Energy Markets and Global Economic Conditions*

Christiane Baumeister
University of Notre Dame
NBER and CEPR

Dimitris Korobilis
University of Glasgow

Thomas K. Lee
U.S. Energy Information Administration

June 22, 2020

Abstract

This paper evaluates alternative indicators of global economic activity and other market fundamentals in terms of their usefulness for forecasting real oil prices and global petroleum consumption. We find that world industrial production is one of the most useful indicators that has been proposed in the literature. However, by combining measures from a number of different sources we can do even better. Our analysis results in a new index of global economic conditions and new measures for assessing future tightness of energy demand and expected oil price pressures. We illustrate their usefulness for quantifying the main factors behind the severe contraction of the global economy and the price risks faced by shale oil producers in early 2020.

*We thank Scott Brave, Nida Çakır Melek, Todd Clark, Olivier Coibion, Ferre De Graeve, Pierre Guérin, James Hamilton, Francesco Ravazzolo, James Stock, and anonymous referees for helpful comments. Special thanks to James Ng and Mark Robison at the Hesburgh Library for their valuable assistance with obtaining access to the shipping data. Aram Derdzyan and Qian Li provided excellent research assistance. The work of Christiane Baumeister was supported by grant DE-EI0003240 from the U.S. Energy Information Administration. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the U.S. Energy Information Administration. Corresponding author: Christiane Baumeister, Department of Economics, 3060 Jenkins Nanovic Hall, Notre Dame, IN 46556, email: cjsbaumeister@gmail.com

1 Introduction

What are the key drivers of world energy markets? This question is of vital interest not just to academic researchers but also to business and government planners around the globe, especially in light of the recent Covid-19 fueled turmoil in energy markets. Financial analysts, investment banks, energy companies, budget agencies, central banks, and organizations like the International Monetary Fund (IMF), the International Energy Agency (IEA), and the U.S. Energy Information Administration (EIA) devote a considerable amount of resources in an effort to assess the current and future outlook for production, consumption, and prices of major sources of energy.

A large academic literature has sought to contribute to these efforts by developing models of energy market dynamics that generate usable forecasts of energy prices. Prominent contributions include Alquist et al. (2013), Alquist et al. (2020), Baumeister and Kilian (2014a, 2015), Baumeister et al. (2017), Bernard et al. (2018), Ferrari et al. (2019), and Manescu and van Robays (2016). This literature has concluded that although a random walk is hard to beat in out-of-sample oil-price forecasting exercises, careful attention to the economic fundamentals that are driving energy markets can lead to practical improvements in forecasts.

A key step in this effort is to find a useful summary of the global economic conditions that influence energy demand.[1] One of the promising early proposals for this purpose was a measure of dry-cargo shipping rates developed by Kilian (2009). This measure is available monthly in real time, is forward looking, and was found by Alquist et al. (2013) and Baumeister and Kilian (2012) to produce promising forecasts of the U.S. refiner acquisition cost (RAC) of crude oil imports. However, since these studies were published, there has been tremendous turbulence in the shipping index that does not seem to reflect changes in world economic activity. Although Kilian and Zhou (2018) have tried to defend continuing the use of shipping costs in modeling commodity price dynamics, they

[1] Studies stressing the importance of the measure of global economic conditions used include Alquist et al. (2019), Baumeister and Kilian (2014a), Delle Chiaie et al. (2017), Ferrari et al. (2019), and Manescu and van Robays (2016).

do not provide any statistical evidence or formal criteria in support of that conclusion. A growing number of researchers are suggesting alternative measures based on world industrial production (Baumeister and Hamilton, 2019; Hamilton, 2019), commodity prices more broadly (Alquist et al., 2020; Delle Chiaie et al., 2017; West and Wong, 2014), or global steel production (Ravazzolo and Vespignani, 2019).

One of the objectives of our paper is to revisit this evidence using updated data and to compare the measures that have been proposed by other researchers with those developed from a broad set of observations on global variables that we assembled for the purpose of this study. We begin by reproducing the success of early models at forecasting real RAC over the period 1992-2010, but document how these break down badly for subsequent data. We note that they perform even more poorly for forecasting alternative measures of oil prices such as Brent. We find that models based on alternative measures of global economic conditions such as world industrial production or a common factor extracted from a panel of real commodity prices lead to substantially better forecasts, even for forecasting RAC over the original sample period.

We also examine the usefulness of Bayesian shrinkage priors, allowing for time-varying volatility, and pooling multiple sources of information. We find that Bayesian shrinkage is beneficial in every specification, and introducing stochastic volatility improves long-horizon forecasts substantially. Although many researchers have reported that pooling leads to superior forecasts in a number of different settings, we do not find evidence of that here.

One of the features that distinguishes our effort from earlier studies is that we investigate potential measures of energy market fundamentals not only in terms of their ability to predict prices but also to predict changes in world oil consumption. We find that none of our baseline specifications perform well for forecasting global petroleum consumption. We investigate alternative measures that add additional determinants of energy demand including measures of geopolitical risk, developments in transportation, oil price uncertainty, and weather-related indicators. We find that constructing an indicator of global economic conditions using these variables along with world industrial production helps improve the forecast accuracy for petroleum consumption considerably.

Our analysis results in some new measures for characterizing energy demand and quantifying oil price risks. We use real-time joint density forecasts obtained from our model of oil prices and petroleum consumption to construct an energy demand indicator that signals market tightness and anticipated future demand pressures. We complement this analysis with measures that signal the likelihood of a build-up of upward or downward oil price pressures relative to the recent past and forecast the probabilities that the oil price will remain within the range of values experienced recently over a two-year horizon. Our analysis suggests that these measures may be very helpful for analysts who are trying to assess the implications of current developments in energy markets for purposes of making budgeting and planning decisions. We illustrate their usefulness for risk assessment from the perspective of shale oil producers in the turbulent environment of early 2020.

The remainder of the paper is structured as follows. Section 2 provides a systematic comparison of the usefulness of alternative indicators of global economic activity based on their forecasting performance for the real price of oil. This evaluation takes place within existing models of the global oil market as well as a new set of forecasting models that focus more on the demand side of the market and allow for time variation in volatilities. Section 3 extends the analysis to global petroleum consumption to gain a more complete understanding of future developments in energy markets. Section 4 proposes a new indicator of global economic conditions that covers a diverse range of variables tied to future energy demand. Section 5 illustrates how price and consumption forecasts can be used to gauge the current and expected state of energy markets by introducing some new real-time monitoring tools. Specifically, we develop measures that provide policymakers and markets with a quantitative assessment of future energy demand conditions and expected oil price risks. Section 6 offers some concluding remarks.

2 Forecasting Oil Prices

2.1 Forecasting with Vector Autoregressive Models

A widely followed approach to forecasting oil prices is the dynamic model of the global oil market proposed by Alquist et al. (2013) and Baumeister and Kilian (2012). Their analysis is based on the following reduced-form vector autoregression (VAR) that contains the fundamental drivers of the real price of oil,

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\mathbf{y}_t$ is a 4×1 vector of monthly data, $\mathbf{c}$ is a 4×1 vector of intercepts, $\mathbf{\Phi}_i, i = 1, ...p$, are 4×4 coefficient matrices with p indicating the number of lags, and $\boldsymbol{\varepsilon}_t$ are white-noise innovations. The four variables included in their VAR are the percent change in global crude oil production, an estimate of the change in global crude oil inventories, the log of the real price of crude oil as measured by the RAC of oil imports deflated by the U.S. consumer price index, and an index of global real economic activity (REA) developed by Kilian (2009).[2] This measure of real economic activity is based on single-voyage dry-cargo freight rates. The idea behind the Kilian index is that changes in real shipping costs expressed in deviations from a linear time trend capture the cyclical component of demand for industrial commodities. Given that shipping of raw industrial materials is linked to future production of manufacturing goods, Kilian (2009) and subsequent researchers have treated this index as a proxy for the state of the global business cycle. Baumeister and Kilian (2012) found this model did an excellent job of predicting oil prices over the period 1992.1 to 2010.6.

2.1.1 Evaluating Forecasts Based on the Kilian Index

Our first step is to reproduce the results in Alquist et al. (2013) and Baumeister and Kilian (2012). In doing so, we use the global real activity measure now recommended by Kilian (2019) which corrects a coding error in the calculation of his original index noted by Hamilton (2019).[3] We

[2] See online appendix A for a description of the data and their sources.

[3] In order to linearly detrend the deflated index only using data available to the forecaster at the time the forecast

set the lag length $p = 12$, which has been shown to deliver the most accurate out-of-sample forecasts for the real RAC (see Baumeister and Kilian, 2015). We estimate the VAR parameters recursively and evaluate the mean-squared prediction error (MSPE) of the oil price forecasts in levels for horizons $h = 1, 3, 6, 12, 24$ months ahead. We first estimate the parameters using data from 1973.2 to 1991.12 to forecast the log of the oil price for 1992.1 for $h = 1$, 1992.3 for $h = 3$, and so on, and then exponentiate to get a forecast of the level of the real oil price. We then re-estimate parameters using data through 1992.1 to forecast 1992.2 for $h = 1$, 1992.4 for $h = 3$, and so on. Following Alquist et al. (2013), we use the random walk without drift as the benchmark for evaluating the forecasting ability of alternative models. All MSPE results are normalized relative to the no-change forecast. A ratio below 1 indicates that the model does better than a random walk, while a value above 1 indicates that it does worse. To gauge the statistical significance of differences in forecasting performance, we follow Carriero et al. (2015) who suggest using the Diebold and Mariano (1995) test for equal MSPE, compared against standard normal critical values, even for nested models.

Table 1, panel (a) presents the recursive MSPE ratios for the same evaluation period as in Baumeister and Kilian (2012) which runs from 1992.1 to 2010.6. Column 1 reproduces Baumeister and Kilian's conclusion that the VAR offers better forecasts at near horizons of the real RAC than does a random walk, with MSPE reductions of 32% at the 1-month horizon and 22% at the 3-month horizon. However, since 2010 the Kilian index has exhibited some erratic behavior that is difficult to attribute to the overall level of global economic activity. Figure 1A in the online appendix shows that in early 2016 the index reached an all-time record low of 159% below trend, suggesting a far weaker global economy than at the trough of the financial crisis in 2009, when the real shipping index was only 75% below trend. After a recovery back to trend the index again dropped sharply to 88% below trend in February 2019 and to 115% in February 2020. As discussed by Hamilton (2019), these sharp contractions in the Kilian index in the post-financial crisis period are at odds

is made, we use the series of the log nominal shipping index provided on Jim Hamilton's webpage.

with common understanding and other available measures of recent fluctuations in global economic activity. The excessive swings and increased volatility of this index have raised concerns about its reliability as an indicator of world economic activity and its usefulness for forecasting. For example, Hamilton (2019) provides in-sample evidence that the Kilian index has little predictive power for a range of real commodity prices and no statistically significant correlation with annual world real GDP growth rates, casting doubts on its ability to identify shifts in demand in industrial commodity markets. Our focus here is on out-of-sample forecasts of real RAC.

In panel (b) of Table 1 we update the analysis using data through 2018.8. The first column shows that the model does not do quite as well when the evaluation period is extended to include more recent data, which may reflect the fact that oil prices were harder to forecast over 2011-2018 as well as problems with using the detrended shipping cost index as a measure of real economic activity (see Hamilton, 2019). Notwithstanding, the VAR continues to beat the random walk for near-term forecasts of real RAC but with much smaller reductions in MSPE. The improvements one and three months ahead are only 13% and 4%, respectively.

For many oil-market participants, predicting other oil prices like Brent may be a higher priority than predicting RAC. In fact, the Brent price has evolved into the global benchmark for oil and oil products with about two-thirds of oil purchases worldwide using it as a reference price according to the Intercontinental Exchange.[4] It is also closely followed by policymakers and frequently referred to in the media. For these reasons, we evaluate the usefulness of the VAR for forecasting the Brent price, replacing RAC with Brent as a more relevant measure for the global price of crude oil. The first column of Table 2(a) shows that if our goal is to forecast the real price of Brent, the VAR is completely unsuccessful. The VAR forecast does not beat the random walk at any horizon.

2.1.2 Alternative Indicators of Global Real Economic Activity

We next explore alternative monthly measures of global real economic activity that have been

[4] See https://www.theice.com/article/brent-crude/the-worlds-leading-crude-oil-benchmark.

proposed in the literature. Details on the different measures we investigate are summarized in Table 1A in the online appendix.

Real shipping cost factor. One possibility is that composite measures of shipping costs other than that proposed by Kilian (2009) may provide better forecasts. Hamilton (2019) argues that removing a deterministic linear time trend is a poor way to isolate the cyclical component in real shipping costs and is not supported by the data. A natural alternative is to use the unbalanced panel of disaggregated shipping costs underlying the Kilian index and extract a common factor.[5] The dataset consists of a cross-section of 61 real freight rates for individual shipping routes for industrial commodities such as coal and iron ore which we manually digitized from Drewry's *Shipping Insight* up to August 2018. Changes in shipping routes and in the composition of freight lead to missing observations which we fill by recursively applying the expectation-maximization (EM) algorithm of Stock and Watson (2002).[6] The real shipping cost factor for 1973.2 to 2018.8 is shown in the top panel of Figure 2A. Visually this series appears to be a far more plausible proxy for global activity than the corrected Kilian index in Figure 1A. This alternative measure also addresses some of the weaknesses of Kilian's REA index summarized by Kilian and Zhou (2018). For example, they document that freight rates are increasingly subject to idiosyncratic shocks in the markets for commodities shipped as dry bulk cargo. By contrast, the factor approach filters out commodity-specific noise and may provide a better characterization of the cyclical component of global activity.

Column 2 of Table 1 shows that for the original evaluation period (panel a) the real shipping cost factor is not quite as accurate as the Kilian index for RAC at short horizons but it does better

[5] Before conducting the principal component analysis, we compute growth rates and standardize each growth rate by subtracting the mean and dividing by its standard deviation.

[6] The algorithm is initialized by replacing missing values with the unconditional mean of the observations available for each series before extracting the first K principal components where K is determined by the Bai and Ng (2002) information criterion. We use the estimated factors to impute the missing observations and repeat the factor analysis with the updated values until the estimates do not change.

at longer horizons. For the extended sample (panel b) the shipping factor is substantially better at any horizon. It is also substantially better for forecasting Brent at all horizons (Table 2(a), column 3). Thus this alternative approach to summarizing shipping costs not only gives a more plausible proxy for global activity since 2010 but also offers a significant improvement for purposes of forecasting oil prices.

World industrial production. Column 3 of Table 1 repeats the analysis replacing the Kilian index with the index of world industrial production (WIP) developed by Baumeister and Hamilton (2019). Their measure remains closer to the traditional concept of economic activity as measured by the physical volume of output generated in the industrial sector. They constructed an updated version of a monthly index of industrial production covering OECD countries and six major emerging markets (Brazil, China, India, Indonesia, the Russian Federation and South Africa) that was originally reported in the OECD Main Economic Indicator database from 1958.1 to 2011.10 by applying the same methodology used by the OECD.[7] The WIP index is plotted in growth rates in the second panel of Figure 2A. The VAR using WIP does considerably better at predicting real RAC over every evaluation period compared to either of the shipping-based indicators. Using WIP also leads to notable improvements in forecast accuracy for the real Brent price at the shortest and longest horizons relative to the no-change forecast (Table 2(a), column 5). In particular, it reduces the MSPE by about 5% at horizons 1 and 3, and by 6% at horizon 24.

Real commodity price factor. Alquist et al. (2020), Delle Chiaie et al. (2017), and West and Wong (2014) extract a global factor from a large cross-section of monthly real commodity prices. The idea is that the source of common variation in commodity prices stems from demand-induced changes in economic activity which tend to move all prices in the same direction, while supply-side developments in specific commodity markets show up as idiosyncratic shocks that are unlikely to have pervasive effects. Our commodity price dataset consists of 23 basic industrial and agricultural

[7] This series is regularly updated and can be downloaded at https://sites.google.com/site/cjsbaumeister/research.

commodities whose markets are sensitive to changes in global economic conditions. The selection of the set of commodities is guided by the same criteria as in Alquist et al. (2019).[8] The third panel of Figure 2A shows the real commodity price factor constructed from the first principal component of the balanced panel of percent changes of real commodity prices. Column 4 of Table 1 reports the results for RAC and column 7 of Table 2(a) for Brent when this measure is used as the global economic activity indicator. The real commodity price factor does almost as well as WIP for short horizons and somewhat better at longer horizons.

Global steel production factor. Ravazzolo and Vespignani (2019) suggest using monthly world steel production as an indicator of global real activity. They argue that steel is an important input for many industries and that it is a relatively homogenous commodity that is traded freely worldwide. The World Steel Association provides an aggregate measure of the level of steel production reported by member countries. One drawback is that a consistent global series is only available since 1994 due to changes in the number of reporting countries. Since this measure aggregates the physical amount of steel produced, an increase in the number of reporting countries leads to discrete jumps, as pointed out by Kilian and Zhou (2018). We propose to construct an alternative measure based on an unbalanced panel of steel production in member countries (see appendix A for details). This enables us to extend the series all the way back to 1973 without encountering the problem of structural breaks due to aggregation, while preserving the broadest possible coverage. We extract the common component using the EM algorithm to obtain a global steel production factor which is plotted in the bottom panel of Figure 2A. Aggregating information in this way also takes care of the concern of Kilian and Zhou (2018) that world steel production is prone to idiosyncratic supply shocks in steel-producing countries and of other potential sources of noise such as measurement error and small data revisions. The last column of Table 1 shows that at short horizons the steel index does not perform quite as well as WIP and the commodity price factor for the extended evaluation

[8]See online appendix A for details on the selection criteria.

period for real RAC, but it performs much better than the Kilian index. At longer horizons its forecasting performance is comparable to that of WIP with MSPE reductions of up to 6%.

Summary. Taking stock, any of the three alternative measures for global real economic activity do significantly better than the Kilian index for forecasting either real RAC or Brent. While the forecasts obtained with the alternative indicators achieve gains in average forecast accuracy of up to 23% in the near term and up to 8% in the long term relative to the no-change forecast for the longer evaluation period, none of these models beats the random walk at the 1-year horizon. Overall, forecasting oil prices with VAR models has become more difficult since 2010 and forecasting the real price of Brent poses additional challenges.[9]

2.2 The Role of Bayesian Shrinkage

All VAR models so far have been estimated by unrestricted least squares. It is widely known that the proliferation of parameters in VARs tends to hurt the out-of-sample forecasting performance of these models. One remedy is to apply Bayesian shrinkage which has been shown to lead to more precise forecasts across a wide range of applications in macroeconomics and finance (see, e.g., Carriero et al., 2015). We examine whether Bayesian methods help reduce the MSPE of our VAR forecasts by using informative priors that shrink our unconstrained models toward a parsimonious benchmark, and thus reduce estimation uncertainty. As in Baumeister and Kilian (2012), we rely on the data-based procedure proposed in Giannone et al. (2015) for selecting the optimal degree of shrinkage in our recursively estimated Bayesian VARs (BVARs) based on the marginal data density.

Panel (a) of Table 2 compares VAR and BVAR forecasts for the real price of Brent obtained with the same five models that we considered before.[10] We find that Bayesian shrinkage leads to substantial improvements for every specification with additional MSPE reductions of up to 7% relative to the unrestricted VAR. Among the set of alternative activity indicators, columns 6 and

[9] Online appendix B illustrates where the differences across evaluation periods come from.

[10] Results on the forecast accuracy of BVAR models for real RAC can be found in Table 3A, panel (a).

8 show that the BVAR forecasts based on WIP and the commodity price factor both outperform the no-change forecast at every horizon. A comparison of columns 2 and 4 reveals that the BVAR with the real shipping cost factor does much better than the one with the Kilian index which is dominated by the random walk except at the 1-month horizon. Bayesian shrinkage also helps the accuracy of the steel factor, though it is still dominated by WIP and the commodity price factor.

While it is customary to model the global oil market from a supply-side perspective, given our interest in developing an energy demand indicator, our goal is to develop a forecasting model that emphasizes the final demand for petroleum products. For this purpose we replace world oil production with a measure of petroleum consumption. The broadest available measure at the global level is monthly total world consumption of liquid fuels provided in the EIA's *Short-Term Energy Outlook* database. Since this time series is only available from 1990.1 onward, we extend it back to 1982.1 using the growth rate of OECD petroleum consumption and further back until 1973.1 at the rate of change of world oil production. With the shift in focus from oil production to petroleum consumption, the pertinent measure for inventories is OECD petroleum stocks which are backcast before 1988 with the growth rate of U.S. petroleum stocks. The most important factor determining prices of petroleum products is the price of crude since it accounts for the largest share of production costs. Thus, the real Brent price remains the relevant price measure given that it serves as a global reference for pricing petroleum products. This consumption-based VAR model is new to the oil price forecasting literature and adds to the suite of models that are derived based on economic grounds. The notion is that fluctuations in the demand for refined products will translate into changes in the demand for crude oil and thus have predictive power for the future path of the real price of crude oil. We evaluate the forecasting performance of this new model using all five indicators of global real economic activity.

Table 2(b) reports the MSPE results for the consumption-based models estimated by unrestricted least squares and Bayesian shrinkage methods. Using fuel consumption instead of oil production leads to better forecasts of oil prices for most indicators and most horizons. Comparing different indicators and estimation methods, the overall pattern is similar to the production-based models. The BVAR forecasts dominate the VAR forecasts across almost all models and horizons, and WIP is a very useful indicator of global economic activity for purposes of any forecast, with significant gains in forecast accuracy of 12% at the 1-month horizon, 7% at the 6-month horizon, and 8% at the 24-month horizon. Forecasts using the commodity price factor are almost as good as WIP for short horizons and a little better for long horizons. Except for the Kilian index, all BVAR models consistently beat the random walk at all forecast horizons (see columns 4, 6, 8, and 10).

We conclude that Bayesian shrinkage can help improve forecasts in this setting and that using petroleum consumption in place of oil production is promising. World industrial production and the real commodity price factor are the most useful indicators of global economic activity.

2.3 The Role of Time Variation in Volatilities

Another important consideration is that energy markets and the global economy have undergone substantial transformations over time that can affect the forecasting ability of our models. Examples include shifts in the energy intensity of production and consumption, capacity constraints in inventory holdings, changes in the composition of oil producers, and market turmoil that induces bouts of volatility which are all likely to influence the models' predictive accuracy. In addition, Baumeister and Peersman (2013) document that heteroskedasticity is pervasive in oil market data.

The importance of modeling time variation in the volatilities for the forecasting performance of macroeconomic variables is well established in other data sets (see, e.g., Carriero et al., 2019; Clark and Ravazzolo, 2015). Clark and Ravazzolo (2015) provide extensive empirical evidence that models with stochastic volatility increase the accuracy of point forecasts relative to models assuming homoskedasticity. Earlier evidence for forecasting energy prices is more mixed. Baumeister and Kilian (2014a) find that adding time-varying parameters and stochastic volatility to the 4-variable

VAR model in column 1 of Table 1 did not improve forecasts of quarterly RAC compared to the monthly no-change forecast. On the other hand, Baumeister et al. (2017) find that the unobserved component stochastic volatility model does great for forecasting retail gasoline prices, especially at longer horizons. The question is thus whether a more accurate modeling of time-varying uncertainty leads to improvements in forecasting oil prices. As pointed out by Primiceri (2005), stochastic volatility is meant to capture possible heteroskedasticity of the shocks as well as nonlinearities in the dynamic relationships of the variables, which are related to low-frequency changes in volatility.

To allow for time variation in the variance of the VAR residuals, we postulate that the error term $\boldsymbol{\varepsilon}_t$ in equation (1) is normally distributed with mean zero and time-varying covariance matrix $\boldsymbol{\Omega}_t$. We factor the latter as $\boldsymbol{\Omega}_t = \mathbf{A}^{-1}\boldsymbol{\Sigma}_t(\mathbf{A}^{-1})'$ where $\mathbf{A}^{-1}$ is a lower triangular matrix with ones on the main diagonal and $\boldsymbol{\Sigma}_t$ is a diagonal matrix that contains the stochastic volatilities such that $\boldsymbol{\varepsilon}_t = \mathbf{A}^{-1}\mathbf{u}_t$ with $\mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_t)$. Carriero et al. (2019) show these assumptions allow the VAR to be written as a system of n univariate equations with the i^{th} equation taking the form:

$$y_{it} = c_i + \sum_{j=1}^{p} \boldsymbol{\Phi}_{i,j}\mathbf{y}_{t-p} + \sum_{\ell=1}^{i-1} a^*_{i,\ell}u_{\ell t} + u_{it}, \qquad u_{it} \sim N(0, \sigma^2_{it}) \tag{2}$$

where $\boldsymbol{\Phi}_{i,j}$ is the i^{th} row of the matrix $\boldsymbol{\Phi}_j$, $a^*_{i,1}, ..., a^*_{i,(i-1)}$ denotes the parameters in the i^{th} row of the triangular matrix $\mathbf{A}^{-1}$ for $i = 2, ..., n$, $u_{\ell t}$ are the residuals from the previous $i-1$ equations, and σ^2_{it} are equation-specific time-varying variances. The benefit of this reparametrization is that we can estimate the model equation by equation which is convenient for modeling stochastic volatility and allows to specify independent priors for the reduced-form coefficients across equations. The law of motion for the stochastic volatilities is $\ln \sigma_{it} = \ln \sigma_{it-1} + \eta_{it}$ with the vector of innovations $\eta_t \sim N(0, \boldsymbol{\Lambda})$ where $\boldsymbol{\Lambda}$ is a full covariance matrix as in Primiceri (2005). For details on the choices of priors and hyperparameters that control the amount of shrinkage, evidence on the convergence of the algorithm, and construction of the forecasts, see online appendix C.

Table 3 compares the forecasting performance of specifications that allow for stochastic volatility

(SV-BVAR) with the homoskedastic BVAR from Section 2.2. Allowing for stochastic volatility substantially improves the forecasting performance of the models based on shipping costs, but only the one using the shipping cost factor is competitive with any of the other three indicators of global economic activity. The most striking result is that stochastic volatility achieves impressive gains in forecast accuracy at longer horizons for the four competitive models. Carriero et al. (2019) make the case that time-varying volatilities should improve the point forecasts especially at longer horizons. The reason is that the heteroskedastic model will provide more efficient estimates given that the predictive means will gradually deviate from their homoskedastic counterparts as the predictive densities cumulate nonlinearly with the forecast horizon. This is consistent with what we find. The reductions in MSPE range between 10-14% 1-year-ahead and 23-29% 2-years-ahead compared to only 2-3% and 6-10% in models without stochastic volatility. These large MSPE reductions for long-run forecasts come at the expense of a small loss of at most 3% in predictive accuracy at near horizons for the models using WIP and the commodity price factor relative to forecasts with constant variance. In contrast, the forecasts based on the steel factor benefit at all horizons from adding stochastic volatility. We also see from Table 3 that consumption-based models still generally outperform production-based models for purposes of forecasting oil prices. We conclude that stochastic volatility is an important ingredient for long-horizon forecasts of the real Brent price.

2.4 Pooling Forecasts and Information

Another approach to guard against forecast failures due to structural change is to pool forecasts. It has long been known that combining forecasts not only can lead to superior forecasting performance but also hedges against varying accuracy of individual forecasting models over time (see, e.g., Timmermann, 2006). There is ample evidence that forecast combinations work well for oil price forecasting and help improve forecast accuracy especially at longer horizons (see, e.g., Baumeister and Kilian, 2014a, 2015; Bernard et al., 2018; Garrett et al., 2019; Manescu and van Robays, 2016). This section explores the benefits of pooling in our setting in four ways.

First, we consider an equal-weighted average of the forecasts from the five consumption-based SV-BVARs whose individual results were reported in Table 3(b). The MSPE ratios for this forecast combination can be found in the first column of Table 4A. For short horizons, these pooled forecasts are always dominated by the individual BVAR forecasts coming from either WIP or the commodity factor. For long horizons, they are always dominated by the individual SV-BVAR forecasts coming from either WIP or the commodity factor. We conclude that although simple averaging has often proved a useful strategy, its success in the current setting is not convincing. One possible explanation is that the models only differ in the measure of global economic activity, while existing evidence on the superiority of forecast combinations is based on a more diverse set of forecasting models.

Since the predictive power of different models likely changes over time, a second strategy is to single out the best model at each point in time. We use the dynamic model selection approach of Koop and Korobilis (2014) with a decay factor $\alpha = 0.99$ to determine the optimal model. Column 2 shows that selecting the single best model is never more accurate than the SV-BVAR using WIP.

A third way we tried to pool information is by extracting the first principal component from the unbalanced panel of 93 variables underlying the different global activity indicators in Table 1A using the EM algorithm. We include this factor directly in the SV-BVAR in place of the individual indicators. Column 3 of Table 4A shows that combining information sets turns out to be dominated by the BVAR or SV-BVAR based on WIP alone. The interpretation may again be that WIP is itself a broad aggregate with weights guided by the importance of different sectors in different countries. Using those weights for aggregation may be superior to simple principal components.

Fourth, we explore a market-based approach to pooling information. Some might argue that the futures market is already pooling in a rational, optimal way all the information that could be relevant for forecasting oil prices. We follow Baumeister and Kilian (2012) and use the following futures-spread model: $R_{t+h|t} = R_t\ (1 + f_t^h - s_t - E_t(\pi_{t+h}))$ where R_t denotes the current level of the real Brent price, f_t^h denotes the log of the current Brent futures price for a contract with maturity

h, s_t denotes the log of the Brent spot price, and $E_t(\pi_{t+h})$ denotes the expected inflation rate over the next h periods (see appendix A). Column 4 shows that the Brent futures spread actually does a worse job at forecasting the real price for near horizons than does a random walk. The futures market does a reasonable job at longer horizons, but still is outperformed by the other approaches.

This evidence suggests that some popular ideas about pooling information from different sources using either a model-based or a market-based approach do not work particularly well in our context. We will explore in Section 4 some alternative approaches to pooling that may hold more promise. Before doing so, however, we first evaluate our set of models in terms of an alternative dimension that has received little attention in the literature, which is forecasting petroleum consumption.

3 Forecasting Global Petroleum Consumption

Up to this point in the paper we have been considering using models like the SV-BVAR for purposes of forecasting a single variable: the real price of oil. However, price forecasts are only one aspect of future developments in energy markets. Government agencies like the U.S. EIA, intergovernmental organization like the IEA and OPEC, and oil companies such as BP also regularly release short-term and long-term projections for oil and liquid fuels consumption as a complement to their price outlook and key component of their overall assessment of global energy demand. For example, the EIA publishes monthly consumption forecasts in its *Short-Term Energy Outlook* but the length of the forecast horizon varies each month from a maximum of 24-months-ahead to a minimum of 13-months-ahead. However, there is no publicly available documentation on how these forecasts are generated. To the best of our knowledge, there are currently no model-based forecasts for global petroleum consumption available. We use the set of VAR models that have been shown to work well for forecasting the real Brent price to also produce forecasts for petroleum consumption. While it would seem natural to evaluate the accuracy of our forecasts against those of the EIA, this is not possible since there are no publicly available records of the EIA's historical forecasts of monthly global fuel consumption for our entire evaluation period that would enable such a comparison. Moreover, the maximum forecast horizon that is consistently available each month from October 2007 onward is one year, while we focus on a horizon of two years. Absent an institutional forecast as benchmark, we follow the macroeconomic forecasting literature and consider a univariate linear autoregressive (AR) model which is the standard when evaluating forecasts of real economic variables (see, e.g., Chauvet and Potter, 2013; Alquist et al., 2013). We set the lag length to $p = 12$ to be consistent with our VAR models. While all models are estimated with petroleum consumption entering in growth rates, we evaluate the forecasts in levels since both the EIA and the IEA report their consumption forecasts in terms of million barrels per day (mbd).

Table 4 presents the recursive MSPE ratios for forecasts of global petroleum consumption obtained with the BVAR models with and without stochastic volatility using the same five economic activity indicators as before and the three model-based pooling methods. Panel (a) shows that none of the BVAR models beats the AR(12) benchmark. All the MSPE ratios are above 1, and their performance quickly deteriorates as the forecast horizon lengthens. Adding stochastic volatility improves the forecast accuracy considerably but the models still only outperform the benchmark at the 1-month horizon with MSPE reductions between 3% and 7% as summarized in panel (b). The only model showing any promise beyond the 1-month horizon is the one that features WIP which consistently produces the lowest MSPE ratios being essentially tied with the benchmark model for horizons 3-12 months. The WIP model is also more accurate than any of the pooling approaches (panel c).[11] This disappointing forecasting performance suggests that our VARs may be missing important predictor variables that carry useful information for future global petroleum consumption, a possibility that we explore in the next section.

4 Towards A New Indicator of Global Economic Conditions

So far the literature has focused on developing indicators that capture cyclical variation in global real economic activity. These measures are rather limited in scope since they are all constructed

[11] Linearly detrending the three global activity indicators does not improve the forecast accuracy (see Table 5A).

based on a single category of variables such as freight rates, commodity prices, steel production or industrial production. The question to which we now turn is whether global economic conditions as they relate to energy markets can be represented by any narrow set of variables or whether there is value in diversifying the basket of variables to include new categories that cover additional dimensions of the global economy. While the forecasting success for the real price of Brent suggests that the information contained in existing activity measures is sufficient, there is reason to believe that considering other types of data might help improve the forecast accuracy for global petroleum consumption. The next section describes the set of variables that we investigate for whether they could capture important information about the global economy as it relates to energy markets.

4.1 A Multi-Dimensional Approach

We compile the set of 16 indicators summarized in Table 6A that cover a broad range of variables tied to energy demand. The variable selection is guided by four principles. First, the variables should represent different data categories to span multiple dimensions of the global economy. We broadly define eight categories: real economic activity, commodity prices, financial indicators, transportation, uncertainty, expectations, weather, and energy-related measures. Second, each individual variable should matter for energy demand on economic grounds. Third, it should have the broadest possible coverage geographically, conceptually, and in time. Fourth, the number of variables should be kept at a manageable size to ensure that the dataset can be easily updated in real time.

Real economic activity. As discussed above, current and future economic activity are a key determinant of global economic conditions and energy demand. We include WIP as the broadest measure of real output in the industrial sector at a global scale. WIP is also important because it covers manufacturing, mining, and utilities, sectors that are closely tied to energy in that they use oil and refined petroleum products in the production process. We also include the Conference Board Leading Economic Index. This is a closely watched leading indicator with a proven track record of signaling peaks and troughs in the business cycle. While it is U.S.-specific, it consists of a range of measures that capture future economic trends. A third measure is the OECD consumer confidence index, which is the broadest available measure to gauge the outlook for households' consumption spending. This survey-based indicator summarizes whether the attitude of consumers is optimistic or pessimistic by asking households about their expected financial situation, unemployment, capability of saving, and sentiment about the general economic situation.

Commodity prices. Among commodity prices, copper stands out in its importance in manufacturing, construction, and infrastructure. Copper prices have been used in a number of studies as a representative commodity price and barometer of future global growth (see, e.g., Hamilton, 2015; Bernanke, 2016). The nominal price of copper is deflated by the U.S. consumer price index.

Financial indicators. Our two main financial indicators are foreign exchange and stock returns. Exchange rate fluctuations reflect trade and financial flows and changes in economic activity. They also are tightly linked to energy demand (see, e.g., De Schryder and Peersman, 2015). We select the broad real trade-weighted U.S. dollar index not only because this series extends furthest back in time, but also because oil prices are quoted in dollars and changes in the exchange rate often translate into changes in petroleum consumption in oil-importing countries. Our measure for stock returns is based on the MSCI world index which contains stocks from companies worldwide and represents a broad cross-section of global markets. We use a third financial indicator that is more specific to energy demand, which is the excess return earned on the Fama-French portfolio for the transportation sector. The transportation sector is obviously the most energy-intensive sector, so excess returns in this sector should provide forward-looking information for energy consumption.

Transportation. We also use two real indicators of transportation demand. Registrations of vehicles are indicative of the future demand for fuel. The longest available series with the broadest coverage is for passenger cars in OECD countries which comprises newly registered private cars and commercial vehicles. Given that the automobile industry is a large sector in many major economies, car sales which precede registration also matter for aggregate fluctuations. We complement this stock variable with a flow measure of traffic volume represented by U.S. total vehicle miles traveled.

Uncertainty measures. Oil production and prices are often driven by geopolitical events. These events matter not just for energy markets but also influence the global economy more broadly. We use the geopolitical risk index developed by Caldara and Iacoviello (2018). This should reflect supply disruption risks and translate into concerns about the future availability of oil which will influence energy demand behavior worldwide. Long-run oil price uncertainty is another important determinant of energy-related spending (see, e.g., Bernanke, 1983; Jo, 2014). It is measured as realized volatility computed based on daily returns for WTI futures contracts with a maturity of 12 months.

Expectations measures. We include the index of consumer expectations from the Michigan Survey which aggregates households' short-term and long-term outlook of the general economy. We also construct a measure of oil price expectations based on the difference between WTI futures prices with 3 and 12 months to maturity. This market-based measure should signal the direction of expected price changes which will likely influence spending on energy-dependent goods and services.

Weather indicators. A key global weather-related variable is El Niño. We use the Oceanic Niño Index (ONI) which is available from the National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center.[12] Cashin et al. (2017) discuss the importance of this complex weather phenomenon for global macroeconomic performance. It also has direct implications for energy use as higher temperatures associated with El Niño episodes lead to more fuel demand for power generation. Another weather-related indicator provided by NOAA is the Residential Energy Demand Temperature Index which is based on heating and cooling degree days in the US, and as such, is a useful measure of fluctuations in demand for residential heating and cooling.

[12] The term El Niño refers to the large-scale ocean-atmosphere climate interaction linked to a periodic warming in sea surface temperatures across the central and east-central Equatorial Pacific. The ONI is NOAA's primary indicator for monitoring El Niño and La Niña, which are opposite phases of this climate pattern.

Energy-related indicators. The broadest energy-specific measure is energy production and electricity distribution for the EU28. This is not only directly tied to energy demand but is also an indicator for the overall intensity of economic activity since the production of most goods and services requires electricity (see Arora and Lieskovsky, 2014).

We extract the first principal component from this unbalanced panel of 16 variables by applying the EM algorithm recursively and use this estimated factor to replace the economic activity measure in our 4-variable consumption-based SV-BVAR model. Row 1 of Table 5 shows that this factor-augmented model produces forecasts of the real Brent price that are marginally less accurate than forecasts from the set of models with existing measures of global real activity.[13] MSPEs are at most 3% higher compared to the best-performing individual model in Table 3(b) with the factor-augmented VAR being the most accurate 24-months-ahead. This small loss in forecasting performance for oil prices is more than made up for by the dramatic improvement in forecasts of petroleum consumption. Row 6 of Table 5 shows that this model outperforms the AR(12) benchmark at all horizons with impressive MSPE reductions between 6% and 13%.

We conclude that while the forecasting performance for the real Brent price is comparable across indicators, the forecasting success for global petroleum consumption confirms our earlier conjecture that existing global economic activity measures miss information that is relevant for determining energy demand. The obvious next question is whether including additional variables from each data category can improve the joint forecast accuracy further.

4.2 Is More Information Better?

To address this question, we collect an additional 150 variables listed in Table 8A to form a very large dataset, which greatly expands the coverage of each of the eight broad categories. We also add all the disaggregated data from the existing real economic activity indices, bringing the total up to 256 variables. Row 2 of Table 5 shows that substantially increasing the size of the panel does

[13] Upon a referee's suggestion, Table 7A shows that this finding is robust to starting the evaluation period in 2000.1.

not improve the usefulness of the first principal component for forecasting the real Brent price, and row 7 shows that it leads to a deterioration of the forecast performance for petroleum consumption for all but the 1-month-ahead forecast with MSPE increases of 6% at longer horizons. Adding more factors worsens the forecast accuracy further. For example, the model with three factors implies a loss of 15% in MSPE reductions for the Brent price at horizon 24 (row 4). Rows 8 and 9 show that for consumption the forecasting gains vanish from horizon 6 onward with all MSPE ratios above 1. These findings are in line with Boivin and Ng (2006) who show that factors extracted from a larger panel can lead to inferior forecasting performance. The reason is that as more series from the same data category are added, the possibility of cross correlation in the idiosyncratic errors increases which leads to a loss in forecast accuracy. We conclude that what matters most for the joint forecasting success is the composition not the size of the dataset.

The 16 variables above were chosen based on their economic relevance and broadness of coverage to be representative of different aspects of the global economy. Our large dataset offers an alternative for selecting the most relevant variables using statistical criteria, namely the 16 variables with the highest loadings in absolute value on the first principal component (see appendix A). Row 5 shows that the statistical variable selection leads to additional gains for the Brent price yielding further MSPE reductions of up to 3% at long horizons. However, these gains come at the cost of the forecast accuracy for consumption (row 10). This model has much higher MSPE at all horizons and from horizon 6 onward it no longer beats the AR(12) benchmark. We conclude that our economically-motivated set of 16 variables gives the most reliable signal and results in the best overall forecasts.

4.3 A Global Economic Conditions Indicator

Panel A of Figure 1 displays the first principal component of the 16 variables in Table 6A over the period 1973.2 to 2020.5. We will refer to this series as our global economic conditions indicator (GECON).[14] The series is normalized so that it has a mean of zero and a standard deviation of one.

[14] This series is available at https://sites.google.com/site/cjsbaumeister/research.

Thus a value of zero corresponds to economic conditions characterized by normal trend growth. To get a better idea of the persistent movements in the global economy we report the 3-month moving average. The indicator tracks known episodes of worldwide contractions and expansions well. The downturn related to the 2008-09 financial crisis is more severe than the worldwide recession of 1974-75 with the indicator being five standard deviations below its long-run average. The second half of the eighties and the mid-2000s are periods of strong economic conditions. The indicator signals an improvement in global economic conditions in 2013 followed by a period of sluggish growth in 2015-17. The most extreme event is the collapse of the global economy in early 2020 as a result of the coronavirus pandemic with the indicator falling almost seven standard deviations below trend.

What are the key drivers of current global economic conditions? Panel B shows the breakdown into the main contributing factors since June 2019. The steep decline in economic activity since the outbreak of the crisis accounts for about half of the deep downturn. The dramatic reduction in traffic volumes as a result of widespread lockdown measures from March to May explains another substantial share of the sharp drop. Rising uncertainty contributes to a further deterioration of global economic conditions. Subdued expectations and weak financial indicators also play some role in the severe contraction. This is in contrast to the Great Recession, the early 2000s slowdown, and the 1974-75 downturn where real activity and financial indicators were the main sources of slack. Figure 3A shows the decomposition of movmements in the GECON indicator into the most important data categories for the entire period to fully appreciate the value of this diverse dataset beyond current events. As can be seen, the relative importance of different categories varies over time. For example, most of the boom in the mid-80s is attributed to financial indicators, while uncertainty is a major factor in the early 1990s slump. Transportation matters more in the early part of the sample but has regained importance in the current environment.

4.4 Density Accuracy and Real-Time Performance

Since our ultimate goal is to use this forecasting model to derive a set of monitoring tools to summarize expected future conditions in energy markets and to quantify oil price risks, two aspects deserve further consideration. First, our proposed risk measures are based on the entire distribution of forecasts for each horizon. To ensure their reliability, we evaluate the joint performance of the density forecasts for price and consumption by computing the log predictive score as in Carriero et al. (2015). The average log score is negative and a less negative value indicates that the model provides a better characterization of the true conditional density. To get an idea of the statistical significance of differences in log scores, we use the Model Confidence Set (MCS) approach of Hansen et al. (2011) that identifies the subset of models $\widehat{M}^*_{1-\alpha}$ that contains the best model with a specified level of confidence α. Table 6 reports the average log scores and the MCS p-values for the entire set of models at each horizon h. The BVAR models are always excluded from the MCS for $\alpha = 0.25$. Modeling stochastic volatility improves not only point forecasts but also provides a better description of the conditional density. The implied ranking of the SV-BVAR models is in line with our conclusions about predicitive ability based on point forecasts.[15] GECON is always among the top three models except for horizon 24.

Second, we examine to what extent real-time data constraints in the form of data being published with a lag and preliminary data being revised for some time after the first release affect the forecasting performance and hence the value of our monitoring tools for real-time analysis.[16] Table 11A compares the recursive MSPE ratios for forecasts of the real Brent price and petroleum consumption using observed and nowcasted values at the end of the sample for all forecast horizons from 1 to 24 months. Taking data delays into account makes little difference overall. The forecast accuracy for the real Brent price suffers slightly for horizons up to 9 months when nowcasted data are used, but is basically unaffected at longer horizons. Interestingly, the real-time nature of the data positively affects consumption forecasts with some gains in MSPE ratios for short horizons.

[15] In Tables 9A and 10A we report MCS p-values for the MSPEs of the Brent price and consumption separately.

[16] Online appendix D provides full details on the real-time setting.

The likely reason for this finding is that in the case of consumption, the gaps in data availability not only affect the SV-BVAR model forecast but also the AR benchmark so that it is not clear a priori which way the MSPE ratios go. From the results it is clear that the nowcasted data for consumption are putting the benchmark model at a slight disadvantage in the short run. In the longer term, this ranking is reversed and the pseudo real-time model performs slightly better.

5 Assessing Energy-Market Conditions

5.1 A Barometer for Future Energy Demand

An important concern for industry analysts, policymakers, and government agencies like the EIA is how demand for energy will evolve in the near term. How can we gauge what demand conditions are expected to prevail in the future? We propose to look at the difference between the 13-month-ahead and the 1-month-ahead forecasts of the level of petroleum consumption. This measure summarizes the slope of the forecasts as a function of horizon and is thus independent of the level. An increase in this forecast measure can be viewed as indicating rising demand pressures over the next year. Figure 2, panel A shows the 6-month moving average of our energy demand indicator for 1992.1 to 2020.5. Each point in the graph tells us whether demand conditions are expected to be tight or loose and should be interpreted relative to the past. The indicator signals strong anticipated growth in demand after the East Asian crisis in 1997. A sharp tightening in demand conditions is also evident before the Great Recession. Demand pressures ease with the recession but energy demand is predicted to pick up again afterwards. Demand conditions are anticipated to loosen as a result of the European double-dip recession. Expected demand pressures have been mounting again since 2012 with steep predicted growth in demand in mid-2014 when oil prices fell dramatically.

In the wake of the coronavirus pandemic, fuel demand collapsed worldwide from 100 mbd at the beginning of 2020 to 89 mbd by the end of March. This unprecedented reduction of demand raised the question of whether and when petroleum consumption will return to its pre-crisis level. Panel B shows the annual growth rate of consumption and its predicted path as of May 2020. Fuel consumption dropped precipitously and is expected to gradually rebound. Panel C indicates that the probability for consumption to surpass its pre-pandemic level rises from 20% by year end to 40% two years out suggesting ongoing weakness in demand. The energy demand indicator signals an expected demand increase of about 2 mbd over the next 12 months.

5.2 Measuring Oil Price Risks

Policymakers and other users of oil price forecasts are typically not only interested in the expected future path of the price of oil but would also like to assess the likelihood that prices exceed a certain upper threshold or fall below a certain lower threshold. This matters because a predicted increase in the probability of higher or lower prices relative to past experience can affect firms' and consumers' spending plans with the potential to influence macroeconomic performance in the short run (see, e.g., Hamilton, 2003). While oil companies, households, and firms may differ in the thresholds they care about, we first calculate probabilities that the oil price will fall outside a common price range but then also ask by how much the forecast exceeds specific upper and lower bounds taking the risk preferences of various agents into account.[17]

5.2.1 Oil Price Pressure Measures

We propose a new measure that signals the likelihood of significant upward and/or downard oil price moves relative to the recent past. Specifically, we use the predictive density to compute the probabilities that the expected price will rise above or fall below the maximum and minimum values of oil price levels over the past year. The resulting price pressure measures thus indicate the likelihood of unusually high or low prices.

[17] Baumeister and Kilian (2014b) also conduct an analysis of the risks embodied in oil price forecasts, but their risk measures are based on a structural model of the global oil market. In particular, they focus on risks associated with hypothetical events about future oil supply and demand conditions summarized in probability-weighted densities.

Our model implies a probability based on information I_t in month t that oil prices h months from now will exceed the highest value seen over the last year: $\Pr([p^{oil}_{t+h} > \max\{p^{oil}_t, p^{oil}_{t-1}, ..., p^{oil}_{t-11}\}]|I_t]$. Following Jackson et al. (2015), we calculate an average value of this probability over 12 months,

$$PPM^{oil,+}_t = (1/12)\sum_{h=1}^{12} \Pr([p^{oil}_{t+h} > \max\{p^{oil}_t, p^{oil}_{t-1}, p^{oil}_{t-2}, ..., p^{oil}_{t-11}\}]|I_t].$$

This measure is plotted in the upper panel of Figure 3. Similarly, we calculate a probability that prices on average over the next year will fall below the lowest value seen over the last 12 months:

$$PPM^{oil,-}_t = (1/12)\sum_{h=1}^{12} \Pr([p^{oil}_{t+h} < \min\{p^{oil}_t, p^{oil}_{t-1}, p^{oil}_{t-2}, ..., p^{oil}_{t-11}\}]|I_t].$$

This measure is plotted in the lower panel of Figure 3.

To validate the usefulness of these measures we examine a number of well-known historical episodes. In the aftermath of the Asian financial crisis, the probability that oil prices would fall below the lowest price over the past 12 months remained consistently high at 40% over a period of two years before plummeting to zero in early 2000. After oil prices reached a record low in December 1998, the likelihood of upward price pressures spiked resulting in a 55% probability that the Brent price will on average surpass its highest value during the past year over a 12-month horizon. At the onset of the Great Recession the average probability that the Brent price would exceed the previous year's price maximum over the next 12 months dropped from around 40% to essentially zero, while chances of prices falling below the lower threshold jumped up to 50% followed by an all-time low of around 20% of moving outside of the recent range in mid-2009. From 2012 onward, the odds that the Brent price would drop below its most recent lower bound gradually increased reaching a peak of 70% in early 2015. As of May 2020, the price pressure measures indicate on average a heightened probability of 44% of Brent falling below the recent lower threshold of $18 but only a 20% probability of exceeding its recent upper threshold of $67 in the period up to May 2021.

5.2.2 Risk Assessment as of March 2020

The market turmoil triggered by the coronavirus crisis and the Saudi-Russia oil price war has taken a major toll on the oil industry in the first quarter of 2020. At the end of March many shale oil producers had to assess the risks associated with future price developments to decide whether to keep producing or to shut in wells. This decision not only depends on the likelihood of reaching the profit zone but also on producers' risk preferences. Following Baumeister and Kilian (2014b), we compute the upside and downside price risks as follows:

$$\begin{aligned} UPR_t^{oil,\alpha} &= \int_{\overline{p}}^{\infty} (p_{t+h}^{oil} - \overline{p})^{\alpha} dF_t(p_{t+h}^{oil}) \ , \ \alpha \geq 0 \\ DPR_t^{oil,\beta} &= -\int_{-\infty}^{\underline{p}} (\underline{p} - p_{t+h}^{oil})^{\beta} dF_t(p_{t+h}^{oil}) \ , \ \beta \geq 0 \end{aligned}$$

where α and β are parameters of risk aversion. The upper threshold $\overline{p}$ for shale producers is the breakeven price which we set at \$55. We assume that what they care about is the expected excess which implies risk neutrality ($\alpha = 1$) after passing the threshold. On the downside instead shale producers are risk-averse ($\beta = 2$) should the price drop under the threshold $\underline{p}$ of the March low of \$32 which already strains the finances of many shale oil producers.

Figure 4 shows that as of March 2020 the expected excess was small and only gradually building up to a maximum of \$64 over a 12-month horizon suggesting a dire outlook for shale producers. Downside risks were mounting fast putting the shale industry in a tight spot. Over the past couple of months we did observe a rise in shutdowns as prices remained well below breakeven levels.

5.2.3 A Historical Perspective on Expected Oil Price Risks

It is also useful to look at the risk assessment implied by the model for selected historical episodes. Figure 5 displays various risk measures based on the forecast distributions for the real Brent price up to 24 months ahead. The upper row of Panel A shows the expected price pressures as of June 2014. In the year before the onset of the oil price slump, Brent prices were pretty stable and fluctuated narrowly between \$107 and \$112. The predictive probabilities signal a rapid build-up of downward price pressures which surpasses 50% after the first six months and reaches a probability as high as 75% at the end of the 2-year forecasting horizon. The probability for oil prices continuing their calm ride subsides quickly falling from 60% at the one-month horizon to below 10% in the long run. The odds for upward price pressures remain fairly constant at around 20%. In the bottom row of Panel A we assess the risks for oil consumers and oil producers assuming quadratic preferences for both. The first plot shows the expected deviations over the next two years conditional upon Brent falling outside the price band. The squared deviations from the lower threshold are much larger than those from the upper threshold. Weighted by the corresponding tail probabilities this implies very low upside price risks for consumers but rapidly rising downside risks for producers. The observed price for Brent crude in June 2016 turned out to be $48.

After bottoming out in early 2016, the Brent price showed strong signs of recovery which in June 2016 led to the question of how likely it was that oil prices would continue to climb, hold steady around that new level, or begin another slide. In the year preceding this episode, the Brent price ranged from $31 to $57. As can be seen in Panel B, the chances for upward oil price pressures increase steadily reaching about 40% after one year. The probability of oil prices falling below $31 is predicted to be tiny for the first couple of months but rises to a robust 20% in the medium to long run. Assuming risk-neutral consumers and risk-averse producers, the last plot shows that the risk of rising prices is a growing concern, while downside risks are expected to be minimal. In June 2018 the Brent price reached $74.

This analysis illustrates that the model can be used to derive accurate and timely measures of risks related to future oil price developments and of the expected tightness of energy demand conditions. These measures should help guide policymakers, market analysts, firms, and consumers in their assessment of the likely future state of energy markets.

6 Conclusions

Global economic conditions are a key driver of energy markets. In this paper we evaluated the usefulness of several existing measures of global real economic activity that have been proposed in the literature in terms of their out-of-sample forecasting performance for the real price of oil and global petroleum consumption. We also compared them to alternative measures derived from a diverse set of global variables that influence energy demand collected specifically for this study. Our analysis has the following main takeaways. First, for short-horizon oil price forecasts consumption-based models using WIP perform best. Second, for long-horizon oil price forecasts allowing for stochastic volatility leads to considerable improvements in forecast accuracy across all indicators. Third, for forecasting the real Brent price and fuel consumption jointly, the most accurate model uses our newly-developed global economic conditions indicator based on 16 variables that cover multiple dimensions of the global economy. We show how the real-time forecasts for price and consumption generated by this model can be used to derive measures that provide policymakers and markets with a quantitative assessment of expected oil price pressures and future energy demand conditions.

References

Alquist, R., S. Bhattarai, and O. Coibion (2020). "Commodity-Price Comovement and Global Economic Activity," *Journal of Monetary Economics* 112: 41-56.

Alquist, R., L. Kilian, and R. Vigfusson (2013). "Forecasting the Price of Oil," in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2A, Amsterdam: North-Holland, 427-507.

Arora, V., and J. Lieskovsky (2014). "Electricity Use as an Indicator of U.S. Economic Activity," mimeo, EIA.

Baumeister, C., and J.D. Hamilton (2019). "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks," *American Economic Review* 109(5): 1873-1910.

Baumeister, C., and L. Kilian (2012). "Real-Time Forecasts of the Real Price of Oil," *Journal of Business and Economic Statistics* 30(2): 326-336.

Baumeister, C., and L. Kilian (2014a). "What Central Bankers Need to Know About Forecasting Oil Prices," *International Economic Review* 55(3): 869-889.

Baumeister, C., and L. Kilian (2014b). "Real-Time Analysis of Oil Price Risks Using Forecast Scenarios," *IMF Economic Review* 62(1): 119-145.

Baumeister, C., and L. Kilian (2015). "Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach," *Journal of Business and Economic Statistics* 33(3): 338-351.

Baumeister, C., L. Kilian, and T.K. Lee (2017). "Inside the Crystal Ball: New Approaches to Predicting the Gasoline Price at the Pump," *Journal of Applied Econometrics* 32(2): 275-295.

Baumeister, C., and G. Peersman (2013). "The Role of Time-Varying Elasticities in Accounting for Volatility Changes in the Crude Oil Market," *Journal of Applied Econometrics* 28(7): 1087-1109.

Bai, J., and S. Ng (2002). "Determining the Number of Factors in Approximate Factor Models," *Econometrica* 70(1): 191-221.

Bernanke, B.S. (1983). "Irreversibility, Uncertainty, and Cyclical Investment," *Quarterly Journal of Economics* 98(1): 85-106.

Bernanke, B.S. (2016). "The Relationship between Stocks and Oil Prices," https://www.brookings.edu/blog/ben-bernanke/2016/02/19/the-relationship-between-stocks-and-oil-prices/

Bernard, J.-T., L. Khalaf, M. Kichian, and C. Yelou (2018). "Oil Price Forecasts for the Long Term: Expert Outlooks, Models, or Both?" *Macroeconomic Dynamics* 22(3): 581-599.

Boivin, J., and S. Ng (2006). "Are More Data Always Better for Factor Analysis?" *Journal of*

Econometrics 132(1): 169-194.

Caldara, D., and M. Iacoviello (2018). "Measuring Geopolitical Risk," mimeo, FRB.

Carriero, A., T.E. Clark, and M. Marcellino (2015). "Bayesian VARs: Specification Choices and Forecast Accuracy," *Journal of Applied Econometrics* 30(1): 46-73.

Carriero, A., T.E. Clark, and M. Marcellino (2019). "Large Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors," *Journal of Econometrics* 212(1): 137-154.

Cashin, P., K. Mohaddes, and M. Raissi (2017). "Fair Weather or Foul? The Macroeconomic Effects of El Niño," *Journal of International Economics* 106: 37-54.

Chauvet, M., and S. Potter (2013). "Forecasting Output," in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2A, Amsterdam: North-Holland, 141-194.

Clark, T.E., and F. Ravazzolo (2015). "Macroeconomic Forecasting Performance Under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics* 30(4): 551-575.

Delle Chiaie, S., L. Ferrara, and D. Giannone (2017). "Common Factors of Commodity Prices," mimeo, FRBNY.

De Schryder, S., and G. Peersman (2015). "The U.S. Dollar Exchange Rate and the Demand for Oil," *Energy Journal* 36(3): 263-285.

Diebold, F.X., and R.S. Mariano (1995). "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13(3): 253-263.

Ferrari, D., F. Ravazzolo, and J. Vespignani (2019). "Forecasting Energy Commodity Prices: A Large Global Dataset Sparse Approach," CAMP Working Paper Series No 11/2019.

Garratt, A., S.P. Vahey, and Y. Zhang (2019). "Real-Time Forecast Combinations for the Oil Price," *Journal of Applied Econometrics* 34(3): 456-462.

Giannone, D., M. Lenza, and G. Primiceri (2015). "Prior Selection for Vector Autoregressions," *Review of Economics and Statistics* 97(2): 436-451.

Hamilton, J.D. (2003). "What is an Oil Shock?" *Journal of Econometrics* 113(2): 363-398.

Hamilton, J.D. (2015). "What's Driving the Price of Oil Down?" https://econbrowser.com/archives/2015/01/whats-driving-the-price-of-oil-down-2

Hamilton, J.D. (2019). "Measuring Global Economic Activity," *Journal of Applied Econometrics*, forthcoming.

Hansen, P.R., A. Lunde, and J.M. Nason (2011). "The Model Confidence Set," *Econometrica* 79(2): 453-497.

Jackson, L.E., K.L. Kliesen, and M.T. Owyang (2015). "A Measure of Price Pressures," *Federal Reserve Bank of St. Louis Review*, First Quarter, 25-52.

Jo, S. (2014). "The Effects of Oil Price Uncertainty on Global Real Economic Activity," *Journal of Money, Credit, and Banking* 46(6): 1113-1135.

Kilian, L. (2009). "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review* 99(3): 1053-1069.

Kilian, L. (2019). "Measuring Global Economic Activity: Do Recent Critiques Hold Up to Scrutiny?" *Economics Letters* 178: 106-110.

Kilian, L., and X. Zhou (2018). "Modeling Fluctuations in the Global Demand for Commodities," *Journal of International Money and Finance* 88: 54-78.

Koop, G., and D. Korobilis (2014). "A New Index of Financial Conditions," *European Economic Review* 71: 101-116.

Manescu, C., and I. van Robays (2016). "Forecasting the Brent Oil Price: Addressing Time-Variation in Forecast Performance," mimeo, ECB.

Primiceri, G.E. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies* 72(3): 821-852.

Ravazzolo, F., and J. Vespignani (2019). "World Steel Production: A New Monthly Indicator of Global Real Economic Activity," *Canadian Journal of Economics*, forthcoming.

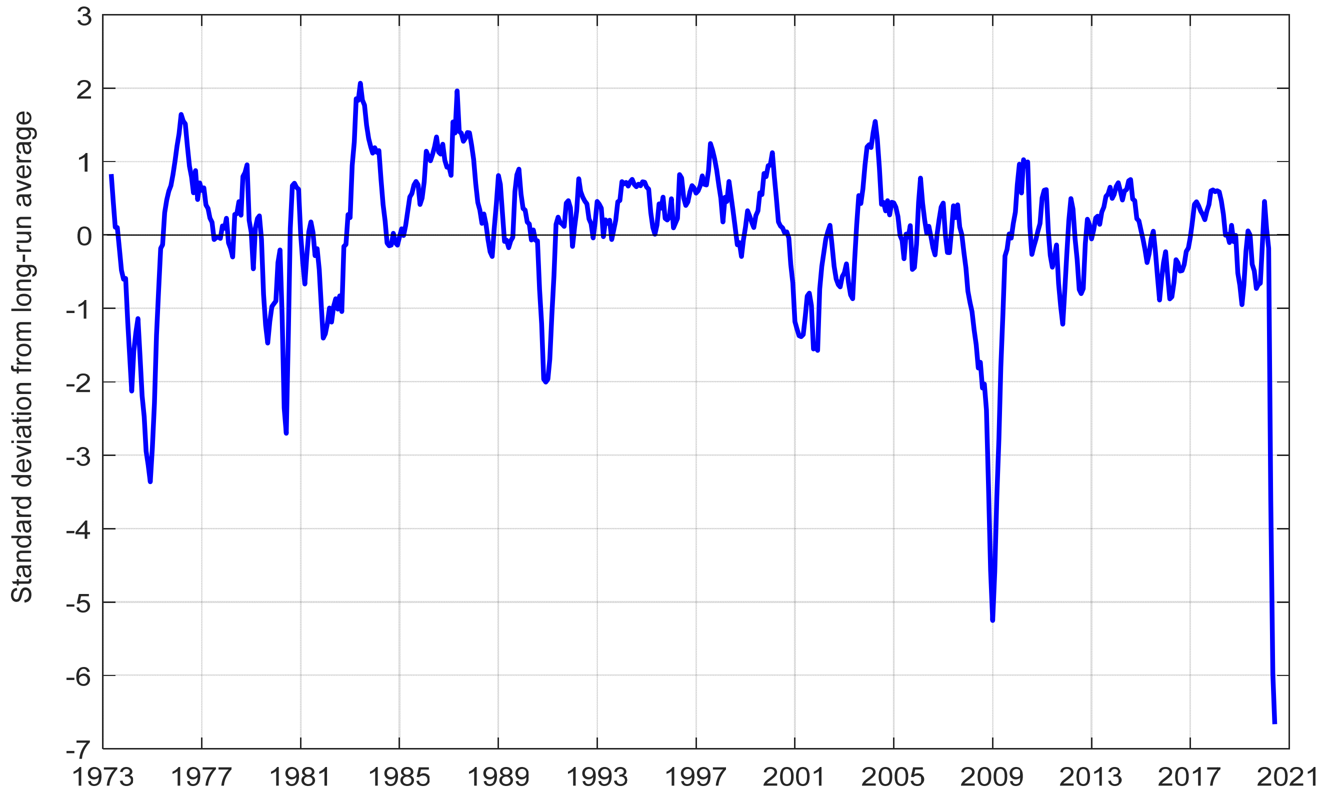
Stock, J.H., and M.W. Watson (2002). "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics* 20(2): 147-162.

Timmermann, A. (2006). "Forecast Combinations," in: G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 1, Amsterdam: North-Holland, 135-196.

West, K.D., and K.-F. Wong (2014). "A Factor Model for Co-movements of Commodity Prices," *Journal of International Money and Finance* 42: 289-309.

Figure 1: Global Economic Conditions Indicator, 1973.2-2020.5 (panel A) and its main contributors over the past year (panel B), 3-month moving average

Panel A



Panel B

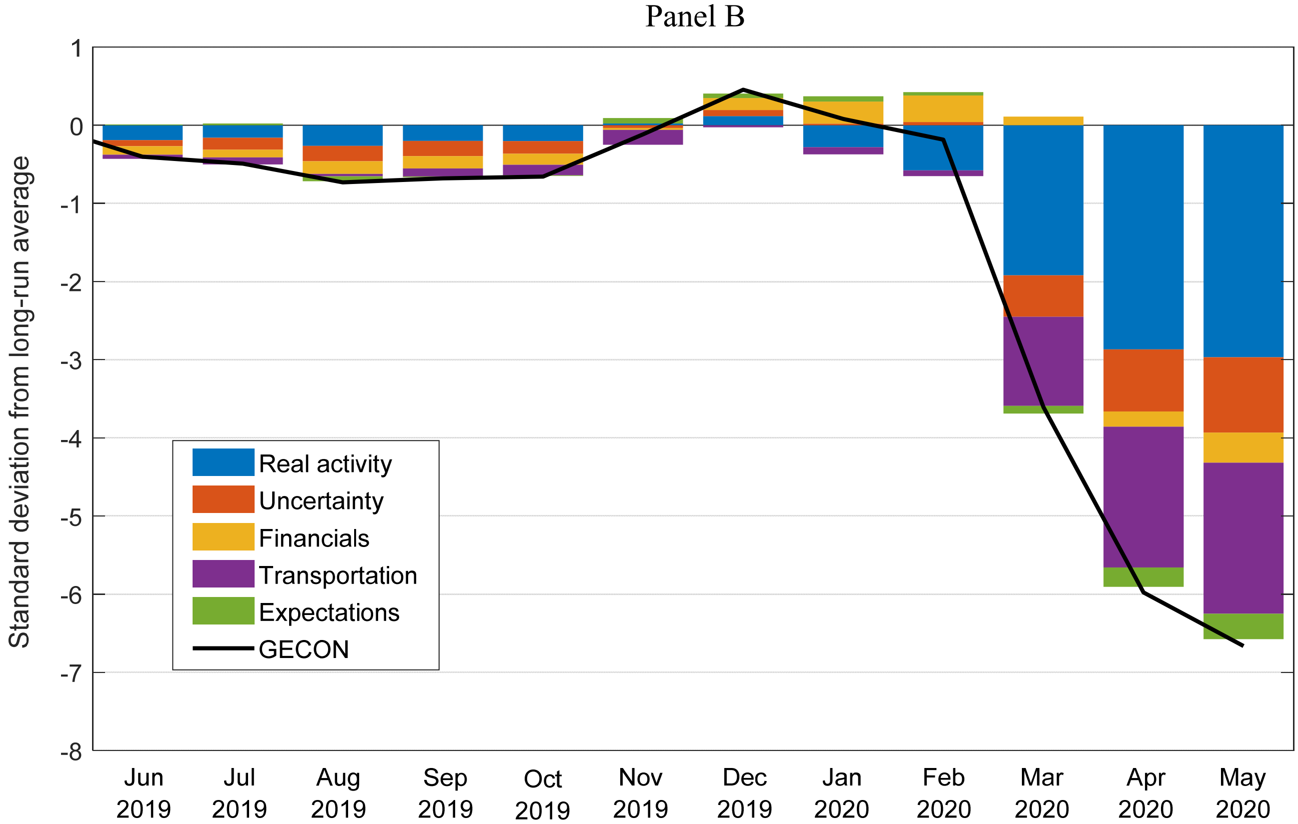
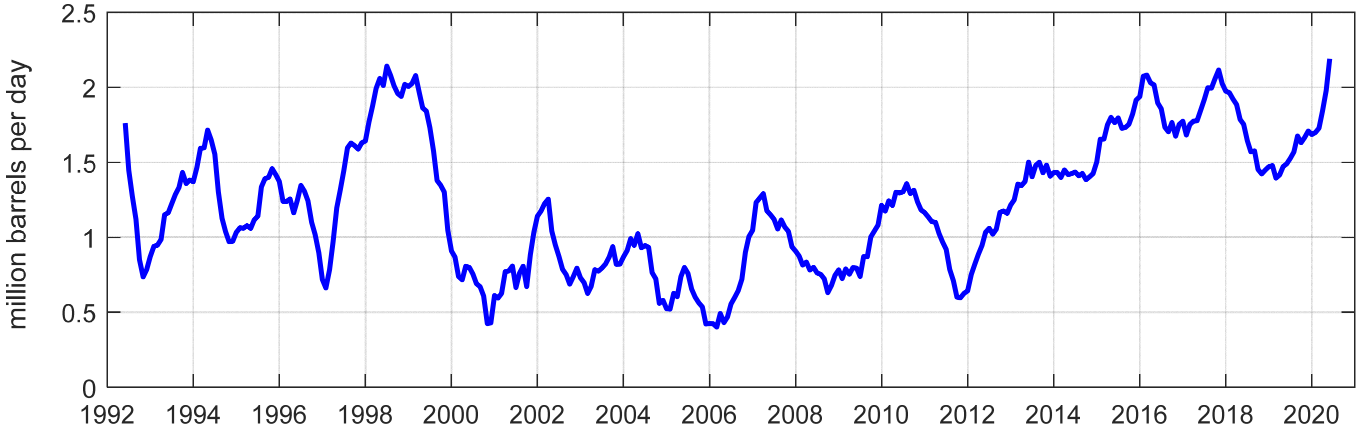
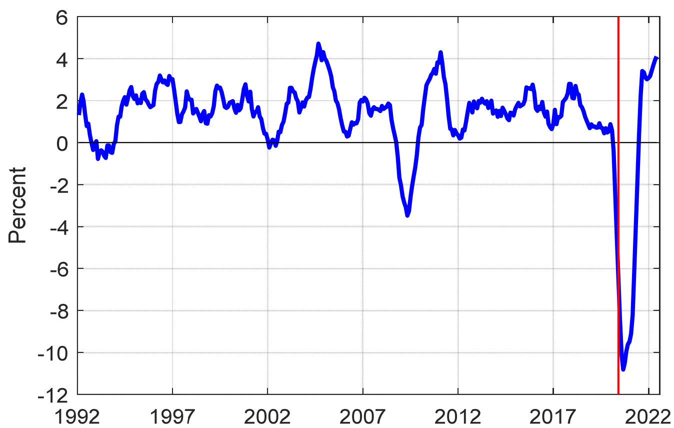


Figure 2: Energy demand assessment

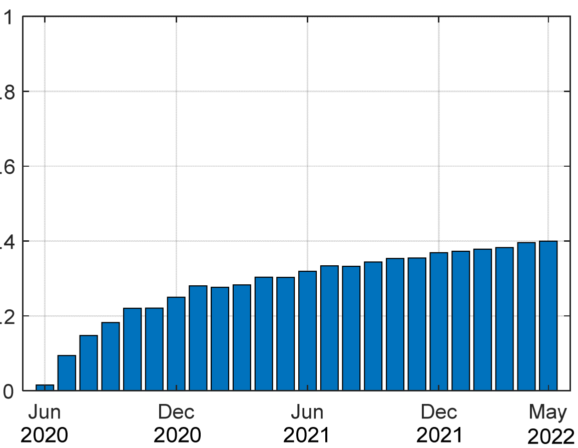
Panel A: Energy demand indicator, 1992.1-2020.5



Panel B: Annual growth rate of global petroleum consumption, 1992.1-2022.5

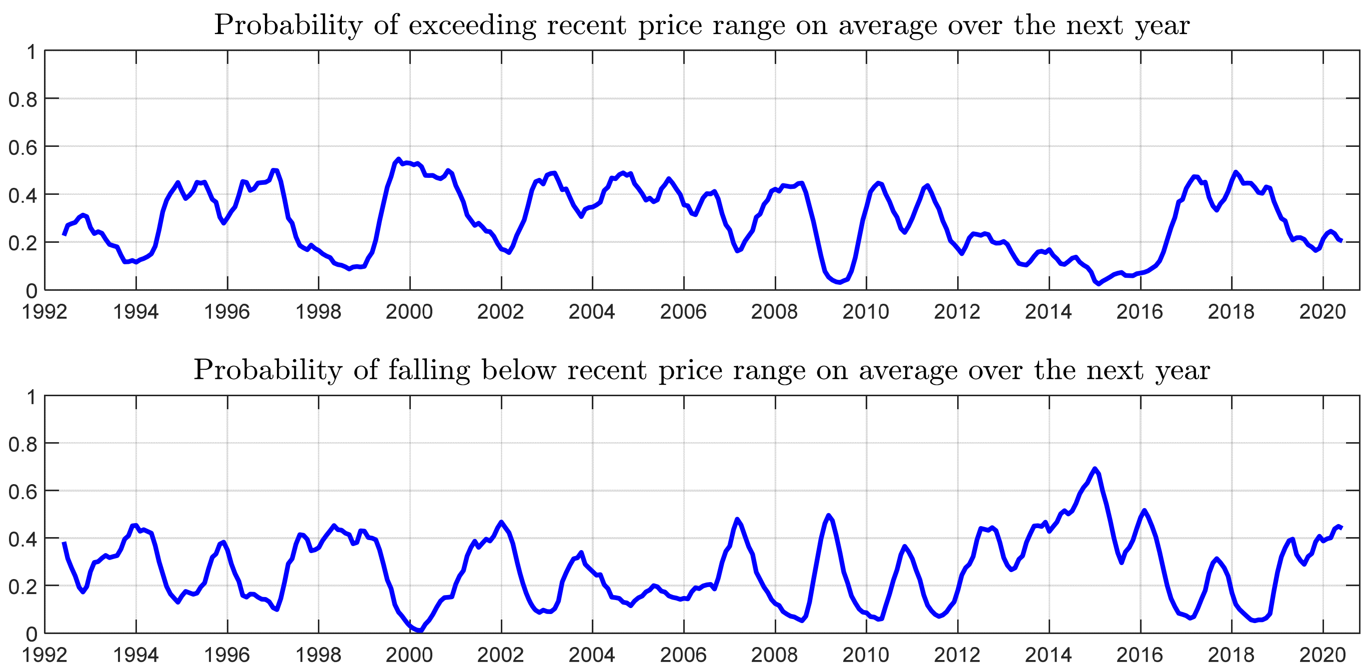


Panel C: Probability that global petroleum consumption exceeds pre-crisis level as of May 2020



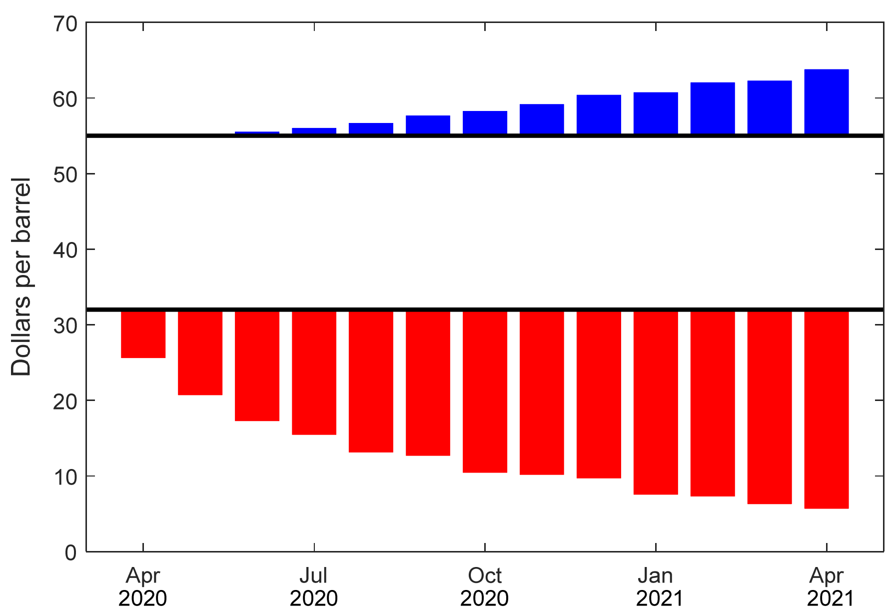
NOTES: Panel A: The indicator of expected demand pressures is computed as the difference between the 13-month-ahead and the 1-month-ahead forecast of global petroleum consumption. The plot shows the six-month moving average. Panel B: The vertical line indicates May 2020. The plot shows the six-month moving average.

Figure 3: Oil price pressure measures, 1992.1-2020.5



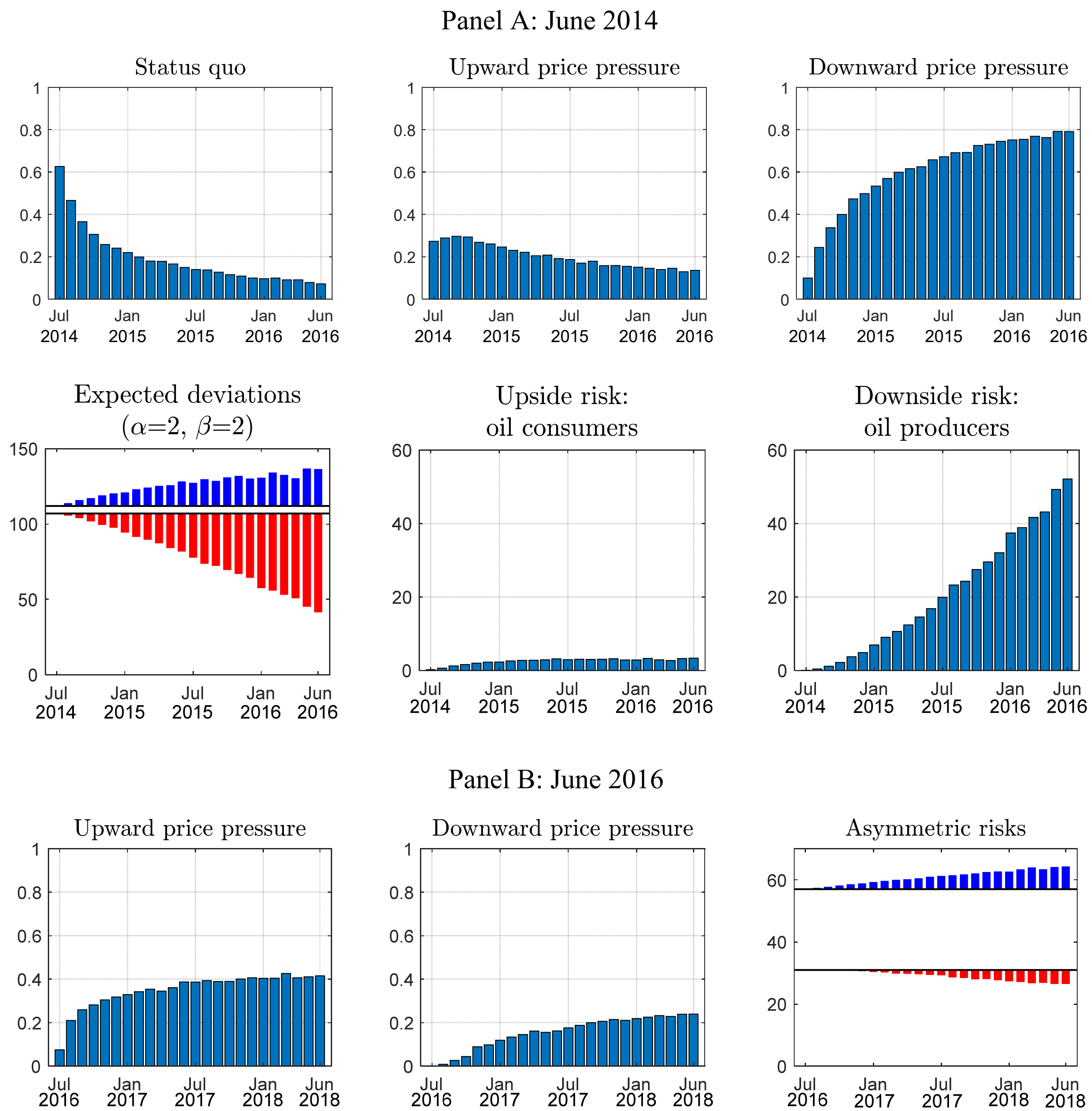
NOTES: The recent price range is determined by the minimum and maximum Brent price during the past year. The plots show the six-month moving average.

Figure 4: Risk assessment of shale oil producers as of March 2020



NOTES: The upper blue bars show the expected excess ($\alpha = 1$) and the lower red bars show the semi-variance downside risk ($\beta = 2$). The black horizontal lines indicate the breakeven price (\$55) and the March 2020 price (\$32).

Figure 5: Expected oil price pressures and risks over 24 months for selected historical episodes



NOTES: Status quo refers to the probability of the oil price staying within the price range over the past year, upward price pressure indicates the probability of exceeding the maximum price over the past year, and downward price pressure indicates the probability of falling below the minimum price over the past year. Asymmetric risks indicate different risk aversion parameters ($\alpha = 1$ for upside risk and $\beta = 2$ for downside risk).

Table 1. Recursive MSPE Ratios Relative to No-Change Forecast of Real Oil Prices in VAR(12) Models with Alternative Monthly Indicators of Global Real Economic Activity

Monthly horizon	**Kilian index (REA)**	**Real shipping cost factor**	**World IP index (WIP)**	**Real commodity price factor**	**Global steel production factor**
	(1)	(2)	(3)	(4)	(5)
	(a) Real refiner acquisition cost of crude oil imports: 1992.1-2010.6				
1	**0.679****	**0.703****	**0.657****	**0.723****	**0.694****
3	**0.779**	**0.806**	**0.728**	**0.780**	**0.776**
6	**0.989**	**0.956**	**0.867**	**0.912**	**0.925**
12	1.099	**0.992**	**0.975**	1.021	**0.940**
24	1.122	1.071	1.095	**0.995**	1.054
	(b) Real refiner acquisition cost of crude oil imports: 1992.1-2018.8				
1	**0.865**	**0.804***	**0.765****	**0.781****	**0.795****
3	**0.955**	**0.911**	**0.852**	**0.841**	**0.911**
6	1.074	1.008	**0.972**	**0.933**	1.025
12	1.159	1.031	1.069	1.038	1.045
24	1.045	1.000	**0.941**	**0.925**	**0.946**

NOTES: Boldface indicates improvements relative to no-change forecast. ** denotes significance at the 5% level and * at the 10% level based on the Diebold-Mariano test. Red indicates the best model among the shipping-based indices and blue the best model among the three alternative indicators.

Table 2. Recursive MSPE Ratios Relative to No-Change Forecast of Real Brent Price in VAR(12) and Bayesian VAR(12) Models

Evaluation Period: 1992.1-2018.8

Monthly horizon	**Kilian index (REA)**		**Real shipping cost factor**		**World IP index (WIP)**		**Real commodity price factor**		**Global steel production factor**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Production-based models									
	VAR	BVAR	VAR	BVAR	VAR	BVAR	VAR	BVAR	VAR	BVAR
1	1.075	**0.983**	**0.998**	**0.930**	**0.946**	**0.893***	**0.961**	**0.896***	**0.997**	**0.934**
3	1.072	1.063	1.027	**0.965**	**0.953**	**0.910**	**0.970**	**0.918**	1.044	**0.983**
6	1.172	1.158	1.087	1.003	1.060	**0.972**	1.021	**0.967**	1.105	1.032
12	1.215	1.237	1.045	**0.974**	1.070	**0.971**	1.037	**0.968**	1.052	**0.983**
24	1.095	1.092	1.019	**0.927**	**0.938**	**0.922**	**0.922**	**0.898**	**0.944**	**0.927**
	(b) Consumption-based models									
	VAR	BVAR	VAR	BVAR	VAR	BVAR	VAR	BVAR	VAR	BVAR
1	1.078	**0.964**	**0.986**	**0.918***	**0.932**	**0.884****	**0.943**	**0.888****	**0.951**	**0.904***
3	1.075	1.045	**0.984**	**0.942**	**0.892**	**0.888**	**0.938**	**0.906**	**0.950**	**0.943**
6	1.164	1.138	1.019	**0.966**	**0.962**	**0.932**	**0.978**	**0.938**	**0.984**	**0.978**
12	1.287	1.267	1.041	**0.979**	1.033	**0.979**	1.061	**0.985**	**0.980**	**0.977**
24	1.152	1.114	1.018	**0.938**	**0.911**	**0.919**	**0.934**	**0.903**	**0.928**	**0.932**

NOTES: See Table 1. Green indicates whether the VAR or the Bayesian VAR (BVAR) performs better.

Table 3. The Role of Stochastic Volatility for the Accuracy of Recursive Forecasts of the Real Brent Price in BVAR(12) Models

Evaluation Period: 1992.1-2018.8

Monthly horizon	**Kilian index (REA)**		**Real shipping cost factor**		**World IP index (WIP)**		**Real commodity price factor**		**Global steel production factor**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Production-based models									
	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR
1	**0.983**	**0.911****	**0.930**	**0.913****	**0.893***	**0.905****	**0.896***	**0.919****	**0.934**	**0.924***
3	1.063	**0.972**	**0.965**	**0.942**	**0.910**	**0.942**	**0.918**	**0.954**	**0.983**	**0.952**
6	1.158	1.024	1.003	**0.949***	**0.972**	**0.966**	**0.967**	**0.963**	1.032	**0.972**
12	1.237	1.046	**0.974**	**0.899****	**0.971**	**0.910***	**0.968**	**0.913****	**0.983**	**0.906***
24	1.092	**0.827**	**0.927**	**0.765****	**0.922**	**0.768****	**0.898**	**0.767****	**0.927**	**0.762****
	(b) Consumption-based models									
	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR	BVAR	SV-BVAR
1	**0.964**	**0.920***	**0.918***	**0.907****	**0.884****	**0.905****	**0.888****	**0.911****	**0.904***	**0.912****
3	1.045	**0.958**	**0.942**	**0.909**	**0.888**	**0.918***	**0.906**	**0.939***	**0.943**	**0.929**
6	1.138	1.018	**0.966**	**0.910****	**0.932**	**0.926***	**0.938**	**0.943***	**0.978**	**0.925***
12	1.267	1.072	**0.979**	**0.874****	**0.979**	**0.869****	**0.985**	**0.887****	**0.977**	**0.864****
24	1.114	**0.821**	**0.938**	**0.719****	**0.919**	**0.710****	**0.903**	**0.718****	**0.932**	**0.710****

NOTES: See Table 1. Green indicates whether the BVAR or the BVAR with stochastic volatility (SV-BVAR) performs better.

Table 4. Recursive MSPE Ratios Relative to AR(12) Forecast for Global Petroleum Consumption

Evaluation Period: 1992.1-2018.8

	Monthly horizon				
	1	3	6	12	24
(a) BVAR(12)					
Kilian index (REA)	1.103	1.445	2.062	3.131	4.122
Real shipping cost factor	1.046	1.134	1.310	1.521	2.055
World IP index (WIP)	1.002	1.252	1.511	1.635	2.043
Real commodity price factor	1.054	1.173	1.335	1.523	1.977
Global steel production factor	1.255	1.609	1.794	2.007	2.553
(b) SV-BVAR(12)					
Kilian index (REA)	**0.968**	1.073	1.213	1.369	1.645
Real shipping cost factor	**0.962***	1.036	1.064	1.114	1.208
World IP index (WIP)	**0.933****	**0.984**	1.009	1.013	1.052
Real commodity price factor	**0.949****	1.017	1.070	1.154	1.304
Global steel production factor	**0.954****	1.026	1.059	1.095	1.141
(c) Model-based pooling using SV-BVAR(12)					
Equal-weighted forecast combination (5 models)	**0.946****	1.016	1.066	1.128	1.258
Dynamic model selection (5 models)	**0.946****	1.014	1.059	1.112	1.212
Factor from disaggregated data of all existing indices	**0.955***	1.028	1.072	1.129	1.266

NOTES: Boldface indicates improvements relative to AR(12) forecast. ** denotes significance at the 5% level and * at the 10% level based on the Diebold-Mariano test.

Table 5. The Role of Different Information Sets for the Accuracy of Recursive Forecasts of the Real Brent Price and Global Petroleum Consumption in Bayesian VAR(12) Models with Stochastic Volatility Evaluated over 1992.1-2018.8

		Monthly horizon				
		1	3	6	12	24
		(a) Real Brent Price				
(1)	Global Economic Conditions Indicator	**0.918****	**0.930**	**0.940**	**0.876***	**0.704****
	Large dataset					
(2)	1 factor	**0.922****	**0.939**	**0.947**	**0.886***	**0.715****
(3)	2 factors	**0.939***	**0.947***	**0.955**	**0.934**	**0.756****
(4)	3 factors	**0.939***	**0.967**	**0.982**	**0.976**	**0.857****
(5)	Statistical variable selection	**0.915****	**0.925**	**0.930***	**0.854***	**0.686****
		(b) Global Petroleum Consumption				
(6)	Global Economic Conditions Indicator	**0.934****	**0.945**	**0.926**	**0.875**	**0.884**
	Large dataset					
(7)	1 factor	**0.921****	**0.973**	1.005	**0.936**	**0.949**
(8)	2 factors	**0.926****	**0.981**	1.048	1.156	1.461
(9)	3 factors	**0.937****	**0.985**	1.076	1.180	1.530
(10)	Statistical variable selection	**0.946****	**0.997**	1.004	1.023	1.045

NOTES: Boldface indicates improvements relative to no-change forecast (panel a) or AR(12) forecast (panel b). ** denotes significance at the 5% level and * at the 10% level based on the Diebold-Mariano test. The Global Economic Conditions Indicator is based on 16 variables covering different dimensions of the global economy as they relate to energy markets (see Table 6A). The large dataset contains 256 variables (see Tables 1A, 6A and 8A for details). The statistical variable selection uses the 16 variables with the highest loadings on the factor extracted from the large dataset (see appendix A for the list of selected variables).

Table 6. Model Confidence Sets for the Joint Density Forecasts of Real Brent Price and Global Petroleum Consumption

Evaluation Period: 1992.1-2018.8

No.	Models	Monthly horizon									
		1		3		6		12		24	
		ALS	p_{MCS}	ALS	p_{MCS}	ALS	p_{MCS}	ALS	p_{MCS}	ALS	p_{MCS}
	BVAR(12)										
(1)	Kilian index	-3.578	0.000	-4.730	0.000	-5.556	0.030	-6.517	0.038	-7.217	0.034
(2)	Shipping factor	-3.559	0.000	-4.665	0.000	-5.433	0.039	-6.283	0.057	-6.973	0.047
(3)	WIP	-3.529	0.007	-4.634	0.002	-5.443	0.067	-6.287	0.050	-6.941	0.080
(4)	Commodity factor	-3.540	0.001	-4.635	0.000	-5.412	**0.100***	-6.293	0.095	-6.955	**0.130***
(5)	Steel factor	-3.561	0.000	-4.646	0.000	-5.415	0.047	-6.248	0.069	-6.940	0.059
	SV-BVAR(12)										
(6)	Kilian index	-3.289	**0.234***	-4.300	0.040	-5.092	**0.302****	-5.815	**0.574****	-6.521	**0.916****
(7)	Shipping factor	-3.274	**0.561****	-4.269	0.073	-5.057	**0.347****	-5.812	**0.349****	-6.550	**0.429****
(8)	WIP	-3.258	**0.881****	-4.249	**0.509****	-5.049	**0.663****	-5.796	**0.574****	-6.535	**0.727****
(9)	Commodity factor	-3.262	**0.855****	-4.251	**0.509****	-5.044	**0.706****	-5.816	**0.424****	-6.550	**0.663****
(10)	Steel factor	-3.278	**0.377****	-4.269	0.056	-5.061	**0.302****	-5.820	**0.217***	-6.560	**0.278****
(11)	Equal-weighted pooling	-5.437	0.000	-5.060	0.020	-5.835	**0.216***	-6.594	**0.179***	-7.091	**0.262****
(12)	Dynamic model selection	-3.296	**0.124***	-4.312	0.020	-5.102	**0.268****	-5.817	**0.445****	-6.523	**0.916****
(13)	Factor from data of existing indices	-3.265	**0.828****	-4.268	**0.105***	-5.058	**0.325****	-5.821	**0.286****	-6.555	**0.549****
(14)	GECON	-3.255	**0.881****	-4.239	**1.000****	-5.033	**0.722****	-5.786	**0.574****	-6.541	**0.688****
(15)	1 factor	-3.249	**1.000****	-4.250	**0.473****	-5.028	**1.000****	-5.766	**1.000****	-6.508	**1.000****
(16)	2 factors	-3.255	**0.881****	-4.262	**0.473****	-5.056	**0.706****	-5.804	**0.574****	-6.539	**0.859****
(17)	3 factors	-3.266	**0.813****	-4.274	**0.157***	-5.110	**0.257****	-5.863	**0.286****	-6.596	**0.337****
(18)	Statistical selection	-3.267	**0.738****	-4.260	**0.175***	-5.051	**0.429****	-5.800	**0.445****	-6.550	**0.429****

NOTES: Average log score (ALS) and MCS p-values (p_{MCS}). The MCS p-values are computed with 10,000 block bootstrap replications using a block size of 10.

The forecasts in the MCS are in bold and $\widehat{M}^{*}_{90\%}$ and $\widehat{M}^{*}_{75\%}$ are identified by one and two asterisks, respectively. Red indicates the highest-ranked models.