Are published oil price forecasts efficient?

Hussein Moghaddam, Jalal Dehnavi and Franz Wirl

Department of Industry, Energy and Environment, Faculty of Business, Economics and Statistics, University of Vienna, Oscar Morgenstern Platz, 1, A- 1090 Vienna, Austria. Email: hussein.moghaddam@univie.ac.at, Email: jalal.dehnavi@univie.ac.at, Email: franz.wirl@univie.ac.at

Abstract

Oil prices are crucial for a wide area of economic decisions ranging from households over business to economic policymaking. Therefore, oil price projections are a crucial input to economic models and to economic activity forecasts. This paper assesses the accuracy and efficiency of crude oil price forecasts published by different organisations, think tanks and companies. Since the sequence of published forecasts appears as smooth, the weak efficiency criterion is clearly violated. Even combining forecasts, cannot increase efficiency due to high correlation among various forecasts. This pattern of oil price forecasts can be attributed to combining myopia (use current oil price levels as a basis) with Hotelling-type exponential growth. Another behavioural explanation in source of inefficiencies is that forecasters prefer to harmonise their forecasts with other forecasters in order to be not an outlier.

1. Introduction

Decisions at the macro (governments) and micro (individuals and firms) levels depend on expectations about the future economic developments, which are often based on published forecasts. Therefore, the efficiency and accuracy of forecasts is of crucial concern for many decision makers (Mamatzakis and Koutsomanoli-Filippaki, 2014) given the garbage in, garbage out (GIGO) principle. This includes governments when planning their budgets, managers and individuals about investment projects involving billions of dollars or less in case of individuals, e.g. buying hybrid car. Demand, supply and price forecasts are crucial for all energy market participants (Sanders *et al.*, 2008, 2009) yet oil price forecasts of particular importance due to implications of oil prices on all aspects of energy markets, on commodity markets and on overall economic activity, GDP-growth and inflation (Mamatzakis and Koutsomanoli-Filippaki, 2014).

This importance has led to the development of different and efficient forecasting techniques and to the efforts of statisticians, economists and market players (companies, governments and international organisations) to forecast. Two techniques have gained popularity due to their accuracy and efficiency: econometric models (both parametric and non-parametric) and more recently computational approaches like neural networks (Liu *et al.*, 2002).

Many studies have employed econometric models to forecast oil prices, Kaufmann (1995), Alquist *et al.* (2011), Baumeister and Kilian (2013) and Manescu and Van Robays (2014); Kaboudan (2001), Yu *et al.* (2008) and Gabralla and Abraham (2013) have applied computational techniques. However, there is no consensus on the most reliable method (Liu *et al.*, 2002) and it is close to a religious war between the two camps. Hence, scholars have begun, in parallel, to evaluate forecasts (Mamatzakis and Koutsomanoli-Filippaki, 2014). Consequently, they have introduced various performance measures, such as forecast accuracy (Clement, 1999; Manescu and Van Robays, 2014), bias (Artis and Marcellino, 2001), combination and encompassing (Bates and Granger, 1969; Clemen and Winkler, 1986), efficiency (Mincer and Zarnowitz, 1969; Nordhaus, 1987) and sign predictability (Nyberg, 2011).

Forecast efficiency has been investigated in Cooper and Nelson (1975), Nelson (1984), Fair and Shiller (1988, 1990), Dovern and Weisser (2008), Allan (2012), and Genberg and Martinez (2014) not only for macroeconomic variables but also for football (Sillanpää and Heino, 2013), baseball (due to excellent performance statistics, Silver (2012)), agricultural commodity markets (Von Bailey and Brorsen, 1998), managers' forecasts (Khan *et al.*, 2013), demographic variables (Tayman *et al.*, 2011), etc. This study focuses on the forecast efficiency of oil prices, more precisely on weak efficiency, which requires that forecasts contain the information sets of all past forecasts (Nordhaus, 1987). Although this is only a necessary criterion for efficient forecasts and other tests will be included, it is of particular practical relevance as it allows forecasters to check and to learn from their past revisions and to assess the efficiency of a forecast prior to its realisation.

The energy sector receives more than its fair share of forecasts (Ahlstrom *et al.*, 2013). A number of studies investigate their efficiency focusing on oil supply (Floris *et al.*, 2001; Lynch, 2002 and Sanders *et al.*, 2009) and oil consumption (Shlyakhter *et al.*, 1994), however, there is relatively little research about the efficiency of oil price forecasts in spite of many and even regular forecasts. Sanders *et al.* (2008 and 2009) investigate the efficiency of price forecasts for several energy commodities (including crude oil) published by the United States Department of Energy (DOE hereafter). They find that the price forecasts for gasoline, diesel fuel, natural gas and electricity are efficient in the long term,[1] that the price forecasts for crude oil provide incremental information for up to three quarters. Similarly, Mamatzakis and Koutsomanoli-Filippaki (2014) examined the rationality of DOE price forecasts for energy commodities, including crude oil. They opt for an asymmetric underlying loss function with respect to positive versus negative forecast errors. The above-mentioned studies have assessed the rationality of DOE's forecasts using quarterly data, while the earlier, and in this area, seminal paper of Nordhaus (1987) finds DOE's oil price are forecasts inefficient because of the autocorrelation[2] problem; this criterion will applied in this paper too.

The price of oil decreased sharply since August 2014, actually almost collapsed from US$100 per barrel in July 2014 (and $140 in summer 2008) to less than $50 in July 2017. The various energy market outlooks in 2013 all failed to foresee the price fall towards the end of 2014 in spite of applying different forecasting techniques. Of course, forecasting is difficult, in particular, if it concerns the future as Yogi Berra famously quipped or as John Kenneth Galbraith believed that "The only function of economic forecasting is to make astrology look respectable".[3] Therefore, our study tries to assess the informational efficiency of real oil price forecasts issued regularly by different institutions.

2. Forecast efficiency

Two main concepts of efficiency (rationality) dominate in the literature: (i) strong efficiency and (ii) weak efficiency. A forecast exhibits strong efficiency when the sum of squared errors is minimised subject to all available information; with weak efficiency, the minimisation is subject to all past forecasts (Nordhaus, 1987).[4] Since it is almost impossible to test strong efficiency in practice [alone data sets, models and assumptions used by forecasters are not publicly available (Dovern *et al.*, 2013)], the focus is on weak efficiency as the necessary but insufficient condition for strong efficiency. This concept has been applied to two different categories of forecasts: rolling event and fixed event forecasts, (REFs and FEFs henceforth) (Nordhaus, 1987; Clements and Hendry, 2008). FEFs involve a series of forecast for a target date at varying dates prior to the forecast event (the forecast event is kept fixed, while the forecasting horizon shrinks). In contrast, the horizon, h, is fixed in REFs and the forecast event is always separated by h periods from the forecast origin (Bernanke, 2007). Given our data set, we analyse the efficiency of fixed event forecasts, e.g. of oil prices in 2010.

To date, several tests and techniques have emerged to examine the efficiency of FEFs (Berger and Krane, 1985; Nordhaus, 1987; Clements, 1997; Davies and Lahiri, 1999; Reis, 2006 and Coibion and Gorodnichenko, 2010).[5] Nonetheless, the most common method is the one suggested by Nordhaus (1987) and it will be applied in this paper.[6]

2.1. Data and methodology

Our data set consists of forecasts from 13 institutes, universities, banks, oil companies and organisations (see **Table 1**). Their names will be abbreviated and the years written in the second column of the Table 1 that shows the periods in which they forecast oil prices, or those years to which we had access.

Table 1 Forecasters and period of forecasts

Institute (Abbreviation)	Forecasting years
Deutsche Banc Alex. Brown (DB)	1999–2010
International Energy Outlook and Annual Energy Outlook (DOE/EIA)	1979–2014
Standard & Poor's Platt's, formerly DRI/McGraw-Hill (DRI)	1981–2001
Energy Ventures Analysis (EVA)	2005–2014
Gas Research Institute (GRI)	1989–2001
International Energy Agency, World Energy Outlook (IEA)	1994–2014
IHS/Global Insight (IHSGI)	2008–2012
The Inter-industry Forecasting Project at the University of Maryland (INFORUM)	2008–2012
Petroleum Economics, Ltd. (PEL)	1994–2005
Petroleum Industry Research Associates (PIRA)	1994–2005
Strategic Energy and Economic Research (SEER)	2003–2012
Wharton Econometric Forecasting Associates (WEFA)	1986–2001
World Bank Group (WB)	2001–2014

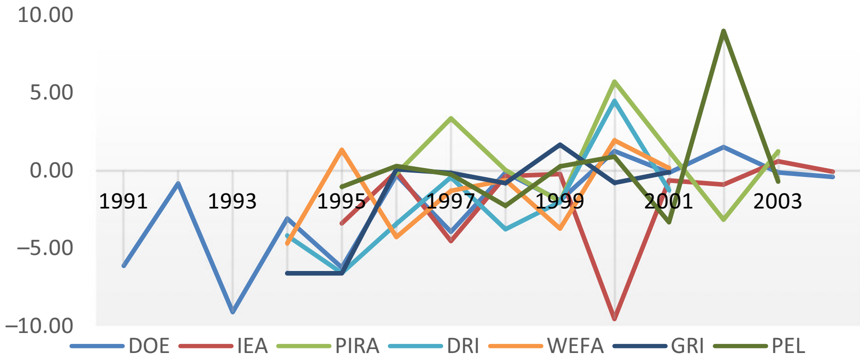
We focus on the forecasts for the now historical (target) years 2005, 2010 and 2015 although the forecasts made between 1979 and 2014 are typically available at 5-year intervals and extend to 2035 and beyond; interesting results for other target years are also reported. Some of the forecasts refer to different crudes, e.g. West Texas Intermediate (WTI), US average prices for crude oil, Brent or a basket of crude oil.[7] However, these differences are irrelevant for long-term forecasting since there exists, by and large, a 'common oil pool' opposite to natural gas with substantial regional price differences, which seem puzzling, see Dehnavi *et al.* (2015).

2.2. Graphical analysis

Given a sequence of forecasts for the same event but made at different points in time, efficient forecasts should appear jagged (Nordhaus, 1987). This means that forecast revisions should be unpredictable. Technically, a sequence of forecasts for the same event is called *weak efficient* if it satisfies the criterion for a random walk (Clements,

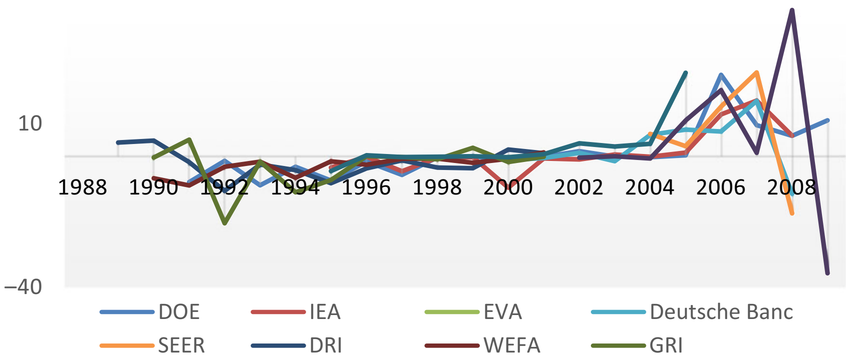
Figure 1 (a) Forecasts by different institutions for 2005 (Real US $/barrel). (b) Forecasts by different institutions for 2010 (Real US $/barrel). (c) Forecasts by different institutions for 2015 (Real US $/barrel).
Sources a to c: Authors' calculations based on different sources. Reference year is 2010.
Note: Similar patterns were observed for the forecasts of other institutions (IHSGI, PIRA and EVA) and for other target years. [Colour figure can be viewed at wileyonlinelibrary.com]

(a)



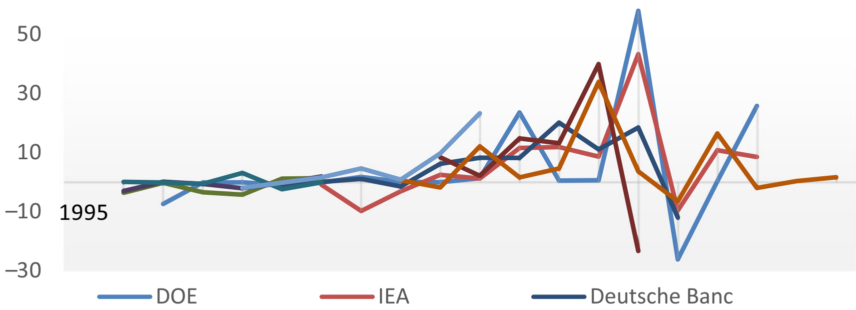
Forecasts by different institutions for 2005- (Real US $/barrel)

(b)



Forecasts by different institutions for 2010- (Real US $/barrel)

(c)



Forecasts by different institutions for 2015- (Real US $/barrel)

1997; Stekler, 2002; Dovern *et al.*, 2013). In contrast, smooth forecasts point towards inefficiency as they incorporate new information in their revisions sluggishly.

Figure 1 compares the forecasts by different institutions for three target years (2005, 2010 and 2015). For 2005, forecasts are characterised by downward revisions until 2000, followed by upward and downward revisions. This continuation of revisions in the same direction is known as 'hot hand' in gambling (compare with Croson and Sundali (2005) for an empirical investigation). A similar observation holds for the forecasts of 2010 up to 2000 but which are then followed by upward revisions until 2008. In comparison, the revisions for 2015 lack a pattern. As we will demonstrate below, many of the substantial revisions are due to changes in the current oil price, providing the starting point for the extrapolation.

3. Empirical results

3.1. Testing (Weak) forecast efficiency

In the following equations, p_{it}^T denotes a forecast of forecaster i at time t for the oil price in the target year (T). Weak efficiency requires that any forecast revision, $\Delta p_{it}^T := p_{it}^T - p_{it-1}^T$, is a martingale and thus independent of past information, in particular, of past forecasts and their revisions. This complies with common sense: if we can learn or even improve the forecaster's own forecast by extrapolating, e.g. expecting a downward revision after a series of downward revisions (Nordhaus (1987) uses the example of the nuclear power forecasts by the IEA, the GDP projections after Lehman crisis followed a similar pattern in many countries), then it cannot be efficient. Therefore, we run the following regression:

$$\Delta p_{it}^T = \rho_i \Delta p_{it-1}^T + \varepsilon_{it} \tag{1}$$

Weak efficiency, our hypothesis, corresponds to random walk and thus we test H_0 of $\rho_i = 0$ against its alternatives $\rho_i = 0$). Therefore, a rejection of the null-hypothesis implies that forecasts do not reflect all new information. The results obtained from this weak efficiency test are presented in **Table 2**. In spite of a sequence of revisions of the same sign, only one out of seven forecasts is inefficient according to this criterion for 2005; based on the graphical analysis, the efficiency of all forecasts for 2015 except one is less surprising. However, only three out of nine institutions could meet this efficiency criterion for 2010. An important feature of these kinds of efficiency tests is that they enable us to evaluate the efficiency of future forecasts ex ante and all the forecasts for 2020 look efficient so far.

3.2. Second test for forecast efficiency

Any forecast is the sum of past revisions. The forecast cannot be weakly efficient if positive (negative) revisions tend to be followed by further positive (negative) revisions.

Table 2 Testing the random walk hypothesis for forecast revisions

	Forecasts for			
	2005	2010	2015	2020
Institution	ρ_i			
IEA	0.08 (0.34)	0.60** (0.24)	0.09 (0.32)	0.15 (0.29)
DOE	0.31 (0.24)	0.47** (0.22)	−0.30 (0.27)	−0.25 (0.26)
DRI	0.41 (0.34)	0.05* (0.25)	0.29 (0.40)	0.33 (0.58)
WEFA	−0.16 (0.32)	0.62*** (0.20)	0.07 (0.39)	−0.01 (0.89)
GRI	0.44* (0.23)	0.02 (0.32)	−0.56 (0.41)	n.a.
PEL	−0.39 (0.34)	2.02* (0.90)	2** (0.52)	n.a.
PIRA	−0.21 (0.40)	0.66*** (0.43)	0.76 (0.70)	0.11 (0.55)
WB	n.a.	−0.51 (0.45)	0.07 (0.32)	−0.48 (0.58)
DB	n.a.	0.09 (0.43)	0.45 (0.32)	−0.13 (0.34)

Source: Authors' estimations.

*, ** and *** indicate statistical significant difference from $\rho_i = 0$ at the 10%, 5% and 1% levels, respectively. The numbers in parentheses are standard deviations. N.a., the small sample size did not allow us to run the regression for all of the target years and institutions.

This suggests a second test of weak efficiency in which the total revision from date t onward is considered a function of the cumulative revisions up to time $t-1$. Hence, when a forecast is efficient, the sum of future forecast revisions $(\sum_{t=\theta}^{T} \Delta p_{it}^{T})$ should be independent of cumulative past revisions $(\sum_{t=0}^{\theta-1} \Delta p_{it}^{T})$. Therefore, we run the following regression for three target years 2005, 2010 and 2015;

$$\sum_{t=\theta}^{T-1} \Delta p_{it}^{T} = \alpha_i \sum_{t=0}^{\theta-1} \Delta p_{it}^{T} + v_{it} \tag{2}$$

Weak efficiency corresponds to the H_0 that $\alpha_i = 0$; **Table 3** reports the results for the estimation of the equation (2).

The results obtained from the second forecast efficiency test for 2005 and 2010 (but not for 2015) are mainly in line with the first test. Five out of seven forecasts are inefficient in 2015. The advantage of the second test is it reveals a higher degree of inefficiency, as it works with the aggregation of forecast revisions instead of annual revisions. This adjusts for outliers, e.g. the forecast revisions for 2015 were mainly smooth except for two considerable revisions in 2009 and 2010; see Fig. 1. Comparing the results in Tables 1 and 2, it is worth mentioning that informational efficiency varies across the three different target years for each of the forecasts. However, we refrain from

Table 3 Second test of forecast efficiency

	Forecasts for		
	2005	2010	2015
Institution	α_i		
IEA	0.24 (0.24)	−1.99*** (0.25)	−1.55*** (0.09)
DOE/EIA	0.01 (0.07)	−1.51*** (0.03)	−1.51*** (0.08)
DRI	0.54 (0.74)	1.03*** (0.16)	0.15 (0.30)
WEFA	0.05 (0.13)	1.03*** (0.21)	0.06 (0.26)
GRI	−0.01 (0.17)	0.78*** (0.09)	−1.12*** (0.18)
PEL	0.62** (0.20)	0.87*** (0.24)	—
PIRA	0.06 (0.17)	0.27 (0.49)	0.91 (0.41)
WB	—	−0.23 (0.34)	0.20** (0.09)
DB	—	−0.10 (0.21)	0.12 (0.14)

Source: Authors' estimations.
*, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. The numbers in parentheses are standard deviations. Because of a lack in data, we were not able to conduct the second forecast efficiency test for 2020.

validating the various forecasts inter alia, because the number of observations vary substantially across forecasters (see Table 1).

3.3. Testing for unbiased forecasts

Optimal forecasts should be unbiased and efficient (Diebold and Lopez, 1996). This suggests to test the unbiasedness of forecasts, however, this can be only performed ex post after the realisation of the forecast variable (for further details see: Sanders *et al.*, 2009). That is, the mean forecasting error, $e_{it} = Aoil^T - Foil_{it}^T$, where $Foil_t^T$ is oil price forecast for the target year T made at time t, has to be zero. The hypothesis for an unbiased forecast is tested in **Table 4**.

All forecasts are biased upwards for 2005 and 2010, i.e. they overestimate the oil price. However, three forecasts for 2015 are unbiased, namely IEA, DOE and WB. We try to explain the origin of this bias in section five. Combining the results in Tables 2 and 4 for weakly efficient and unbiased results (thus ignoring 2020), only 3 out of 21 forecasts, namely IEA, DOE and WB, all for 2015, meet these two criteria.

Another interesting observation is that this bias shrinks, if we use the year prior to the forecast one for our comparison, in particular, for oil price forecasts in 2010 and 2015. Given the forecast trends in **Fig. 2**, the real oil price was \$60 in 2009 and about \$83 in

Table 4 Forecast Bias Test

Forecasts for			
Institution	2005	2010	2015
IEA	−21.25*** (2.02)	−35.99*** (5.18)	6.48 (9.17)
DOE/EIA	−19.35*** (2.71)	−32.11*** (3.88)	4.71 (7.58)
DRI	−26.02*** (2.25)	−38.80*** (3.94)	−20.12*** (1.59)
GRI	−29.16*** (1.66)	−39.97*** (4.38)	−26.97*** (0.37)
PIRA	−28.10*** (1.12)	−45.51*** (2.16)	−26.00** (7.98)
WB	−31.01*** (1.07)	−31.60*** (8.81)	−0.21 (7.65)
DB	−31.84*** (0.31)	−42.85*** (4.19)	−8.77* (7.79)

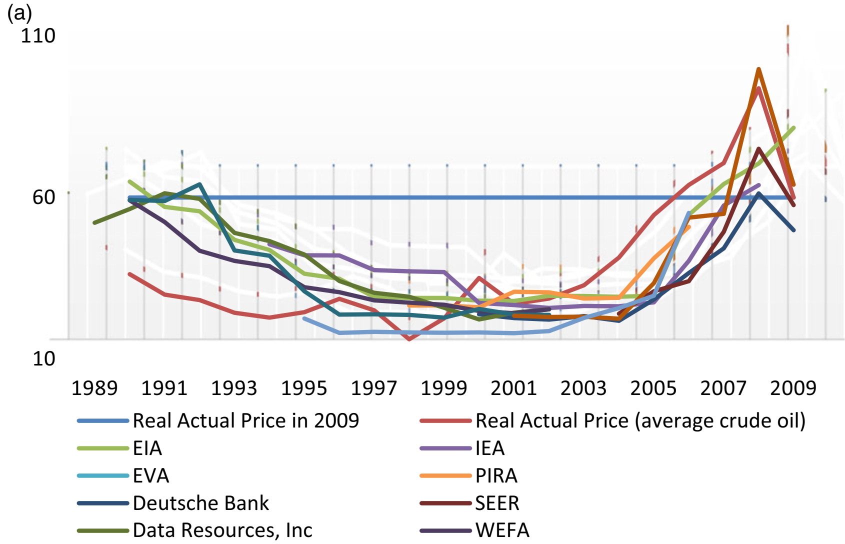
Source: Authors' estimations.
** and *** indicate statistical significance at the 5% and 1% levels, respectively. The numbers in parentheses are standard deviations.

2014. When forecasters approach the target years, their biases begin to decline if the target years were 2009 and 2014.

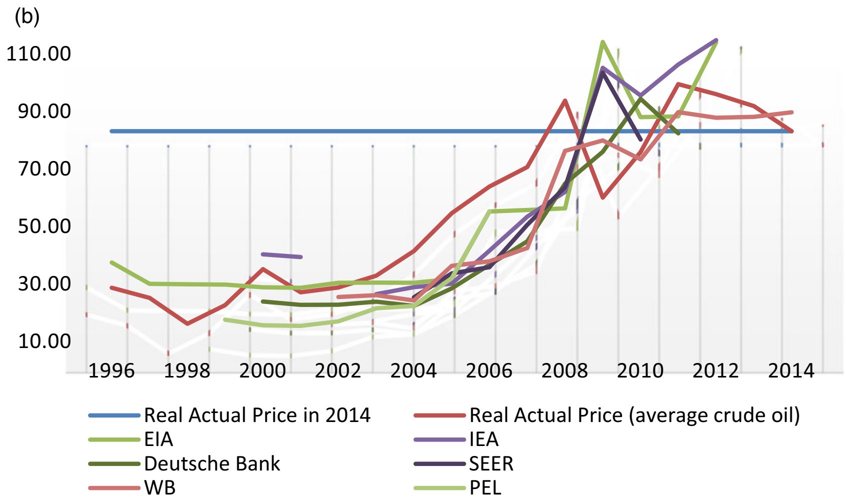
Given this observation and the oil price volatility, it seems too demanding to make a 10-year forecast (or more) and to hit the price in the target year exactly. Therefore, we assess the degree of forecast bias using a simple 3-year moving average (centred for 2005 and 2010, but not for 2015) of actual oil prices. Doing so, the forecast biases are reduced considerably for about 20 per cent of all target years, except 2005. The reason is that in each of the above-mentioned target years, the oil market has experienced a shock (either upward or downward) that forecasters could not predict. Indeed, as Wirl (2008) argues, erratic ups and downs in oil prices are a puzzling feature of oil markets. They seem to surprise not only consumers but also experts who forecast oil markets. After each positive (negative) shock, oil price sluggishly moves down (up) towards a new equilibrium price. By studying the history of oil markets, one can say that since 2000 the oil market has experienced a kind of shock that forecasters who sluggishly adjust their forecasts could not (and did not) revise early and often enough.

3.4. Sources of inefficiency

Another class of theories investigates the sources of forecast inefficiency and suggest explanations for forecast rigidity (see Dovern *et al.*, 2012 for details). Zellner (1986) points out that inefficient forecasts may be optimal from the forecaster's perspective (e.g. to get a project going) and thus biased forecasts are not necessarily evidence of irrational behaviour. For example, Nordhaus (1987) argues that their forecasts might be



Oil Price forecasts' biases (Target year: 2010)- (Real US $/Barrel)



Oil Price forecasts' biases (Target year: 2015)- (Real US $/Barrel)

Figure 2 (a) Oil Price forecasts' biases (Target year: 2010)—(Real US $/Barrel). (b) Oil Price forecasts' biases (Target year: 2015)—(Real US $/Barrel).
Source a and b: Authors' calculations based on different sources: Reference year is 2010. [Colour figure can be viewed at wileyonlinelibrary.com]

efficient, but not necessarily what they publish. Tversky and Kahneman (1981) state that the tendency to smooth one's forecasts is rooted in the way people think about the future, and people overestimate the precision of their knowledge. Other explanations can be found in the literature, such as fixed costs of obtaining information (Mankiw and Reis, 2001), noisy signals about the true state of the economy (Sims, 2003) and incomplete understanding of the dynamic nature inherent in the variables (Batchelor and Dua, 1991). Woodford (2003) argues that expectation adjustments appear sluggishly due to the noise in information, which makes it hard to detect systematic changes. Forecast rigidity is observed in forecasts of macroeconomic variables, e.g. inflation (Coibion and Gorodnichenko, 2015) and GDP (Loungani *et al.*, 2013) and here found in the sluggish revisions of oil price forecasts. In other words, Nordhaus' criterion assesses implicitly the rigidity of information. Although this criterion is highly convincing, it may be too critical, because each forecaster faced with new information has to separate this signal into its noise and information content. Several scholars have attempted to develop models assessing the degree of information rigidity. For instance, in a recent study, Coibion and Gorodnichenko (2015) develop a new approach in order to assess how the forecasts are formed and changed. However, this approach cannot be applied to the fixed event forecasts analysed in this paper. Since there have been few empirical investigations that try to understand the source of inefficient oil price forecasts, we try this below.

3.5. Myopia

Some research has been carried out about the source of inefficient forecasts. We test in the following, whether the observed forecast inefficiency can be explained by the myopic behaviour of forecasters (who use current oil price levels) combined with Hotelling-type exponential growth. Hotelling's (1931) and its extensions have attracted much attention, but a significant body of empirical evidence suggests it poorly explains prices of exhaustible resources, in particular, for oil and natural gas (Alquist *et al.*, 2011; Tuo and Yanbing, 2011; Gabralla and Abraham, 2013).[8] Nevertheless, a forecaster may still adhere to Hotelling's (1931) rule (an alternative explanation is that to keep the growth in oil demand due to rising incomes in check, requires ever-growing real prices). In order to test this explanation, we use a modified version of Hotelling's rule as:

$$p_{it}^{T} = p_t e^{\gamma i(T-t)} \tag{3}$$

It turns out that the fitting of (3) is poor because it is too demanding (see the Appendix 1). Therefore, we allow for reactions to current oil prices that are less than 100% and include instead an elasticity how each forecaster reacts to the changes in oil prices. This leads to estimating:

$$\text{In}p_{it}^{T} = \beta_i \text{In}p_t + \gamma_i (T - t) + \varepsilon_{it} \tag{4}$$

and the results are presented in **Table 5**.

Therefore, all estimates of β are statistically significant yet whether they are statistically different from myopia representing one to one myopic adjustment to the current price is crucial. The growth rate must be positive ($\gamma > 0$) and if $\beta = 1$, forecasters are myopic and applies Hotteling type extrapolation. Reactions to the current oil price are strong, albeit most of them are below 1 and thus below full myopia. This degree of myopia seems to be declining for IEA, WEFA and GRI and increasing for DOE, PEL, WB and DB from 2005 over 2010 to 2015. Forecasts are revised in accordance with oil price changes but they show different levels of elasticity ranging from 0.54 for IEA 2015 to 1.06 for SEER 2015. Positive, partially large and significant growth rates dominate, but a few are even negative for 2015. An alternative explanation of the observed exponential growth patterns is that they result from assuming an ever-growing GDP and, consequently, a higher demand that must be met with a production profile of an exhaustible resource. Ignoring backstops, this requires ever-growing prices to avoid diverging demand and supply. Summarising, a very simple model captures the sources of inefficiency and bias for most of the forecasts.

3.6. Can combining improve forecasts?

Given the dissatisfaction with individual forecasts, the question is how to improve. One of the most popular techniques is to combine and encompass different available forecasts, and application of the *wisdom of the crowds*. This technique was introduced in Bates and Granger (1969). If alternative forecasts are available, Clements and Hendry (2008) suggest that combining them can be useful in order to obtain more accurate and robust forecasts (Timmermann and Granger, 2004). Combinations of forecasts may provide some insurance against possible individual model misspecifications and smooth structural changes (Baumeister and Kilian, 2013), and allow forecasters to hedge against model uncertainty (Elliott and Timmermann, 2008).[9] However, as Fair and Shiller (1988) highlight, there is no guarantee that combined forecasts perform better.[10] The problem is that forecasts should only be combined if they contain different information sets (Stekler, 2002).

Therefore, we investigate whether the combination of oil price forecast are more efficient than any forecast alone. From Fig. 1 it is apparent that, all model builders' use very similar information sets, and the forecasts do not contain much independent information as all forecasts are similar and revised smoothly. Thus, a combination of smooth forecasts would likewise be smooth and inefficient according to the weak efficiency criterion of Nordhaus. Additionally, there is a high correlation between the

Table 5 Estimation of Hotelling-type forecasts

Institutions	Forecasts for					
	2005		2010		2015	
	β_i	γ_i	β_i	γ_i	β_i	γ_i
IEA	0.86*** (0.03)	0.10*** (0.01)	0.74*** (0.03)	0.07*** (0.01)	0.54*** (0.05)	0.09*** (0.02)
DOE	0.90*** (0.02)	0.06*** (0.01)	0.90*** (0.02)	0.045*** (0.01)	0.95*** (0.03)	0.01 (0.01)
WEFA	0.84*** (0.05)	0.07*** (0.02)	0.77*** (0.05)	0.01*** (0.01)	0.62*** (0.01)	0.07*** (0.04)
GRI	0.84*** (0.06)	0.06*** (0.02)	0.68*** (0.06)	0.08*** (0.01)	0.71*** (0.16)	0.05 (0.03)
PEL	0.88*** (0.04)	0.02 (0.01)	0.92*** (0.04)	0 (0.01)	0.95*** (0.11)	−0.01 (0.02)
PIRA	0.98*** (0.05)	0.008 (0.02)	0.91*** (0.04)	0.03*** (0.01)	0.93*** (0.06)	0.02 (0.01)
SEER	—	—	0.94*** (0.05)	−0.01 (0.04)	1.06*** (0.04)	−0.04** (0.02)
WB	0.86*** (0.01)	0.05*** (0.01)	0.98*** (0.03)	−0.03 (0.02)	1*** (0.02)	−0.03*** (0.01)
DB	0.85*** (0.06)	0.06 (0.05)	0.88*** (0.02)	0.01 (0.01)	1.01*** (0.04)	−0.02 (0.01)

Source: Authors' estimations.

** and *** indicate statistical significance at the 5% and 1% levels, respectively. The numbers in parentheses are standard deviations. Small sample size did not allow us to run the regression for all of the target years and institutions.

different forecasts as summarised in **Table 6**. The main implications are: (i) According to Fig. 1, it is apparent that there is a high correlation between oil price forecasts and the current oil prices. Therefore, it is not surprising that forecasts are also highly correlated. The maximum number of correlations can be observed for 2010, after oil prices

Table 6 Correlations among forecasts for several target years

2005	DOE/EIA	IEA	PIRA	WEFA	GRI	WB	PEL
DOE/EIA	1						
IEA	0.683*	1					
PIRA	−0.283	−0.755*	1				
WEFA	0.897**	0.916**	−0.534	1			
GRI	0.909**	0.600	−0.250	0.744*	1		
WB	0.110	0.163	−0.526	−1.000**	1.000**	1	
PEL	0.226	−0.025	−0.034	0.685	0.489	0.778	1

2010	DOE/ EIA	IEA	PIRA	DB	DRI	WEFA	GRI	WB	PEL
DOE/ EIA	1								
IEA	0.884**	1							
PIRA	0.851**	0.371	1						
DB	0.936**	0.970**	0.952**	1					
DRI	0.954**	0.860**	−0.115	−0.999*	1				
WEFA	0.983**	0.867**	−0.584	−0.180	0.926**	1			
GRI	0.968**	0.575	−0.251	0.974	0.952**	0.925**	1		
WB	0.870**	0.928**	0.978**	0.982**	−1.000**	−1.000**	1.000**	1	
PEL	0.888**	0.209	0.928**	0.947**	0.802*	0.666	0.951**	0.970**	1

2015	DOE/EIA	IEA	INFORUM	IHSGI	EVA	DB	SEER	WB
DOE/EIA	1							
IEA	0.966**	1						
INFORUM	−0.780	−0.996	1					
IHSGI	0.931	0.698	−0.974	1				
EVA	0.892*	0.977**	−0.271	−0.433	1			
DB	0.913**	0.949**	1.000**	−0.942	0.964**	1		
SEER	0.963**	0.978**	−1.000**	0.979	0.947*	0.897**	1	
WB	0.894**	0.950**	−0.891	0.363	0.976**	0.947**	0.940**	1

**Correlation is significant at the 0.01 level. *Correlation is significant at the 0.05 level. a. cannot be computed because at least one of the variables is constant. Positive significant correlations are in green.

experienced a relatively stable increasing trend for about two decades; (ii) There are some organisations which are not (or are rarely) significantly correlated with the others, namely IHSGI, PIRA, INFORUM and SEER. However, none of them provide efficient and accurate oil price forecast; (iii) Most of the forecasts are correlated with IEA and DOE/EIA which seem to serve as a reference for other forecasting entities. Furthermore, their forecasts are highly correlated with one another. Given these significant correlations, a combination of oil price forecasts can neither lead to more efficient, nor to more accurate forecasts. However, it appears that DOE/EIA affect or encompass the other forecasts; (iv) However, if we move from 2015 onward (2020, 2025, 2030 and 2035), the cross correlations decrease significantly (see Table 8 in Appendix 2).

4. Conclusion

Oil price forecasts are very important for many economic decisions at the small scale (individual, e.g. should I invest in a hybrid vehicle, or in a passive-heated home) and at the large scale (investments of firms, e.g. should US companies continue investing in shale oil and gas, or should Russia invest in oil exploration in the Arctic, or even globally in the context of global warming). Therefore, the efficiency of these highly relevant forecasts is crucial in order to reduce the range of uncertainty (Akmal *et al.*, 2012). This paper investigates whether oil price forecasts, made by different researchers and institutions, are (informationally) efficient. There are two types of forecast efficiency criteria: strong and weak efficiency. As it is impossible to analyse all available information and study whether forecasters implement all existing information, we refrain from studying the strong efficiency of forecasts. Instead two tests for assessing the weak efficiency are implemented.

The first test is that of weak efficiency which requires that “forecast revisions are not serially correlated”. Only one institution fails this test for 2005 and 2015, but many for 2010. To obtain more robust results, we proceed with second test of forecast efficiency, where we find that forecasts for 2015 are mainly inefficient. This means that examined forecasts did not fully incorporate all available information related to the oil price in 2010 and 2015. Complementing, we assess the bias of forecasts. Combining both criteria, efficiency and unbiasedness, only three pass for 2015: IEA, DOE, and WB. Having concluded this from the auto-regressive behaviour of revisions in the third step, this study has raised and attempted to answer important questions about the nature of inefficiency.

Forecast revisions are almost predictable through the examination of previous forecasts. Hence, two hypotheses (based on literature) are raised to explain the reason behind inefficiency of oil price forecasts: (i) the current oil price (myopic expectations) combined with (ii) Hotelling's rule. Several regressions (for each institute and for three target years) were run and estimation results support both hypotheses.

Finally, the high correlation between forecast institutions implies that the combination of oil price forecast cannot lead to more efficient forecasts. The auto-correlation of forecasts has been documented in other contexts as well, such as for inflation forecasts in the United States (e.g. in Coibion and Gorodnichenko, 2015). This apparent failure to react to most recent information is probably due to the difficulty in disentangling a recent signal into its fundamental term and its noise. Using some filtering methods to update the true state of fundamental variables implies some kind of exponential smoothing (even if done by Kalman filtering), which in turn makes the forecast revisions to appear sluggish. This in turn can also explain the correlation across forecasters if all get the same or similar signals. Another potential explanation is that while they could improve their forecasts, they prefer to coordinate their forecasts with other agencies in order to be not be an outlier.

Notes

1. The authors have mentioned that forecasters examine the forecast efficiency of long-term energy price forecasts as they deal with quarterly data. When compared to daily or monthly forecasts of energy prices, this is equivalent to 2, 3 or 4 quarters ahead.
2. Similar results were obtained for macroeconomic and energy-consumption forecasts.
3. http://www.economist.com/node/21685480, accessed at March 2016.
4. Fama *et al.* (1969) introduce another concept as semi-strong form efficiency. It occurs when the information set includes all publicly available information.
5. Although much of the available literature on forecast efficiency deals with fixed event forecasts, several researchers have attempted to develop methods for testing forecast efficiency in rolling- event forecasts (Fair and Shiller, 1988, 1990).
6. Davies and Lahiri (1999) have developed an econometric methodology under which it is possible to simultaneously test the rationality of rolling event and fixed event forecasts.
7. Price projections by DRI, DB, GRI and WEFA are for composite refiner acquisition prices while EVA, IHSGI, PIRA and SEER forecast West Texas Intermediate crude oil.
8. For example, if discounted future oil price is higher than the spot price, it would be more profitable by just leaving oil in the ground and waiting to produce it until the price has risen. As a result, the net present values are calculated using the empirical distribution of oil prices under a range of different discount rates and production periods.
9. Combination will provide insurance against smooth structural changes.
10. They examine the informational content of three sets of ex ante forecasts: the American Statistical Association and National Bureau of Economic Research Survey (ASA), DRI and WEFA. The authors conclude that both DRI and WEFA seem to use very similar information sets and do not contain much independent information.

References

Ahlstrom, M.L., Bartlett, D., Collier, C., Duchesne, J., Edelson, D., Gesino, A. and de la Torre Rodríguez, M., 2013. Knowledge is power: efficiently integrating wind energy and wind forecasts. *Power and Energy Magazine, IEEE* 11, 6, 45–52.

Akmal, S., Riaz, M. and Ameer, M.W., 2012. Rationality analysis for ARIMA forecasts. *Research Journal of Finance and Accounting* 3, 5, 168–177.

Allan, G. (2012). Evaluating the usefulness of forecasts of relative growth. Discussion Papers in Economics, No. 12-14.

Alquist, R., Kilian, L. and Vigfusson, R. 2011. Forecasting the price of oil. Forthcoming: G. Elliott and A. Timmermann, eds., *Handbook of Economic Forecasting 2*. Amsterdam: North-Holland.

Artis, M. and Marcellino, M., 2001. Fiscal forecasting: The track record of the IMF, OECD and EC. *The Econometrics Journal* 4, 1, 20–36.

Batchelor, R. and Dua, P. 1991. Blue Chip rationality tests. *Journal of Money, Credit and Banking* 23, 4, 692–705.

Bates, J.M. and Granger, C.W.J., 1969. The combination of forecasts. *Operations Research Quarterly*, 20, 451–468.

Baumeister, C. and Kilian, L., 2013. Forecasting the real price of oil in a changing world: A forecast combination approach (No. 2013/11). CFS Working Paper.

Berger, A.N. and Krane, S.D. 1985. The informational efficiency of econometric model forecasts. *The Review of Economics and Statistics* 67, 128–134.

Bernanke, B.S., 2007. Inflation expectations and inflation forecasting. In *Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Mass.*, July (Vol. 10).

Clemen, R.T. and Winkler, R.L., 1986. Combining economic forecasts. *Journal of Business & Economic Statistics* 4, 1, 39–46.

Clement, M.B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27, 3, 285–303.

Clements, M.P., 1997. Evaluating the rationality of fixed-event forecasts. *Journal of Forecasting* 16, 4, 225–239.

Clements, M.P. and Hendry, D.F. (Eds.). 2008. *A companion to economic forecasting*. John Wiley & Sons, USA, Chapter 10.

Coibion, O. and Gorodnichenko, Y., 2010. Information rigidity and the expectations formation process: a simple framework and new facts. *National Bureau of Economic Research*, No. w16537.

Coibion, O. and Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: a simple framework and new facts. *American Economic Review* 105, 8, 2644–2678.

Cooper, J.P. and Nelson, C.R., 1975. The ex-ante prediction performance of the St. Louis and FRB-MIT-PENN econometric models and some results on composite predictors. *Journal of Money, Credit and Banking* 7, 1, 1–32.

Croson, R. and Sundali, J., 2005. The gambler's fallacy and the hot hand: empirical data from casinos. *Journal of Risk and Uncertainty* 30, 3, 195–209.

Davies, A. and Lahiri, K., 1999. Re-examining the rational expectations hypothesis using panel data on multi-period forecasts. *Analysis of Panels and Limited Dependent Variable Models* Cambridge Univ. Press, ISBN 0521631696, 226–254.

Dehnavi, J., Wirl, F. and Yegorov, Y., 2015. Arbitrage in natural gas markets? *International Journal of Energy and Statistics* 3, 04, 1550018.

Diebold, F.X. and Lopez, J.A., 1996. Forecast evaluation and combination. In Maddala, G.S. and Rao, C.R. (eds.) *Handbook of Statistics*, North-Holland, Amsterdam, pp. 241–268.

Dovern, J. and Weisser, J., 2008."Are they really rational?: Assessing professional macro-economic forecasts from the G7-countries. Kiel working paper, No. 1447.

Dovern, J., Ulrich, F., Prakash, L. and Natalia, T., 2012. Information rigidity and herding in economic growth forecasts: stylized facts from a large international panel. IMF Working paper.

Dovern, J., Fritsche, M.U., Loungani, M.P. and Tamirisa, M.N.T. 2013. Information rigidities in economic growth forecasts: evidence from a large international panel. *International Monetary Fund*, No. 13-56.

Elliott, G. and Timmermann, A., 2008. Economic forecasting. *Journal of Economic Literature* 46, 1, 3–56.

Fair, R.C. and Shiller, R.J., 1988. The informational content of ex ante forecasts. *National Bureau of Economic Research*, No. w2503, 325–331.

Fair, R.C. and Shiller, R.J., 1990. Comparing information in forecasts from econometric models. *The American Economic Review* 80, 3, 375–389.

Fama, E.F., Fisher, L., Jensen, M.C. and Roll, R., 1969. The adjustment of stock prices to new information. *International Economic Review* 10, 1, 1–21.

Floris, F.J.T., Bush, M.D., Cuypers, M., Roggero, F. and Syversveen, A.R., 2001. Methods for quantifying the uncertainty of production forecasts: a comparative study. *Petroleum Geoscience*, 7(S), S87–S96.

Gabralla, L.A. and Abraham, A., 2013. Computational modeling of crude oil price forecasting: a review of two decades of research. *International Journal of Computer Information Systems and Industrial Management Applications* 5, 729–740.

Genberg, H. and Martinez, A., 2014. *"On the Accuracy and Efficiency of IMF Forecasts: A Survey and Some Extensions". IEO Background Paper, No. BP/14/04*. Independent Evaluation Office of the IMF, Washington.

Hotelling, H., 1931. The economics of exhaustible resources. *The Journal of Political Economy* 39, 2, 137–175.

Kaboudan, M.A., 2001. Compumetric forecasting of crude oil prices. In Evolutionary Computation, Proceedings of the 2001 Congress on (Vol. 1), 283-287, IEEE.

Kaufmann, R.K., 1995. A model of the world oil market for project LINK Integrating economics, geology and politics. *Economic Modelling* 12, 2, 165–178.

Khan, U., Rozenbaum, O. and Sadka, G., 2013. "Bias and Efficiency: A Comparison of Analyst Forecasts and Management Forecasts." Available at SSRN 2177992.

Liu, F., Smallwood, G.J. and Gulder, O.L., 2002. An accurate efficient and flexible SNBCK-based unified band model for calculations of spectrally resolved and integrated quantities in participating media containing real-gases. *Heat Transfer* 1, 663–668.

Loungani, P., Stekler, H. and Tamirisa, N., 2013. Information rigidity in growth forecasts: some cross-country evidence. *International Journal of Forecasting* 29, 4, 605–621.

Lynch, M.C., 2002. Forecasting oil supply: theory and practice. *The Quarterly Review of Economics and Finance* 42, 2, 373–389.

Mamatzakis, E. and Koutsomanoli-Filippaki, A., 2014. Testing the rationality of DOE's energy price forecasts under asymmetric loss preferences. *Energy Policy* 68, 567–575.

Manescu, C. and Van Robays, I., 2014. Forecasting the Brent oil price: addressing time-variation in forecast performance. ECB Working Paper, No. 1735.

Mankiw, N.G. and Reis, R., 2001. *Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve*. National Bureau of Economic Research, No, p. w8290.

Mincer, J.A. and Zarnowitz, V., 1969. The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, 3-46, NBER.

Nelson, C.R., 1984. A benchmark for the accuracy of econometric forecasts of GNP. *Business Economics* 19, 3, 52–58.

Nordhaus, W.D., 1987. Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics* 69, 4, 667–674.

Nyberg, H., 2011. Forecasting the direction of the US stock market with dynamic binary probit models. *International Journal of Forecasting* 27, 2, 561–578.

Reis, R., 2006. Inattentive producers. *The Review of Economic Studies* 73, 3, 793–821.

Sanders, D.R., Manfredo, M.R. and Boris, K., 2008. Accuracy and efficiency in the US Department of energy's short-term supply forecasts. *Energy Economics* 30, 3, 1192–1207.

Sanders, D.R., Manfredo, M.R. and Boris, K., 2009. Evaluating information in multiple horizon forecasts: The DOE's energy price forecasts. *Energy Economics* 31, 2, 189–196.

Shlyakhter, A.I., Kammen, D.M., Broido, C.L. and Wilson, R., 1994. Quantifying the credibility of energy projections from trends in past data: The US energy sector. *Energy Policy* 22, 2, 119–130.

Sillanpää, V. and Heino, O., 2013. Forecasting football match results-A study on modeling principles and efficiency of fixed-odds betting markets in football, Master's thesis, Aalto University, School of Business, Finland.

Silver, N., 2012. *The Signal and the Noise: The art and science of prediction*. Penguin, UK.

Sims, C.A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50, 3, 665–690.

Stekler, H.O., 2002. The rationality and efficiency of individuals' forecasts. A companion to economic forecasting, 222–240.

Tayman, J., Smith, S.K. and Rayer, S., 2011. Evaluating population forecast accuracy: a regression approach using county data. *Population Research and Policy Review* 30, 2, 235–262.

Timmermann, A. and Granger, C.W., 2004. Efficient market hypothesis and forecasting. *International Journal of Forecasting* 20, 1, 15–27.

Tuo, J. and Yanbing, S., 2011, October. "Summary of World Oil Price Forecasting Model." In *Knowledge Acquisition and Modeling (KAM)*, 2011 Fourth International Symposium, 327-330. IEEE.

Tversky, A. and Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 4481, 453–458.

Von Bailey, D. and Brorsen, B.W., 1998. Trends in the accuracy of USDA production forecasts for beef and pork. *Journal of Agricultural and Resource Economics* 23, 2, 515–525.

Wirl, F., 2008. Why do oil prices jump (or fall)? *Energy Policy* 36, 3, 1029–1043.

Woodford, M., 2003. Imperfect Common Knowledge and the Effects of Monetary Policy. In Ph. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (eds) Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps."

Yu, L., Wang, S. and Lai, K.K., 2008. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics* 30, 5, 2623–2635.

Zellner, A., 1986. Biased predictors, rationality and the evaluation of forecasts. *Economics Letters* 21, 1, 45–48.

Appendix 1

Extreme myopia

The full myopia was examined for each institution and for three target years by running an Ordinary Least Squares (OLS) regression:

Table 7 Estimation of Hotelling's pricing rule; H_0: $\gamma > 0$ against H_1: $\gamma < 0$ and $\beta = 1$.

	Forecasts for		
	2005	2010	2015
Institution	γ	γ	γ
IEA	0.05*** (0.01)	0.02* (0.01)	−0.006 (0.009)
DOE	0.03*** (0.006)	0.02*** (0.005)	0.003 (0.006)
WEFA	0.01* (0.009)	0.01*** (0.004)	0.006 (0.005)
GRI	0.009 (0.01)	0.01*** (0.005)	−0.002 (0.005)
PEL	−0.01 (0.01)	−0.02*** (0.005)	−0.02** (0.007)
PIRA	0 (0.01)	0.008 (0.006)	0.009** (0.003)
SEER	—	−0.05** (0.01)	−0.016* (0.008)
WB	−0.09 (0.04)	−0.04*** (0.01)	−0.02*** (0.007)
DB	−0.04 (0.02)	−0.03** (0.01)	−0.017** (0.006)

Source: Authors' estimations.
** and *** indicate statistical significance at the 5% and 1% levels, respectively. The numbers in parentheses are standard deviations.

Estimates of equation 5 produced insignificant coefficients and a weak fitting (see Table 7) as in this equation the coefficient of current oil prices is assumed to be equal to 1 according to the Hotelling's rule.

$$\mathrm{In}p_{it}^{T} = \beta \mathrm{In}p_t + \gamma_i \cdot (T - t) + \varepsilon_i t \tag{5}$$

Appendix 2

Correlations among forecasts from 2015 onward.

Table 8 Correlations among forecasts from 2015 onward

	DOE/ EIA	IEA	INFORUM	IHSGI	EVA	WB	DB	SEER
2020								
DOE/EIA	1							
IEA	0.977**	1						
INFORUM	−0.215	−0.639	1					
IHSGI	−0.813	−0.606	−0.141	1				
EVA	0.965**	0.964**	−0.715	0.280	1			
WB	0.085	0.037	−0.809	−0.407	0.735	1		
DB	0.878**	0.927**	1.000**	0.593	0.822	−0.840	1	
SEER	0.945**	0.967**	−1.000**	−0.925	0.933*	0.583	0.822*	1
2025								
DOE/EIA	1							
IEA	0.980**	1						
INFORUM	0.506	−0.522	1					
IHSGI	0.278	0.638	0.384	1				
EVA	0.872**	0.884**	−0.466	−0.608	1			
WB	−0.970	−1.000**	−0.650	−1.000**	−0.429	1		
DB	0.843**	0.910**	1.000**	0.300	0.788	.[a]	1	
SEER	0.880**	0.928**	0.866	0.876	0.817*	−1.000**	0.815*	1
2030								
DOE/EIA	1							
IEA	0.409	1						
INFORUM	−0.946	−0.416	1					
IHSGI	−0.086	0.960*	−0.609	1				
EVA	−0.198	−0.795	0.999*	−0.804	1			
DB	0.028	0.808	−1.000**	0.708	−0.970	1		
SEER	0.735	0.775	−1.000**	0.855	−0.475	0.661	1	
2035								
DOE/EIA	1	−0.221	0.043	−0.312	−0.597			
IEA	−0.221	1	−0.899	0.949	1.000**			
INFORUM	0.043	−0.899	1	−0.944	−0.812			
IHSGI	−0.312	0.949	−0.944	1	1.000**			
EVA	−0.597	1.000**	−0.812	1.000**	1			

**Correlation is significant at the 0.01 level, *Correlation is significant at the 0.05 level.