INTERNATIONAL ECONOMIC REVIEW
Vol. 55, No. 3, August 2014

# WHAT CENTRAL BANKERS NEED TO KNOW ABOUT FORECASTING OIL PRICES*

By Christiane Baumeister and Lutz Kilian[1]

*Bank of Canada, Canada; University of Michigan, U.S.A.*

Central banks routinely use short-horizon forecasts of the quarterly price of oil in assessing the global and domestic economic outlook. We address a number of econometric issues specific to the construction of quarterly oil price forecasts in the United States and abroad. We show that quarterly forecasts of the real price of oil from suitably designed vector autoregressive models estimated on monthly data generate the most accurate real-time forecasts overall among a wide range of methods, including quarterly averages of forecasts based on monthly oil futures prices, no-change forecasts, and forecasts based on regression models estimated on quarterly data.

## 1. INTRODUCTION

Forecasting the real price of oil is an important but difficult exercise. It is widely accepted that changes in the cost of imported crude oil are an important determinant of economic activity, which is why central banks worldwide and international organizations such as the International Monetary Fund (IMF) routinely rely on real-time forecasts of the real price of oil in assessing the economic outlook. For example, the *IMF World Economic Outlook* of April 2012 concludes that "risks through 2013 remain to the upside for oil prices and thus to the downside for global growth" (p. 14). Perhaps the most common approach to forecasting the price of oil has been to employ forecasts based on oil futures prices, but such forecasts usually are not significantly more accurate than simple no-change forecasts.[2] Our analysis takes a different approach. We propose forecasting methods that exploit insights from recently developed economic models of the determination of the real price of oil to generate more accurate real-time forecasts of the quarterly real price of oil. Economic theory suggests that a number of global economic aggregates such as the global business cycle or changes in global crude oil inventories should contain information about future oil prices. We combine these predictors in the form of a vector autoregressive (VAR) model and compare this forecasting model to a wide range of alternative approaches. Our analysis takes account of the fact that the real-time oil price forecasts central banks and international organizations require as an input for policy decisions differ in several dimensions from currently available forecasts.

First, central banks rely on forecasts of the quarterly real price of oil instead of forecasts of the monthly real price of oil because their macroeconomic models tend to be specified at quarterly frequency. This distinction raises several important practical questions for real-time forecasters. For example, is it better to average monthly forecasts of the real price of oil or to generate forecasts from a model estimated at quarterly frequency? Is the appropriate benchmark for forecast comparisons the most recent quarterly average of the real price of oil or the most recent monthly observation? How does time aggregation to quarterly frequency affect

*Manuscript received February 2013; revised March 2013.

[1] The views in this article are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Bank of Canada. We thank the editor, two anonymous referees, Francesco Ravazzolo, and Rob Vigfusson for detailed comments. David Finer provided excellent research assistance. Please address correspondence to: Lutz Kilian, Department of Economics, 611 Tappan Street, Ann Arbor, MI 48109-1220, USA. E-mail: *lkilian@umich.edu*.

[2] Oil futures contracts allow traders to lock in today a price at which to buy or sell a fixed quantity of crude oil at a predetermined date in the future. Futures prices refer to the prices of futures contracts with different maturities.

the specification of forecasting models and how does time aggregation affect the properties of conventional central bank oil price forecasts based on oil futures prices?

Second, many central banks are concerned with the problem of forecasting the Brent price of crude oil or the price of West Texas Intermediate (WTI) crude oil. Brent and WTI refer to different grades of crude oil traded in different locations. Given the recent instability in the spread of the Brent price over the WTI price and given the increasing importance of the Brent price as a benchmark for global oil markets, this raises the question of how to model and how to forecast the real price of Brent crude oil in particular. This task is further complicated by the fact that Brent prices are available only back to mid-1987, whereas VAR models require larger samples. Possible modeling choices include, for example, backcasting the Brent price on the basis of alternative oil price series (such as the U.S. refiners' acquisition cost for crude oil imports, which is commonly viewed as the best proxy for global oil prices) or modeling the spread relative to this alternative oil price series as a random walk.

Third, central banks are interested in forecasting the real price of oil measured in domestic consumption units because this price captures the true cost of oil for domestic consumers. For example, the European Central Bank requires forecasts of the real price of oil in European consumption units and the Bank of Canada in Canadian consumption units. The fact that crude oil is traded in U.S. dollars necessitates the inclusion of the real exchange rate in the forecasting model when forecasting the real price of oil for countries other than the United States. One option is to simply augment the forecasting model by the real exchange rate; another is to treat the quarterly real exchange rate as a random walk.

Our objective in this article is to develop new forecasting methods that help to address these questions. We explore the relative merits of a range of alternative methods and provide guidance on how to construct the most accurate forecast of the quarterly real price of oil in real time, reflecting the data constraints under which central bankers operate. Such forecasts are of interest not only to policymakers, but also to energy economists modeling the demand for energy-using durables and to companies in the transportation and energy sector whose investment decisions directly depend on the expectation of the real price of oil (see, e.g., Bernanke, 1983; Busse et al., 2013). Finally, many of the tools and insights developed in this article are relevant for forecasting other commodity prices as well.

In the course of our analysis, we also investigate several other important modeling issues. The first issue is how to approximate changes in global real economic activity in oil market VAR forecasting models. It is now widely understood that the state of the global economy matters for the real price of oil and should be incorporated in forecasting models. One common choice is the monthly global shipping index originally developed in Kilian (2009); an alternative measure is the OECD measure of monthly industrial production for the OECD economies and six emerging economies. The set of possible proxies becomes even wider once we allow for the use of quarterly data. It includes not only quarterly aggregates of conventional monthly measures, but also quarterly real GDP data for the OECD economies. To date nothing is known about the relative merits of these proxies for forecasting the real price of oil. Nor is it known which transformation of the OECD real activity measures (e.g., percent changes, detrended log data) is most useful in the out-of-sample setting.

A second important concern is that the predictive relationships in global oil markets may be subject to smooth structural change. Models of structural change have received increasing attention in recent years among forecasters of inflation and economic activity (see, e.g., Stock and Watson, 2003). Although the use of time-varying parameter vector autoregressive (TVP-VAR) models is not feasible when working with large-dimensional monthly VAR forecasting models, it becomes feasible when working with quarterly VAR models. An obvious question is whether modeling possible smooth structural change as a quarterly TVP-VAR model improves the accuracy of oil price forecasts relative to constant-parameter VAR models. An alternative response to potential structural change in the literature has been to generate VAR forecasts based on rolling instead of recursive windows of data. We explore both modeling choices.

A third and related conjecture is that, given the inevitable misspecification of all forecasting models, central banks may be better off relying on forecast combinations instead of one forecasting model only. For example, an important question faced by all central banks is whether a combination of forecasts based on oil futures prices and forecasts based on econometric models is more accurate than model-based forecasts alone.

Our analysis shows that the monthly random walk model is considerably more accurate than the quarterly random walk model, reflecting the informational advantages of using the most recent monthly observation for a given quarter. In fact, it is easy to find short-horizon forecasts that seemingly outperform the random walk benchmark by a wide margin, if the random walk benchmark is based on the most recent quarterly average. The mean squared prediction error (MSPE) rankings are reversed when using the monthly random walk as the benchmark, however. This result also provides intuition for our finding that quarterly VAR models tend to be less accurate than monthly VAR(12) models in forecasting the quarterly real price of oil. The latter result holds true regardless of which real activity variable, lag order, or estimation method is used and regardless of how the real activity variable and the real price of oil are specified. It also holds after allowing for smooth structural change in the quarterly model.

In addition, the monthly VAR(12) specification for the real refiners' acquisition cost for crude oil imports is more accurate than quarterly forecasts of the real price of oil based on monthly oil futures prices at horizons of one and two quarters. This result also holds when forecasting the real price of Brent crude oil and the real price of WTI crude oil. The best way to forecast the latter oil prices is to augment the baseline monthly VAR forecasting model for the real refiners' acquisition cost for crude oil imports with a no-change forecast for the spread of these prices over the U.S. refiners' acquisition cost. Likewise, the most accurate forecasts for the real price of oil in non-U.S. consumption units are obtained by scaling the forecast of the U.S. real price of oil by the most recently observed monthly real exchange rate. At horizons of three and four quarters, in contrast, forecasts based on oil futures prices have lower MSPE (but not systematically higher directional accuracy) than forecasts from monthly VAR models.

Finally, there is little evidence that forecast combinations improve forecast accuracy compared with the best-performing monthly VAR forecast. Only combinations of forecasts based on oil futures prices and the monthly VAR(12) model show any promise at all. In the latter case, some MSPE reductions are obtained at horizons of three and four quarters, whereas at shorter horizons the VAR model forecast alone tends to be more accurate than the combined forecast.

The remainder of the article is organized as follows: Section 2 describes the forecasting environment and the construction of the real-time data underlying our analysis. Section 3 focuses on the problem of forecasting the real price of oil in U.S. consumption units. We first discuss the choice of the random walk benchmark for the quarterly real price of oil. Then we show how to generate quarterly forecasts from monthly oil market VAR models and compare the results to conventional forecasts based on oil futures prices and to forecasts from quarterly VAR models. Next we evaluate alternative approaches to extending the baseline results for the U.S. refiners' acquisition cost for crude oil imports to other crude oil benchmarks such as WTI crude oil and Brent crude oil. Finally, we explore the merits of alternative proxies for global real activity in monthly and quarterly VAR models, and we investigate the costs and benefits of allowing for time variation in the VAR model parameters. Section 4 discusses how to adapt the analysis to the problem of forecasting the real price of oil in the consumption units of other countries. We focus on the examples of Canada, Norway, and the Euro area. In Section 5, we examine the question of whether central banks should rely on forecast combinations instead of relying on one forecasting method only. In particular, we ask whether a combination of forecasts based on oil futures prices and forecasts based on econometric models is more accurate than model-based forecasts alone. Section 6 discusses several extensions. The concluding remarks are in Section 7.

## 2. THE FORECASTING ENVIRONMENT

2.1. *Background.* The objective throughout this article is to forecast the quarterly average of the real price of oil at horizons of up to one year. Our focus on the average quarterly price is consistent with the fact that agencies such as the U.S. Energy Information Administration (EIA) produce forecasts of the quarterly price of oil. There are two basic approaches to constructing forecasts of the quarterly real price of oil. One option is to forecast the monthly real price of oil for each horizon between 1 and 12 months and to convert these monthly forecasts into quarterly averages. This approach allows one to rely on well-established methods of forecasting the real price of oil at monthly frequency (see Baumeister and Kilian, 2012, 2014). The other option is to respecify the forecasting model for quarterly data. We consider both approaches.

Our analysis covers a wide range of alternative forecasting methods for the quarterly real price of oil, including many monthly and quarterly reduced-form VAR forecasting models, monthly and quarterly no-change forecasts, forecasts based on oil futures prices, and forecast combinations. Our simulated out-of-sample exercise mimics the real-time constraints on the availability and reliability of the data used by each forecasting method, providing an indication of how each method would have fared on the last 25 years of data. We consider two loss functions. First, we construct the recursive MSPE of each forecasting method relative to the MSPE of a random walk benchmark model. A ratio below unity indicates a reduction in the MSPE relative to the random walk benchmark model. The choice of the random walk benchmark is conventional. It does not affect the ranking of the forecasting methods but facilitates the comparison with earlier results in the literature. Second, we compute the success ratio corresponding to the fraction of times the recursive forecast correctly predicts the direction of change in the real price of oil. The loss functions are evaluated in terms of the level of the real price of oil instead of its log-level because it is the level that matters for policy discussions at central banks.

2.2. *The Real-Time Data Set.* The use of real-time data raises two distinct complications. One is that even preliminary data often become available only with a lag; the other complication is that data will be continuously revised after they become available. These features of the data require us to keep track of each *vintage* of data, containing the data actually known to real-life forecasters at each point in time, when evaluating each method's simulated out-of-sample accuracy.[3] Our real-time data set builds on the existing real-time database documented in Baumeister and Kilian (2012), which consists of monthly vintages of real-time data for 1991.1–2011.12, each of which extends back to 1973.1.

The series included in this database allow the real-time construction of the four model variables underlying the oil market VAR model of Kilian and Murphy (2014): (i) the growth rate of global oil production, (ii) the Kilian (2009) shipping index of global real economic activity, (iii) the real price of oil (obtained by deflating the nominal price of oil by the U.S. Consumer Price Index [CPI]), and (iv) the change in global crude oil inventories obtained by scaling U.S. crude oil inventories by the ratio of OECD petroleum inventories over U.S. petroleum inventories. The database also allows the construction of real-time forecasts of the real price of oil based on WTI futures prices.[4] For the purpose of our analysis, we extend this real-time database to include several additional time series that allow the construction of real-time real exchange rates for selected countries and that allow us to explore alternative oil price measures and alternative measures of global real economic activity.

[3] For a more detailed exposition of the real-time data problem, the reader is referred to Croushore (2011).

[4] Specifically, we use the following monthly series from this database: (i) the average daily WTI spot price of crude oil, (ii) averages of daily WTI oil futures prices at maturities between 1 and 12 months, (iii) an index of bulk dry cargo ocean shipping freight rates, (iv) the nominal U.S. refiners' acquisition cost for crude oil imports, (v) world crude oil production, (vi) U.S. crude oil inventories, (vii) U.S. petroleum inventories, (viii) OECD petroleum inventories, (ix) the U.S. consumer price index for all urban consumers, and (x) the U.S. producer price index for crude oil. The nominal shipping rate data are obtained from Kilian (2009) for 1973.1–84.12 and are extrapolated through 2010.12 using the Baltic Dry Cargo Index (BDI) from Bloomberg.

2.2.1. *Brent oil prices.* One addition is spot and futures prices for Brent crude oil traded on the Intercontinental Exchange. The Brent spot price is provided by Reuters (as reported by the EIA). Monthly averages of daily Brent oil prices are only available back to 1987.6. We extrapolated the monthly Brent spot price series back to 1973.1 at the rate of change in the U.S. refiners' acquisition cost for crude oil imports, which is widely considered a good proxy for the global price of crude oil. We also added the corresponding futures price of Brent crude oil from Bloomberg at maturities of 1–9 months. Maturities of 10–12 months are only available starting in April 1994. The Brent data are available in real time and not subject to revisions.

2.2.2. *Nominal exchange rates for Canada, Norway, and the Euro area.* Monthly averages of the daily nominal spot exchange rates for the Canadian dollar, the Norwegian kroner, and the Euro with respect to the U.S. dollar were obtained from the Federal Reserve Board database. The Euro exchange rate prior to January 1999 was constructed based on the U.S. dollar/ECU exchange rate. The nominal exchange rate data are available in real time by construction.

2.2.3. *Consumer price indices for Canada, Norway, and the Euro area.* For the construction of real-time real exchange rates, monthly consumer price indices for Canada, Norway, the Euro area, and the United States are obtained from the *Original Release Data and Revisions Database* for the OECD Main Economic Indicators. Vintages start in February 1999 and contain data back to 1973. The missing vintages for 1991.1–99.1 are approximated by mimicking the constraints on the real-time availability of the CPI data. We adjust the ex post revised OECD CPI data to reflect a one-month delay in the real-time availability of the CPI data to the forecaster.[5] Gaps in the availability of these pseudo real-time CPI data are filled by nowcasting the most recent monthly observation based on the average rate of inflation up to that point in time. This simple nowcasting procedure works well for the U.S. data, as shown in Baumeister and Kilian (2012). The resulting pseudo real-time data reflect constraints on the real-time availability of the data but do not reflect data revisions across vintages. There is reason to believe that the latter effect is small, however, given evidence from the United States as well as Germany that consumer prices are rarely revised and, if so, only to a small extent.[6]

2.2.4. *Alternative measures of global real economic activity.* Another addition to the database is a monthly index of industrial production for the OECD economies and six major non-OECD economies including China, India, Brazil, South Africa, Indonesia, and the Russian Federation (abbreviated as OECD+6). The index is available back to 1973.1 from the OECD's *Main Economic Indicator* database. Yet another addition is the quarterly real GDP index for all OECD economies from the same source, which is based on purchasing power parity weights as of 2005 and spans the period 1973Q1–2011Q3. Although it is not possible to construct a true real-time version of these global real activity measures that captures data revisions over time, we can construct pseudo real-time data that account for delays in the real-time availability of these data. In constructing these pseudo real-time data, we impose a delay of three months for the OECD+6 industrial production data and of two quarters for the OECD real GDP data based on the observable delay at the end of 2011. We nowcast the resulting gaps in the global real activity data by extrapolating the most recent observation at the average rate of growth in the earlier data. Our analysis also accounts for the fact that all subsequent data transformations for these real activity data must be applied in real time.

[5] Although in the OECD real-time database there tends to be a two-month lag in the availability of CPI data for the countries in question, it can be shown that national central banks provide data on consumer prices with a delay of only one month, making it reasonable to impose a one-month delay.

[6] For the Euro area, matters are more complicated than for Canada and Norway because the real-time data for the harmonized consumer price index (HICP) for the Euro area compiled by the OECD are not compatible with the U.S. CPI data due to benchmark revisions. This fact matters for the construction of real exchange rates (also see Giannone et al., 2012, for related discussion). In constructing the euro–dollar real exchange rate, we therefore rely on pseudo real-time equivalents for the HICP and U.S. CPI data based on the December 2011 vintage throughout.

2.3. *The Construction of the Quarterly Real-Time Data.* Quarterly real-time data, as required by quarterly VAR forecasting models, may be constructed from this monthly real-time data set as follows: The quarterly growth rate of global oil production is constructed as the log difference of the last monthly observation for each quarter. Similarly, changes in global crude oil inventories are computed as the difference of the last monthly observation for each quarter. Quarterly averages of the (suitably updated) Kilian (2009) index of global real activity and of the real price of oil are constructed as the average of the three monthly real-time observations for each quarter. The same approach applies to the construction of quarterly averages of and quarterly growth rates of industrial production in the OECD+6 economies.

2.4. *The Construction of the Ex Post Revised Quarterly Real Price of Oil.* The objective throughout this article is to forecast the final release of the real price of oil after all revisions have taken place. Allowing for presample observations, the estimation period for monthly data starts in 1974.2. The evaluation period extends from 1992.1 to 2011.6. For quarterly models, the estimation period starts in 1974.II and the evaluation period covers 1992.I–2011.II. When evaluating the forecasting methods, we treat the data up to 2011.6 in the December 2011 vintage as our proxy for the ex post revised data. The implicit premise is that all data revisions underlying the real oil price data have taken place within half a year, which is consistent with evidence presented in Baumeister and Kilian (2012).

## 3. QUARTERLY FORECASTS FOR THE UNITED STATES

The problem of forecasting the quarterly real price of oil has not been studied to date, making it necessary to discuss a number of methodological issues and modeling choices.

3.1. *The Random Walk Benchmark.* As in the literature on forecasting asset prices, the traditional benchmark in forecasting the real price of oil has been the random walk forecast or no-change forecast (see, e.g., Alquist et al., 2013). The highest frequency at which a no-change forecast of the real price of oil can be constructed is monthly because CPI data are not available at higher frequency. Thus, when forecasting the monthly real price of oil, the random walk benchmark is clearly defined. When forecasting the quarterly real price of oil, in contrast, there are two ways of constructing a no-change forecast. We may rely on the most recent quarterly real price of oil ("quarterly no-change forecast") or we may use the last monthly observation for the real price of oil in the most recent quarter ("monthly no-change forecast"). Which random walk benchmark will have the lower MSPE in practice is not clear ex ante. On the one hand, to the extent that the real price of oil moves up (or down) persistently, one would expect the monthly no-change forecast to have lower squared errors except in the rare event of a turning point in the data. On the other hand, monthly no-change forecasts are noisier by construction than forecasts based on quarterly averages, which will inflate their MSPE.

The first column of Table 1 shows that at forecast horizons of one and two quarters, the monthly no-change forecast is clearly more accurate than the quarterly no-change forecast. For example, the one-quarter-ahead MSPE of the quarterly no-change forecast is 68% higher than that of the monthly no-change forecast. Two quarters ahead the difference is still 11%. At longer horizons, in contrast, the differences are minimal, with little to choose between the two random walk forecasting models. We conclude that the monthly no-change forecast is the only credible benchmark for judging the forecasting ability of alternative forecasting methods. Hence, all MSPE results in this article are normalized relative to the MSPE of the monthly no-change forecast of the quarterly real price of oil.

3.2. *Using the Monthly Oil Market VAR Model to Generate Quarterly Forecasts.* Although the random walk is a tough benchmark to beat, Baumeister and Kilian (2012) showed that it is possible to beat the no-change model for the monthly real price of oil in real time at horizons of up to 12 months. The most successful forecasting models in that study are reduced-form VAR

TABLE 1
REAL-TIME ACCURACY OF RECURSIVE FORECASTS OF THE QUARTERLY REAL U.S. REFINERS' ACQUISITION COST FOR OIL IMPORTS

| Quarterly Horizon | Quarterly No-Change Forecast | Monthly VAR(12) | Oil Futures | Quarterly VAR($p$) $p=4$ | $p=6$ | $p=8$ | Quarterly BVAR($p$) $p=4$ | $p=6$ | $p=8$ |
|---|---|---|---|---|---|---|---|---|---|
| MSPE ratio | | | | | | | | | |
| 1 | 1.66 | **0.80** | **0.97** | 1.59 | 1.86 | 2.18 | 1.65 | 1.58 | 1.62 |
| 2 | 1.10 | **0.93** | **0.98** | 1.17 | 1.32 | 1.40 | 1.13 | 1.10 | 1.13 |
| 3 | **0.98** | 1.02 | **0.96** | 1.05 | 1.09 | 1.13 | 1.02 | **0.98** | 1.03 |
| 4 | **0.99** | 1.01 | **0.93** | 1.02 | 1.06 | 1.18 | 1.01 | **0.99** | 1.05 |
| Success ratio | | | | | | | | | |
| 1 | – | **0.69*** | **0.69*** | **0.55** | **0.62*** | **0.56** | **0.56** | **0.63*** | **0.67*** |
| 2 | – | **0.58*** | 0.49 | **0.53** | **0.58** | **0.53** | **0.55** | **0.64*** | **0.64*** |
| 3 | – | **0.57** | **0.51**** | 0.45 | 0.50 | **0.51** | **0.53** | **0.58** | **0.54** |
| 4 | – | **0.60*** | **0.56*** | 0.48 | **0.61** | **0.55** | **0.52** | **0.60** | **0.57** |

NOTES: All MSPE ratios have been normalized relative to the monthly no-change forecast. Boldface indicates an improvement on the monthly no-change forecast. With the exception of the oil futures forecast and the quarterly random walk forecast, the statistical significance of the real-time recursive MSPE ratio cannot be assessed because no valid statistical tests are available in the literature. None of the improvements, if any, produced by the oil futures forecast and the quarterly random walk forecast in this and subsequent tables are statistically significant. For the success ratio, improvements that are statistically significant at the 5% (10%) level based on test of the null of no directional accuracy in Pesaran and Timmermann (2009) are marked * (**).

models of the real price of oil containing data on global real activity, global oil production, and global oil inventories that matter for the determination of the real price of oil according to economic theory. A natural starting point for our analysis thus is the type of monthly VAR forecasting model found to work well in the analysis in Baumeister and Kilian (2012):

$$y_t^M = \nu + B_1 y_{t-1}^M + \cdots + B_{12} y_{t-12}^M + u_t^M, \quad u_t^M \sim (0, \Omega),$$

where $y_t^M$ is a $4 \times 1$ vector of monthly model variables containing the growth rate of global crude oil production, the Kilian (2009) business cycle index of global real economic activity, the real U.S. refiners' acquisition cost of crude oil imports, and the change in global above-ground crude oil inventories. Here, $\nu$ refers to the vector of intercepts, $B_i, i = 1, \ldots, 12$, denotes the $4 \times 4$-dimensional matrices of slope parameters, and $\Omega$ is the variance–covariance matrix of the innovations. The monthly VAR model is estimated by the method of least squares.[7] Forecasts of the quarterly real price of oil are constructed as the average of the forecasts for each month contained in a given quarter.

It should be noted that, notwithstanding the favorable performance of this forecasting model for the monthly real price of oil in earlier research, there is no a priori reason why this model should be equally accurate at forecasting the quarterly real price of oil. Not only did Baumeister and Kilian only report results for a subset of the relevant monthly horizons, but the monthly forecasts in question will be correlated. This makes it impossible to infer from the MSPE ratio of a forecast at the 12-month horizon, for example, the MSPE ratio for the quarterly average over the horizons of 10, 11, and 12 months.

Table 1 shows that this forecasting approach is remarkably accurate nevertheless. The monthly VAR(12) model yields MSPE reductions relative to the monthly no-change forecast of 20% one quarter ahead and of 7% two quarters ahead. At longer horizons, the monthly VAR(12) model has about the same MSPE as the benchmark model. Moreover, unlike the

[7] Adding more lags is not advisable unless Bayesian estimation methods are used. A monthly BVAR(24) model, for example, yields slightly higher directional accuracy, but also higher MSPEs, especially at longer horizons. Moreover, Bayesian estimation does not improve the accuracy of the VAR(12) model, which is why we focus on the unconstrained VAR model. We also evaluated univariate monthly AR($p$) models, $p \in \{6, 12, 24\}$, for the real price of oil, but the results were not as accurate as for the monthly VAR(12) model and are not reported.

no-change forecast, the VAR forecast has directional accuracy ranging from 57% to 69%, depending on the horizon. This compares with a success ratio of 50% under the null hypothesis of no directional accuracy. Such success ratios are remarkably high by the standards of the empirical finance literature (see, e.g., Pesaran and Timmermann, 1995). Except for the third quarter, the gains in directional accuracy are statistically significant based on the test of the null of no directional accuracy discussed in Pesaran and Timmermann (2009).[8]

3.3. *Oil Futures-Based Forecasts of the Quarterly Real Price of Oil.* A central banker would raise the obvious question of how these VAR results compare to conventional quarterly forecasts generated on the basis of oil futures prices. Although there are no oil futures prices for the U.S. refiners' acquisition cost for crude oil imports, we may use the expected change in the WTI price of oil to extrapolate from the current real refiners' acquisition cost:

$$R^M_{t+h|t} = R^M_t \left(1 + f_t^{M,WTI,h} - s_t^{M,WTI} - \pi_t^{M,h}\right), \quad h = 1, \ldots, 12,$$

where $R^M_t$denotes the level of the monthly real U.S. refiners' acquisition cost for crude oil imports, $f_t^{M,WTI,h}$ is the log of the monthly average nominal WTI futures price of $h$ months maturity, and $s_t^{M,WTI}$is the corresponding monthly nominal WTI spot price in logs. Under standard assumptions used by practitioners, $f_t^{M,WTI,h} - s_t^{M,WTI}$ may be viewed as the expected change in the nominal WTI spot price over the next $h$ months (see, e.g., Alquist and Kilian, 2010). The term $\pi_t^{M,h}$ denotes the expected cumulative inflation rate over the next $h$ months, which in practice can be approximated based on the cumulative inflation rate over the last $h$ months.[9] The monthly real oil price forecasts generated using this method are averaged by quarter to produce the quarterly forecasts of the real refiners' acquisition cost. Table 1 shows that the MSPE of the oil futures-based forecast does not significantly improve on the monthly no-change forecast at any horizon.[10] In fact, it is about as accurate as the no-change forecast at the one- and two-quarter horizons. Only at the four-quarter horizon does the relative MSPE improve noticeably, but the improvement is not statistically significant. This does not mean that the there is no information in the futures-based forecast, however. Table 1 indicates much higher directional accuracy at some horizons than reported in Baumeister and Kilian (2012) for the same model evaluated at monthly frequency, illustrating our earlier point that the results for quarterly horizons cannot be inferred from existing results in the literature for monthly horizons. The success ratios range from 51% to 69%. They are statistically significant at three of four quarterly horizons, but they are typically lower than for the monthly VAR(12) model.

3.4. *Forecasts from Quarterly Oil Market VAR Models.* As discussed in Section 2, an alternative approach to generating forecasts of the quarterly real price of oil is to respecify the VAR model in question at quarterly frequency:

$$y_t^Q = \nu + B_1 y_{t-1}^Q + \cdots + B_p y_{t-p}^Q + u_t^Q, \quad u_t^Q \sim (0, \Omega),\ \ p \in \{4, 6, 8\},$$

[8] Although we report tests of statistical significance for the success ratios in Table 1, we do not provide measures of statistical significance for the MSPE reductions except in the case of the quarterly no-change forecasts and the futures-based forecasts. The reason is that such tests are not available for iterated forecasts subject to regression estimation error. Existing tests are inappropriate and unreliable (see, e.g., Clark and McCracken, 2012). This problem is compounded by the fact that there are no suitable tests in the literature that allow for real-time data constraints (see Clark and McCracken, 2009). Nor is it possible to rely on bootstrap methods to simulate the critical values of tests of equal predictive accuracy in our iterated real-time setting.

[9] Given the small magnitude of the inflation rate compared with fluctuations in the nominal price of oil over the horizons of interest, this approximation is adequate, as shown in Baumeister and Kilian (2012).

[10] Given the absence of parameter estimation error, this test is conducted based on the *DM*-test statistic of Diebold and Mariano (1995).

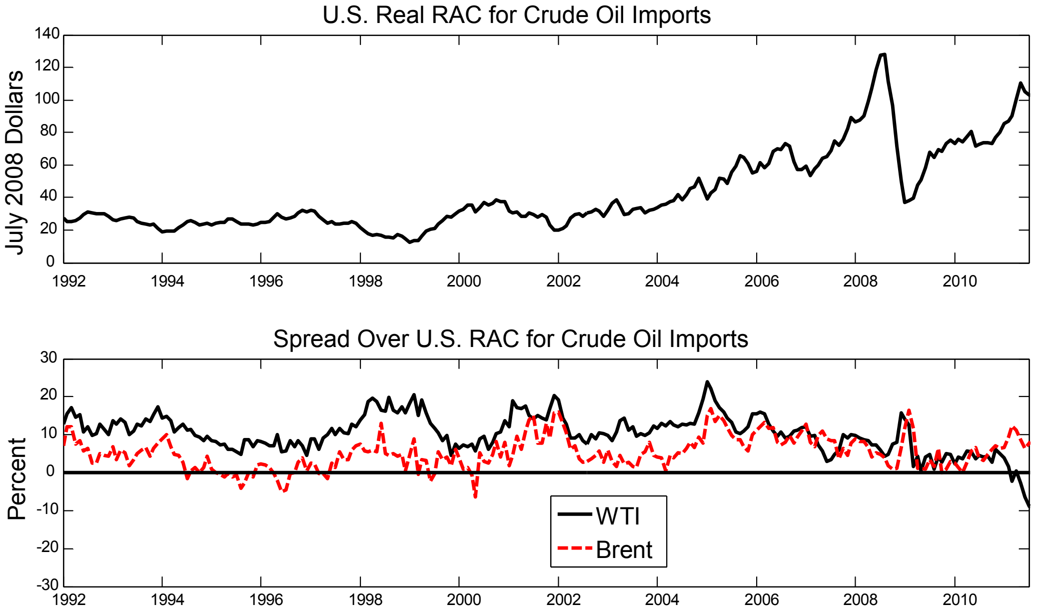


FIGURE 1

ALTERNATIVE OIL PRICES AND THEIR RELATIONSHIP SINCE 1992

where $y_t^Q$ denotes the $4 \times 1$ vector of quarterly model variables obtained by time aggregation from the data used in the monthly VAR models. Time aggregation will affect the dynamics of the VAR model, so we explore three alternative lag order settings. Given that a model with more than four autoregressive lags may be too heavily parameterized for unrestricted estimation, we estimate these VAR models alternatively using the method of least squares and the Bayesian estimation method developed by Giannone et al. (2012) for VAR forecasting models.[11] Models estimated using the latter methods are denoted as BVAR models, whereas models estimated by the method of least squares are denoted as VAR models.

Table 1 shows that the quarterly BVAR(6) model is the most accurate forecasting model among all models estimated on quarterly data. Even the BVAR(6) model, however, tends to have much higher MSPEs than the monthly no-change forecast at short horizons. At the one-quarter horizon, the loss in accuracy is 58%. Compared with the monthly no-change model, the one-quarter ahead MSPE almost doubles. Although the BVAR(6) forecasting model has some directional accuracy, overall, it is clearly dominated by the monthly VAR(12) model. We conclude that quarterly VAR and BVAR models cannot be recommended.

### 3.5. *Alternative Crude Oil Benchmarks.*

3.5.1. *West Texas Intermediate.* So far we have focused on the U.S. refiners' acquisition cost for crude oil imports, which traditionally has been considered a good indicator for the price of oil in global markets. An alternative benchmark that tends to receive more attention in the press is the price of WTI crude oil. The latter price was subject to U.S. government regulation until the early 1980s, making it less suitable for a VAR analysis of the global market for crude oil. Figure 1 illustrates the differences between these oil price series.

[11] Giannone et al. (2012) propose a method for selecting a Gaussian prior for the VAR parameters in real time. Their approach avoids the temptation of searching for priors that ex post generate more accurate forecasts and preserves the real-time nature of the forecasting exercise. The prior mean for all model variables, including the real price of oil, is chosen such that VAR coefficients are shrunk toward independent white noise under the assumption of stationarity. The degree of shrinkage is determined by the marginal data density.

TABLE 2
REAL-TIME ACCURACY OF RECURSIVE FORECASTS OF THE QUARTERLY REAL WTI PRICE

| Quarterly Horizon | Quarterly No-Change Forecast | Monthly VAR(12) | Monthly VAR(12) with No-Change Forecast for Spread | Oil Futures | Quarterly BVAR(6) | Quarterly BVAR(6) with No-Change Forecast for Spread |
|---|---|---|---|---|---|---|
| MSPE ratio | | | | | | |
| 1 | 1.67 | **0.93** | **0.85** | **0.99** | 1.57 | 1.55 |
| 2 | 1.11 | **0.97** | **0.94** | **0.99** | 1.07 | 1.10 |
| 3 | **0.99** | 1.02 | 1.01 | **0.98** | **0.96** | **0.98** |
| 4 | **0.99** | 1.01 | 1.00 | **0.95** | **0.95** | **0.98** |
| Success ratio | | | | | | |
| 1 | – | **0.65*** | **0.69*** | **0.67*** | **0.55** | **0.60*** |
| 2 | – | **0.61*** | **0.61*** | 0.49 | **0.57** | **0.58** |
| 3 | – | **0.51** | **0.58** | **0.53**** | **0.59** | **0.59** |
| 4 | – | **0.56** | **0.60*** | **0.56*** | **0.65*** | **0.63** |

NOTES: See Table 1.

There are two ways of modeling the quarterly real price of WTI. One approach is to replace the real U.S. refiners' acquisition cost by the real WTI price in the VAR(12) baseline model. This approach disregards the fact that the real WTI price strictly speaking is not appropriate for modeling global oil markets. Table 2 shows that this approach still works reasonably well in practice, but the extent of the gains in accuracy relative to the monthly no-change forecast is somewhat diminished. For example, whereas the same type of model in Table 1 produced MSPE reductions of 20% one quarter ahead, the corresponding reduction in Table 2 is only 7%. Likewise, the one-quarter ahead directional accuracy falls from 69% to 65%.

An alternative and more appealing approach is to retain the baseline VAR model of Table 1, but to convert the resulting forecasts to the WTI benchmark. This conversion requires a forecast of the spread of the WTI price over the U.S. refiners' acquisition cost for crude oil imports. A parsimonious way of forecasting this spread is to use a no-change forecast. Table 2 shows that this alternative approach indeed works better. For example, it produces an MSPE reduction of 15% and a success ratio of 69% one quarter ahead. Overall, these results are similar to those for the baseline monthly VAR(12) model in Table 1.

The modified monthly VAR(12) model based on the spread also is much more accurate than forecasts based on WTI futures at short horizons. Although the directional accuracy of the futures-based forecast is higher than that of the monthly no-change forecast, futures-based forecasts have lower directional accuracy than the suitably modified monthly VAR(12) model. Finally, Table 2 shows that the quarterly no-change forecast and a range of quarterly VAR forecasting models including the BVAR(6) model do not perform well by any metric.

3.5.2. *Brent.* Traditionally, the spread between different crude oil benchmarks has been stable over time. An interesting recent development in global oil markets has been that the spread of the price of Brent crude oil has grown disproportionately relative to the price of U.S. benchmarks such as the WTI. Indeed, some observers have suggested that the marginal barrel of oil today is being priced at the Brent benchmark, making the Brent price de facto a measure of the global price of crude oil. The lower panel of Figure 1 shows the evolution of the spread of the Brent price over the U.S. refiners' acquisition cost for crude oil imports and the corresponding spread of the WTI price over the U.S. refiners' acquisition cost for crude oil imports since 1992. Although the implicit spread between Brent and WTI has widened since 2011, the spread of the Brent price over the U.S. refiners' acquisition cost is not unusual by historical standards. As in the case of the WTI price, there are two ways of modeling the real price of Brent crude oil.

One approach is to substitute the real Brent price for the real refiners' acquisition cost in the baseline monthly VAR(12) model. Given that Brent prices do not exist prior to mid-1987,

TABLE 3
REAL-TIME ACCURACY OF RECURSIVE FORECASTS OF THE QUARTERLY REAL BRENT PRICE

| Quarterly Horizon | Quarterly No-Change Forecast | Monthly VAR(12) | Monthly VAR(12) with No-Change Forecast for Spread | Oil Futures | Quarterly BVAR(6) | Quarterly BVAR(6) with No-Change Forecast for Spread |
|---|---|---|---|---|---|---|
| MSPE ratio | | | | | | |
| 1 | 1.68 | **0.92** | **0.89** | 1.11 | 1.73 | 1.61 |
| 2 | 1.11 | **0.98** | **0.98** | 1.06 | 1.15 | 1.12 |
| 3 | **0.98** | 1.01 | 1.04 | 1.00 | 1.00 | 1.00 |
| 4 | **0.98** | 1.01 | 1.03 | - | 1.00 | 1.01 |
| Success ratio | | | | | | |
| 1 | – | **0.72*** | **0.68*** | **0.59*** | **0.59** | **0.59** |
| 2 | – | **0.61*** | **0.62*** | **0.53** | **0.60** | **0.62**** |
| 3 | – | **0.51** | **0.57**** | **0.55**** | **0.54** | **0.57** |
| 4 | – | **0.60*** | **0.57**** | – | **0.52** | **0.56** |

NOTES: See Table 1. Brent futures prices with a maturity of 10–12 months are not available for our evaluation period.

this approach requires backcasting the Brent price at the rate of change in the U.S. refiners' acquisition cost. The other approach is to retain the original model, but to treat the Brent price spread as a random walk in forecasting the quarterly real price of Brent. The latter approach does not require any backcasting of the Brent price data. Table 3 shows that both approaches work well, but the model based on the Brent price spread overall appears slightly more accurate. In particular, the latter model has lower MSPE at short horizons and significant directional accuracy at all horizons. In contrast, forecasts based on Brent futures prices cannot be recommended. The futures-based model has much higher MSPE than the monthly no-change forecast at short horizons and lower directional accuracy.[12] Likewise, the quarterly no-change forecast and the quarterly VAR models do not perform well.

3.6. *Alternative Measures of Global Real Activity.* One of the reasons that VAR forecasting models tend to be more accurate than univariate forecasting models of the real price of oil is the inclusion of a proxy for global real economic activity (see Alquist et al., 2013). In the baseline model, we rely on the monthly global shipping index originally developed in Kilian (2009) that by now has become standard in modeling the real price of oil. This global business cycle index, while appealing for reasons discussed in Kilian (2009), is not the only possible choice, however. One alternative is the use of OECD data on monthly industrial production in the OECD economies and in six emerging economies (China, India, Brazil, Russia, South Africa, and Indonesia). We explore three transformations of these industrial production data: a specification in growth rates, a business cycle index obtained by applying a one-sided HP-filter to these data, and a business cycle index obtained by linear deterministic detrending.[13] All these transformations are implemented in real time, as are the corresponding transformations underlying the shipping index.

For expository purposes, Table 4 focuses on the baseline monthly VAR(12) model. We consider all three oil price specifications. Table 4 shows that using the shipping index always produces lower MSPE ratios than using industrial production regardless of the data transformation. When it comes to directional accuracy, no model uniformly dominates all others, but on average the specification involving the growth rate of industrial production has slightly higher

[12] The performance of the forecast based on Brent futures prices is somewhat sensitive to the definition of the Brent spot price. When replacing our baseline series for the Brent spot price (as reported by the EIA and in the FRED data base of the St. Louis Fed) by an alternative price series with broader coverage obtained from Datastream (and backcast like the baseline price series), the MSPE ratio of the Brent futures-based forecast improves to 1.01 one quarter ahead and 1.03 two quarters ahead, while remaining at 1.00 three quarters ahead. Likewise, the accuracy of the VAR forecasting models does not change systematically.

[13] We implement the one-sided HP filter, as discussed in Stock and Watson (1999).

TABLE 4
REAL-TIME ACCURACY OF RECURSIVE FORECASTS OF THE QUARTERLY REAL PRICE OF OIL: ALTERNATIVE MONTHLY MEASURES OF GLOBAL REAL ACTIVITY IN THE VAR(12) MODEL

| | | | | MSPE Ratio Quarterly Horizon | | | | Success Ratio Quarterly Horizon | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Source | Transformation | Measure | Coverage | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| U.S. refiners' acquisition cost for imports | | | | | | | | | | | |
| Kilian (2009) | - | Shipping index | World | **0.80** | **0.93** | **1.02** | **1.01** | 0.69* | 0.58* | 0.57 | 0.60* |
| OECD | Growth rate | Industrial production | OECD + 6 | 0.83 | 0.96 | 1.06 | 1.06 | **0.72*** | **0.56*** | **0.59*** | **0.61*** |
| OECD | HP filtered | Industrial production | OECD + 6 | 0.88 | 1.01 | 1.15 | 1.19 | 0.68* | 0.55* | 0.49 | 0.47 |
| OECD | Linearly detrended | Industrial production | OECD + 6 | 0.83 | 1.00 | 1.10 | 1.10 | 0.71* | 0.60* | 0.59 | 0.56 |
| WTI price | | | | | | | | | | | |
| Kilian (2009) | - | Shipping index | World | **0.93** | **0.97** | **1.02** | **1.01** | 0.65* | 0.61* | 0.51 | 0.56 |
| OECD | Growth rate | Industrial production | OECD + 6 | 0.93 | 1.00 | 1.05 | 1.03 | **0.67*** | **0.58*** | **0.58*** | **0.63*** |
| OECD | HP filtered | Industrial production | OECD + 6 | 0.94 | 1.01 | 1.10 | 1.11 | 0.71* | 0.55** | 0.53 | 0.51 |
| OECD | Linearly detrended | Industrial production | OECD + 6 | 0.96 | 1.03 | 1.09 | 1.05 | 0.64* | 0.58* | 0.57 | 0.57 |
| Brent price | | | | | | | | | | | |
| Kilian (2009) | - | Shipping index | World | **0.92** | **0.98** | **1.01** | **1.01** | 0.72* | 0.61* | 0.51 | 0.60* |
| OECD | Growth rate | Industrial production | OECD + 6 | 1.01 | 1.07 | 1.11 | 1.09 | **0.64*** | **0.62*** | **0.59*** | **0.63*** |
| OECD | HP filtered | Industrial production | OECD + 6 | 1.03 | 1.10 | 1.17 | 1.21 | 0.65* | 0.56* | 0.50 | 0.52** |
| OECD | Linearly detrended | Industrial production | OECD + 6 | 1.08 | 1.13 | 1.16 | 1.14 | 0.67* | 0.60** | 0.61** | 0.57 |

NOTES: All MSPE ratios have been normalized relative to the monthly no-change forecast. For each oil price series the measure with the lowest average MSPE and the measure with the highest average success ratio is shown in bold. The penalty parameter for the one-sided HP filter was set to 129,600 for all monthly data following Ravn and Uhlig (2002). Statistically significant success ratios at the 5% (10%) level are marked * (**).

directional accuracy than the specification involving the shipping index. This gain in directional accuracy comes at the cost of a higher MSPE, however.

We also examined this same question for the quarterly VAR specification, where, in addition to the quarterly version of the industrial production index for the OECD+6 economies and a quarterly version of the Kilian shipping index, we employed the OECD estimate of quarterly real GDP in the OECD economies.[14] Extensive comparisons of a wide range of models showed that the quarterly version of the shipping index for all three oil price measures tends to yield lower MSPEs and higher directional accuracy than all other real activity measures. None of these quarterly models, however, comes close to matching the accuracy of the monthly VAR models.

3.7. *Does Allowing for Time Variation Help in Forecasting the Quarterly Price of Oil?* There are many economic reasons to expect linear VAR models of the global market for oil to be at best an approximation to a more general TVP-VAR model. For example, capacity constraints in oil production and in oil inventory holdings, delays in oil production responses to investment decisions, and changes in energy intensity over time may cause the dynamic relationship between oil market variables to evolve over time. Indeed, TVP-VAR models have been used to describe the evolution of the global market for crude oil, although those models have been smaller in dimension and simpler than the VAR model underlying our analysis (see, e.g., Baumeister and Peersman, 2013).

The fact that TVP-VAR models seem plausible ex ante does not necessarily mean that TVP-VAR models should replace VAR models in forecasting the real price of oil, however. An obvious concern is that in practice TVP-VAR models may overparameterize the data, resulting in poor out-of-sample accuracy. Although several recent studies have reported that TVP-VAR model forecasts may improve on VAR forecasts of quarterly macroeconomic aggregates (see, e.g., Cogley et al., 2005; D'Agostino et al., 2013), the usefulness of the TVP-VAR model for forecasting the real price of oil has yet to be explored.

One immediate problem with the TVP-VAR approach is that estimation of such models requires computationally intensive nonlinear estimation methods that prohibit applications to models with a large number of variables and/or autoregressive lags. For example, it is not possible in practice to estimate monthly VAR oil market models with 12 or more lags allowing for time variation in the parameters. This means that at best some of the quarterly models of oil markets may be estimated allowing for time variation in the parameters. For expository purposes, we focus on the TVP-VAR(4) model:

$$y_t^Q = \nu_t + B_{1,t}y_{t-1}^Q + \cdots + B_{4,t}y_{t-4}^Q + u_t^Q, \quad u_t^Q \sim N(0, \Omega_t),$$

where $y_t^Q$ is a $4 \times 1$ vector of quarterly model variables containing the growth rate of global crude oil production, the Kilian (2009) real activity index, the real U.S. refiners' acquisition cost of crude oil imports, and the change in global crude oil inventories. Here, $\nu_t$ refers to the vector of time-varying intercepts, and $B_{i,t}$, $i = 1, \ldots, 4$, denotes $4 \times 4$-dimensional matrices of time-varying slope parameters. Given that each model parameter is allowed to evolve according to a random walk, four lags allow for considerable flexibility in fitting the data. The model also allows for time variation in the variance–covariance matrix of the innovations, $\Omega_t$.

The TVP-VAR model is reestimated recursively in real time using Bayesian techniques, as described in Kim and Nelson (1999). For further details, the reader is referred to the technical appendix in Baumeister and Peersman (2013).[15] There are no closed-form

[14] In related work, Baumeister and Peersman (2013) used a measure of quarterly world industrial production from the *United Nations Monthly Bulletin of Statistics*. Although this index is available in real time, it has not been updated since 2008Q3 pending the implementation of ISIC Revision 4. For that reason, this index was not included in our comparison.

[15] We do not refer to this model as a TVP-BVAR model because the priors for this model are diffuse, in contrast to the informative priors developed for the VAR model by Giannone et al. (2012) that we used for our BVAR models.

TABLE 5
REAL-TIME ACCURACY OF RECURSIVE FORECASTS OF THE QUARTERLY REAL U.S. REFINERS' ACQUISITION COST FOR OIL IMPORTS FROM A QUARTERLY TVP-VAR(4) MODEL

| Quarterly Horizon | Posterior Mean | Posterior Trimmed Mean | Posterior Median |
|---|---|---|---|
| MSPE ratio | | | |
| 1 | 1.45 | 1.48 | 1.48 |
| 2 | 1.20 | 1.23 | 1.26 |
| 3 | 1.18 | 1.19 | 1.20 |
| 4 | 1.55 | 1.21 | 1.23 |
| Success ratio | | | |
| 1 | **0.58** | **0.58** | **0.62**** |
| 2 | **0.65**** | **0.61****** | **0.60**** |
| 3 | **0.62** | **0.62** | **0.55** |
| 4 | **0.64****** | **0.63** | **0.56** |

NOTES: All results are obtained by Monte Carlo integration from the pointwise posterior distribution of the TVP-VAR model forecasts. The trimmed mean eliminates the top and bottom 0.5% of the posterior forecasts.

solutions for the forecasts of the real price of oil in the TVP-VAR model. Instead, forecasts are obtained by simulating the posterior predictive density. In practice, at each point in time forecasts are obtained by first randomly drawing 5,000 starting values for the model parameters and simulating for each starting value the future path of the model parameters based on their law of motion. For each such path, we then simulate the evolution of the quarterly price of oil, conditional on the four most recent data points, by randomly drawing from the time-varying distribution of the error terms. This results in 5,000 future paths of the quarterly real price of oil up to the maximum horizon of interest. For each forecast horizon, we compute several summary measures of the central tendency of the simulated forecasts.

A natural starting point is the posterior mean of the forecasts obtained by Monte Carlo integration. Table 5 shows that the TVP-VAR model has somewhat lower MSPE at the one-quarter horizon than both the quarterly VAR(4) model and the quarterly BVAR(6) model, which we showed to be most accurate among all quarterly VAR and BVAR models. This ranking is reversed at horizons of two, three, and four quarters, however. More importantly, from the point of view of applied users, the TVP-VAR(4) model has much higher MSPE than the monthly VAR(12) model at all horizons. For example, its one-quarter-ahead MSPE ratio is 1.45 compared with 0.80 for the linear monthly VAR model. Nor does the TVP-VAR model systematically improve on the directional accuracy of the monthly VAR(12) forecast. Although the TVP-VAR(4) model provides some gains in directional accuracy at longer horizons, its directional accuracy is much reduced at the one-quarter horizon.

An obvious concern is that TVP-VAR(4) forecasts obtained by Monte Carlo integration may be sensitive to outliers. For comparison, we also report forecasts based on the trimmed posterior mean and the posterior median at each horizon. Table 5 shows that controlling for outliers does not alter the substance of the earlier results for the posterior mean forecast. We conclude that allowing for time variation in the VAR parameters does not improve the accuracy of the forecasts of the quarterly real price of oil. Quarterly TVP-VAR models based on conventional prior specifications tend to be much less accurate than standard monthly VAR models. The remainder of the article therefore focuses on linear monthly VAR forecasting models.

Although the common notion that rolling VAR forecasts protect the forecaster against future structural changes is not correct, as demonstrated in Inoue and Kilian (2006), it is useful to investigate whether rolling VAR forecasts would have improved on the MSPE of recursive VAR forecasts. Table 6 demonstrates that—regardless of the length of the rolling window—rolling VAR forecasts are less accurate than recursive VAR forecasts. This reflects the loss in accuracy from estimating high-dimensional models on short windows of data.

TABLE 6
REAL-TIME ACCURACY OF ROLLING FORECASTS OF THE QUARTERLY REAL U.S. REFINERS' ACQUISITION COST FOR OIL IMPORTS: SELECTED VAR MODELS

| | | VAR($p$) | | BVAR($p$) | |
|---|---|---|---|---|---|
| Quarterly Horizon | Monthly VAR(12) | $p=4$ | $p=6$ | $p=4$ | $p=6$ |
| MSPE ratio: 15-year rolling window | | | | | |
| 1 | **0.98** | 2.05 | 2.96 | 1.87 | 1.82 |
| 2 | 1.16 | 1.47 | 1.64 | 1.39 | 1.32 |
| 3 | 1.28 | 1.23 | 1.29 | 1.34 | 1.28 |
| 4 | 1.35 | 1.18 | 1.40 | 1.44 | 1.40 |
| MSPE ratio: 10-year rolling window | | | | | |
| 1 | 1.22 | 2.30 | 5.17 | 1.93 | 1.89 |
| 2 | 1.33 | 1.57 | 3.72 | 1.37 | 1.34 |
| 3 | 1.45 | 1.26 | 1.81 | 1.26 | 1.31 |
| 4 | 1.52 | 1.22 | 2.25 | 1.34 | 1.48 |

NOTES: All MSPE ratios have been normalized relative to the monthly no-change forecast. Boldface indicates an improvement on the monthly no-change forecast. The VAR(8) and BVAR(8) forecasts are omitted given their poor accuracy.

## 4. FORECASTING THE QUARTERLY REAL PRICE OF OIL IN FOREIGN CONSUMPTION UNITS

Oil in global markets is predominantly traded in dollars. When forecasting the real price of oil it is common to deflate the nominal price of oil by the U.S. CPI. This allows one to measure the real cost of purchases of oil in terms of U.S. consumption goods. This practice is perfectly adequate for the Federal Reserve Board, but central banks in other countries are concerned with the real cost of purchasing crude oil in terms of their domestic basket of consumption goods. Moreover, the U.S. benchmarks for crude oil such as the U.S. refiners' acquisition cost or WTI need not be representative for other countries. For example, the European Central Bank views the price of Brent crude oil as the relevant benchmark, as does the Norges Bank in Norway.

In this section, we explore how to adapt our analysis to generate forecasts of the real price of oil of the type required by the Bank of Canada, the Norges Bank, and the European Central Bank, as three representative examples. Whereas the latter two central banks focus on the real price of Brent crude oil, the Bank of Canada has traditionally focused on the real price of WTI crude oil. For Canada, we take as our starting point the best forecasting model for the price of WTI in terms of U.S. consumption goods. Table 2 suggests that we take as our starting point the monthly VAR(12) model for the U.S. refiners' acquisition cost for crude oil imports. That forecast may be converted to the WTI benchmark using a simple no-change forecast of the spread of the WTI price over the U.S. refiners' acquisition cost for crude oil imports. This leaves the conversion from U.S. consumption goods to Canadian consumption goods, which requires a forecast of the Canadian–U.S. real exchange rate. There are two ways of generating a forecast for this real exchange rate. One approach is to rely on a simple no-change forecast of the real exchange rate, given that fluctuations in the real exchange rate are dominated by fluctuations in the nominal exchange rate.[16] The advantage of this approach is that we can rely on the same forecasting model we already showed to work well for the United States.

The other approach is to augment the original global oil market VAR model by the real exchange rate, resulting in the five-variable VAR(12) model. Which approach is more accurate will depend in part on how close the real exchange rate is to a random walk. It also will depend on how much the estimation of the additional model parameters will inflate the MSPE of the model in finite samples. The entries on the right of panel (a) in Table 7 show that the approach of treating the real exchange rate as a random walk works quite well. There is some increase

[16] In practice, we found the last monthly observation of the most recent quarter to be a better predictor than the most recent quarterly average real exchange rate. Only the former results are reported.

TABLE 7

INTERNATIONAL COMPARISON OF THE REAL-TIME ACCURACY OF QUARTERLY FORECASTS OF THE REAL PRICE OF OIL IN DOMESTIC CONSUMPTION UNITS

| Quarterly Horizon | Real Exchange Rate Included in Baseline Monthly VAR(12) Model for RAC and No-Change Forecast of the Spread of the Benchmark Price over the RAC | | Baseline Monthly VAR(12) Model for RAC with No-Change Forecasts of the Real Exchange Rate and of the Spread of the Benchmark Price over the RAC | |
|---|---|---|---|---|
| | MSPE Ratio | Success Ratio | MSPE Ratio | Success Ratio |
| (a) Canada: WTI benchmark | | | | |
| 1 | **0.84** | **0.62*** | **0.93** | **0.73*** |
| 2 | **0.96** | **0.48** | **0.97** | **0.60*** |
| 3 | 1.04 | **0.50** | 1.02 | **0.54** |
| 4 | 1.03 | **0.47** | **1.00** | **0.55** |
| (b) Norway: Brent benchmark | | | | |
| 1 | **0.92** | **0.60*** | **0.90** | **0.65*** |
| 2 | 1.07 | **0.58*** | **0.98** | **0.61*** |
| 3 | 1.15 | **0.53** | 1.07 | **0.55** |
| 4 | 1.15 | **0.53** | 1.05 | **0.60*** |
| (c) Euro area: Brent benchmark | | | | |
| 1 | **0.96** | **0.69*** | **0.90** | **0.68*** |
| 2 | 1.08 | **0.60*** | 1.01 | **0.61*** |
| 3 | 1.17 | **0.57**** | 1.08 | **0.54*** |
| 4 | 1.17 | **0.61*** | 1.06 | **0.59*** |

NOTES: Boldface indicates an improvement on the monthly no-change forecast. For the success ratio, improvements that are statistically significant at the 5% (10%) level are marked * (**). For the real-time recursive MSPE ratio, the degree of statistical significance cannot be reported because no valid statistical tests are available in the literature.

in the one-quarter ahead and two-quarter ahead MSPE ratios compared to the corresponding result in Table 2, but that is to be expected, for now we also must account for uncertainty in the real exchange rate. Nevertheless, Table 7 shows some reductions in the MSPE relative to the monthly no-change forecast one quarter and two quarters ahead. Three and four quarters ahead, the MSPE ratios are about as high as that of the monthly no-change forecast. The directional accuracy of the model actually improves from 69% to 73% at the one-quarter horizon and is statistically significant at the 5% level. At higher horizons, the observed gains in directional accuracy are somewhat lower than for the U.S. model in Table 2 and not statistically significant.

Including the real exchange rate in the VAR model, as shown on the left in panel (a) of Table 7, produces MSPE ratios more in line with Table 2, but at the cost of losing all directional accuracy beyond the first quarter. Likewise, VAR models that replace the real U.S. refiners' acquisition cost for imports with the real WTI price of oil (instead of forecasting the price spread) proved inferior regardless of how the real exchange rate is modeled. The latter results are not shown to conserve space.

Much the same approach also works for Norway (panel b) and the Euro area (panel c) with the difference that now we are relying on the no-change forecast of the spread of the Brent price over the U.S. refiners' acquisition cost. The latter approach also avoids having to backcast the Euro area exchange rate further than the early 1990s. The results in Table 7 are quite similar to those shown for the United States in Table 3. For Norway and the Euro area, the reduction in the MSPE ratio is 10% one quarter ahead. Two quarters ahead, the two models' MSPEs are essentially tied with the monthly no-change forecast, and three and four quarters ahead their MSPE is somewhat higher. On the other hand, there is strong and mostly statistically significant evidence of directional accuracy, even four quarters ahead. For example, the success ratio for the Euro area is 68% one quarter ahead and 59% four quarters ahead.

Even granting that the MSPE reductions for Canada, Norway, and the Euro area are somewhat less pronounced than when forecasting the corresponding real price of oil in U.S. consumption units, these forecasts remain useful for central banks given their directional accuracy.

TABLE 8
REAL-TIME ACCURACY OF FORECAST COMBINATIONS OF THE MONTHLY VAR(12) MODEL FORECAST AND THE FORECAST BASED ON OIL FUTURES

| | U.S. Refiners' Acquisition Cost for Crude Oil Imports | | WTI Price | | Brent Price | |
|---|---|---|---|---|---|---|
| Quarterly Horizon | Equal Weights | Inverse MSPE Weights | Equal Weights | Inverse MSE Weights | Equal Weights | Inverse MSPE Weights |
| MSPE ratio | | | | | | |
| 1 | **0.83** | **0.91** | **0.85** | **0.92** | **0.92** | **1.01** |
| 2 | **0.92** | **0.92** | **0.93** | **0.93** | **0.97** | **0.99** |
| 3 | **0.95** | **0.96** | **0.95** | **0.96** | **0.98** | **0.99** |
| 4 | **0.92** | **0.93** | **0.92** | **0.93** | - | - |
| Success ratio | | | | | | |
| 1 | **0.71*** | **0.71*** | **0.68*** | **0.68*** | **0.67*** | **0.69*** |
| 2 | **0.53** | **0.56*** | **0.56*** | **0.58*** | **0.58*** | **0.53** |
| 3 | **0.57**** | **0.53** | **0.58*** | **0.57**** | **0.53** | **0.53** |
| 4 | **0.53*** | **0.53*** | **0.53*** | **0.53**** | - | - |

NOTES: See notes to Table 7. The VAR forecasts for the real WTI price and real Brent price are obtained from the baseline model for the U.S. refiners' acquisition cost by applying the most recent price spread. Brent futures prices with a maturity of 10–12 months are not available for our evaluation period. The inverse MSE weights are computed as $MSPE_{i,t,h}^{-1}/\sum_{j=1}^{2} MSPE_{j,t,h}^{-1}$, where $MSPE_{i,t,h}$ denotes the recursive MSPE of model $i$at horizon $h$ in period $t$. The procedure is initialized with a weight of 0.5.

## 5. FORECAST AVERAGING

Even if one forecasting model is more accurate than the other, it may still be possible to improve on the forecast accuracy of the more accurate model by taking a weighted average of the two forecasts. Forecast combinations may also help guard against the misspecification of forecasting models due to structural change. A question of obvious interest to central bankers therefore is whether they should—in light of our earlier results—abandon conventional forecasts of the real price of oil based on oil futures prices in favor of VAR forecasts or instead rely on a weighted average of these two forecasts. In response to this question, we now explore the possibility that a linear combination of oil price forecasts is more accurate than any one of the forecasting models alone. Table 8 presents results for several oil price measures. We first focus on the equal-weighted combination of the forecast from the best-performing monthly VAR(12) model and the forecast based on oil futures prices.[17] Table 8 shows that forecast averaging may indeed reduce the MSPE of the forecast of the U.S. real refiners' acquisition cost, especially at longer horizons. Qualitatively similar results hold for the real WTI price and the real Brent price. We also experimented with inverse MSPE weights based on recursive MSPEs of each model, as discussed in Stock and Watson (2004). Table 8 shows that allowing the forecast combination weights to evolve over time produces results broadly similar to those based on equal weights.

## 6. EXTENSIONS

6.1. *Longer Forecast Horizons.* Many central banks are interested in forecasts of the quarterly real price of oil at horizons as high as two years. Our focus so far has been on forecasts at horizons of up to four quarters only. Indeed, one would not expect the forecasting methods we considered to be more accurate than a no-change forecast at longer horizons. Further extensive analysis (not shown to conserve space) indicates that, at the two-year horizon, the monthly VAR

[17] Other forecast combinations (not shown to conserve space) only proved less accurate. For example, combining the quarterly TVP-VAR model with the monthly VAR(12) model systematically worsened the forecast accuracy compared with the best-performing VAR(12) model. Moreover, combinations involving additional forecast methods in various combinations had higher MSPE than the combination shown in Table 8.

TABLE 9
REAL-TIME ACCURACY OF SELECTED FORECASTS AT LONGER HORIZONS: U.S. REFINERS' ACQUISITION COST FOR OIL IMPORTS

| Quarterly Horizon | Monthly VAR(12) | Hybrid Method | Quarterly No-Change Forecast |
|---|---|---|---|
| MSPE ratio | | | |
| 5 | 1.06 | 1.07 | **0.97** |
| 6 | 1.12 | 1.13 | **0.95** |
| 7 | 1.15 | 1.13 | **0.95** |
| 8 | 1.14 | 1.07 | **0.97** |
| Success ratio | | | |
| 5 | **0.58*** | **0.54** | - |
| 6 | **0.52** | 0.47 | - |
| 7 | **0.50** | 0.49 | - |
| 8 | **0.52** | **0.52** | - |

NOTES: The hybrid method treats the four-quarter forecast from the monthly VAR(12) model as the forecast for horizons 5–8. Boldface indicates an improvement on the monthly no-change forecast.

forecast of the quarterly real price of oil has an MSPE ratio of about 1.1 relative to the monthly no-change forecast. Although forecasts from the monthly VAR model retain some directional accuracy at these longer horizons, the success ratios are at most 0.58 and never statistically significant beyond the five-quarter horizon. On the other hand, the quarterly no-change forecast yields MSPE reductions between 2% and 6% relative to the monthly no-change forecast. This finding suggests that we replace the monthly VAR forecast by the quarterly no-change forecast at horizons beyond one year.

The latter proposal is not without limitations, however. In policy settings, the path of oil prices receives much attention. One potential concern is that replacing the monthly VAR model forecasts at horizons beyond four quarters by the quarterly no-change forecast may introduce a discontinuity in the forecast path between the four- and five-quarter horizons. One way of addressing this concern is to treat the monthly VAR model forecast for the fourth-quarter horizon as the forecast for the remaining quarters, which ensures a smooth forecast path. Table 9 shows that this alternative proposal results in similar forecast accuracy at horizons beyond one year than the baseline monthly VAR model for the U.S. refiners' acquisition cost for crude oil imports. In other words, the cost of insisting on a smooth forecast path is an increase in the MSPE of between 10% and 15%.

6.2. *The Link between Forecasts of the Real Price of Oil and the Nominal Price of Oil.* Based on economic theory, a model involving only real variables is the natural framework in which to forecast the real price of oil. One potential concern is that our forecasting success for the real price of oil may simply reflect the monthly VAR model's ability to forecast inflation at short horizons. It can be shown that this is not the case (see Alquist et al., 2013). Indeed, this point is quite obvious because much of the variability in the nominal price of oil stems from variation in the real price of oil instead of variation in inflation. Nor is it necessary to develop a separate forecasting model for the nominal price of oil. Given the CPI forecasts routinely generated by central banks, it is straightforward to generate the implied forecast path for the nominal price of oil from our VAR forecasts.

## 7. CONCLUSION

Central banks rely on forecasts of the real price of oil when making policy decisions. We provided strong evidence that the U.S. real price of oil may be forecast several quarters ahead, provided suitable forecasting methods are employed. For monthly VAR(12) models containing data on global crude oil production, global real economic activity, global crude oil inventories, and the U.S. real price of oil, for example, we obtained real-time reductions in the MSPE between 7% and 20% at the one-quarter horizon and between 2% and 7% at the two-quarter horizon, depending on the choice of the oil price series. At longer horizons, the MSPEs of these models are similar to those of the monthly no-change forecast. These models are much more accurate at the one-quarter and two-quarter horizons than conventional central bank forecasts based on oil futures prices. In addition, the same monthly VAR(12) forecasting models yield strong and often statistically significant gains in directional accuracy. For example, at the one-quarter horizon the success ratio ranges from 69% to 72%, and even at the four-quarter horizon the success ratio is between 56% and 60%. Such directional accuracy is high by the standards of the empirical finance literature (see, e.g., Pesaran and Timmermann, 1995). It is also higher than the directional accuracy of forecasts based on oil futures spreads, especially at short horizons. These accuracy gains of the VAR forecasting models reflect predictable variation in economic fundamentals over our evaluation period. In the absence of persistent changes in economic fundamentals, the VAR forecast would still be expected to be about as accurate as but no more accurate than the no-change forecast, as illustrated in Baumeister and Kilian (2012).

Central banks outside the United States face the more complicated problem of forecasting the quarterly real price of oil in terms of domestic consumption units. This involves a forecast of the real exchange rate in addition, making it even more challenging to generate accurate real-time forecasts. For the examples of Canada, Norway, and the Euro area, we compared several alternative forecasting approaches and showed that even for those countries real-time MSPE reductions between 7% and 10% are possible at the one-quarter horizon and up to 3% at the two-quarter horizon. At longer horizons, the MSPEs of the most accurate forecasting models are about as good as the monthly no-change forecast in the case of Canada and slightly worse in the case of Norway and of the Euro area. Although the MSPE reductions are at best modest, the directional accuracy of these forecasts remains consistently high in all three cases. It ranges from 65% to 73% at the one-quarter horizon, for example, and lies between 55% and 60% at the four-quarter horizon.

Much of our analysis focused on comparing alternative approaches to forecasting the quarterly real price of oil. The forecasting methods we discussed in this article were specifically designed to reflect the forecasting environment faced by central bankers. First, we demonstrated that, among the no-change forecasts, a forecast based on the most recent monthly observation is far more accurate in the short run than the no-change forecast based on the most recent quarterly average. At longer horizons, there is little to choose between these methods. Thus, the choice of the random walk benchmark is crucial, and choosing an inferior benchmark can make alternative forecasting models look spuriously accurate.

Second, we showed that VAR forecasting models based on monthly data are far more accurate in all dimensions than the corresponding VAR forecasting models estimated on quarterly data. This result is robust to various changes in the specification of the quarterly model and in the estimation method. Third, when modeling the WTI price or the Brent price, working with a baseline VAR model for the U.S. real refiners' acquisition cost for crude oil imports and treating the spread of the WTI and Brent prices over the U.S. refiners' acquisition cost for crude oil imports as a random walk without drift yields more accurate forecasts than replacing the real oil price series in the original VAR model. Likewise, relying on a no-change forecast of the real exchange rate in conjunction with the original VAR forecasting model for the U.S. real price of oil is more accurate than augmenting the original VAR model by the real exchange rate. This is true even when using Bayesian estimation methods. Fourth, we provided evidence that the shipping index of global economic activity proposed by Kilian (2009) indeed lowers the MSPE of VAR forecasting models of the quarterly real price of oil compared with alternative measures, including world real GDP or industrial production for the OECD and six emerging economies. This result does not depend on how these alternative measures of real activity are transformed during the analysis. Although the Kilian (2009) global shipping index also has high directional accuracy, specifications based on the growth rate of OECD+6 industrial production in some cases yield even higher directional accuracy by a small margin, but at the expense of a lower MSPE.

Fifth, we found that allowing for time variation in the VAR parameters does not improve forecast accuracy. Finally, we demonstrated that forecast combinations do not systematically improve the accuracy of the forecast of the quarterly real price of oil and may worsen it. The most interesting finding was that suitably weighted combinations of the VAR(12) model forecast and the forecast based on oil futures prices tend to have lower MSPE than the VAR(12) model alone at longer horizons, but not at short horizons.

There may be alternative forecasting methods that could further improve the accuracy of short-horizon forecasts, however. One question left unexamined in this article is whether factor-augmented VAR forecasting models—or alternatively large-scale Bayesian VAR forecasting models of the type discussed in Banbura et al. (2010)—would be able to improve on existing VAR forecasting models of the quarterly real price of oil. One problem with the use of such large-scale models is the difficulty of obtaining suitable real-time data. Another possible extension would involve the use of mixed-frequency forecasting methods in the tradition of the Mixed Data Sampling (MIDAS) model or mixed-frequency VAR models (see Schorfheide and Song, 2011; Andreou et al., forthcoming).

## REFERENCES


ALQUIST, R., AND L. KILIAN, "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics* 25 (2010), 539–73.

———, ———, AND R. J. VIGFUSSON, "Forecasting the Price of Oil," in G. Elliott and A. Timmermann, eds., *Handbook of Economic Forecasting*, Volume 2 (Amsterdam: North-Holland, 2013) 427–507.

ANDREOU, E., E. GHYSELS, AND A. KOURTELLOS, "Forecasting with Mixed Frequency Data," in M. P. Clements and D. F. Hendry, eds., *Oxford Handbook of Economic Forecasting*, forthcoming.

BANBURA, M., D. GIANNONE, AND L. REICHLIN, "Large Bayesian VARs," *Journal of Applied Econometrics* 25 (2010), 71–92.

BAUMEISTER, C., AND L. KILIAN, "Real-Time Forecasts of the Real Price of Oil," *Journal of Business and Economic Statistics* 30 (2012), 326–36.

———, AND ———, "Real-Time Analysis of Oil Price Risks Using Forecast Scenarios," *IMF Economic Review*, 62 (2014), 119–45.

———, AND G. PEERSMAN, "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market," *Journal of Applied Econometrics*, 28 (2013), 1087–109.

BERNANKE, B. S., "Irreversibility, Uncertainty, and Cyclical Investment," *Quarterly Journal of Economics* 98 (1983), 85–106.

BUSSE, M., C. KNITTEL, AND F. ZETTELMEYER, "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review* 103 (2013), 220–56.

CLARK, T. E., AND M. W. MCCRACKEN, "Tests of Equal Predictive Ability with Real-Time Data," *Journal of Business and Economic Statistics* 27 (2009), 441–54.

———, AND ———, "Nested Forecast Model Comparisons: A New Approach to Testing Equal Accuracy," mimeo, Federal Reserve Bank of St. Louis, 2012.

COGLEY, T., S. MOROZOV, AND T. J. SARGENT, "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System," *Journal of Economic Dynamics and Control* 29 (2005), 1893–925.

CROUSHORE, D., "Frontiers of Real-Time Data Analysis," *Journal of Economic Literature* 49 (2011), 72–100.

D'AGOSTINO, A., L. GAMBETTI, AND D. GIANNONE, "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics* 28 (2013), 82–101.

DIEBOLD, F. X., AND R. S. MARIANO, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13 (1995), 253–63.

GIANNONE, D., J. HENRY, M. LALIK, AND M. MODUGNO, "An Area-Wide Real-Time Database for the Euro Area," *Review of Economics and Statistics* 94 (2012), 1000–13.

———, M. LENZA, AND G. PRIMICERI, "Prior Selection for Vector Autoregressions," mimeo, Department of Economics, Free University of Brussels, 2012.

INOUE, A., AND L. KILIAN, "On the Selection of Forecasting Models," *Journal of Econometrics* 130 (2006), 273–306.

KILIAN, L., "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review* 99 (2009), 1053–69.

———, AND D. P. MURPHY, "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," *Journal of Applied Econometrics*, 29 (2014), 454–78.

KIM, C. J., AND C. R. NELSON, *State Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications* (Cambridge, MA: MIT Press, 1999).

PESARAN, M. H., AND A. TIMMERMANN, "Predictability of Stock Returns: Robustness and Economic Significance," *Journal of Finance* 50 (1995), 1201–28.

———, AND ———, "Testing Dependence among Serially Correlated Multicategory Variables," *Journal of the American Statistical Association* 104 (2009), 325–37.

RAVN, M. O., AND H. UHLIG, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics* 84 (2002), 371–80.

SCHORFHEIDE, F., AND D. SONG, "Real-Time Forecasting with a Mixed Frequency VAR," mimeo, University of Pennsylvania, 2011.

STOCK, J. H., AND M. W. WATSON, "Forecasting Inflation," *Journal of Monetary Economics* 44 (1999), 293–335.

———, AND ———, "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature* 41 (2003), 788–829.

———, AND ———, "Combination Forecasts of Output Growth in a Seven-Country Data Set," *Journal of Forecasting* 23 (2004), 405–30.