CAHIER DE RECHERCHE #1510E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER #1510E
Department of Economics
Faculty of Social Sciences
University of Ottawa

Oil Price Forecasts for the Long-Term:

Expert Outlooks, Models, or Both?[*]

Jean-Thomas Bernard[†], Lynda Khalaf[‡],

Maral Kichian[§] and Clement Yelou[**]

November 2015

[*] Thanks to the editors, three anonymous referees, participants of the Bank of Canada 2015 commodities workshop, and to Christiane Baumeister for comments and suggestions. This work was supported by the Institut de Finance Mathématique de Montréal [IFM2], the Social Sciences and Humanities Research Council of Canada, and the Fonds FQRSC (Government of Québec).
[†] Department of Economics, University of Ottawa, 120 University Private, Ottawa, Ontario, Canada, K1N 6N5; e-mail: jbernar3@uOttawa.ca.
[‡] Carleton University.
[§] Graduate School of Public and International Affairs, University of Ottawa, 120 University Private, Ottawa, Ontario, Canada, K1N 6N5; e-mail: maral.kichian@uOttawa.ca.
[**] CREATE, Laval University.

Abstract

Expert outlooks on the future path of oil prices are often relied on by industry participants and policymaking bodies for their forecasting needs. Yet little 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 accuracy of annual recursive oil price forecasts generated by the National Energy Modeling System model of the Agency for forecast horizons of up to 15 years. Our results reveal that the EIA model is quite successful at beating the benchmark random walk model, but only at either end of the forecast horizons. We also show that, for the longer horizons, simple econometric forecasting models often produce similar if not better accuracy than the EIA model. Among these, time-varying specifications generally also exhibit stability in their forecast performance. Finally, while combining forecasts does not change the overall patterns, some additional accuracy gains are obtained at intermediate horizons, and in some cases forecast performance stability is also achieved.

Key words: *Oil price, expert outlooks, long run forecasting, forecast combinations.*

JEL Classification: Q47, C20.

Résumé

Les perspectives d'experts sur la trajectoire future des prix du pétrole sont souvent utilisées par les participants de l'industrie et les entités politiques pour leurs prévisions. Malgré cela, peu de travaux analysent si ces perspectives sont précises. Nous nous basons sur les publications régulières de l'Energy Information Administration (EIA) pour analyser la précision des prévisions annuelles du prix récursif du pétrole générées par le modèle du National Energy Modeling System de l'Agence pour les prévisions avec un horizon jusqu'à 15 ans. Nos résultats révèlent que le modèle EIA est assez performant afin de battre le modèle de marche aléatoire de référence, mais seulement à chaque extrémités des horizons de prévision. Nous montrons également que les modèles simples d'économétrie produisent souvent des résultats similaires sinon meilleurs en précision que le modèle EIA pour les horizons longs. Parmi ceux-ci, les spécifications variant dans le temps produisent également de la stabilité dans leurs performances de prédictions. Finalement, bien que combiner plusieurs prédictions ne change pas les tendances générales, quelques gains de précisions additionnels sont obtenus pour les horizons de temps intermédiaires. Dans certains cas, la stabilité des prédictions de performance est aussi atteinte.

Mots clés : *Prix du pétrole, perspectives d'experts, prévision de long terme, combinaisons de prévision.*

Classification JEL : Q47, C20.

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 effects[1]. This also means that oil price is among a handful of indicators that can to a certain extent predict future movements in 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 often paid to medium and long term forecasting. For instance, government agencies and international organizations predicate their policy decisions and regulatory recommendations on macroeconomic projections that directly depend on a given assumed multi-year future path for the real price of oil. Similarly, industry participants make use of forecasts of the real price of oil 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.

To obtain their forecast paths, industry and policymaking bodies have traditionally (and predominantly) resorted to futures prices for horizons of up to two years, and to survey or expert outlook forecasts for longer horizons[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, Alquist, Kilian, and Vigfusson (2013) and Bernard, Khalaf, Kichian, and MacMahon (2015)), little attention has been paid to the extent to which expert

[1] There is a large literature on this topic; see, for example, Hamilton (2009), Kilian and Vigfusson (2011), Hamilton (2011), Ravazzolo and Rothman (2013), Kilian and Vigfusson (2013), as well as the surveys by Kilian (2008) and Kilian (2014).

[2] While futures prices for crude oil do exist for maturities of as long as seven years, the market is much less liquid at maturities longer than two years; see Alquist et al. (2013)

and/or survey approaches are helpful in predicting oil prices[3].

In this paper, we examine the accuracy of expert outlook forecasts over the medium and long run. In particular, we focus on the Energy Information Administration (EIA) agency forecasts. The agency adopted a formal econometric model, National Energy Modeling System (NEMS), in 1991, and has been publishing annual forecasts for up to 25 years ahead based on this model. In addition to studying the extent to which NEMS forecasts are accurate, this paper also fills a gap in the literature regarding the medium and long run forecasting performance of some simple econometric models, some purely statistical, and others motivated by economic rationale. Furthermore, given that some models are likely to forecast better over particular horizons, we also assess the potential for improvements in forecasts resulting from forecast combination approaches.

More specifically, our analysis considers out-of-sample forecasts for the real price of oil, one to fifteen years ahead, using annual frequency data, and for the 142-year period extending from 1870 to 2011. Our forecast evaluation period starts in 1995, given that NEMS underwent some important updates until 1994. We use two different forecast criteria for our comparisons. These are the mean square forecast error (MSFE) and the mean absolute percentage 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. As for our forecast combinations, three different methods are examined, applied to various model groupings. These are discussed in later sections with more detail.

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. NEMS is also found to have an advantage (of about 9 per cent) over the no-change forecast two-years-ahead, but only based on the MAPE. Interestingly, the model cannot beat the no-change

[3]There are a few exceptions, including Alquist, Kilian, and Vigfusson (2013), Baumeister et al. (2014), and Sanders, Manfredo, and Boris (2009). These focus only on short-horizon monthly or quarterly forecasts.

model for forecasts of three to five years ahead according to MAPE, and for forecasts of two to eight years ahead according to MSFE. Yet, at longer horizons, it is once again found to outperform the random walk, sometimes by substantial margins.

The outcomes also show that many of the simple forecasting models that we consider are also successful at horizons of six years and greater, often with much higher forecast accuracy. In addition, specifications that also control for parameter dynamics are also found to exhibit some stability in their forecast performance over successive forecast evaluation periods.

Finally, although combining forecasts does not change overall patterns, (i) some additional accuracy gains are obtained at intermediate horizons, and (ii) remarkable stability of forecast performance is achieved when considering forecasts at the 7-year or 10-year horizons, and when forecasts from economically-founded empirical models are combined.

2. Models

Among the currently available expert outlooks regarding the future path of oil prices, the most popular is the 'Annual Energy Outlook' publication produced by the Energy Information Administration agency of the U.S. Department of Energy. Releases of this report are prompty talked about in the media, and the forecasts that are provided therein are frequently used by both private sector participants and by policymaking bodies. Yet little is known about the accuracy of these forecasts, or how they would compare to ones generated from simple econometric forecasting models[4]. The literature on oil price forecasting has shown 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 over shorter horizons 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

[4]Previous work examining U.S. Department of Energy forecasts, notably those pertaining to energy consumption and production, include Lynch (2002), O'Neil and Desai (2005), Winebrake and Sakva (2006), Fischer, Herrnstadt, and Morgenstern (2008), and Huntington (2011).

[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).

examine whether expert forecasts are more successful than the random walk forecast.

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, based on the supplied information on the EIA website, NEMS is a large and intricate model for the U.S. aimed at integrating a multitude of factors that affect energy price dynamics. Such a model presents an advantage over simplistic specifications in that it aims to incorporate the different supply and demand channels that can impact the price of oil. However, at the same time, the model has a large number of parameters that need to be quantified, and for forecasts to be accurate, this needs to be done with enough precision[7]. This task often proves to be very difficult, notably (i) when data samples are relatively short, (ii) when structural changes occur that affect the strength or nature of specific economic relationships, but they are not modelled (since they can only be recognized as having changed some years later), and (iii) when geopolitical considerations (unrelated to economic fundamentals) drive the price changes[8]. In this case it is natural to ask how expert forecasts compare with forecasts obtained from much simpler and potentially more flexible models. In addition, it is quite conceivable that at times structural models such as NEMS might be more successful with forecast accuracy, while simpler models might be more accurate at other times. Such considerations then leave the door open to combining forecasts from both types of models for potentially superior overall forecast performances[9].

The EIA and the National Energy Modeling System (NEMS)

The EIA was created after the first oil shock in 1973, when concerns were expressed about the quality of the energy statistics that were available at the time to the U.S. government.

[7]In practice, only some of the parameters of such large models can be estimated while others can only be calibrated, often with the aid of judgement

[8]See, for example, Dvir and Rogoff (2010), Bernard, Dufour, Khalaf, and Kichian (2012) and Alquist, Kilian, and Vigfusson (2013) on structural changes in oil prices, and Kilian and Murphy (2014) for geopolitical impacts on oil prices.

[9]See Baumeister and Kilian (2014a) and Baumeister, Kilian, and Lee (2014) for forecast combination outcomes in the case of short-horizon forecasts.

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, as well as for the study of energy policy. Although the 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, the EIA promoted its Annual Energy Outlook every year since 1982. The early release of the Annual Energy Outlook takes place in December, and the forecasts are discussed at a widely attended conference held in Washington D.C. in the following spring. The annual forecasts are for horizons of up to twenty-five years ahead, and include production, imports, exports, consumption, and energy prices. Among these, a key variable that draws considerable interest is the annual oil price forecast.

The model underlying the annual forecasts is **NEMS**. The National Energy Modeling System, one of the EIA's main models, is a large-scale model for the U.S. that is aimed at integrating a myriad of worldwide influences on oil price related to oil production and demand. Further details are provided at the agency's website[10]. The website notably states that 'NEMS is designed to represent the important interactions of supply and demand in U.S. energy markets' and that it 'represents the market behavior of the producers and consumers of energy at a level of detail that is useful for analyzing the implications of technological improvements and policy initiatives'. In addition, model assumptions include 'the estimated size of the economically recoverable resource base of fossil fuels, and changes in world energy supply and demand. The projections are business-as-usual trend estimates, given known technological and demographic trends'. Finally, NEMS has a modular structure, including 'four supply modules (oil and gas, natural gas transmission and distribution, coal market, and renewable fuels); two conversion modules (electricity market and petroleum market); four end-use demand modules (residential demand, commercial demand, industrial demand, and

[10]See http://www.eia.gov/oiaf/aeo/overview/.

transportation demand); one module to simulate energy/economy interactions (macroeconomic activity); one module to simulate international energy markets (international energy); and one module that provides the mechanism to achieve a general market equilibrium among all the other modules ('integrating module)'. In sum, the model thus solves for the prices of each energy type, interacting supply and demand in each case. Furthermore, industry structure and various energy policies and regulations are taken into account.

In this paper we make use of the published annual forecasts generated by NEMS. In the next section we also define a range of econometric forecasting models that we use for comparisons with these EIA forecasts. It is of course impossible to consider for this purpose all of the proposed models in this vast literature, and thus we select some illustrative specifications from various popular model classes. We consider only parsimonious specifications based on single-equations, on the one hand to accommodate the smooth nature of our annual data, and on the other, to control for degrees of freedom. In the final analysis, we make use of purely statistical formulations, as well as simple linear and non-linear equations that were suggested in the literature, and that are motivated by economic arguments.

Alternative Forecasting Models

The overwhelming majority of models that describe the behaviour of oil prices over the long run appeal to theoretical foundations based on the Hotelling (1930) premise regarding the evolution of an exhaustible resource. The models take into account aspects related to either inventory management, uncertainty regarding future oil discovery, or the presence of other energy alternatives. We consider two classes of simple forecasting specifications that were suggested in the literature and that are based on arguments drawn from the above-described models: one class includes specifications that are linear in the parameters, and the other includes specifications where the parameters evolve in a nonlinear fashion. In all the considered cases, the regressors are predetermined.

Within the linear-parameter category we consider the two forecasting equations pro-

posed by Slade (1982). The underlying theoretical setup characterizes the long-run price movements of a non-renewable natural resource accounting for both exogeneous technical change and endogenous change in the grade of the unrefined material. Under some assumptions, the model implies that price will equal marginal extraction cost plus rent, and 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 simply 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 as:

$$P_t = c_1 + c_2 t + c_3 t^2 + \epsilon_t, \quad t = 1, ..., T \tag{1}$$

and its restricted counterpart

$$P_t = c_1 + c_2 t + \epsilon_t, \quad t = 1, ..., T, \tag{2}$$

where P_t refers to the logarithm of real price and ϵ_t to random disturbances. The models are estimated by ordinary least-squares.

For our nonlinear-parameter model category, we consider the specifications suggested by Pindyck (1999) that build on a basic Hotelling model for a depletable resource produced in a competitive market. The latter model is comprised of 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 (which is the long-run marginal cost) that itself fluctuates in the same manner. A class of models which integrates the above features is the generalized Ornstein-Uhlenbeck process, and Pindyck (1999) proposes

a discretized version of this process 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 the trend are potentially time-varying.

We consider the three possible econometric versions of the above[11]. 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. The three specifications are therefore given by:

$$P_t = c_1 + \phi_{1t} + c_2 t + \phi_{2t} t + c_4 P_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T, \tag{3}$$

$$P_t = c_1 + \phi_{1t} + (c_2 + \phi_2)\, t + c_4 P_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T, \tag{4}$$

$$P_t = (c_1 + \phi_1) + c_2 t + \phi_{2t} t + c_4 P_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T. \tag{5}$$

with

$$\phi_{1t} = \phi_{1,t-1} + v_{1t}\,, \qquad t = 1, \ldots, T,$$

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

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 time-varying-parameter (TVP) specifications are estimated by maximum likelihood using the Kalman filter. The estimations for all of the models are conducted on the logarithm of real prices.

[11] Pindyck (1999) shows that, given a quadratically detrended price $\bar{p} = p - \alpha_0 - \alpha_1 trend - \alpha_1 trend^2$, and assuming that the log of this detrended price follows a multivariate Ornstein-Uhlenbeck continuous process, a discrete-time version of the log price is obtained where the equation will contain both constant and time-varying slopes and trends.

Finally, we consider a set of purely statistical specifications with predetermined regressors. Unlike structural models, these specifications make no distinction between economic and non-economic propagating mechanisms. Accordingly, they often better capture price variations in oil price driven by non-fundamental factors[12].

The most general specification that we consider is an autoregressive model with linear and quadratic trends (denoted **AR-QT** model) given by:

$$P_t = c_1 + c_2 t + c_3 t^2 + c_4 P_{t-1} + \epsilon_t, \quad t = 1, ..., T \tag{6}$$

Imposing different restrictions on this specification, we obtain (i) an autoregressive model with linear trend (denoted **AR-LT**), (ii) an autoregressive model without trends (referred to as **AR**), and (iii) a random walk with drift model (**RW-WD**). Thus we have four statistical forecasting models for the logarithm of real prices, all of which are estimated using ordinary least-squares[13].

For all of the models discussed in this section, model evaluations are made in comparison with the random walk benchmark[14]. Throughout the text we also refer to this model as the no-change forecast model. The model is defined as $P_t = P_{t-1} + \epsilon_t$, where P_t refers to the logarithm of the average real price of oil over the year, and ϵ_t refers to random disturbances.

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). The data, which are annual averages of nominal producer prices in the U.S., were later updated by Pindyck (1999) until 1995 using data from

[12] Please refer to Kim and Nelson (1999) for explanations.

[13] As explained above, all model parameters are estimated recursively, and in particular, the drift of the RW-WD, given that the latter can be treated differently (see Alquist, Kilian, and Vigfusson (2013)).

[14] Unit root tests applied to the log price series (not reported) show that the unit root hypothesis cannot be rejected at the 5 percent level.

the EIA and, for 1996, with data from the *Wall Street Journal*[15]. For our part, we further update the data until 2011 using annual averages of U.S. imported Refiner Acquisition Cost of Crude Oil obtained from EIA publications. The series are deflated with the U.S. wholesale price index for all commodities until 1970, and with the producers price index thereafter. Thus we have 142 annual observations in our sample.

The data from 1870 to 1994 are used to conduct initial estimations of our alternative econometric models, and out-of-sample forecasts are generated for each of these specifications one to fifteen years ahead. Thus, we conduct forecasts for $h = 1, 2, \ldots, 15$ years ahead. Where relevant, and for horizons greater than one year, 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. In addition, estimations are conducted recursively, whereby model parameters are updated ahead of a given forecast. Thus, following the initial 1870-1994 estimation sample, an observation is added to the sample, and each model is then re-estimated in order to generate new sets of one to fifteen-year-ahead out-of-sample forecasts. This process is repeated until all of the available data is processed.

As for the NEMS model forecasts, we rely on the publically available EIA annual forecasts of the nominal price of oil made each year from 1995 to 2011, and for forecast horizons of one up to fifteen years. That is, we do not conduct estimations on NEMS ourselves, but use the December release numbers directly. Nominal Prices are deflated using the Implicit Price Deflator[16].

Finally, we report two accuracy criteria for our evaluation period. 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

$$MSFE\left(h\right) = \frac{1}{H-h+1}\left(\sum_{j=1}^{H-h+1}\left(\hat{y}_{T+j-1+h|T+j-1} - y_{T+j-1+h}\right)^2\right) \tag{7}$$

[15]The data were generously provided by Pindyck.

[16]Note that the above will not produce real-time forecasts in the sense of Baumeister and Kilian (2012) since we use neither differing data vintages nor do we use implicit price deflator forecasts.

and the Mean Absolute Percentage 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[17]. 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. The $MAPE$ criterion is also used, for example by EIA in their own publications.

4. Results and Discussion

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 not feasible without knowledge of the underlying models, the orders of magnitude that we obtain reflect the economic importance of the results.

4.1 Baseline Results

Table 1 shows the relative performances of the different models for all of the forecast horizons. The upper panel reports results according to the MSFE criterion, while the lower panel pertains to the MAPE outcomes.

One thing that is immediately apparent from the table is that many of the models including NEMS are able to outperform the no change model (these cases are indicated in bold in the table), sometimes with considerable margins. This result suggests that, despite

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

the extensive volatility in oil prices, longer run systematic dynamics in this variable can be captured to some extent by all of the models examined[18]. Furthermore, while overall patterns are similar based on either MSFE or MAPE, the former favours more strongly the evidence regarding the extent to which the various models outperform the no-change forecast, while the latter favours more the number of times these models outperform the random walk.

Regarding the performance of NEMS, we find that the model is specially successful at forecasting one-year-ahead. Whether based on the MSFE or the MAPE, it is in fact the only model to outperform the random walk at this horizon, and by relatively important margins. This is a noteworthy result for users of these forecasts. 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 larger margins.

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. For instance, Baumeister and Kilian (2014b) show that, for one up to to four quarter forecast horizons, specific monthly-data-based VARs and futures-based models can generate quarterly forecasts of the real price of oil that are more accurate than no-change quarterly forecasts. Similarly, Baumeister, Kilian, and Zhou (2013) show that forecasts obtained from models based on product spreads, where the weights on the different spreads vary over time, outperform the no-change model up to 24 months ahead.[19] In these and subsequent studies it is also shown that how the random walk benchmark is defined matters for the obtained conclusions.

[18]This result is not specific to the particular sample examined. Forecasts pertaining to non-NEMS models for the 1985-2011 period (not reported to save space) generated similar findings.

[19]See also Baumeister and Kilian (2012), Baumeister, Guerin, and Kilian (2015), Chen (2014), and Baumeister, Kilian, and Lee (2014).

As a robustness check, we re-calculated all of our MSFE and MAPE percentages relative to random walk forecasts with the random walk benchmark defined differently. In particular, instead of using annual average values for our annual oil price observations from which random walk forecasts were calculated (as is the case with the results in Table 1), we considered two alternatives: (i) using the December value of a given year to represent the observation for the entire year, and (ii) using the value for last quarter of the year to represent annual observation. Random walk forecasts were then obtained based on these two definitions of our annual data, and relative MSFE and MAPE values were calculated.

Overall we found the results to be qualitatively similar to those obtained in our main comparisons reported in Table 1, and that conclusions did not change. If anything, models were found to outperform the random walk forecasts more often against these alternative random walk benchmarks than against the annual average random walk benchmark[20].

4.2 Combining Forecasts

The vast literature on forecasting has established that no one particular forecasting model can demonstrate superior forecast accuracy at all horizons and under all circumstances. This observation subsequently led to the development of alternative forecast combination strategies. That is, 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) and Baumeister et al. (2014) show that combining forecasts from VAR and futures-based models, as well as from models based on product spreads, perform generally better than no-change forecasts for horizons of one up to 18 months.

In this section, we assess the benefit of using various forecast combination approaches for medium and long term forecasting. Three different forecast combination methods are

[20]Results are available upon request.

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) combining forecasts based on the Akaike information criterion (denoted AIC in the tables of results)[21]. The model weights for the latter are obtained as follows: let us assume that M different models are available and could be used to obtain forecasts. For each model m, $m = 1, ..., M$, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) weights, respectively denoted as $w_{AIC}(m)$ and $w_{BIC}(m)$, are defined as follows. Let n denote the sample size, $k(m)$ the number of unknown parameters of model m, and $\sigma^2(m)$ the estimate of the variance of the error term. The AIC weight is defined as: $w_{AIC}(m) = \exp\left(-\frac{1}{2}\left[n\ln\left(\sigma^2(m)\right) + 2k(m)\right]\right)$, The BIC weight is defined as: $w_{BIC}(m) = \exp\left(-\frac{1}{2}\left[n\ln\left(\sigma^2(m)\right) + k(m)\ln(n)\right]\right)$. The simple forecast average corresponds to giving an equal weight to each of the M models, while the forecast median corresponds to the median of the forecasts obtained from each of the M models. Please see Hansen (2008) for more details.

The forecasts are obtained for three different model groupings. The first includes all of our considered models (denoted ALL), a second includes only our statistical models (denoted Statistical), a third includes only the models by Slade and Pindyck (denoted as 'Suggested' in the text), while a fourth grouping includes NEMS, as well as the best-performing models in the Statistical and Suggested categories. Table 2 reports the results expressed as percentage gains or deteriorations relative to the random walk forecast. The forecast evaluation period is again 1995-2011. The upper panel reports outcomes pertaining to the MSFE criterion, and the lower panel summarizes the results based on the MAPE criterion.

Amongst the four methods of combining forecasts, with the exception of the Statistical model grouping, simple averaging generally produces 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 we had obtained before. In particular, the no-change forecast continues to be

[21] We also calculated the combined forecasts based on the Bayesian information criterion, but since results were generally qualitatively similar to the AIC case, they are not reported here to save space.

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.

4.3 Assessing Directional Accuracy of Forecasts

Another criterion according to which we can evaluate our models is the extent to which, for a given forecast horizon, they can on average correctly predict the sign of forecast change. Knowing how well models perform according to this dimension can add further useful information to the information provided by the MSFE and MAPE criteria. We therefore test the null of no directional accuracy of forecasts for selected models and for specific forecast horizons. For this purpose, and given that multi-period forecasts are serially correlated, we make use of the methods proposed by Pesaran and Timmermann (2009). The test is applied in the context of a linear regression where, for a given model specification, the binary variable consisting of signs of oil price changes relative to the previous period is regressed on a constant and on the binary variable consisting of forecasted signs of oil price changes relative to the previous period. If, for a given model, the coefficient of the regression is found to be significant, then that model is interpreted as being able to, on average, accurately forecast the sign of the actual oil price change.

In Table 3, we report the obtained test p-values. The top panel summarizes the results for specific individual models, while the bottom panel reports the outcomes for various forecast combination options considered. The results indicate that, for certain horizons, some models and some forecast combination strategies are able to produce statistically significant accuracy with respect to predicting on average the direction of forecasts. Interestingly, while none of the considered specifications is able to produce directional accuracy of forecasts at the 2-year forecast horizon, some are successful at longer forecast horizons.

Examining the performance of individual models, we find that a few of the econometric specifications proposed in the literature, namely the QT and TVP-I models, as well as

the statistical autoregressive model with a quadratic trend, are able to produce directional accuracy for two different forecast horizons. The statistical autoregressive model with a linear trend is also able to generate forecast direction accuracy, but only in one case (that is, at the 5-year horizon). However, the remaining models, and in particular NEMS, are not similarly successful.

Turning to the forecast combination cases, we find that nine of the eleven strategies considered are able to predict forecast direction with statistical accuracy (at the 5 % level), but each only for a single forecast horizon. For example, taking the median forecast of all of the considered models proves to be successful at the 9-year forecast horizon, while combining forecasts from all models based on AIC weights is successful at the 5-year forecast horizon. We also find that, regardless of the manner in which forecasts from statistical models are combined, the combination is able to generate directional accuracy at the 3-year horizon. On the other hand, the combination strategy used affects forecast direction accuracy when considering forecasts from the 'Suggested' category of models or when combining the best-performing models from each model class.

At this stage we recommend to interpret these findings with caution, specially for the longer forecast horizon cases, given that the available number of observations for calculating test p-values are often too few to yield reliable conclusions.

5. Stability of Forecast Performance

Until this point in the text, all of the reported results above were obtained for a specific forecast evaluation period, namely 1995 to 2011. As explained, this choice was motivated by the availability of forecasts from the NEMS model. However, for non-NEMS models, it is possible to also produce forecasts over earlier evaluation periods. Doing so would allow us to examine the stability over time of the forecast performances of these specifications. We thus calculate the MSFE criterion for given forecast horizons and for selected models over successive forecast evaluation periods. In each case, forecasts are obtained over a particular

forecast evaluation period and for a given h from a specific model as well as from the random walk specification, and the relative MSFE is calculated.

Specifically, we consider the forecast horizons $h = 3, 5, 7$, and 10, and for our specifications we select two individual models and three forecast combination strategies. We fix each forecast evaluation period to a 17-year time span. The first estimation period is 1870-1984. The corresponding forecast evaluation period over which the various MSFEs are calculated is 1985-2001. The next estimation period is 1870-1985, with a forecast evaluation period of 1986-2002, over which MSFEs are again calculated. This process is continued until the 1870-1994 esimation period, and the related 1995-2011 forecast evaluation period being, are reached. We thus document the evolution of the MSFEs for the different models relative to the no-change MSFE, and we evaluate the relative forecast performances of these models over time.

Figures 1-4 show the evolution of the MSFE ratios for the two individual models and the three forecast combination strategies for the four forecast horizons. The individual models are two of the suggested ones in the literature, namely the linear trend specification of Slade (1982) and the TVP-I model of Pindyck (1999). The considered forecast combinations pertain to the three main model groupings: all models (denoted Comb-All), all of the statistical models (denoted Comb-Stat), and all the econometric models suggested by Slade and by Pindyck (denoted Comb-Sugg). Except for the middle case where AIC weights are applied, forecast combinations were generated by taking the average of forecasts of the models in the group.

We find that forecast performances vary with the considered forecast horizon and with the considered forecast-generating specification. In particular, for a given forecast horizon h, most models perform worse than the random walk at some time periods, and more accurately at other periods. The exception is the forecast combination over the 'Suggested' model grouping case (referred to as Comb-Sugg in the figures) and only for the 10-year forecast horizon, where forecast performance is very stable over time. In contrast, the least stable

forecast performances are obtained with the forecast combination over the 'Statistical' model grouping case (Comb-Stat in the figures). Finally, we note that the majority of the chosen specifications yield more accurate predictions than the random walk when $h = 10$, and that the most consistent and best-performing model across all of the forecast horizons is the TVP-I model. This lends credibility to economically-motivated yet flexible empirical models that also control for parameter dynamics.

6. Conclusion

Expert outlooks regarding the future path of oil prices are often relied on by industry participants and policymaking bodies for their forecasting needs. Yet we know little on how accurate these are. Using the regular publications by the Energy Information Administration, we examined the accuracy of annual real oil price out-of-sample forecasts generated by the NEMS model of the Agency, for horizons of up to fifteen years.

Our results showed that the EIA model is quite successful at beating the benchmark random walk model, but only at the very short and at the longer ends of the forecast horizons. We also showed that, for the longer horizons, a range of simple econometric forecasting models often produced similar if not better accuracy than the EIA model. Among these, specifications that also control for parameter dynamics were, for some forecast horizons, also found to exhibit some stability in their forecast performances over time.

Finally, while we found that combining forecasts did not change the overall patterns, there were nonetheless some advantages to this strategy. In particular, with most of these combinations, additional accuracy gains were obtained at intermediate forecast horizons. Similarly, when forecasts from economically-founded empirical models were combined, considerable forecast performance stability was obtained at the 7-year and 10-year forecast horizons.

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.

Baumeister, C., P. Guerin, and L. Kilian. 2015. "Do High-Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work." *International Journal of Forecasting* 32: 238–252.

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

———. 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. 2014. "Are there Gains from Pooling Real-Time Oil Price Forecasts?" *Energy Economics* 46: S33–S43.

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. 2015. "The Convenience Yield and the Informational Content of the Oil Futures Price." *The Energy Journal* 36: 29–46.

Chen, S. 2014. "Forecasting Crude Oil Price Movements with Oil-Sensitive Stocks." *Economic Inquiry* 52: 830–44.

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.

Fischer, C., E. Herrnstadt, and R. Morgenstern. 2008. *Understanding Errors in EIA Projections of Energy Demand*. Technical report, Resources for the Future Discussion Paper 07-54.

Hamilton, J. 2011. "Nonlinearities and the Macroeconomic Effects of Oil Prices." *Macroeconomic Dynamics* 15: 364–78.

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

Hansen, B.E. 2008. "Least-squares forecast averaging." *Journal of Econometrics* 146(2): 342–350.

Huntington, H. 2011. "Backcasting US Oil Demand Over a Turbulent Decade." *Energy Policy* 39: 5674–80.

Kilian, L. 2008. "The Economic Effects of Energy Price Shocks." *Journal of Economic Literature* 46: 871–909.

———. 2014. "Oil Price Shocks: Causes and Consequences." *Annual Review of Reseource Economics* 6: 133–54.

Kilian, L. and D. Murphy. 2014. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil." *Journal of Applied Econometrics* 29: 454–78.

Kilian, L. and R. Vigfusson. 2011. "Nonlinearities in the Oil Price-Output Relationship." *Macroeconomic Dynamics* 15: 337–63.

———. 2013. "Do Oil Prices Help forecast U.S. Real GDP? The Role of Nonlinearities and Asymmetries." *Journal of Business and Economic Statistics* 31: 78–93.

Kim, C.J. and C.R. Nelson. 1999. *State-Space Models with Regime-Switching*. USA: The MIT Press.

Lynch, M. 2002. "Forecasting Oil Supply: Theory and Practice." *The Quarterly Revoiew of Economics and Finance* 42: 373–89.

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

O'Neil, B. and M. Desai. 2005. "Accuracy of Past Projections of US Energy Consumption." *Energy Policy* 33: 979–93.

Pesaran, M.H. and A. Timmermann. 2009. "Testing Dependence Among Serially Correlated Multicategory Variables." *Journal of the American Statistical Association* 104(485): 325–337.

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

Ravazzolo, F. and P. Rothman. 2013. "Oil and U.S. GDP: A Real-Time Out-of-Sample Examination." *Journal of Money, Credit, and Banking* 45: 449–63.

Sanders, D., M. Manfredo, and K. Boris. 2009. "Evaluating Information in Multiple-Horizon forecasts: The DOE's Energy Price Forecasts." *Energy Economics* 31: 189–96.

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.

Winebrake, J. and D. Sakva. 2006. "An Evaluation of Errors in US Energy Forecasts: 1982-2003." *Energy Policy* 34: 3475–83.

Fcst Hzn	RW-WD	AR	AR-LT	AR-QT	LT	QT	TVP-IS	TVP-I	TVP-S	NEMS
	Relative MSFE									
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
5	14.98	69.70	22.04	**-16.12**	31.56	**-10.42**	27.65	14.78	25.26	27.36
7	1.44	35.57	**-7.67**	**-32.33**	**-14.42**	**-44.85**	3.68	**-10.04**	1.40	11.26
9	**-4.34**	20.15	**-20.23**	**-41.36**	**-30.54**	**-63.79**	**-8.08**	**-20.50**	**-10.72**	**-9.59**
11	2.26	24.04	**-20.64**	**-49.97**	**-30.50**	**-74.80**	**-8.56**	**-21.61**	**-11.24**	**-11.15**
13	**-13.69**	8.48	**-28.27**	**-51.81**	**-37.98**	**-81.82**	**-15.60**	**-27.44**	**-18.30**	**-29.50**
15	**-2.17**	22.64	**-25.58**	**-65.42**	**-33.09**	**-85.36**	**-11.22**	**-25.92**	**-14.02**	**-30.53**
	Relative MAPE									
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
5	6.02	33.99	4.63	**-2.16**	4.57	33.43	6.98	2.20	5.77	7.67
7	**-1.31**	21.92	**-14.93**	**-16.62**	**-16.50**	**-3.38**	**-4.55**	**-16.33**	**-7.15**	**-2.68**
9	**-3.43**	14.33	**-16.73**	**-34.81**	**-25.32**	**-40.90**	**-5.99**	**-16.79**	**-8.08**	**-12.65**
11	**-1.92**	9.53	**-15.13**	**-35.81**	**-22.94**	**-63.95**	**-7.08**	**-15.24**	**-8.82**	**-12.84**
13	**-6.09**	7.36	**-16.00**	**-38.57**	**-22.08**	**-68.40**	**-7.62**	**-15.70**	**-9.26**	**-16.09**
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 MSFE or 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 1: Relative MSFE and MAPE, Evaluation Period: 1995-2011

h	All			Statistical			Suggested			Best	
	AVE	MED	AIC	AVE	MED	AIC	AVE	MED	AIC	AVE	MED
	Relative MSFE										
1	35.44	32.40	30.09	11.81	17.46	30.30	64.95	42.34	41.24	**-18.06**	10.32
2	8.00	10.32	3.13	**-16.29**	**-12.63**	**-6.46**	36.03	22.58	19.59	**-7.67**	**-1.36**
3	54.11	53.22	50.29	28.65	31.97	44.59	70.44	53.04	53.88	29.32	29.47
5	11.84	18.15	16.54	2.62	1.28	3.21	11.40	17.37	27.16	6.36	4.99
7	**-10.76**	**-7.16**	**-4.97**	**-10.91**	**-17.44**	**-6.47**	**-7.73**	3.18	**-25.42**	**-13.32**	**-11.07**
9	**-21.56**	**-17.29**	**-14.38**	**-18.21**	**-16.73**	**-27.76**	**-30.80**	**-20.50**	**-8.47**	**-33.23**	**-26.52**
11	**-22.97**	**-17.67**	**-14.13**	**-1.95**	**-15.31**	**-31.48**	**-33.65**	**-21.53**	**-8.86**	**-40.42**	**-30.58**
13	**-30.28**	**-23.28**	**-19.80**	**-25.24**	**-21.43**	**-36.78**	**-40.55**	**-27.44**	**-15.83**	**-59.29**	**-46.54**
15	**-29.71**	**-19.85**	**-18.13**	**-26.42**	**-15.04**	**-41.63**	**-40.61**	**-25.92**	**-11.50**	**-73.39**	**-51.46**
	Relative MAPE										
1	6.52	2.57	-0.63	-0.41	4.85	9.49	19.24	7.55	4.25	**-20.87**	**-5.37**
2	**-10.34**	**-11.48**	**-11.50**	**-15.88**	**-15.04**	**-12.94**	2.62	**-8.33**	**-6.22**	**-15.80**	**-15.36**
3	6.93	2.66	5.97	**-1.78**	**-7.11**	3.39	15.04	4.31	9.86	3.99	**-0.33**
5	2.25	1.16	2.96	**-0.16**	**-3.72**	1.29	1.63	2.35	6.78	1.63	**-0.80**
7	**-15.83**	**-13.97**	**-10.71**	**-15.69**	**-13.78**	**-17.43**	**-16.09**	**-15.62**	**-4.90**	**-16.82**	**-16.21**
9	**-18.94**	**-13.49**	**-10.81**	**-15.77**	**-12.84**	**-25.25**	**-26.30**	**-16.79**	**-6.25**	**-29.84**	**-24.58**
11	**-16.63**	**-12.37**	**-10.12**	**-13.22**	**-10.64**	**-21.94**	**-24.60**	**-15.22**	**-7.25**	**-27.306**	**-21.34**
13	**-17.82**	**-12.71**	**-10.79**	**-15.17**	**-11.39**	**-23.66**	**-24.99**	**-15.70**	**-7.77**	**-30.64**	**-22.85**
15	**-20.22**	**-13.49**	**-12.00**	**-17.23**	**-10.47**	**-28.00**	**-28.66**	**-17.38**	**-8.20**	**-43.40**	**-35.36**

Note: Numbers shown are the percent improvements or deteriorations in MSFE or 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 2: Forecast Combinations, Relative MFSE and MAPE; 1995-2011

	Forecast Horizon					
Individual Models	2	3	5	8	9	11
AR-LT	0.428	1.000	**0.041**	-	0.127	0.802
AR-QT	0.831	**0.000**	-	-	-	**0.001**
LT	-	-	-	-	-	-
QT	-	-	**0.035**	-	-	**0.000**
TVP-IS	0.429	0.432	0.561	0.332	0.156	0.863
TVP-I	0.824	1.000	**0.015**	-	**0.000**	-
TVP-S	0.405	0.563	0.587	0.332	0.156	0.863
NEMS	1.000	0.334	0.720	-	-	-
Forecast Combinations						
All Average	0.824	0.125	0.099	-	-	-
All Median	0.482	0.164	0.635	-	**0.000**	-
All AIC	0.364	0.810	**0.015**	-	-	-
Statistical Average	0.120	**0.000**	-	-	-	-
Statistical Median	0.741	**0.000**	-	-	-	-
Statistical AIC	0.890	**0.000**	-	-	-	-
Suggested Average	0.482	0.305	**0.000**	-	-	-
Suggested Median	0.496	0.775	0.886	-	**0.000**	-
Suggested AIC	0.429	0.432	0.561	0.332	0.156	0.863
Best Average	0.967	**0.000**	-	-	-	-
Best Median	0.859	1.000	**0.015**	-	0.127	0.802

Note: A dash implies that the p-value could not be obtained due to collinearity

Table 3: Direction of Forecast Test P-values, Selected Horizons

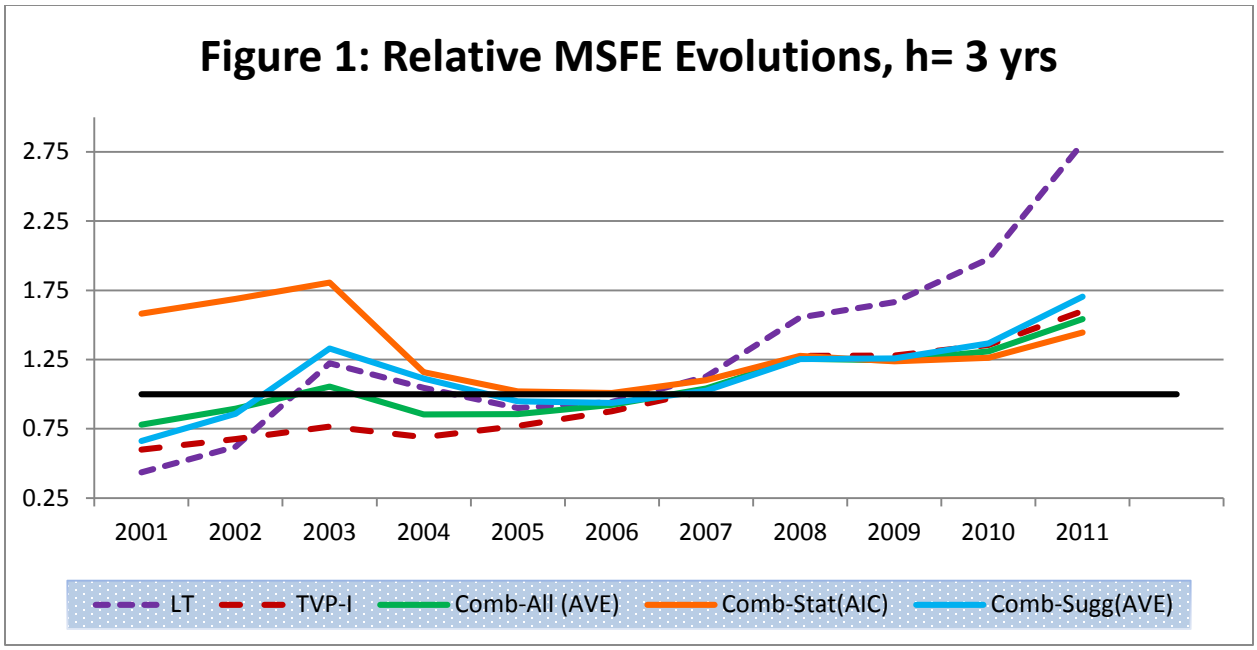
Figure 1: Relative MSFE Evolutions, h= 3 yrs
2.75
2.25
1.75
1.25
0.75
0.25
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
LT
TVP-I
Comb-All (AVE)
Comb-Stat(AIC)
Comb-Sugg(AVE)

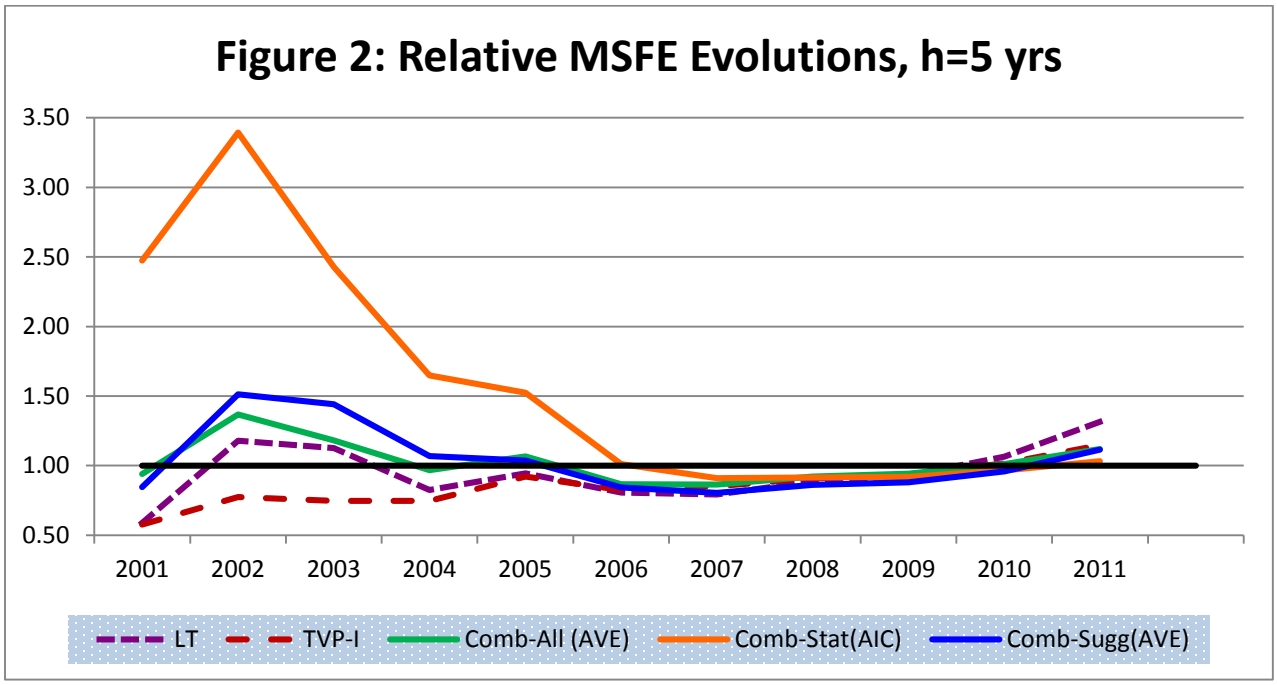
Figure 2: Relative MSFE Evolutions, h=5 yrs
3.50
3.00
2.50
2.00
1.50
1.00
0.50
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
LT
TVP-I
Comb-All (AVE)
Comb-Stat(AIC)
Comb-Sugg(AVE)

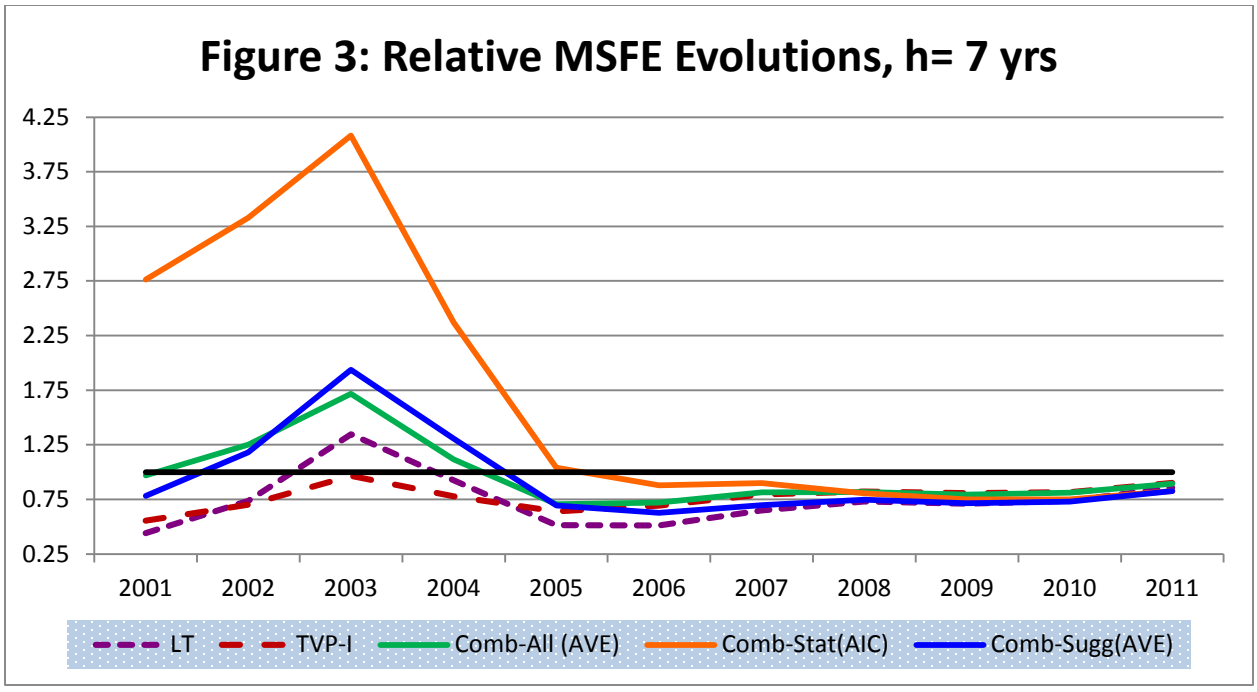
Figure 3: Relative MSFE Evolutions, h= 7 yrs
4.25
3.75
3.25
2.75
2.25
1.75
1.25
0.75
0.25
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
LT
TVP-I
Comb-All (AVE)
Comb-Stat(AIC)
Comb-Sugg(AVE)

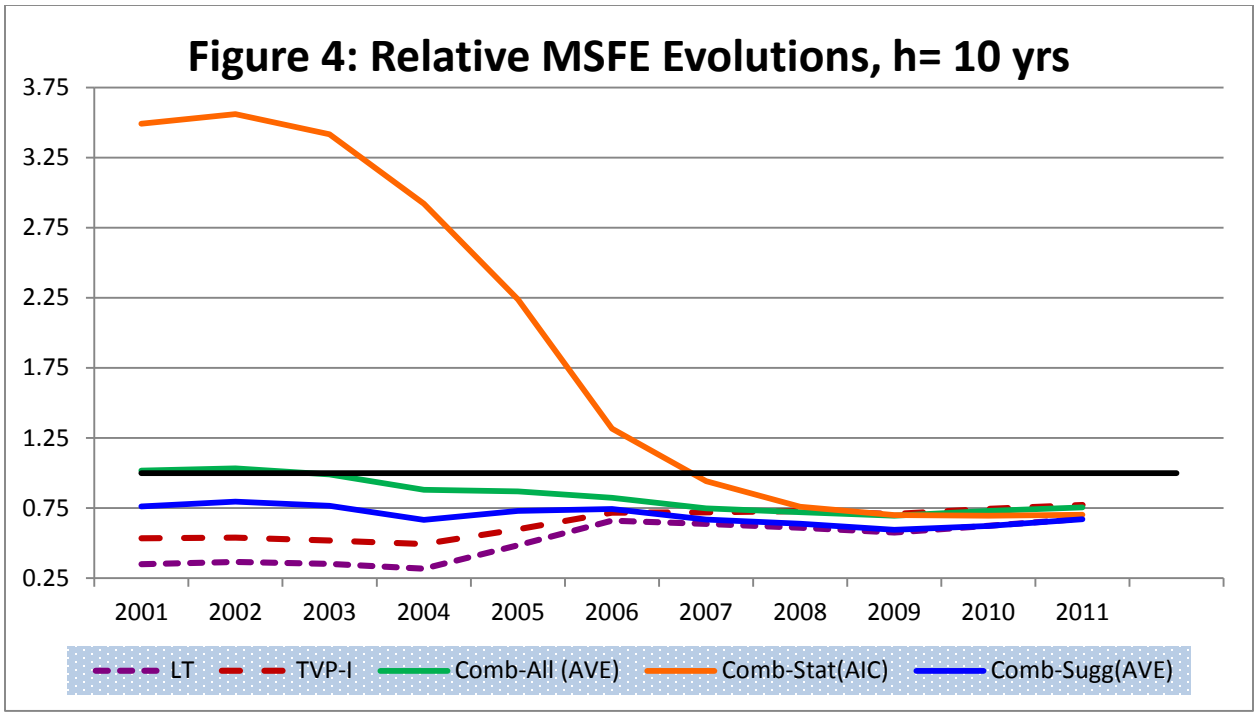
Figure 4: Relative MSFE Evolutions, h= 10 yrs
3.75
3.25
2.75
2.25
1.75
1.25
0.75
0.25
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
LT
TVP-I
Comb-All (AVE)
Comb-Stat(AIC)
Comb-Sugg(AVE)