Accuracy of the real crude oil price forecast for different specification of VAR models

Sławomir Śmiech[1]

Abstract

The aim of the paper is to examine the forecasting ability of the real price of crude oil Brent. In order to forecast the crude oil price a large set of predictors including a variable describing a real and financial process of economy; prices of other commodity and supply of crude oil are collected. The analysis follows a recursive scheme and forecasts are generated for the period between January 2005 and October 2014. In the study all possible combinations of predictors are used for different specification of four-dimensional VAR models. Forecast accuracy of VAR models is compared with the naïve forecast. The results obtained indicate that, at short horizons, certain models generate more accurate forecasts than the benchmark models. The comparison of various specifications of the VAR models reveals that the most accurate forecasts are generated by the VAR(2) models.

Keywords: *forecasting, crude oil prices, real economy, financial market*

JEL Classification: C53, Q47

1. Introduction

The oil price is focusing the attention of not only drivers worldwide but also of economists and politicians. The oil price directly affects the income of both its exporters and importers, indirect impacts the inflation levels (Rothemberg and Woodford, 1996), investments (Elder and Serletis, 2009). It is also source of information for financial investors as determine the value of futures contracts (Baumeister et al., 2013).

Many efforts have been made in order to propose forecasting models and variables, with which it is possible to improve the accuracy of forecasts of the price of oil. Some paper refer to the predictive accuracy of oil futures prices (e.g. Knetsch, 2007; Alquist and Kilian, 2010; Reeve and Vigfusson, 2011; Alquist et al., 2013). Others investigate the forecasting efficiency of professional and survey forecasts (see Sanders et al., 2009; Alquist et al., 2013; Bernard et al., 2013). The last strand considers the forecasting power of variables related to supply and demand in the global oil market, it is: crude oil inventories, production as well as macroeconomic fundamentals (Kilian and Vega, 2011; Kilian and Hicks, 2013).

[1] Corresponding author: Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland, e-mail: smiechs@uek.krakow.pl.

The main aim of our analysis is to evaluate forecast accuracy of crude oil prices by four dimensional VAR models, with different specification and different set of predictors. To do so, we collected large data set (including 23 time series variables) and examined VAR models for all possible combination of predictors. The specifications of the VAR models we take into consideration differ with respect to the number of lags and the occurrence of deterministic components. As a result of the analysis it was possible to point, which combinations of variables are most effective in terms of the oil price forecasting as well as which specifications of models provide the lowest forecast errors.

The real price of crude oil Brent is chosen for forecasting. Forecasting models are developed on the basis of monthly data from the period between January 1995 and October 2014, and they generate one-month ahead forecasts. Our models follow a recursive scheme, presented in the studies by Inoue and Kilian (2006), Alquist et al. (2013) and Baumeister and Kilian (2014). The data used to build the models cover observations from the period between January 1995 and September 2014, while forecasts cover the period between January 2005 and October 2014. This way of forecasting yields a time series consisting of 118 elements $\{\hat{y}_t\}_{t=1}^{118}$, which is then further evaluated.

The evaluation of forecast accuracy is performed by comparison forecast errors of the best our (i.e. for models generating the lowest prediction error) models and the benchmark models (the random walk without drift). It is obtained by using Diebold and Mariano test (1995). To compare all forecasting models analysed in the study we also use forecast accuracy measures: the root mean square error (RMSE) and the mean absolute percentage error (MAPE).

2. Data

2.1 Dependent variables

The Brent spot price of crude oil is chosen for the verification of the possibility of forecasting prices of fossil fuels. This price, together with WTI, is considered the world benchmark (see e.g. Baumeister and Kilian, 2014). The International Monetary Fund serves as a source of data. Crude oil spot price is expressed as real, in constant prices in 2010. The consumer price index in the USA CPI_{US} is used as the GDP deflator. The analysis is based on monthly data from the period between January 1995 and October 2014. This means that the whole sample period contains 238 monthly observations.

2.2 Predictors used in forecasting real crude oil prices

According to reach literature in the study we use 23 predictors for crude oil price[2]. The whole variable set is divided into three following subsets: macroeconomic variables, financial variables, energy prices.

Variables describing real economy include: the global industrial production index (*IP_W*) and in the euro area (*IP_EA*) (e.g. Akram, 2009) and variables referring to the economic activity (see Kilian and Hicks, 2013): the ISM manufacturing index in the U.S. (*ISM_US*), the PMI manufacturing index in the euro area (*PMI_EA*) (*Purchasing Managers Index - Markit Eurozone Manufacturing* PMI), and the German Ifo index (*IFO*) for the business climate among entrepreneurs in trade and industry published by the Ifo Institute; the Baltic Dry Index (*BDI*) (see Baumeister and Kilian, 2012) and the global real economic activity index (*IK*) proposed by Kilian (2009).

The remaining variables refer to the global oil market: the global crude oil production (*PR_OIL*) (see e.g. Kilian, 2009 and Kilian and Murphy, 2014) and the world crude oil inventories (*INV*)[3] (Kilian and Murphy, 2014).

Variables in the second set include: the real short-term interest rates in the U.S. (*IR_US*) and in the euro area (*IR_EA*); the real money supply M1 in the U.S. (*M1_US*) and in the euro area (*M1_EA*) (see, e.g. Anzuini et al., 2013; Gargano and Timmermann, 2014); the real effective exchange rates deflated by the consumer price index (narrow index) (2010 = 100) (*RN_US*) published by the Bank of International Settlements; the US dollar-euro exchange rate (*REX*) (see Chen et al., 2010); the Standard and Poor's 500 (SP500) stock price index, the German stock index – DAX (DAX) (Schalck and Chenavaz, 2015); the Chicago Board Options Exchange Market Volatility Index (VIX) (Issler et al., 2014).

The third group of variables contains the following energy prices: WTI crude oil spot price (*WTI*), steam coal price in Australian ports (*NEWC*), steam coal price in Richards Bay port (the Republic of South Africa) (*RB*), Russian natural gas border price in Germany (*NG_RUS*) and natural gas spot price in the U.S. (*NG_US*). The data are taken from the International Monetary Fund (IMF). All prices are expressed as real, in constant prices in 2010. The consumer price index in the USA CPI_{US} is used as a deflator.

---

[2] The same data sets is used in Papież (2015).

[3] Given the lack of data on crude oil inventories for countries other than the U.S., its world inventories will be approximated following the suggestions from Hamilton's (2009) paper. Crude oil inventories in the U.S. will be scaled by the ratio of OECD crude oil stocks over the U.S. on the basis of data provided by the EIA. (http://www.eia.gov). In the period we analyse this ratio range between 2,31 and 2,61.

3. Empirical results

The proposed approach is based on statistical evaluation of forecast error distributions of VAR models obtained for all possible combinations of variables that affect the price of crude oil. We take into account the four-dimensional VAR models assuming different number of lags and possible deterministic components. Forecasts are determined using a recursive prediction schemes. Part of the analysis is dedicated to identify optimal, from the forecasting point of view, specification of VAR models. What's more, as the result it is possible to point out the combination of variables, for which given models generate minimal forecasts errors.

3.1 Recursive scheme of forecasting

The recursive scheme applied in our study indicates that the number of observations used to evaluate the parameters of the model increases by one observation for each consecutive model estimated. We assume that, during the first step, the estimation period will cover first 120 observations ($y_1,\ldots,y_{120}$), that is between January 1995 and December 2004, and the forecast is generated for one period ahead $(T+1)$, that is $\hat{y}_{121}$ (January 2005). During the second step, the parameters are estimated on the basis of first 121 observations ($y_1,\ldots,y_{121}$), and the forecast is generated for the period $\hat{y}_{122}$, etc. During the last step the parameters of the model are estimated on the basis of observations $y_1,\ldots,y_{237}$ (that is between January 1995 and September 2014), and the forecast is generated for the $\hat{y}_{238}$ period (October 2014).

The application of this forecasting scheme makes it possible to develop a time series $\{\hat{y}_t\}_{t=120}^{238}$, with forecast values generated by consecutively estimated models. The values of this series will be further used to evaluate forecast accuracy. This means that the out-of-sample period used to evaluate forecast performance contains 118 observations (from January 2005 till October 2014).

3.2 The evaluation of forecast generated by four-dimensional VAR models with all possible combinations of predictors

This part presents the evaluation of forecast accuracy generated by four-dimensional VAR models estimated for variables $[z_t, x_{1,t}, x_{2,t}, x_{3,t}]'$, where z_t is either real Brent crude oil price, (or the first difference of crude oil price), $[x_{1,t}, x_{2,t}, x_{3,t}]$ is a set of variables selected from all variables described in section 2. The number of the VAR models is large enough to contain all possible subsets of variables. The VAR models taken into consideration include models with a constant and with trend and without trend. The number of lags is 1,2,3, and 12 (as Hamilton and Herrera, 2004 show the importance of allowing for long lags in the crude oil price models). Additionally, the VAR models with one lag for the first difference of crude oil prices are evaluated. Taking into consideration the number of variables in the set, each VAR specification (after establishing the lag order and deterministic components) require estimating 1540 models, covering all possible combinations of 22 variables, which yields 13860 forecasting models. Due to the number of models considered in the paper, the assessment of their in sample properties is omitted. In all these models the vector of forecast is computed following the recursive scheme, which is later assessed with the use of forecast accuracy measures (the RMSE, the MAPE, the R^2_{OS} [4]) and the Diebold-Mariano statistics[5].

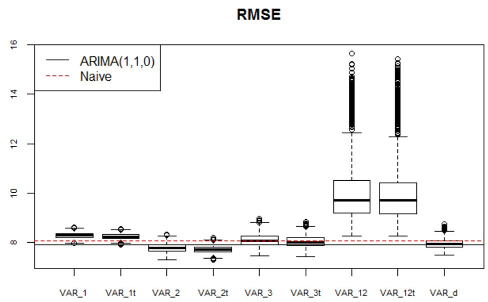

Fig. 1. Distribution of the root mean squared error (RMSE) for all VAR specifications considered.

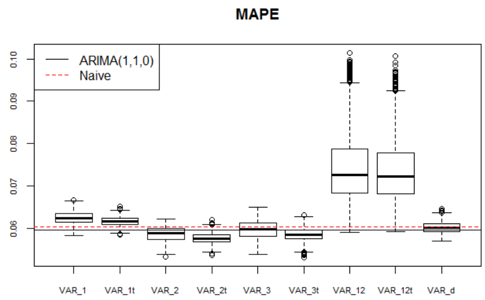

Fig. 2. Distribution of the mean absolute percentage error (MAPE) for all VAR specifications considered.

Fig. 1 presents the distribution of RMSE for all specifications of the VAR models considered in the study. Symbol 't' added to the name of the model indicates that a deterministic trend is used in this model (e.g. VAR(p)_t). The models for first differences are marked as VAR(1)_d. The smallest value of RMSE is obtained for the VAR with two lags with or without trend. For most models of this class errors are smaller than in case of the naïve model and the ARIMA(1,1,0) model, which serves as the second benchmark model in section 4.2 (8.040 and 7.905, respectively). The VAR(2) model without trend with the smallest RMSE (7.289) includes variables *IP_W*, *VIX*, *RN_US*. The best VAR(2) model with

[4] A positive (negative) value of indicates by how many percent forecast errors of the model are more (less) accurate than forecasts generated by the benchmark model Campbell and Thompson (2008).

[5] In this part we compare forecasts generated in our study with the naïve forecasts Diebold and Mariano test (1995).

trend includes the same set of variables, and its RMSE is 7.278. The RMSE distribution for VAR(3), VAR(3) with trend and VAR(1) for first differences are shifted up and, in general, have higher values. However, among models in these specifications it is possible to find such for which the RMSE is relatively small. For example, the smallest values of RMSE for VAR(3) and VAR(3) with trend are less than 7.420 and for VAR(1)_d it is 7.488. Most VAR(1) models with trend and without trend have the RMSE larger than the benchmark models. The largest errors measured by the RMSE are obtained using VAR(12) models with or without trend. There is no combination of variables for VAR(12) which yields a model with errors smaller than in the benchmark models.

The distributions of the MAPE for all specifications are presented in Fig. 3. Similarly to the RMSE, the smallest MAPE errors are obtained for VAR(2) with trend and without trend. Most combinations of variables generate more accurate forecasts than two benchmark models. The lowest mean absolute percentage error for VAR(2) without trend – 5.332% – is obtained through the combinations of variables (NEWC, IP_W, SP500), and for VAR(2) model with trend the MAPE equals 5.372% for variables (*IP_W*, *VIX*, *M1_EA*). Similar forecast accuracy measured by the MAPE is generated by most VAR(3) models with trend. The best model of this type is built for variables (*INV, VIX, RN_US*) and its MAPE equals 5.303%. In VAR(3) models without trend and VAR(1)_d model the MAPE lower than for the naïve forecast is obtained for about a half of variable combinations. The remaining specifications of VAR(1) and VAR(12) models usually generate forecasts of lower accuracy measured by the MAPE. In VAR(12) models, the range of MAPE is the largest (the best models in this category have MAPE lower than 0.06, the worst almost 0.10).

Table 1 presents ten VAR models with the smallest RMSE and MAPE. It also contains the class of the models and the variables used for their development. The most accurate forecasts for such a criterion are generated by the VAR(2) models with or without trend.

The majority of most accurate forecasting models contain variables describing both the global conjuncture (*IP_W*, *ISM_US*) and the financial variables describing the U.S. stock exchange (i.e. *SP500*, *VIX*) and the real effective exchange rates (*RN_US*). In two models the variable describing the global oil inventories (*INV*) is included.

A similar level of forecast errors is found in all models presented in Table 2. The RMSE error ranges between 7.278 and 7.429, while the MAPE for all models does not exceed 5.490%. Similar values are obtained for the DM statistics, with p-value for one and two-sided test, which leads to the conclusion that forecasts generated by given models are not more accurate than the naïve forecast.

	The name of the model	Variables in the VAR model
M1	VAR(2)	*BRENT, IP_W, INV, SP500*
M2	VAR(2)	*BRENT, IP_W, ISM_US, RN_US*
M3	VAR(2)	*BRENT, IP_W, SP500, DAX*
M4	VAR(2)	*BRENT, IP_W, SP500, VIX*
M5	VAR(2)	*BRENT, IP_W, SP500, RN_US*
M6	VAR(2)	*BRENT, IP_W, VIX, RN_US*
M7	VAR(2)	*BRENT, IFO, M1_EA, RN_US*
M8	VAR(2) trend	*BRENT, IP_W, ISM_US, RN_US*
M9	VAR(2) trend	*BRENT, IP.W, VIX, RN_US*
M10	VAR(2) trend	*BRENT, INV, VIX, RN_US*

Table 1. The variables and the specifications of the models with the smallest RMSE and MAPE.

Model	RMSE	MAPE[%]	R^2_{OS}	DM	p-value (one sided)
M1	7.426	5.464	0.147	1.141	0.127
M2	7.323	5.420	0.170	1.292	0.098
M3	7.394	5.452	0.154	1.061	0.144
M4	7.408	5.402	0.151	1.038	0.150
M5	7.336	5.490	0.167	1.273	0.101
M6	7.289	5.406	0.178	1.193	0.117
M7	7.429	5.487	0.146	1.028	0.152
M8	7.331	5.481	0.169	1.233	0.109
M9	7.278	5.405	0.181	1.212	0.113
M10	7.423	5.487	0.148	1.201	0.115
naïve	8.040	6.028	0.000		

Table 2. The evaluation of forecasts with the best combination of variables for the MAPE and the RMSE.

Conclusion

Comparing the results generated by the benchmark models and the VAR models developed for all combinations of predictors does not lead to conclusive implications. On the one hand, we have identified such combinations of variables (not used in previous studies) which yield forecasting models with smaller forecast errors measured by the MAPE and the RMSE. On the other hand, we cannot reject the hypothesis that forecasts generated by these models do not differ from the naïve forecast. Models that generate forecasts with the smallest errors contain variables describing economic activity (the global manufacturing, the conjuncture) and the climate on stock markets, (*SP500, DAX, VIX*) as well as the real effective exchange rates (*RN_US*). The sets of variables which yielded the smallest forecast errors should be taken into consideration when forecasting the real price of crude oil at short horizons.

The comparison of the VAR models with a different number of lags allows us to conclude that the best results, that is the most accurate forecasts, are generated by the VAR(2) models with and without trend. Less accurate results are obtained for the VAR(3) models with trend. The largest forecast errors, usually larger than the errors generated by the naïve forecast, are generated by VAR(12) models.

Acknowledgements

Supported by the grant No. 2012/07/B/HS4/00700 of the Polish National Science Centre.

References

Akram, Q. F. (2009). Commodity prices, interest rates and the dollar. *Energy Economics, 31*(6), 838-851.

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

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the Price of Oil. In: Elliott, G., Timmermann, A. (eds.), *Handbook of Economic Forecasting*, *2*, 427-507.

Anzuini, A., Lombardi, M. J., & Pagano, P. (2013). The Impact of Monetary Policy Shocks on Commodity Prices. *International Journal of Central Banking*, *9*(3), 125-150.

Baumeister, C., & Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business & Economic Statistics*, *30*(2), 326-336.

Baumeister, C., Kilian, L., & Zhou, X. (2013). Are product spreads useful for forecasting? An empirical evaluation of the Verleger hypothesis, No. 2013-25. Bank of Canada Working Paper.

Bernard, J. T., Khalaf, L., Kichian, M., & Yelou, C. (2013). *On the Long-Term Dynamics of Oil Prices: Learning from Combination Forecasts.* mimeo, Carleton University.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, *21*(4), 1509-1531.

Chen, Y. C., Rogoff, K. S., & Rossi, B. (2010). Can exchange rates forecast commodity prices? *Quarterly Journal of Economics*, *125*(3), 1145-1194.

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

Elder, J., & Serletis, A. (2010). Oil price uncertainty. *Journal of Money, Credit and Banking*, *42*(6), 1137-1159.

Gargano, A., & Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting*, *30*(3), 825-843.

Hamilton, J. D. (2009). *Causes and Consequences of the Oil Shock of 2007-08*, No. w15002. National Bureau of Economic Research.

Hamilton, J. D., & Herrera, A. M. (2004). Comment: oil shocks and aggregate macroeconomic behavior: the role of monetary policy. *Journal of Money, Credit and Banking*, 265-286.

Inoue, A., & Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, *23*(4), 371-402.

Issler, J. V., Rodrigues, C., & Burjack, R. (2014). Using common features to understand the behavior of metal-commodity prices and forecast them at different horizons. *Journal of International Money and Finance*, *42*, 310-335.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *The American Economic Review*, 1053-1069.

Kilian, L., & Hicks, B. (2013). Did unexpectedly strong economic growth cause the oil price shock of 2003–2008? *Journal of Forecasting*, *32*(5), 385-394.

Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, *29*(3), 454-478.

Kilian, L., & Vega, C. (2011). Do energy prices respond to US macroeconomic news? A test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics*, *93*(2), 660-671.

Knetsch, T. A. (2007). Forecasting the price of crude oil via convenience yield predictions. *Journal of Forecasting*, *26*(7), 527-549.

Papież, M. (2015). Is it possible to successfully forecast the real price of crude oil at short horizons? In: Papież, M., Śmiech, S. (eds.), *The 9th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*. *Conference Proceedings.* Cracow: Foundation of the Cracow University of Economics, 155-164.

Reeve, T. A., & Vigfusson, R. (2011). Evaluating the Forecasting Performance of Commodity Futures Prices. *FRB International Finance Discussion Paper*, (1025). Board of Governors of the Federal Reserve System.

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

Schalck, C., & Chenavaz, R. (2015). Oil Commodity Returns and Macroeconomic Factors: A Time-Varying Approach. *Research in International Business and Finance*, *33*, 290-303.