SPECIAL FOCUS

Forecasting Industrial Commodity Prices: An Assessment
The Special Focus of this edition evaluates the performance of five well-known approaches to forecasting the prices of three key industrial commodities—aluminum, copper, and crude oil—over the period 2015Q1 to 2022Q1. High short-term volatility and significant longer-term movements in commodity prices—both features of commodity markets in recent years—present major challenges for policymakers in commodity-exporting EMDEs. Such challenges are easier to meet the more accurately price changes can be forecast. The evaluation reveals four main results. First, there is no “one-approach-beats-all” for commodity price forecasting, as the forecast accuracy of approaches varies significantly across commodities and time horizons. Second, macroeconometric models tend to be more accurate at longer horizons, partly because they can incorporate the effects of structural changes on prices. Third, it is critical to complement forecasts by incorporating judgment (information that cannot be accounted for by statistical approaches), especially when confronted by unusual or unprecedented events. Finally, these results underscore the value of employing a range of approaches in forecasting commodity prices.

Introduction

Almost two-thirds of emerging market and developing economies (EMDEs) heavily rely on commodities for export, fiscal revenue, and economic activity (figure 18.A-B). Among commodity-exporting EMDEs, resource sectors, on average, account for nearly 40 percent of exports of goods and non-factor services, 31 percent of goods exports, and 10 percent of value added. In some commodity-importing EMDEs, commodities account for a large share of imports and, in the presence of subsidies, fiscal spending. The substantial volatility of commodity prices exacerbates revenue management challenges in EMDEs (figure 18.C-D). Large and persistent price shocks caused by commodity price volatility weaken fiscal and external positions and lower economic growth in commodity-exporting EMDEs (IMF 2015, Richaud et al. 2019).

This Special Focus evaluates the performance of five well-known approaches to forecasting the prices of three key global commodity exports: aluminum, copper, and crude oil (figure 19.A). These three commodities account for about half of global commodity exports. The evaluation focuses on four model-based approaches (bivariate regression models; Bayesian vector autoregression models; a macroeconometric model; and a machine-learning model) and forecasts from Consensus Economics.

FIGURE 18 Commodity dependence and commodity price volatility

Many EMDEs depend heavily on commodity exports, with oil exporters typically more reliant than metals exporters. These export proceeds, in turn, provide a sizeable share of fiscal revenues. Price volatility is high, as is evident in the size and amplitude of cycles, with price booms more pronounced than price slumps, on average. The speed of commodity price rises in booms is also much faster than that for declines in slumps, especially for crude oil.

A. Share of EMDE exports

B. Resource revenues

C. Amplitude of booms and slumps

D. Average speed of booms and slumps

Sources: International Monetary Fund; UNComtrade (database); UNU-Wider (database); World Bank.

A. Figure shows the median share of exports accounted for by oil, copper, and aluminum for EMDEs that are commodity exporters. Oil includes 20 EMDEs, copper 6, and aluminum 4. Blue bars show medians and orange whiskers show interquartile ranges.

B. Unweighted average of resource revenues as a share of fiscal revenues for EMDE commodity exporters of oil (25 countries), copper (4 countries), and aluminum (3 countries). Countries relying on the export of multiple commodities are included in the averages for each commodity. Orange whiskers indicate the range between the minimum and maximum values.

C. Data from January 1970 to October 2021. Amplitude measures the average real price change (in percentage terms) from trough to peak for booms and from peak to trough for slumps.

D. Data from January 1970 to October 2021. Slump refers to the average monthly amplitude (in this case, amplitude divided by the duration).
The benefits of employing a menu of approaches to forecast commodity prices, rather than attempting to identify a single “best” approach, have been discussed by earlier studies (for example, Baumeister and Kilian 2015). Specifically, compared to a single approach, a variety of approaches can enhance the reliability of forecasts by taking into account a broader view of possible price outcomes. A previous edition of the Commodity Markets Outlook (World Bank 2023) featured a Special Focus that presented a brief literature review on various approaches for forecasting industrial commodity prices. It found that the five approaches examined here have performed relatively well for selected commodities. Building on that review, this Special Focus presents new empirical evidence by comprehensively evaluating these approaches.

**Forecasting Approaches**

**Bivariate regression models**

Bivariate regression models are useful for capturing the basic relationships between macroeconomic variables and commodity prices. They are used to forecast industrial commodity prices by simply regressing the change of a commodity price on a “past” value of an explanatory variable. Six explanatory variables are employed in regressions: the Commodities Research Bureau Raw Industrial Commodity Index; U.S. M1 growth; U.S. Treasury Bill 10-year interest rates; China’s manufacturing purchasing managers’ index (PMI); and global composite and manufacturing PMIs. The first three variables are often found to be useful predictors of crude oil prices (Alquist, Kilian, and Vigfusson 2013). The last three reflect the importance of the outsized share of China in global metal markets and the importance of global activity in driving industrial commodity prices. This approach yields six forecasted prices for the three commodities and for each forecast horizon, reflecting the number of independent variables. The final forecast is the average of all estimations with statistically significant coefficients produced by bivariate regression models.

**Bayesian vector autoregressive models**

The Bayesian vector autoregressive (BVAR) model, a multivariate VAR model, estimates the relationships among two or more variables. It differs from the standard multivariate VAR model in that the model parameters are treated as random variables—with prior probabilities—rather than fixed values. A BVAR model with sign restrictions is employed here to forecast commodity prices following techniques developed by Kilian and Murphy (2014). The dependent variables in the model include industrial commodity prices, log differences in metal production, and the global GDP growth rate. The Bayesian estimation approach allows for structural identifications such as elasticity and sign restrictions and prior beliefs about future economic events. The estimated impulse response functions that satisfy the sign restrictions are used to forecast industrial commodity prices.

**Macroeconometric model**

For this exercise, we utilized the Oxford Economics Global Economic Model (OEM). The OEM is a large-scale, cross-country, semi-structural projection model well suited to the analysis of alternative projections for the global economy (Oxford Economics 2019). It includes 81 countries, 6 regional blocs, and the Eurozone. Most have data available quarterly. Behavioral equations governing domestic economic activity, monetary and fiscal policy, global trade, and commodity prices are used. The model combines short-run momentum factors and long-term demand and supply fundamentals. In the short run, shocks to demand drive business cycles that can be influenced by fiscal and monetary policies. Over the long run, output is determined by supply-side factors such as investment, labor force participation, and productivity. The resulting dynamics of short-run fluctuations and long-run trends yield quarterly commodity price forecasts.

**Machine-learning model**

The machine-learning model combines algorithms and econometric methods to learn patterns and
estimate relationships in data. It then makes projections based on the patterns and estimated relations without imposing any theoretical prior. Following Zhang et al. (2015), a hybrid machine-learning model is employed here that comprises an empirical mode decomposition, and a generalized autoregressive conditional heteroskedasticity model. The hybrid model approach used in this exercise separates commodity price series into different nonlinear and time-varying components. Commodity price forecasts are constructed by adding forecasts of these components.

Forecasts from Consensus Economics

Consensus forecasts are published by Consensus Economics (CE), a service that surveys several forecasters for their projections of future output growth, commodity prices, current account balance, and other major macro variables. CE forecasts of commodity prices are drawn from “Energy & Metals Consensus Forecasts” reports, which are based on a monthly survey of up to 40 leading private-sector commodity forecasters covering 50 individual commodities. These forecasts are a simple compilation, which means they do not consider the systematic consistency of methodologies used by different forecasters.

Data

The data frequency and sample periods used for estimation vary across the approaches due to data availability. Bivariate models are based on monthly data for the period 2004Q1-2022Q1. For the BVAR models, quarterly averages of monthly data between 1995Q1-2022Q1 are used. The quarterly GDP growth rates are drawn from Haver Analytics, with forecasts based on the World Bank’s June 2022 Global Economic Prospects report (World Bank 2022). Production data for aluminum and copper are drawn from the World Bureau of Metal Statistics, and for oil from the International Energy Agency. The machine learning approach is estimated using monthly data from 1995Q1 to 2022Q1.

FIGURE 19 Directional accuracy of commodity price forecasts

While industrial commodity prices exhibit volatility that makes forecasting challenging over the short and longer terms, most evaluated forecasting approaches accurately predicted the direction of price changes for the three industrial commodities. BVAR and Consensus Forecasts had lower accuracy in predicting the direction of price changes for copper over longer horizons. BVAR exhibited less directionally accuracy than other approaches for oil. Bivariate regressions consistently demonstrated high directional accuracy across the three commodities.
The five forecasting approaches are evaluated in terms of three well-known statistical criteria:

- **Directional accuracy.** The directional accuracy simply assesses the likelihood of forecasts and actual prices moving in the same direction.

- **Forecast bias.** The forecast bias, defined as the mean of the difference between the actual and forecasted prices (mean forecast error), evaluates whether forecasts systematically over- or under-predict their realized values.

- **Forecast accuracy.** The accuracy of model forecasting is evaluated by the Diebold and Mariano (DM) test (Diebold and Mariano 1995). The DM test checks whether a particular model is more accurate than another by assessing the statistical significance of the difference in forecast errors between model pairs. The DM test is implemented for each model against all other approaches.

The forecast evaluation covers price forecasts for each of the three industrial commodities, ranging from one to eight quarters ahead, for the period 2015Q1-2022Q1. As bivariate regressions are statistically significant only for horizons up to one year ahead, other horizons are excluded. Historical OEM forecasts are only available semi-annually; hence, the forecast accuracy tests are adjusted for the fewer degrees of freedom.

**Results**

**Directional accuracy of forecasts**

The directions of price changes are often accurately predicted by most models for most commodities (figure 19.B-F). However, there are few exceptions, notably lower directional accuracy of the BVAR compared to other approaches, particularly for oil. For copper, the BVAR and Consensus Forecasts correctly predict the direction of price changes less frequently. Bivariate regressions tend to produce directionally accurate forecasts at shorter horizons.

**Forecast bias**

Forecast bias does not differ significantly across models for most forecast horizons and commodities. While there are some exceptions, they are not statistically significant. Bivariate regressions lead to forecasts with a higher bias for aluminum and copper for horizons up to one year. BVARs produce forecasts with a greater bias for oil beyond the one-year horizon (table 2).
Forecast accuracy

The forecast accuracy of models differs significantly across commodities and forecast horizons. All models are compared against each other for all horizons using the DM test. Table 3 highlights the two forecasting approaches that perform better for each commodity for forecast horizons 3-12 months (short-term) and 15-24 months ahead (medium-term).

- For aluminum prices, bivariate regression models and the machine learning approach are the most accurate for forecast horizons up to one year. In contrast, the OEM and the machine learning approach performed better than other approaches over the medium-term horizons.

- For copper prices, the OEM approach is more accurate than other approaches for all forecast horizons. CE forecasts and machine learning approaches had similar forecast accuracy compared to other approaches for forecast horizons up to one year.

- For oil prices, three approaches—bivariate regressions, CE forecasts, and the machine learning approach—produce forecasts with similar accuracy compared to other approaches for forecast horizons up to one year. CE forecasts and the OEM are the two most accurate approaches at forecast horizons beyond one year.

Additional considerations

Adding modeler priors. None of the forecasting approaches here have pre-designed scenarios or priors. However, in reality, it is often of interest to condition the forecasts using different scenarios (for example, different trajectories for the world economy, a shock to the supply of oil, or changes in policies). The OEM is particularly useful for scenario exercises that consider changes in policy variables, global growth, inflation, and structural variables. Another purpose of conditional forecasts is to incorporate information from higher frequency data or judgment into the model (Karlsson 2013). This underscores the main advantage of the BVAR model, which allows the forecaster to simulate scenarios or test priors while maintaining the statistical properties of the model. The forecaster can then make inferences about the posterior forecast, conditional on the prior. The version of the BVAR model used in this exercise leads to a larger oil price forecast bias than other models. However, in practice, the forecaster could subsequently adjust the parameters to reflect changes in priors to arrive at more informed forecasts.

Adjusting to shocks and incorporating judgment. Commodity markets are subject to a wide range of shocks affecting demand and supply dynamics, ultimately leading to sharp price movements. For example, the COVID-19 pandemic triggered a global recession in 2020, led...
FIGURE 20 Forecasts and realizations: 2015Q1-2022Q1

The approaches often lead to better forecast accuracy during the pre-pandemic period. This is in part driven by the disruption caused by the pandemic to the standard relationships between global macroeconomic conditions and commodity prices. Among the five approaches, the expert-and modeler-centric approaches (CE forecasts and the macroeconomic model) produced more accurate forecasts in the post-pandemic period, in part because they were better able to incorporate this information.

A. Aluminum prices and range of forecast results from 3 approaches
B. Copper prices and range of forecast results from 3 approaches
C. Crude oil prices and range of forecast results from 2 approaches
D. Crude oil prices and range of forecast results from 5 approaches

Notes: Grey shaded area denotes trough of the COVID-19 pandemic (2020Q2-Q3). Data shows 1-year-ahead rollover over a 12-month horizon forecast period. The range includes the forecast outcomes of the five approaches examined (Bivariate regressions, BVAR, Consensus Forecasts, the machine learning approach, and the Oxford Economic Model).
C.D. Measures average crude oil prices (unweighted average of Brent, West Texas Intermediate, and Dubai benchmarks) in U.S. dollars per barrel.
D. Statistical forecast approaches refer to the Bivariate regressions, BVAR, and the machine learning approach.

Utilizing multiple approaches. Each forecasting approach has unique strengths that should be considered in practice. For instance, the BVAR and OEM approaches excel in scenario analysis, while machine learning methods demonstrate proficiency in uncovering complex patterns in time series data that traditional statistical models may overlook. Bivariate regressions are particularly valued for their simplicity and ability to identify the most influential explanatory variables. Meanwhile, CE forecasts serve as a valuable sentiment indicator, offering a robust benchmarking alternative with the additional advantage of being timely and accessible.

Considering other industrial commodities. While the results are not reported here, lead, nickel, tin, and zinc prices are also evaluated using the same forecast evaluation exercise (Arroyo-Marioli et al. 2023). These four industrial metals have recently seen increasing demand in part due to their use in clean energy technologies, though they are less systemically important to the global economy compared to the three studied here. For these four industrial commodities, most approaches performed well, with the OEM forecast showing the lowest bias and forecast error for forecast horizons of 12 months or more.

Using futures prices. Prices of futures are often used for forecasting purposes by many organizations (Nixon and Smith 2012; World Bank 2023). They are simple to utilize as they reflect market expectations of future spot prices. Earlier studies report that futures prices are often unbiased but inefficient forecasts (with large forecast errors in either direction) compared to forecasts produced by other approaches, including to gyrations in commodity prices, and, for a brief period, even upended the standard relationships among macroeconomic variables. Forecasting performance of all approaches naturally suffered during the pandemic. The purely statistical forecasting approaches (bivariate regressions, BVAR, and, to a lesser extent, machine learning) were unable to quickly adjust to extraneous factors outside of the model. Past relationships embedded into the models during the training period temporarily broke down (figure 20.A-C). In contrast, forecasts based on the OEM displayed a smaller difference in relative performance before and after the pandemic, as the model had been adapted to accommodate the pandemic shock (figure 20.D). CE forecasts also performed relatively better as they reflected the aggregate views of many forecasters who could adjust their projections to account for the pandemic shock. Though the pandemic period was highly unusual, the behavior of different approaches serves as a good illustration of the need to incorporate judgment and other factors outside the model specification when forecasting.
VARs, machine learning techniques, and univariate time series models. Despite their relative underperformance, some studies find that futures prices do contain important predictive information, and the financialization of commodity markets may have helped improve their predictive power over time (Arroyo-Marioli et al. 2023; Ellwanger and Snudden 2023).

**Conclusions**

This Special Focus evaluates the performance of five widely used approaches to forecasting the prices of industrial commodities. The evaluation focuses on the prices of aluminum, copper, and crude oil as these commodities account for almost half of global commodity exports. It examines four model-based approaches (bivariate regressions; Bayesian vector autoregression models; a macroeconometric model; and a machine learning technique) and CE forecasts. These approaches are evaluated in terms of their performance with respect to directional accuracy, forecast bias, and forecast accuracy over the period 2015Q1-2022Q1. The evaluation finds four major results.

**No “one-approach-beats-all” for commodity price forecasting.** Most approaches produce directionally accurate forecasts at horizons of less than one year. Forecast bias does not differ significantly across approaches for most forecast horizons and commodities. However, the forecast accuracy of approaches varies significantly across commodities and time horizons. Since market conditions often change, it is not possible for one approach to systematically outperform the others across different commodities, particularly over shorter horizons.

**Macroeconometric models are better for longer-term forecasts.** These models tend to be more accurate at horizons of one year or more, partly because they can incorporate the impact of structural changes on prices. These models are also useful for conducting scenario analyses and considering forecasts conditional on certain outcomes.

**Need for incorporating judgment.** It is critical to complement mechanistic forecasts with judgment and information that cannot be accounted for by these approaches. Commodity prices are driven by forces that may not be captured by backward-looking statistical techniques. These techniques can be improved with reference to events or information known to the modeler but not yet incorporated in the data.

**Importance of multiple approaches.** A single approach can sometimes produce large forecast errors. Moreover, forecast accuracy varies significantly across approaches. These results collectively emphasize the importance of employing a rich menu of approaches in forecasting commodity prices.

For policymakers, these results underscore the uncertainty around commodity price forecasts and the need to develop contingency plans for alternative outcomes, particularly for economies heavily dependent on commodities for revenues. The usefulness of forecasting is sometimes less about predicting the future with accuracy and more about looking at how changes in certain assumptions might lead to different outcomes, as well as the risk associated with those outcomes. In practice, it is crucial to use various models, each with its strengths, coupled with an informed assessment of potential changes in commodity markets.

**References**


