable is a proxy for the state of the Chinese economy in general and the health of its industrial sector in particular. China is the largest global importer and one of the biggest producers of iron ore, as a result of which its economic development is expected to impact world iron ore prices. The other endogenous variables are the same as in the first two models, except for the exclusion of Brazil's export variable, and the optimal testing-based number of lags for such a specification equals five. In contrast, the fourth model is more of a combination of the aforementioned specifications. This two-lag VAR model consists of the real price of iron ore, changes in steel production in China, and changes in the Baltic Dry Index, China's Manufacturing PMI, Brazil's iron ore exports, and in the U.S. CPI. Exogenous variables include constants (as in other models) and the recession dummy for the U.S., i.e. an NBER-based recession indicator, is available from the FRED database of the FRB of St. Louis. The recession dummy takes the value of 1 from January 2008 through June 2009 and from March 2020 through April 2020, which represents the recessionary periods in the U.S. Again, where necessary, the variables are transformed to log-difference or first difference forms or are winsorized to smooth out the outliers (like the change in Brazil's iron ore exports). Depending on data availability, and after some adjustments to the series, the estimation samples start either in April 2004, March 2005, or June 2005 and run through February 2023, as in all studied commodity groups.

3.5. Steel Price Forecasting

In order to model and predict global steel prices based on market fundamentals, we apply the VECM methodology. Error-correction is needed since there are cointegration relationships between the variables. As mentioned above, the global steel and iron ore markets are very interrelated, so the explanatory variables for steel prices are very similar to those used in iron ore forecasting.

As Ukraine used to be among the global top ten largest steel exporters (before russia's full-scale invasion of Ukraine and the port blockade), we are interested in forecasting prices for Ukrainian steel. As a benchmark, we use the monthly averages of daily steel billet prices, FOB Ukraine. For a period after 24 February 2022 (the start of the full-scale war), we use proxy prices for Ukrainian steel calculated based on either Turkey C&F steel billet prices (up until October 2022) or the Black Sea billet FOB UA prices. After calculating the monthly average steel price, we then take the log of it as we do for other price variables.

Explanatory variables in the first VECM include the price of iron ore fines (CFR China 62% Fe) and the price of coal (Australian thermal coal, FOB Newcastle, 6,000 kcal/kg futures price from the World Bank database) as inputs used to produce steel. Moreover, we use the Brent oil price (expressed in logs) as a determinant of global steel prices. There is a growing bulk of literature suggesting that there are cross-market price links between various commodity groups that can be used to predict prices (see, e.g., Campiche et al., 2007, Nazlioglu and Soytas, 2012, Ding and Zhang, 2020). More specifically, there is empirical evidence of crude oil prices having long-term cointegration relationships with other commodity prices and having an impact on their prices. Moreover, it is believed that there is a unidirectional relationship between crude oil and steel as suggested, for instance, by Chou et al. (2012) and Asmoro (2017).

Given the interrelations between the iron ore and steel markets, we also include changes in the Baltic Dry Index and China's Manufacturing PMI into this model, just as we did in the VARs for iron ore prediction. This VECM has one lag determined by the lag length selection tests and also incorporates a recession dummy in its exogenous variables. The second model is simplified for analytical purposes – it only includes the prices of steel and iron ore and the Baltic Dry Index. The optimal number of lags equals one, and there are no exogenous variables included. The third VECM is the price-only model as it includes the prices of steel, iron ore, coal and crude oil, and no other explanatory variables. This specification requires two lags, as chosen by the AIC and BIC statistical criteria and lag exclusion tests. Finally, the fourth model is a two-lag VECM incorporating steel, iron ore and coal prices, changes in the Baltic Dry Index and the manufacturing PMI for China, as well as a recession binary variable. After making necessary adjustments to the data and taking into account the time span of the data, the estimation samples start either in August 2008 or September 2008 and last until February 2023.

4. RESULTS

In this section, we provide the results of the forecasting performance of the models used to predict commodity prices six months ahead. We run the models to forecast oil and natural gas prices from the beginning of the models' estimation samples until the end of 2015 and then do out-of-sample six-month-ahead forecasting simulations starting from January 2016 through February 2023. For the models that forecast iron ore and steel prices, the out-of-sample exercise starts in January 2018 to better adjust the models to changes in the pricing regime of iron ore prices. Next, we calculate the root mean square errors (RMSEs) of all individual models based on forecast simulations and divide them by the RMSEs of the respective random walk (RW) models. Figures 2–5 (Supplementary Materials) depict the results of the out-of-sample forecast simulations six months ahead for the nominal prices of the four commodities of interest. Tables 1–4 provide a summary of the relative RMSEs of the models for one- to six-month-ahead horizons. Values below 1, highlighted in green, mean that the RMSEs of the given models are lower than the random walk RMSEs. This means that the forecasting power of a given

Table 1. The RMSEs of Individual Models Relative to RW RMSEs – Crude Oil

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
VAR_1	6	**0.939**	0.776	**0.701**	0.737	0.823	0.803
VAR_2	6	1.067	0.826	0.827	0.816	0.896	0.875
VAR_3*	6	1.052	0.818	0.816	0.819	0.912	0.907
VAR_4*	3	1.005	**0.772**	0.747	**0.735**	**0.775**	**0.775**

* VAR models 3 and 4 use a different U.S. CPI index (non-seasonally adjusted index, 2010=100), which also causes variations in real prices. Therefore, the RMSEs of these models are compared to the RMSEs of a RW model, which also uses real prices calculated on the basis of a non-seasonally adjusted CPI index, whereas models 1 and 2 are compared to a RW model based on comparable real prices (where the CPI index, s.a., 1982-1984=100 is a deflator).

model is higher than that of an RW benchmark. Values above 1, highlighted in red, indicate that the given models fail to outperform the respective RW models over the given forecast horizons. The lowest relative RMSEs over each of the six horizons are presented in bold, indicating the best-performing models.

The models that predict oil prices demonstrate good forecasting performance over almost all forecasting horizons (Table 1). It is only in the one-month-ahead forecast that the random walk outperforms three out of four of the selected models. Note that models VAR_3 and VAR_4 use real oil prices calculated based on a different CPI index compared to models 1 and 2 (see the Data and Methodology section). Therefore, in order to make fair conclusions, we divided their RMSEs by the RMSEs of a different random walk model, which used comparable real prices.

The first model, which uses oil production, petroleum inventories, Kilian's index, and the CPI index as endogenous oil price determinants, improves benchmark RW forecasts over all forecast horizons, and is the best-performing model to predict Brent prices one and three months ahead. VAR model number 4, despite having the smallest number of lags – i.e. 3 lags, has the smallest number of forecast errors over two-, four-, five- and six-month-ahead horizons among all the other models. This may suggest that, in the case of the short-term forecasting of monthly oil prices, the information contained in only three lags of fundamentals might be enough to predict the future movement of oil prices. This finding is new and adds to the modern techniques of oil price forecasting, which normally rely on up to 12 lags of information. Models 2, 3 and 4 also have an estimation sample that is almost five years smaller than the sample of the first model due to the limited availability of the manufacturing PMI index that they use instead of Kilian's IGREA.

All of the models used to predict TTF gas prices in the short run outperform the RW benchmark model during all forecast periods, as shown in Table 2. Figure 3 (Supplemenary Materials) presents the results of the out-of-sample forecasting exercise.

The most consistent results are produced by the two-lag VECM model, which uses the weather conditions variable – degree days – as an exogenous one. This once again proves the importance of air temperature conditions for gas price forecasting. It is also worth noting that models 3 and 4, which use crude oil-related variables (such as the change in oil production and petroleum inventories) and general demand-side variables (such as Kilian's index and the PMI) and do not use gas-specific market fundamentals, also show good results and even outperform other models over some forecast horizons. These results add to the empirical evidence of the crude oil and natural gas markets being highly interrelated, making room for further investigation of cross-market energy price predictions.

Tables 3 and 4 present the relative RMSEs of the models that predict the prices of iron ore and steel, respectively, while Figure 4 and 5 (Supplemenary Materials) show the forecasting simulations of nominal prices. For these two commodities, we also test the models' performance compared to a random walk process. However, we include simple AR(1) models for comparison, since they are often used as a benchmark to test the forecasts of metal prices, or are extended to ARIMA models as standalone forecasting techniques (Pincheira and Hardy, 2019).

All of the selected models that forecast global iron ore prices in general produce better forecasts than both RW and AR(1) models. VAR_1 and VAR_2, which have the same specifications but a different lag number, are the best-performing approaches. This can be explained by these models having the most comprehensive set of supply and demand variables, including China's steel production, the Baltic Dry Index of freight rates and Brazil's exports of iron ore. As expected, the model with two lags produces better results over longer horizons. Model 4, which incorporates the recession dummy, also proves reliable. At the same time, model 3, which excludes Brazil's ore exports, performs worse than other models, despite having the highest lag number. However, the chart of out-of-sample simulations for model 3 indicates that it might have a better predictive power for more unstable periods, similar to those that we are currently observing on the markets, while other models are relatively better for stable periods (Figure 4 in Supplemenary Materials).

VECM number 1, which is used to predict steel prices, outperforms other approaches over half of the forecast periods. The model includes the prices of iron ore, coal and oil, the freight rates index, China's manufacturing sector proxy, as well as the recession dummy. It is also noteworthy that this VAR model has the smallest possible number of lags and shows better results than the two-lag models. AR(1) only slightly outperforms model 1 over one- and two-month horizons, but lags behind over longer forecast periods. Interestingly, the two-lag price-only model (VECM_p_3) which only uses the prices of iron ore, coal, and crude oil to predict steel prices, has the smallest number of RMSEs six months ahead (Table 4). This finding can be further developed and tested in future research into the longer-term forecasting of steel prices.

The results of individual models indicate that our choice of forecasting techniques and explanatory variables is reliable and the models can be used to predict commodity prices, at least in the short run. Moreover, given the generally high performance of these models and their varying forecast

Table 2. The RMSEs of Individual Models Relative to RW RMSEs – Natural Gas

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
VAR_1	3	0.840	0.689	0.840	0.917	0.855	0.916
VAR_2	3	0.845	**0.681**	0.820	0.927	0.836	0.897
VAR_3	12	**0.791**	0.684	0.863	0.861	0.855	0.770
VAR_4	6	0.823	0.691	0.850	0.834	0.826	**0.766**
VECM_DD_5	2	0.829	0.695	**0.802**	**0.826**	**0.826**	0.850

Table 3. The RMSEs of Individual Models Relative to RW RMSEs – Iron Ore

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
AR(1)	1	0.7893	**0.7449**	0.6906	0.6296	0.6801	0.6932
VAR_1	2	0.8190	0.7719	0.6916	**0.6260**	**0.6747**	**0.6926**
VAR_2	1	**0.7856**	0.7451	**0.6895**	0.6273	0.6794	0.6931
VAR_3	5	0.9374	0.8373	0.7445	0.6957	0.7141	0.7178
VAR_4	2	0.8414	0.7772	0.7084	0.6378	0.6840	0.6956

Table 4. The RMSEs of Individual Models Relative to RW RMSEs – Steel

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
AR(1)	1	**0.8324**	**0.6943**	0.6855	0.7357	0.7443	0.7126
VECM_1	1	0.8918	0.7014	**0.6802**	**0.7275**	**0.7390**	0.7073
VECM_2	1	0.9001	0.7521	0.7620	0.8047	0.7801	0.7122
VECM_p_3	2	0.9592	0.9286	0.8478	0.8052	0.7456	**0.7031**
VECM_4	2	1.0598	0.8798	0.8586	0.9515	0.8573	0.7536

precision over different periods, it makes sense to apply a combination approach and to merge the models' benefits to generate a combined baseline forecast for each of the commodities.

5. CONCLUSIONS

This study offers a relatively simple hands-on approach to forecasting the global prices of crude oil, natural gas, iron ore, and steel. In line with the modern literature, we apply VAR and VECM approaches based on demand and supply factors to forecast commodity prices over the short term period. This paper adds to the literature in a few ways.

First, unlike most other similar papers, the forecasting models in this paper focus on predicting monthly rather than quarterly prices (while being developed from the central bank's perspective). The rationale behind this is that monthly time series are more detailed and contain more information about price movements. Thus, generating monthly rather than quarterly forecasts increases forecast precision. Moreover, generated monthly price forecasts can then be used to construct more reliable quarterly projections than those derived from smoothed quarterly data. This, in turn, could improve the forecast performance of other central banks' macroeconomic quarterly projection models that use commodity price forecasts as inputs or assumptions.

Second, in our models we forecast real prices as well as inflation indices in order to construct forecasts of nominal commodity prices, which are of greater interest to us. Furthermore, this study focuses on spot prices, and does not include futures-based predictions, which are still popular among many central banks and forecasters, despite their being rather inaccurate under the current conditions. There is no need to include such models in the set of our forecasting techniques as reliance on futures prices does not necessarily provide robust outcomes for forecasting spot prices.

Third, our findings suggest that, among the individual models in each of the four commodity groups, the models with the most balanced and comprehensively chosen fundamental explanatory variables, which cover supply and demand fundamentals equally, prove the most reliable in terms of forecasting. These fundamentals, which are important for commodity price prediction, include, inter alia, stock changes, changes in commodity production volumes, export volumes by the largest players, changes in the manufacturing sector of the largest consumers, the state of global real economic activity, freight rates, and a recession indicator. Seasonal factors play an important role in shaping commodity prices as well. Moreover, Kilian's index of global real economic activity is found to be a useful proxy for global demand and a reliable input in forecasting both energy and metal prices. In the case of iron ore and steel prices, developments in the Chinese economy prove to be essential inputs.

Furthermore, we demonstrate that when predicting energy and metal prices in the short run, the models with smaller lag orders tend to outperform those with a higher number of lags, which is a new finding in the literature, to our knowledge. While literature usually suggest using up to 12 lags (as a rule of thumb in vector models to predict monthly commodity prices), we show that the most important information in terms of short-term price prediction can be found in the most recent historical data, and there is no need to overload the models. The conducted lag length selection tests and information criteria demonstrate that our models require no more than six lags. Moreover, the models with a smaller number of lags in general show higher forecast accuracy, as can be seen from the RMSE tables. This finding can be used and further developed by both researchers and forecasters with the purpose of finding the best-fitting forecasting techniques.

Finally, we conclude that selected individual models, while showing standalone high performance, have varying forecast accuracy over different periods. Our findings show that no individual model outperformed the others consistently throughout the forecast horizon. Thus, it might make sense to apply a combination approach to merge the models' benefits and generate a combined baseline forecast for each of the commodities.

The methodology used in this study is a hands-on approach to forecasting commodity prices in the short run. That notwithstanding, there is room for further research and improvement of the models, for example by using Bayesian techniques or applying non-equally weighted model combinations based on forecast errors. Moreover, it can be observed from the results that the forecasting accuracy of some of the models deteriorated during the most recent periods, in particular for natural gas (see, for example, the out-of-sample simulations). Therefore, under the current conditions of high uncertainty and abnormal movements in commodity markets, further development of crisis-time forecasting techniques might be needed.

REFERENCES

Alquist, R., Kilian, L., Vigfusson, R. J. (2011). Forecasting the price of oil. Working Paper, 2011–15. Ottawa: Bank of Canada. Retrieved from https://www.bankofcanada.ca/wp-content/uploads/2011/06/wp2011-15.pdf

Ambya, A., Gunarto, T., Hendrawaty, E., Kesumah, F. S. D., Wisnu, F. K. (2020). Future natural gas price forecasting model and its policy implication. International Journal of Energy Economics and Policy, 10(5), 64–70. https://doi.org/10.32479/ijeep.9676

Asmoro, T. H. (2017). A VAR model to investigate the volatility of line-pipe steel prices using oil price as a referred Currency. PM World Journal, 6(11), 1–12. Retrieved from https://pmworldlibrary.net/wp--content/uploads/2017/11/pmwj64-Nov2017-Asmoro-VAR-Model-Line-Pipe-Oil-Price-featured-paper.pdf

Baffes, J., Nagle, P. (eds.) (2022). Commodity Markets: Evolution Challenges and Policies. Washington, DC: World Bank. http://doi.org/10.1596/978-1-4648-1911-7

Baumeister, C., Kilian, L. (2013). Forecasting the real price of oil in a changing world: a forecast combination approach. Working Paper, 2013–28. Ottawa: Bank of Canada. Retrieved from https://www.bankofcanada.ca/wp-content/uploads/2013/08/wp2013-28.pdf

Baumeister, C., Kilian, L. (2014). What central bankers need to know about forecasting oil prices. International Economic Review, 55(3), 869-889. https://doi.org/10.1111/iere.12074

Beckers, B., Beidas-Strom, S. (2015). Forecasting the nominal Brent oil price with VARs – one model fits all? IMF Working Paper, 15/251. Washington: International Monetary Fund. Retrieved from https://doi.org/10.5089/9781513524276.001

Berrisch, J., Ziel, F. (2022). Distributional modeling and forecasting of natural gas prices. Journal of Forecasting, 41(6), 1065–1086. https://doi.org/10.1002/for.2853

Bowman, C., Husain, A. (2004). Forecasting commodity prices: futures versus Judgment. IMF Working Paper, 04/41. Washington: International Monetary Fund. Retrieved from https://doi.org/10.5089/9781451846133.001

Campiche, J. L., Bryant, H. L., Richardson, J. W., Outlaw, J. L. (2007). Examining the evolving correspondence between petroleum prices and agricultural commodity prices. American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007. https://doi.org/10.22004/ag.econ.9881

Chernenko, S. V., Schwarz, K. B., Wright, J. H. (2004). The information content of forward and futures prices. International Finance Discussion Paper, 808. Board of Governors of the Federal Reserve System. Retrieved from https://doi.org/10.17016/ifdp.2004.808

Chou, M.-T., Yang, Y.-L., Chang, S.-C. (2012). A study of the dynamic relationship between crude oil price and the steel price index. Review of Economics and Finance, 2, 30–42. Retrieved from http://www.bapress.ca/Journal-7/A%20Study%20of%20the%20Dynamic%20Relationship%20between%20Crude%20Oil%20Price%20and%20the%20Steel%20Price%20Index.pdf

Ding, S., Zhang, Y. (2020). Cross market predictions for commodity prices. Economic Modelling, 91, 455–462. https://doi.org/10.1016/j.econmod.2020.06.019

ECB (2015). Forecasting the price of oil. Economic Bulletin, 4(3), 87–98. Frankfurt: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/other/art03_eb201504.en.pdf

Eurostat (2023). Heating and cooling degree days at EU level. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Heating_and_cooling_degree_days_-_statistics#Heating_and_cooling_degree_days_at_EU_level

Gao, S., Hou, C., Nguyen, B. H. (2021). Forecasting natural gas prices using highly flexible time-varying parameter models. Economic Modelling, 105, 105652. https://doi.org/10.1016/j.econmod.2021.105652

Hamie, H. (2020). Econometric modeling of natural gas prices : different modeling approaches and case studies related to gas markets (dissertation). Technische Universität Wien. https://doi.org/10.34726/hss.2020.41262

Hamilton, J. D. (1994). Time Series Analysis. Princeton, New Jersey: Princeton University Press.

Hosseinipoor, S., Hajirezaie, S., Hadjiyousefzadeh, N. (2016). Application of ARIMA and GARCH models in forecasting the natural gas prices. Retrieved from https://ou.edu/content/dam/cas/economics/Student%20Journal%20of%20Economics%20publications/Saied%20Hosseinipoor_AppJOE.pdf

Hulshof, D., van der Maat, J.-P., Mulder, M. (2016). Market fundamentals, competition and natural gas prices. Energy Policy, 94, 480–491. https://doi.org/10.1016/j.enpol.2015.12.016

Jin, J., Kim, J. (2015). Forecasting natural gas prices using wavelets, time series, and artificial neural networks. PLoS ONE, 10(11). https://doi.org/10.1371/journal.pone.0142064

Kilian L. (2009). Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. American Economic Review, 99, 1053–1069. https://doi.org/10.1257/aer.99.3.1053

Kilian, L., Murphy, D. P. (2014). The Role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics, 29(3), 454–478. https://doi.org/10.1002/jae.2322

Kilian, L., Zhou, X. (2018). Modeling fluctuations in the global demand for commodities. Journal of International Money and Finance, 88, 54–78. https://doi.org/10.1016/j.jimonfin.2018.07.001

Kriechbaumer, T., Angus, A., Parsons, D., Rivas Casado, M. (2014). An improved wavelet-ARIMA approach for forecasting metal prices. Resources Policy, 39, 32–41. https://doi.org/10.1016/j.resourpol.2013.10.005

Li, D., Moghaddam, M. R., Monjezi, M., Armaghani, D. J., Mehrdanesh, A. (2020). Development of a group method of data handling technique to forecast iron ore price. Applied Sciences, 10(7), 2364. https://doi.org/10.3390/app10072364

Ma, Y., Zhen, W. (2020). Market fundamentals and iron ore spot prices. Economic Record, 96(315), 470–489. https://doi.org/10.1111/1475-4932.12564

Manescu, C., Van Robays., I. (2014). Forecasting the Brent oil price. Addressing time-variation in forecast performance. Working paper series, 1735. Frankfurt: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1735.pdf

Mei, F., Chen, P. (2018). VAR analysis of the factors influencing overcapacity in the iron and steel industry. Journal of Scientific and Industrial Research, 77, 566–569. https://nopr.niscpr.res.in/bitstream/123456789/45115/1/JSIR%2077%2810%29%20566-569.pdf

Nademi, A., Nademi, Y. (2018). Forecasting crude oil prices by a semiparametric Markov switching model: OPEC, WTI, and Brent cases. Energy Economics, 74, 757–766. https://doi.org/10.1016/j.eneco.2018.06.020

Nazlioglu, S., Soytas, U. (2012). Oil price, agricultural commodity prices, and the dollar: a panel cointegration and causality analysis. Energy Economics, 34, 1098–1104. https://doi.org/10.1016/j.eneco.2011.09.008

Nick, S., Thoenes, S. (2014). What drives natural gas prices? – A structural VAR approach. Energy Economics, 45, 517-527. https://doi.org/10.1016/j.eneco.2014.08.010

Nixon, D., Smith, T. (2012). What can the oil futures curve tell us about the outlook for oil prices? Quarterly Bulletin, Q1. London: Bank of England. Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/quarterly-bulletin/2012/what-can-the-oil-futures-curve-tell-us-about-the-outlook-for-oil-prices.pdf

Obadi, S. M., Korcek, M. (2020). Driving fundamentals of natural gas price in Europe. International Journal of Energy Economics and Policy, 10(6), 318–324. https://doi.org/10.32479/ijeep.10192

Ouliaris, S., Pagan, A. R., Restrepo, J. (2018). Quantitative Macroeconomic Modeling with Structural Vector Autoregressions – An EViews Implementation. https://www.eviews.com/StructVAR/structvar.pdf

Pincheira, P., Hardy, N. (2019). Forecasting aluminum prices with commodity currencies. MPRA Paper, 97005. https://mpra.ub.uni-muenchen.de/97005/

Reichsfeld, D. A., Roache, S. K. (2011). Do commodity futures help forecast spot prices? IMF Working Paper, 11/254. Frankfurt: European Central Bank. Retrieved from https://doi.org/10.5089/9781463923891.001

Sharma, V., Cali, U., Sardana, B., Kuzlu, M., Banga, D., Pipattanasomporn, M. (2021). Data-driven short-term natural gas demand forecasting with machine learning techniques. Journal of Petroleum Science and Engineering, 206, 108979. https://doi.org/10.1016/j.petrol.2021.108979

Tularam, G. A., Saeed, T. (2016). Oil-price forecasting based on various univariate time-series models. American Journal of Operations Research, 6(3), 226–235. http://dx.doi.org/10.4236/ajor.2016.63023

Wårell, L. (2014). The effect of a change in pricing regime on iron ore prices. Resources Policy, 41, 16–22. https://doi.org/10.1016/j.resourpol.2014.02.002

Wårell, L. (2018). An analysis of iron ore prices during the latest commodity boom. Mineral Economics, 31, 203–216. https://doi.org/10.1007/s13563-018-0150-2

West, K. D., Wong, K.-F. (2014). A factor model for co-movements of commodity prices. Journal of International Money and Finance, 42, 289–309. http://doi.org/10.1016/j.jimonfin.2013.08.016

Zivot, E., Wang, J. (2003). Modeling Financial Time Series with S-PLUS. New York: Springer-Verlag.

APPENDIX A. TABLES

Table 5. Augmented Dickey-Fuller (ADF) Test Results

Variable	tau-stat (p-value)
Brent real price 1	-2.840 (0.185)
Brent real price 1, dlog	-10.048*** (0.000)
Brent real price 2	-2.811 (0.195)
Brent real price 2, dlog	-10.061*** (0.000)
Petroleum inventories, dlog	-3.374*** (0.001)
Change in petroleum inventories, % mom	-3.135*** (0.002)
Crude oil production, dlog	-13.359*** (0.000)
Change in crude oil production, % mom	-11.583*** (0.000)
Manufacturing PMI, Global	-4.490*** (0.002)
U.S. CPI_1 (s.a. index, 1982-1984=100)	0.723 (1.000)
U.S. CPI_1 (s.a. index, 1982-1984=100), dlog	-9.165*** (0.000)
U.S. CPI_2 (n.s.a. index, 2010=100)	-1.329 (1.000)
U.S. CPI_2 (n.s.a. index, 2010=100), dlog	-10.061*** (0.000)
Kilian index (IGREA)	-2.819*** (0.010)
TTF natural gas, nominal price	-4.528*** (0.000)
TTF natural gas, nominal price, dlog	-3.000*** (0.003)
TTF natural gas, real price	-6.047*** (0.000)
TTF natural gas, real price, dlog	-2.887*** (0.004)
Gas stock changes, Netherlands	-4.597*** (0.001)
Supply of natural gas, Netherlands	-2.357 (0.401)
Supply of natural gas, Netherlands, first difference	-8.593*** (0.000)
Iron ore real price	-3.032 (0.126)
Iron ore real price, dlog	-9.827*** (0.000)
Steel nominal price	-4.445*** (0.003)

Table 5 (continued). Augmented Dickey-Fuller (ADF) Test Results

Variable	tau-stat (p-value)
Coal, Australian, nominal price	-4.309*** (0.004)
Change in China's crude steel production, % mom	-5.229*** (0.000)
Manufacturing PMI, China	-7.096*** (0.000)
Baltic Dry Index	-7.065*** (0.000)
Change in Baltic Dry Index, % mom	-13.306*** (0.000)
Change in Brazil's iron ore exports, winsorized, % mom	-15.659*** (0.000)

*** denote significance at the 1% level.