Oil Price Forecasts for Macroeconomic Projections: Experts Outlooks, Models, or Both?

Jean-Thomas Bernard
University of Ottawa

Lynda Khalaf
Carleton University

Maral Kichian*
University of Ottawa

Clement Yelou
Statistics Canada

12 August, 2014

Abstract

Releases of expert outlook publications regarding the future path of oil prices attract much media attention and are considered highly useful by industry participants and policymaking bodies who often rely on these sources for their forecasting needs. Yet little or no attention has been paid to the extent to which these are accurate. Using the regular publications by the Energy Information Administration (EIA), we examine the worth of annual oil price forecasts over the medium and long runs generated by the formal model (NEMS) of the Agency.

Our results reveal that NEMS is quite successful at beating the benchmark random walk model, but only at either end of the fifteen year forecast horizon. Based on the traditional mean square error forecast criterion, the EIA model outperforms the no-change model at the one-year-ahead horizon, and at horizons of nine-years-ahead or more. We also show that simple statistical or equilibrium-based models often produce similar if not better forecast performances than the EIA model, specially for forecast horizons of six years and greater. Finally, we find that combining forecasts does not change the overall obtained patterns of results from individual comparisons, but that some gains are made over the medium term forecasts in the extent to which they outperform the random walk.

JEL classification:

Keywords:

*This work was supported in part by the Social Sciences and Humanities Research Council of Canada.

1. Introduction

It is difficult to overstate the widespread and heavy reliance on oil by developed and developing countries around the world. Oil use permeates practically every sector of the economy, affecting both consumers and producers. As a result, fluctuations in the price of oil influence the economy as a whole, sometimes with large negative effects[1]. This also means that oil price is among a handful of indicators that can be used to predict future movements in sectoral and total real GDP. Consequently, industry participants, policy-makers, and various international organizations scrutinize the oil price and invest considerable resources on forecasting its future evolution.

In this respect, particular attention is paid to medium and long term forecasting. For instance, central Banks, government agencies and international organizations all predicate their policy decisions and regulatory recommendations on macroeconomic projections that directly depend on a given assumed multi-year future path for real oil prices. Similarly, industry participants make use of forecasts of real oil prices five, ten, or even fifteen years ahead when analyzing medium or long-term strategies and investments. Consequently, it is desirable to have the most accurate forecasts possible, as inaccurate assessments of the oil price path could mean wrong predictions about key macroeconomic outcomes, as well as wasteful and unproductive investments, both leading to potentially large welfare costs to society.

To obtain their forecast paths, industry and policymaking bodies have traditionally (and predominantly) resorted to futures prices and to survey or expert outlook forecasts[2]. So it is important to know how accurate these forecasts have been. Yet while the forecasting worth of futures prices has been extensively studied (see, for instance, Bernard, Khalaf, Kichian, and MacMahon (2014) for a recent reference), little or no attention has been paid to the extent to which expert and/or survey approaches are helpful in predicting oil prices[3].

In this paper, we therefore examine the worth of expert outlook forecasts over the medium and long terms. For concreteness we focus on the Energy Information Administration (EIA) agency forecasts. The agency introduced and adopted a formal model, National Energy Modeling System (NEMS), and has, since 1995, been publishing annual forecasts based on this model for up to (at least) fifteen years ahead[4]. NEMS provides

[1]There is a large literature on this topic; examples of studies documenting the impact of oil prices on the recent 2007-2008 recession include Hamilton (2009) and Edelstein and Kilian (2009).

[2]See Alquist et al. (2013) for some examples.

[3]There are a few exceptions, including Alquist et al. (2013) and Baumeister and Kilian (2014a), but these focus only on short term and high-frequency data forecasts.

[4]The EIA stemmed from the first oil shock in 1973 when concerns were expressed about the quality of the energy statistics that were available to the U.S. government to support analysis and formulation of energy policies. Building on the organization created in 1974, the Department of Energy Organization Act of 1977 set up EIA as the main office that is responsible for the collection and analysis of energy data and the study of energy policy. Although EIA is part of the U.S. Department of Energy, it is independent of the Administration with respect to publication and analysis and it does not support nor make policy recommendations. Data and publications are readily available on its website. After a first publication in 1979, EIA promoted its Annual Energy Outlook every year since 1982. The early release takes place in December and a widely attended conference is held in Washington D.C. in the spring. Forecasts are presented on production, imports, consumption and energy prices over the next fifteen or twenty-five

a natural reference benchmark for our question of interest. In addition to studying the extent to which NEMS forecasts are accurate, this paper also fills a gap in the literature regarding the worth of simple equilibrium-based and statistical models for generating useful forecasts over the medium and long terms.

Specifically, our analysis considers out-of-sample forecasts of the real oil price one to fifteen years ahead using annual frequency data. For the latter we update the dataset originally used by Manthy (1978) and later by Pindyck (1999). This data starts in 1870, providing us with the longest dataset available. Since NEMS forecasts are available only since 1995, our forecast evaluation period for NEMS starts in 1995. For completion, we also analyze the forecasting performance of some much simpler models, both statistically and structurally-motivated, including the random walk. These are examined alongside NEMS for the same evaluation period, but also for a pre-NEMS evluation sample. We use two different forecast criteria to make our comparisons. These are the mean square forecast error (MSFE) and the mean absolute percent error (MAPE); the former being the criterion routinely-used in the literature, and the latter being the U.S. Department of Energy EIA agency's criterion of choice. Finally, given that some models are likely to forecast better over more immediate horizons, while others would do better over further ones, we also document the worth of using forecast combination approaches for medium and long term forecasting. Four different forecast combination methods are examined for various model groupings.

Our results reveal that NEMS is particularly successful at beating the random walk model at the one-year-ahead horizon, with a 49 per cent gain in forecasting accuracy according to the MSFE, and a 26 per cent gain according to the MAPE. Indeed, it is the only model among our many alternatives to outclass the random walk at this forecast horizon. Moreover, based on the MAPE criterion, NEMS also seems to have an advantage over the no-change forecast (of about 9 per cent) two-years-ahead, although this is reversed when the MSFE is used instead. We also find that the model is able to outperform the random walk at much longer forecast horizons (starting at nine-years-ahead horizon according to the MSFE, and at six-years-ahead based on the MAPE), but that many of the simple models that we consider are also successfu at these horizonsl, often with much better forecast performances. Furthermore, NEMS cannot beat the no-change model for forecasts of three to five years ahead according to MAPE, and for forecasts of two to eight years ahead according to MSFE. Finally, we find that combining forecasts does not change the overall obtained patterns of results from individual comparisons, but that some improvement is made over the medium term forecasts in the extent to which they outperform the random walk.

years. A key variable that draws a lot of interest is the oil price forecast. This is a rather unique exercise conducted by an independent and stable organization that has delivered a sequence of annual oil price forecasts over a period of time. A brief description of NEMS is provided below; the agency's website http://www.eia.gov/oiaf/aeo/overview/ provides continuously updated information.

2. Model Considerations

Nowadays, expert outlooks on oil prices are abundant, sponsored by *e.g.* OPEC, British Petroleum, and the U.S. Department of Energy's Energy Information Administration agency. These are frequently consulted and talked about in the media, and the forecasts that are provided are often used by private sector participants and by policymaking bodies. Yet little is known about the worth of these forecasts or how they compare to ones generated from simple economic or statistical models. From the literature on oil price forecasting, we know that it is difficult, although not impossible, to find forecasting models that can outperform a naive no-change forecast[5]. Indeed even futures prices, which are the most commonly-used forecasts in both the private sector and policy-making circles, often come second compared to the no-change forecasting model[6]. Hence, it is important to first examine whether expert outlooks are more successful than the random walk model forecasts.

Moreover, it is reasonable to ask how these expert forecasts fare against other model formulations. While the details are unavailable to the public at large, and based on the supplied information on the website, NEMS appears to be a rich and seemingly complex model that is aimed at integrating a myriad of worldwide influences on oil price related to oil production and demand. Such a model presents clear advantages over simplistic specifications that can capture only a few channels of impact on the price. Yet these advantages may quickly turn to disadvantages when the large number of models parameters need to be quantified from relatively short samples of data, when there are unknown (and therefore unmodelled) structural changes affecting some of the postulated relationships, or when various geopolitical considerations, unrelated to economic fundamentals, actually drive the price. In this case it is natural to wonder how expert forecasts compare with forecasts obtained from much simpler and potentially more flexible models. Indeed, depending on world circumstances, fully structural models such as NEMS might be more successful a proportion of the time, with the simpler models doing better at other times. Such considerations then leave the door open to forecast combination possibilities, based on predictions from both types of models, for superior overall performances.

In this section we define the models that we use for the above-mentioned comparisons. It is of course impossible to consider all of the proposed models in this vast literature, and thus we can select some illustrative specifications from various popular model classes. In all cases, we consider only parsimonious and single-equation-based specifications, on the one hand to accommodate the smooth nature of our annual data (which would be conceivably better represented with simple formulations), and on the other to minimize the number of regressors (as the latter themselves need to be forecasted to be able to forecast the oil price). Finally, simple models are likely to be relatively more robust to structural breaks than more parameterized highly-restricted models.[7]. In the final analysis, we consider structurally-motivated linear and non-linear equations, as well as purely statistical formulations.

[5] See, notably, the survey provided by Alquist et al. (2013) on the topic.

[6] See, for instance, Alquist and Kilian (2010) and Chinn and Coibion (2013).

[7] See, for example, Dvir and Rogoff (2010) and Bernard, Dufour, Khalaf, and Kichian (2012) on structural changes in oil prices.

Structurally-Motivated Forecasting Models

The overwhelming majority of models that describe the behaviour of oil prices over the long run are based on structural foundations, generally stemming from the Hotelling (1930) premise regarding the evolution of an exhaustible resource. The models are thus based on worldwide fundamental supply and demand conditions, and often also take into account the respective roles of inventory management, uncertainty regarding future oil discovery, and the presence of other energy alternatives. We thus consider two classes of structural-based models for forecasting, one that includes models that are linear in the parameters (denoted Structural-Linear), and another where the model parameters evolve in a nonlinear fashion (denoted Structural-NonLinear).

Within the linear-parameter structural-based category we consider two forecasting equations based on Slade (1982). The theoretical setup underlying these specifications describes long-run price movements for a non-renewable natural resource accounting for exogeneous technical changes and endogenous changes in the grade of the unrefined material. Under some assumptions, the model implies that price will equal marginal extraction cost plus rent, where the rate of change of the price is equal to the rate of change of marginal cost due to changes in technology plus the discount rate times rent. Without technical change, prices can thus increase with time but when the rate of technical change is sufficiently large, prices can fall. Slade (1982) proposes two econometric versions of this model, with one specification that includes a constant and a linear trend (denoted **LT**), and another that also adds a quadratic trend to the previous formulation (denoted **QT**). The equations are described below (starting with the more general specification), and where P_t refers to the logarithm of real price and ϵ_t to random disturbances.

$$P_t = c_1 + c_5 t + c_6 t^2 + \epsilon_t, \quad t = 1, ..., T \tag{1}$$

and its restricted counterpart (**LT**)

$$P_t = c_1 + c_5 t + \epsilon_t, \quad t = 1, ..., T \tag{2}$$

obtained by imposing

$$H_{LT} : c_6 = 0, \tag{3}$$

As for the structurally-motivated nonlinear-parameter model category, we consider the specifications proposed by Pindyck (1999) which builds on a basic Hotelling model for a depletable resource produced in a competitive market. The model postulates a price equation based on a constant marginal cost of extraction and a unit elasticity for isoelastic demand, and where changes in demand, extraction costs, and reserves all affect the slope of the price level. Pindyck (1999) argues that these factors fluctuate in a continuous and unpredictable manner over time, implying that long-run energy prices should revert to a trend (and the long-run marginal cost) which itself fluctuates in the same fashion. A class of models which integrates the above features is the generalized Ornstein-Uhlenbeck process, and Pindyck (1999) proposes a discretized version of this model as a suitable framework for analyzing long-run energy prices. The resulting econometric specifications are autoregressive models with trend, where model parameters on the constant and trend are potentially time-varying.

We thus consider three versions of this econometric framework. The first, denoted **TVP-IS**, is the most general model, allowing both the intercept and the slope to evolve over time according to random walk processes. The second, denoted the **TVP-I** model, restricts the trend coefficient to be fixed, but continues to allow the intercept to vary over time according to a random walk. Finally, the **TVP-S** model restricts the intercept to a constant, but continues to allow the trend coefficient to vary over time according to a random walk. The three specifications are given below as follows:

$$P_t = c_1 + \phi_{1t} + c_2 P_{t-1} + c_5 t + \phi_{2t} t + \epsilon_t, \qquad t = 1, \ldots, T, \tag{4}$$
$$P_t = c_1 + \phi_{1t} + (c_5 + \phi_2)\, t + c_2 P_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T, \tag{5}$$
$$P_t = (c_1 + \phi_1) + c_5 t + \phi_{2t} t + c_2 P_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T. \tag{6}$$

and where

$$\phi_{1t} = c_3 \phi_{1,t-1} + v_{1t}\,, \qquad t = 1, \ldots, T,$$
$$\phi_{2t} = c_4 \phi_{2,t-1} + v_{2t}\,, \qquad t = 1, \ldots, T.$$

The disturbances ϵ_t, v_{1t}, and v_{2t}, $t = 1, \ldots, T$, are assumed to be independently and identically normally distributed with zero means and covariances, and variances σ^2_ϵ, $\sigma^2_{v_1}$, and $\sigma^2_{v_2}$, respectively.

The models in the structural linear-parameter category are estimated by ordinary least-squares, while the time-varying-parameter (TVP) specifications of the structural non-linear-parameter category are estimated by maximum likelihood using the Kalman filter. For all cases, estimation is conducted on the logarithm of real prices.

As for the NEMS model of the EIA, we rely on the publicly available annual forecasts of the nominal price of oil. To be clear, we do not conduct estimations on NEMS ourselves, which to some extent restricts our choice of forecast combination methods. Indeed, any method that requires re-running the model to derive forecasting weights is beyond the scope of this paper. While possibly restrictive from a statistical perspective, our analysis of these public forecasts does not conflict with the agency's mandate to *collect, analyze and disseminate independent and impartial energy information.* NEMS, developed in line with this mandate, is a set of integrated modules representing the energy-economy structural relationships of the U.S. and it projects up to 25 years ahead production, imports, consumption and energy prices subject to macroeconomic, demographic, technological and behavioral assumptions. Examples of data inputs include energy production, stocks, demand, imports, exports, and prices. Its main use is to support, since 1995, the flagship EIA publication Annual energy Outlook.

Statistical Forecasting Models

According to Pindyck (1999), long-run forecasting models should ideally be explained in structural terms, meaning that they should be based on supply and demand. However, he also notes that explanatory variables of such models (which are essentially the determinants of supply and demand) need to be themselves forecasted as well. An alternative approach that avoids this complication is to rely on a set of simple statistical specifications. Moreover, such models might actually be better suited to account for the various geopolitical factors that are generally not accounted for in structural models but that also influence oil price. We consider five such models in this section. These are simple autoregressive models with and without trend, and random-walk models with and without drift.

The most general specification, which we refer to as the **AR-QT** model, is the autoregressive model with linear and quadratic trends, and it is given by:

$$P_t = c_1 + c_2 P_{t-1} + c_5 t + c_6 t^2 + \epsilon_t, \quad t = 1, ..., T \tag{7}$$

Imposing, in turn, the following restrictions, we obtain two additional models. One is an autoregressive model with linear trend, which we denote as **AR-LT**, and the other is an autoregressive model without trends referred to as **AR-NT**. Formally, we have,

$$H_{AR/LT} : c_6 = 0, \tag{8}$$
$$H_{AR/NT} : c_5 = c_6 = 0. \tag{9}$$

The remaining models in this category are the random walk with drift model (**RW-WD**) and its counterpart imposing a zero drift, denoted as **ND**. Forecasts based on the latter model are referred to as no-change forecasts. These models can be obtained by imposing different restrictions on our most general statistical model (7). Thus, we have:

$$H_{RW/WD} : c_5 = 0 = c_6 = 0, \quad c_2 = 1, \tag{10}$$
$$H_{RW/ND} : c_1 = c_5 = c_6 = 0, \quad c_2 = 1. \tag{11}$$

All the statistical models considered here are linear in the parameters, however nonlinearity is present in the most general specification via the squared trend term.

Thus we have in total five statistical forecasting models, including the no-change forecast model, all of which are estimated using ordinary least-squares. Estimation is conducted on the logarithm of real prices in all cases.

3. Data and Forecast Assessment Criteria

We consider the annual dataset for the nominal price of crude oil that goes back to 1870 and that was originally used by Manthy (1978) and later updated until 1995 by Pindyck (1999) using data from the EIA and, for 1996, with data from the *Wall Street Journal*[8]. The series are deflated using the U.S. wholesale price index until 1970, and with the producers price index thereafter. We extend this series until 2011 using the EIA publications and adopting the same definitions as Pindyck. Thus we have 151 annual observations in our sample.

We place ourselves at the end of the 1994, and for each of our eleven models, we calculate out-of-sample annual forecasts from one up to fifteen years ahead. The relevant estimation sample thus covers the period 1870-1994, while the evaluation period is 1995-2011. But while the latest version of the NEMS model forecasts only start in 1995, the alternative specifications can be evaluated over a longer period. Thus for completion, and to get a sense of the consistency of the performance of such models over time, we also consider the 1985-2011 estimation period for these models, with the corresponding evaluation sample being 1985-2011. The two end of estimation sample dates correspond to very different sets of economic circumstances in the world, with very different implications for oil prices. Thus, 1984 is the year that follows a period of deep recession in many countries, at a time when growth was finally starting to gain some momentum, while 1994 is a year that preceded a period of relative tranquility in oil prices that would continue for close to a decade.

We conduct recursive estimations, whereby model parameters are updated ahead of a given forecast. In addition, where relevant, dynamic forecasts are made. That is, the forecasted values for $T+h$ are used to produce forecasts for $T+h+1$, where T is the last observation of a given estimation sample and h is the forecasting horizon. As noted above, we conduct forecasts for $h = 1, 2, \ldots, 15$ years ahead.

We also report two accuracy criteria for each of our two evaluation periods. Let $H = 2011 - T$ be the total number of years in a given evaluation period. Then, for a given forecast horizon, h, the Mean Squared Forecast Error (MSFE) is given by

$$MSFE\,(h) = \frac{1}{H-h+1}\left(\sum_{j=1}^{H-h+1}\left(\hat{y}_{T+j-1+h|T+j-1} - y_{T+j-1+h}\right)^2\right) \tag{12}$$

and the Mean Absolute Percent Error (MAPE) is given by

$$MAPE\,(h) = 100 \times \frac{1}{H-h+1}\left(\sum_{j=1}^{H-h+1}\left(\frac{\left|\hat{y}_{T+j-1+h|T+j-1} - y_{T+j-1+h}\right|}{y_{T+j-1+h}}\right)\right).$$

In the above, y stands for the observed price and $\hat{y}$ for the forecasted one[9].

The MSFE, which is routinely used in the literature, can be linked to a quadratic loss principle and has been analyzed in statistics from an inferential perspective. While less common, the $MAPE$ criterion has the advantage that the errors can directly be interpreted as percentage errors. It is used, for example, by EIA (notably in the EIA Annual Energy Outlook publications).

[8]The data were generously provided by Pindyck.

[9]For our econometric models, we forecast the logarithm of the price and then transform it into the price level prior to calculating forecast errors.

4. Empirical Analysis

Having calculated the MSFE and the MAPE over a given evaluation period for each of the models and for each forecast horizon, the percentage deterioration or improvement is calculated for every model relative to the no-change case. Negative values indicate better forecasting performance compared to the random walk model, while positive values point to worse outcomes. While statistical comparisons are unfortunately infeasible for NEMS, the orders of magnitude that we obtain nevertheless reflect the economic importance of the relative results. The 1995-2011 results are collected in Tables 1 and 2, showing the relative performances for the MSFE, and the MAPE, respectively.

One thing that is immediately apparent from the Tables is that many of the models, including NEMS, are able to outperform the no change model (these cases are indicated in bold in the Tables), sometimes with considerable margins. This result suggests that, despite the extensive volatility in oil prices, there is a fairly important systematic component in the dynamics of this variable that can be captured to different extents by all the classes of models examined. We will see later on that this result is not specific to the particular sample examined, but that it is generally upheld for the sample starting in 1985 as well.

Furthermore, while similar overall patterns can be observed looking at either the MSFE or the MAPE results, the former stresses more the evidence regarding the extent to which the various models outperform the no-change forecast, while the latter stresses their frequency of success over the fifteen forecast horizons.

As for the performance of NEMS in particular, we find that the model is specially successful at forecasting one-year-ahead. Whether based on the MSFE or the MAPE, indeed it is the only model to outperform the random walk at this horizon, and by relatively important margins. This outcome could be due to a number of factors, including the fact that fundamental demand for and supply of oil are already largely fixed, and therefore known, some months ahead of delivery dates, and because judgment can be applied as needed to temper these numbers when short term geopolitical disruptions with largely foreseeable impacts occur. According to the MAPE criterion, NEMS also outperforms the no-change forecast two-years-ahead, though we find that a number of other models are also able to do the same and to a similar extent. Interestingly, the model is not particularly useful at the medium forecasting horizons, notably for three to about eight-year-ahead forecasts, whereas some of the other models fare better, specially towards the upper end of this range. Beyond these horizons, we find that NEMS is able to outperform the random walk again, but so do all the other models that we considered, sometimes by a considerable margin.

Interestingly, the fact that the random walk model can be beaten has also been found in the context of short-term forecasting and with higher frequency data, though only for specific models. For instance, Baumeister and Kilian (2014b) show that, to obtain quarterly forecasts of the real price of oil from one up to to four quarters, it is preferable to use monthly-data-based VARs and futures-based models since these can outperform no-change quarterly forecasts. Similarly, Baumeister, Kilian, and Zhou (2013) show that, forecasts obtained from some models based on product spreads, with weights on the different spreads that are allowed to vary over time, are able to outperform the no-change model up to 24 months ahead.[10]

[10]See also Baumeister, Guerin, and Kilian (2014a) and Baumeister, Kilian, and Lee (2014b).

Table 1: Relative MSFE, Evaluation Period: 1995-2011

Fcst Hzn	Statistical Models				Structural-Linear		Structural-Nonlinear			
	RW-WD	AR-NT	AR-LT	AR-QT	LT	QT	TVP-IS	TVP-I	TVP-S	NEMS
1	40.96	69.53	52.47	18.54	254.18	137.27	43.76	32.75	21.23	**-48.65**
2	**-16.36**	31.72	15.21	**-19.76**	148.87	70.13	19.89	11.34	22.85	5.20
3	45.63	125.43	81.67	20.96	181.28	97.35	54.16	60.02	45.53	55.62
4	19.40	84.64	40.73	**-3.72**	75.56	25.35	40.06	31.40	36.99	36.29
5	14.98	69.70	22.04	**-16.12**	31.56	**-10.42**	27.65	14.78	25.26	27.36
6	11.46	53.53	6.16	**-25.12**	3.17	**-32.78**	14.89	1.98	12.58	15.00
7	1.44	35.57	**-7.67**	**-32.33**	**-14.42**	**-44.85**	3.68	**-10.04**	1.40	11.26
8	3.21	30.11	**-13.02**	**-36.47**	**-22.56**	**-54.29**	**-2.09**	**-16.08**	**-4.57**	1.83
9	**-4.34**	20.15	**-20.23**	**-41.36**	**-30.54**	**-63.79**	**-8.08**	**-20.50**	**-10.72**	**-9.59**
10	**-10.38**	17.05	**-23.25**	**-46.18**	**-32.69**	**-70.58**	**-9.45**	**-22.72**	**-12.14**	**-13.75**
11	2.26	24.04	**-20.64**	**-49.97**	**-30.50**	**-74.80**	**-8.56**	**-21.61**	**-11.24**	**-11.15**
12	**-3.60**	13.89	**-25.70**	**-50.50**	**-36.37**	**-79.00**	**-14.41**	**-26.00**	**-17.15**	**-22.18**
13	**-13.69**	8.48	**-28.27**	**-51.81**	**-37.98**	**-81.82**	**-15.60**	**-27.44**	**-18.30**	**-29.50**
14	**-9.41**	16.40	**-25.11**	**-58.49**	**-32.61**	**-83.33**	**-11.58**	**-24.80**	**-14.15**	**-26.64**
15	**-2.17**	22.64	**-25.58**	**-65.42**	**-33.09**	**-85.36**	**-11.22**	**-25.92**	**-14.02**	**-30.53**

Note: Numbers shown are the percent improvements or deteriorations in MSFE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

Table 2: Relative MAPE, Evaluation Period: 1995-2011

Fcst Hzn	Statistical Models				Structural-Linear		Structural-Nonlinear			
	RW-WD	AR-NT	AR-LT	AR-QT	LT	QT	TVP-IS	TVP-I	TVP-S	NEMS
1	17.72	17.01	9.62	10.76	57.51	113.29	4.47	0.46	4.75	**-26.21**
2	**-10.33**	3.39	**-7.84**	**-9.91**	28.55	74.61	**-6.06**	**-11.38**	**-3.89**	**-9.18**
3	**-2.99**	25.88	10.98	**-0.54**	42.96	94.83	10.01	7.43	10.49	6.62
4	12.86	38.47	20.58	10.76	34.12	80.68	14.49	15.40	12.78	17.99
5	6.02	33.99	4.63	**-2.16**	4.57	33.43	6.98	2.20	5.77	7.67
6	0.84	27.30	**-10.65**	**-16.14**	**-14.20**	2.81	**-0.81**	**-10.22**	**-2.75**	**-1.59**
7	**-1.31**	21.92	**-14.93**	**-16.62**	**-16.50**	**-3.38**	**-4.55**	**-16.33**	**-7.15**	**-2.68**
8	**-4.51**	16.03	**-17.71**	**-28.08**	**-23.30**	**-23.57**	**-6.29**	**-18.75**	**-8.65**	**-10.07**
9	**-3.43**	14.33	**-16.73**	**-34.81**	**-25.32**	**-40.90**	**-5.99**	**-16.79**	**-8.08**	**-12.65**
10	**-2.12**	12.30	**-15.25**	**-38.70**	**-23.75**	**-54.82**	**-6.26**	**-15.50**	**-8.13**	**-11.98**
11	**-1.92**	9.53	**-15.13**	**-35.81**	**-22.94**	**-63.95**	**-7.08**	**-15.24**	**-8.82**	**-12.84**
12	**-5.22**	6.60	**-16.44**	**-36.30**	**-23.54**	**-68.89**	**-8.43**	**-16.21**	**-10.12**	**-15.67**
13	**-6.09**	7.36	**-16.00**	**-38.57**	**-22.08**	**-68.40**	**-7.62**	**-15.70**	**-9.26**	**-16.09**
14	**-3.48**	9.83	**-14.94**	**-41.87**	**-20.19**	**-66.66**	**-6.55**	**-14.96**	**-8.13**	**-17.48**
15	**-3.10**	10.07	**-17.35**	**-48.47**	**-23.36**	**-69.59**	**-8.03**	**-17.38**	**-9.81**	**-24.95**

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

Now that we know how the NEMS model forecasts compare to the random walk, and also to our simple illustrative models, it is important to examine whether results are sample driven or whether they are reflective of a more general pattern over time. While we cannot do so specifically for NEMS, we can for our other models. Below (see Tables 3 and 4) we report results for model estimations for the period ending in 1984, and the resulting forecast criteria over the 1985-2011 period. These are interesting in their own right, but by transitivity, and provided relative performances amongst models has not changed dramatically over time, we can even get some sense of how NEMS would have performed had it existed much earlier.

We find that results are generally comparable to the ones obtained over the later sample. In particular, many of the models are able to outperform the random walk. In addition, the models do even better the further out the forecast horizon, although according to the MAPE criterion, the TVP class of models are able to beat the no-change forecasts even for the one to six-year-ahead horizons. Taken together, the results thus suggest that simple models such as the ones we have examined have much to offer in terms of forecasting, and given their ease of use, should not be neglected by the consumers of these forecasts.

Table 3: Relative MSFE, Evaluation Period: 1985-2011

Fcst Hzn	Statistical Models				Structural-Linear		Structural-Nonlinear		
	RW-WD	AR-NT	AR-LT	AR-QT	LT	QT	TVP-IS	TVP-I	TVP-S
1	53.29	56.99	45.72	35.98	199.03	172.42	28.01	21.36	10.09
2	10.17	22.17	12.27	9.03	97.29	102.98	3.05	1.88	5.58
3	59.29	87.83	58.63	47.92	118.63	137.37	26.11	35.91	19.94
4	30.14	59.23	27.44	23.36	47.36	72.11	20.02	17.07	17.92
5	28.48	55.06	17.83	17.37	23.01	50.75	16.05	8.46	14.16
6	24.89	42.48	4.65	10.54	**-0.49**	29.18	6.28	**-1.72**	4.38
7	15.68	25.60	**-9.05**	3.80	**-18.86**	12.28	**-4.93**	**-13.53**	**-6.86**
8	16.68	19.85	**-14.56**	2.85	**-26.84**	5.56	**-10.55**	**-19.33**	**-12.66**
9	9.21	11.16	**-21.32**	0.54	**-34.08**	**-1.65**	**-15.76**	**-23.29**	**-18.03**
10	4.72	9.54	**-24.01**	0.44	**-35.80**	**-2.72**	**-16.54**	**-25.18**	**-18.92**
11	16.23	16.22	**-22.52**	2.48	**-34.47**	**-0.10**	**-15.70**	**-24.88**	**-18.18**
12	10.67	9.00	**-27.61**	0.69	**-40.13**	**-7.73**	**-20.01**	**-29.19**	**-22.69**
13	2.71	5.11	**-30.64**	1.54	**-42.90**	**-11.08**	**-21.33**	**-31.13**	**-24.06**
14	3.37	11.39	**-29.81**	1.48	**-40.40**	**-9.11**	**-18.44**	**-30.20**	**-21.21**
15	5.89	19.94	**-28.78**	**-3.11**	**-37.17**	**-10.70**	**-13.21**	**-29.01**	**-16.35**

Note: Numbers shown are the percent improvements or deteriorations in MSFE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

Table 4: Relative MAPE, Evaluation Period: 1985-2011

Fcst Hzn	Statistical Models				Structural-Linear		Structural-Nonlinear		
	RW-WD	AR-NT	AR-LT	AR-QT	LT	QT	TVP-IS	TVP-I	TVP-S
1	26.61	7.26	11.34	31.66	29.13	138.18	**-7.94**	**-4.86**	**-4.96**
2	16.52	4.92	3.09	28.40	3.11	112.84	**-12.79**	**-7.92**	**-12.95**
3	18.05	8.21	7.33	37.93	6.03	125.21	**-7.16**	**-5.21**	**-5.57**
4	21.84	3.66	7.68	52.27	1.23	124.09	**-9.85**	**-0.75**	**-10.25**
5	29.00	2.00	10.54	68.04	1.24	130.63	**-14.61**	**-1.28**	**-13.93**
6	28.04	**-2.44**	2.84	66.69	**-6.36**	117.50	**-19.92**	**-6.48**	**-19.18**
7	27.23	**-6.58**	**-3.89**	62.77	**-17.17**	102.67	**-25.28**	**-12.73**	**-25.77**
8	19.67	**-9.80**	**-10.22**	57.28	**-24.15**	88.84	**-28.15**	**-17.99**	**-28.55**
9	21.45	**-10.01**	**-9.08**	65.86	**-23.69**	92.45	**-27.63**	**-15.69**	**-28.43**
10	28.24	**-7.60**	**-8.08**	73.99	**-23.19**	97.19	**-26.07**	**-15.17**	**-26.30**
11	29.19	**-11.76**	**-9.99**	77.00	**-25.50**	94.24	**-27.99**	**-16.39**	**-28.75**
12	25.39	**-8.50**	**-13.49**	73.90	**-28.85**	86.46	**-30.52**	**-19.12**	**-30.38**
13	32.21	**-3.50**	**-13.62**	77.45	**-30.30**	82.80	**-27.68**	**-19.19**	**-29.07**
14	20.90	**-7.36**	**-22.17**	63.51	**-37.08**	62.51	**-29.39**	**-24.57**	**-30.62**
15	26.65	0.81	**-20.24**	62.25	**-33.68**	68.72	**-22.09**	**-24.46**	**-24.89**

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

5. Combining Forecasts

An interesting feature that we observe from the whole set of results obtained so far is that, while the quadratic models (linear or equilibrium-based) were generally the best-performing over the more recent forecast evaluation sample, they are the worse-performing over the longer sample. We also find that the performance of the various TVP models differ over these two periods. Moreover, for a given forecast evaluation period, we find that some models yield relatively better results than others for particular forecast horizons.

These observations suggest that there might be a benefit to resorting to forecast combination techniques. Rather than relying on one model that at times might overperform and at other times underperform, it might be preferable to combine forecasts from a multitude of models so as to provide consistently good forecasts over time. A number of studies have found merit in such approaches, notably for short-run forecast horizons. For example, Baumeister and Kilian (2014a) show that combining forecasts from VAR and futures-based models, as well as from models based on other commodity prices and models using the spread between the prices of refined oil product prices and crude oil prices (known as product spreads), in general can outperform no-change forecasts for horizons of one up to 18 months.

In this section, we document the worth of using forecast combination approaches for medium and long term forecasting. Four different forecast combination methods are considered: (i) taking the simple average of forecasts (denoted AVE in the tables of results), (ii) taking the median forecast (denoted MED in the tables of results), (iii) forecast combinations based on the Akaike information criterion (denoted AIC in the tables of results), and finally, (iv) forecast combinations based on the Bayesian information criterion (denoted BIC in the tables of results). The forecasts are obtained for three different groupings of models. One includes all of our available models (denoted ALL), a second includes only our statistical models, and a third includes only our structural models. Table 5 reports the obtained results over the 1995-2011 forecast evaluation period for the MSFE criterion, while Table 6 reports results for the same sample for the MAPE criterion. For completion, we also provide in an Appendix corresponding results for the 1985-2011 evaluation sample.

Amongst the four methods of combining forecasts, with the exception of the statistical model grouping, simple averaging produces overall the best results both for the MSFE and the MAPE. Having said that, we see that the general pattern of results is not much changed from what was obtained before. In particular, the no-change forecast continues to be beaten as the forecast horizon increases beyond six years. Nevertheless, we do find a higher consistency in the quantity of the obtained forecast improvements across the various forecast horizons, with the numbers improving for the six to nine-year-ahead horizons.

Table 5: Forecast Combinations, Relative MFSE; 1995-2011

h	All				Statistical				Structural				Individual Best	
	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	Model	Fcst
1	35.44	32.40	30.09	40.71	11.81	17.46	30.30	37.10	64.95	42.34	41.24	29.67	NEMS	-48.65
2	8.00	10.32	3.13	**-2.19**	**-16.29**	**-12.63**	**-6.46**	**-7.80**	36.03	22.58	19.59	17.43	AR-QT	-19.76
3	54.11	53.22	50.29	62.45	28.65	31.97	44.59	51.51	70.44	53.04	53.88	51.86		
4	26.56	33.22	28.88	33.70	10.48	7.40	16.44	23.35	33.41	32.52	39.66	36.72	AR-QT	-3.72
5	11.84	18.15	16.54	24.48	2.62	1.28	3.21	13.57	11.40	17.37	27.16	23.50	AR-QT	-16.12
6	**-0.16**	5.53	5.24	15.52	**-3.89**	**-3.71**	**-7.61**	4.73	**-5.25**	4.57	14.41	10.86	QT	-32.78
7	**-10.76**	**-7.16**	**-4.97**	4.02	**-11.37**	**-10.91**	**-17.44**	**-6.47**	**-17.42**	**-7.73**	3.18	**-0.49**	QT	-44.85
8	**-15.43**	**-11.60**	**-9.08**	1.72	**-13.31**	**-11.80**	**-21.78**	**-8.83**	**-24.28**	**-14.87**	**-2.57**	**-6.12**	QT	-54.29
9	**-21.56**	**-17.29**	**-14.38**	**-4.67**	**-18.21**	**-16.73**	**-27.76**	**-15.23**	**-30.80**	**-20.50**	**-8.47**	**-11.47**	QT	-63.79
10	**-24.67**	**-20.25**	**-15.95**	**-8.06**	**-21.76**	**-21.00**	**-31.70**	**-19.25**	**-33.59**	**-22.69**	**-9.81**	**-12.68**	QT	-70.58
11	**-22.97**	**-17.67**	**-14.13**	**-1.95**	**-19.06**	**-15.31**	**-31.48**	**-13.24**	**-33.65**	**-21.53**	**-8.86**	**-11.33**	QT	-74.80
12	**-27.50**	**-21.50**	**-18.47**	**-6.82**	**-21.75**	**-17.68**	**-34.41**	**-17.24**	**-38.73**	**-26.00**	**-14.66**	**-16.83**	QT	-79.00
13	**-30.28**	**-23.28**	**-19.80**	**-11.90**	**-25.24**	**-21.43**	**-36.78**	**-22.54**	**-40.55**	**-27.44**	**-15.83**	**-17.86**	QT	-81.82
14	**-28.48**	**-20.28**	**-17.36**	**-8.63**	**-25.47**	**-17.90**	**-37.83**	**-20.01**	**-38.21**	**-24.80**	**-11.83**	**-14.01**	QT	-83.33
15	**-29.71**	**-19.85**	**-18.13**	**-5.60**	**-26.42**	**-15.04**	**-41.63**	**-17.77**	**-40.61**	**-25.92**	**-11.50**	**-13.90**	QT	-85.36

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

Table 6: Forecast Combinations, Relative MAPE; 1995-2011

h	All				Statistical				Structural				Individual Best	
	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	Model	Fcst
1	6.52	2.57	-0.63	10.83	-0.41	4.85	9.49	12.26	19.24	7.55	4.25	3.34	NEMS	-26.21
2	**-10.34**	**-11.48**	**-11.50**	**-9.82**	**-15.88**	**-15.04**	**-12.94**	**-12.08**	2.62	**-8.33**	**-6.22**	**-7.41**	RW-WD	-10.33
3	6.93	2.66	5.97	5.67	**-1.78**	**-7.11**	3.39	1.38	15.04	4.31	9.86	8.75	RW-WD	-2.99
4	15.91	12.46	12.66	17.26	10.15	4.62	14.73	14.99	17.66	13.76	14.33	13.15		
5	2.25	1.16	2.96	7.79	**-0.16**	**-3.72**	1.29	2.83	1.63	2.35	6.78	5.28	AR-QT	2.16
6	**-11.50**	**-9.62**	**-6.25**	1.94	**-11.95**	**-12.53**	**-14.53**	**-9.05**	**-14.10**	**-9.36**	**-1.15**	**-3.45**	AR-QT	-16.14
7	**-15.83**	**-13.97**	**-10.71**	**-1.37**	**-15.69**	**-13.78**	**-17.43**	**-12.02**	**-16.09**	**-15.62**	**-4.90**	**-7.66**	AR-QT	-16.62
8	**-19.17**	**-14.94**	**-11.90**	**-3.98**	**-17.57**	**-14.30**	**-23.38**	**-13.63**	**-23.67**	**-18.45**	**-6.61**	**-9.14**	AR-QT	-28.08
9	**-18.94**	**-13.49**	**-10.81**	**-3.41**	**-15.77**	**-12.84**	**-25.25**	**-12.05**	**-26.30**	**-16.79**	**-6.25**	**-8.33**	QT	-40.90
10	**-17.13**	**-12.63**	**-10.15**	**-2.87**	**-14.02**	**-11.33**	**-23.03**	**-10.44**	**-25.50**	**-15.49**	**-6.47**	**-8.22**	QT	-54.82
11	**-16.63**	**-12.37**	**-10.12**	**-3.33**	**-13.22**	**-10.64**	**-21.94**	**-10.11**	**-24.60**	**-15.22**	**-7.25**	**-8.70**	QT	-63.95
12	**-17.86**	**-13.41**	**-11.19**	**-5.37**	**-14.49**	**-11.67**	**-22.86**	**-12.01**	**-25.44**	**-16.21**	**-8.59**	**-9.92**	QT	-68.89
13	**-17.82**	**-12.71**	**-10.79**	**-5.52**	**-15.17**	**-11.39**	**-23.66**	**-12.34**	**-24.99**	**-15.70**	**-7.77**	**-9.08**	QT	-68.40
14	**-17.27**	**-11.74**	**-10.15**	**-4.00**	**-15.07**	**-9.50**	**-24.20**	**-10.88**	**-24.45**	**-14.96**	**-6.70**	**-8.04**	QT	-66.66
15	**-20.22**	**-13.49**	**-12.00**	**-4.52**	**-17.23**	**-10.47**	**-28.00**	**-12.01**	**-28.66**	**-17.38**	**-8.20**	**-9.66**	QT	-69.59

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

6. Conclusion

The public releases of expert outlooks regarding the future path of oil prices attract much media attention and are considered highly useful by industry participants and policymaking bodies that often rely on these sources for their needs. Yet little or no attention has been paid to the extent to which these are accurate. Using the regular publications by the Energy Information Administration (EIA), in this paper we examined the worth of annual oil price forecasts over the medium and long runs generated by the formal model (NEMS) of the Agency. Two forecast criteria were used for this purpose, the MSFE and the MAPE, and recursive (and where relevant, dynamic) estimations were conducted prior to obtaining each forecast value.

Our results showed that NEMS is quite successful at beating the benchmark random walk model, but only at either end of the fifteen year forecast horizon. In particular, based on the MSFE, the EIA model outperforms the no-change model at the one-year-ahead horizon, and at horizons of nine-years-ahead or more. However, it fails to beat the benchmark model for the interim forecast horizons. In addition, we showed that simple statistical or equilibrium-based models often produce similar if not better forecast performances than the EIA model, specially for forecast horizons of six years and greater. Finally, we calculated forecast combinations over three model groupings to see if forecasts would improve. The results indicated that combining forecasts does not change the overall obtained patterns of results from individual comparisons, but that some gains can be made over the medium term forecasts in the extent to which they outperform the random walk.

References

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

Alquist, R., L. Kilian, and R.J. Vigfusson. 2013. "Forecasting the Price of Oil." In *Handbook of Economic Forecasting, 2*, edited by G. Elliott and A. Timmermann, 1–46. Amsterdam: North-Holland: forthcoming.

Baumeister, C., P. Guerin, and L. Kilian. 2014a. "Do High-Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work." *International Journal of Forecasting* forthcoming.

Baumeister, C. and L. Kilian. 2014a. "Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach." *Journal of Business and Economic Statistics* forthcoming.

———. 2014b. "What Central Bankers Need to Know about Forecasting Oil Prices." *International Economic Review* 55(3): 869–889.

Baumeister, C., L. Kilian, and T. Lee. 2014b. "Are there Gains from Pooling Real-Time Oil Price Forecasts?" *Energy Economics* forthcoming.

Baumeister, C., L. Kilian, and X. Zhou. 2013. *Are Product Spreads Useful for Forecasting? An Empirical Evaluation of the Verleger Hypothesis.* Technical report, Bank of Canada Working Paper2013-25.

Bernard, J.T., J.M. Dufour, L. Khalaf, and M. Kichian. 2012. "An identification-robust test for time-varying parameters in the dynamics of energy prices." *Journal of Applied Econometrics* 27(4): 603–624.

Bernard, J.T., L. Khalaf, M. Kichian, and S. MacMahon. 2014. "The Convenience Yield and the Informational Content of the Oil Futures Price." *The Energy Journal* forthcoming.

Chinn, M. and O. Coibion. 2013. "The Predictive Content of Commodity Futures." *Journal of Futures Markets* forthcoming.

Dvir, E. and K. Rogoff. 2010. *Three Epochs of Oil.* Technical report, NBER Working PaperNo. 14927.

Edelstein, P. and L. Kilian. 2009. "How Sensitive are Consumer Expenditures to Retail Energy Prices?" *Journal of Monetary Economics* 56(6): 766–779.

Hamilton, J.D. 2009. "Understanding crude oil prices." *The Energy Journal* 30: 179–206.

Manthy, R.S. 1978. *Natural Resource Commodities: A Century of Statistics.* Baltimore: John Hopkins Press.

Pindyck, R.S. 1999. "The Long Run Evolution of Energy Prices." *The Energy Journal* 20: 1–27.

Slade, M.E. 1982. "Trends in natural-resource commodity prices: An analysis of the time domain." *Journal of Environmental Economics and Management* 9(2): 122–137.

1 Appendix

Table 7: Forecast Combinations, Relative MFSE; 1985-2011

h	All				Statistical				Structural				Individual Best	
	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	Model	Fcst
1	26.83	23.69	22.34	44.77	16.89	23.83	39.27	45.60	47.26	27.10	25.97	16.69		
2	2.73	4.01	**-2.33***	10.70	**-3.88***	1.02	11.22	11.09	18.25	5.99	2.85	1.48		
3	36.05	33.35	32.02	58.97	30.71	35.01	53.02	56.96	42.60	25.55	25.94	24.76		
4	16.69	19.67	18.12	32.78	13.65	11.84	27.73	29.43	18.89	13.83	19.73	17.66		
5	10.10	12.33	11.62	28.20	9.87	8.82	20.54	24.08	8.14	8.47	15.63	12.65		
6	0.63	2.29	2.73	20.49	4.51	4.17	11.79	16.19	**-5.23** *	**-1.11**	5.87	2.91	TVP-I	-1.72
7	**-9.87**	**-10.08**	**-6.84**	9.97	**-2.58**	**-2.71**	2.84	5.80	**-17.96**	**-13.55**	**-5.35**	**-8.43**	LT	-18.86
8	**-14.29**	**-14.65**	**-10.39**	7.70	**-4.27**	**-3.93**	0.46	3.63	**-24.41**	**-20.59**	**-10.95**	**-13.93**	LT	-26.84
9	**-19.83**	**-19.91**	**-14.80**	1.89	**-8.73**	**-8.76**	**-3.84**	**-2.31**	**-30.26**	**-25.84**	**-16.09**	**-18.61**	LT	-34.08
10	**-22.38**	**-22.32**	**-15.70**	**-0.46**	**-11.38**	**-12.22**	**-5.18**	**-5.12**	**-32.60**	**-27.60**	**-16.85**	**-19.32**	LT	-35.80
11	**-21.37**	**-20.62**	**-14.04**	5.08	**-8.90**	**-7.61**	**-2.25**	**0.59**	**-33.29**	**-26.83**	**-15.97**	**-18.20**	LT	-34.47
12	**-26.23**	**-24.55**	**-18.25**	0.63	**-12.35**	**-10.42**	**-5.76**	**-3.60**	**-38.92**	**-30.86**	**-20.27**	**-22.40**	LT	-40.13
13	**-29.56**	**-26.79**	**-19.16**	**-2.79**	**-15.66**	**-13.56**	**-6.26**	**-7.52**	**-42.07**	**-33.58**	**-21.58**	**-23.71**	LT	-42.90
14	**-30.12**	**-25.98**	**-17.05**	**-1.03**	**-17.27**	**-12.68**	**-4.49**	**-6.80**	**-42.45***	**-33.12**	**-18.73**	**-21.12**	LT	-40.40
15	**-31.30**	**-24.37**	**-18.42**	**-0.78**	**-20.07**	**-11.91**	**-8.95**	**-7.48**	**-43.30***	**-30.59**	**-13.56**	**-16.48**	LT	-37.17

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.

Table 8: Forecast Combinations, Relative MAPE; 1985-2011

h	All				Statistical				Structural				Individual Best	
	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	AVE	MED	AIC	BIC	Model	Fcst
1	6.85	**-0.24**	**-4.02**	16.47	8.66	14.57	23.50	21.99	14.34	**-3.67**	**-8.19***	**-7.70**	TVP-IS	-7.94
2	**-0.36**	**-4.41**	**-7.76**	9.42	4.26	5.47	14.98	11.72	3.59	**-10.71**	**-12.90**	**-13.70***	TVP-S	-12.95
3	4.86	**-3.75**	**-1.11**	12.93	10.20	7.51	24.53	17.50	4.37	**-11.87***	**-7.15**	**-6.98**	TVP-IS	-7.16
4	8.95	**-1.67**	1.40	13.75	17.23	12.20	35.76	22.70	8.52	**-7.97**	**-9.90**	**-10.09**	TVP-S	-10.25
5	14.53	**1.77**	2.51	18.98	24.95	19.09	45.91	29.17	12.07	**-7.41**	**-14.57**	**-13.95**	TVP-IS	-14.61
6	8.79	**-2.88**	**-1.39**	17.21	20.86	15.25	41.33	24.68	5.43	**-12.21**	**-19.85**	**-18.84**	TVP-IS	-19.92
7	2.24	**-8.34**	**-6.05**	14.70	16.42	12.50	36.19	21.42	**-2.20**	**-19.12**	**-25.29**	**-25.29**	TVP-S	-25.77
8	**-3.60**	**-13.69**	**-10.03**	8.80	11.34	7.65	32.44	16.46	**-9.60**	**-25.18**	**-28.17**	**-28.22**	TVP-S	-28.55
9	**-0.93**	**-12.57**	**-6.83**	10.94	16.51	9.71	40.07	20.79	**-7.76**	**-24.31**	**-27.77**	**-27.97**	TVP-S	-28.43
10	1.23	**-11.01**	**-3.26**	14.98	20.39	13.84	47.57	25.60	**-6.17**	**-24.10**	**-26.04**	**-25.74**	TVP-S	-26.30
11	**-0.02**	**-11.53**	**-4.24**	16.24	19.75	13.24	48.99	25.19	**-8.31**	**-24.80**	**-28.01**	**-28.16**	TVP-S	-28.75
12	**-3.32**	**-16.01**	**-5.98**	13.24	17.04	10.32	45.69	23.47	**-12.79**	**-27.03**	**-30.46**	**-29.94**	TVP-IS	-30.52
13	**-2.57**	**-15.64**	**-4.00**	18.34	19.32	13.60	52.60	27.29	**-13.95**	**-29.50**	**-27.81**	**-28.96**	LT	-30.30
14	**-12.22**	**-21.88**	**-5.75**	10.01	8.71	4.32	42.04	16.63	**-23.64**	**-35.03**	**-29.54**	**-30.12**	LT	-37.08
15	**-10.66**	**-20.63**	**-8.17**	14.23	8.85	8.11	38.84	20.50	**-23.13**	**-32.63**	**-22.37**	**-24.79**	LT	-33.68

Note: Numbers shown are the percent improvements or deteriorations in MAPE relative to the no-change model; numbers in bold refer to models which, for a given forecast horizon, improve on the no change forecast.