A COMPARISION OF VARIOUS OIL PRICE FORECASTING METHODS WITH A LARGE NUMBER OF VARIABLES

Krzysztof Drachal

Faculty of Economic Sciences, Univeristy of Warsaw (Poland)

kdrachal@wne.uw.edu.pl

Research funded by the Polish National Science Centre grant under the contract number DEC-2015/19/N/HS4/00205.

Oil spot price forecasting

- Hard task
- It’s hard to produce more accurate forecast than ARIMA or the naive method
- It is interesting to compare forecasts from several models, for the same sample
- Narrow the considerations to the methods dealing with variable (model) uncertainty problem

Data

- Juvenal, L. and Petrella, I. (2015). “Speculation in the oil market”, Journal of Applied Econometrics, 30: 612-649.
- Quarterly data between 1971 and 2009
- Time-series with missing observations excluded
- Finally, 127 explanatory variables: several macroeconomic variables from different countries, as well as, data derived from stock markets, economic activity index, etc.

Methodology: models

- DMA.DOW is the forecast from Dynamic Model Averaging with the dynamic Occam's window,
- BMA.DOW is the forecast from Bayesian Model Averaging with the dynamic Occam's window,
- DMA.1V is the forecast from Dynamic Model Averaging applied only to one-variable models,
- BMA.1V is the forecast from Bayesian Model Averaging applied only to one-variable models,
- DMS.1V is the forecast from Dynamic Model Selection applied only to one-variable models,
- BMS.1V is the forecast from Bayesian Model Selection applied only to one-variable models,
- TVP is the forecast from Time-Varying Parameters regression,
- LASSO is the forecast from LASSO regression,
- RIDGE is the forecast from RIDGE regression,
- DYN.EL.NET is the forecast from the elastic net regression, with the elastic net mixing parameter changing with time index,
- LARS is the forecast from the least-angle regression,
- B.LASSO is the forecast from the Bayesian LASSO regression,
- B.RIDGE is the forecast from the Bayesian RIDGE regression.

Methodology: models - literature

- Friedman, J., Hastie, T. and Tibshirani, R. (2010). “Regularization paths for generalized linear models via coordinate descent”, Journal of Statistical Software, 33: 1-22.
- Gramacy, R.B. (2017). monomvn: Estimation for multivariate normal and Student-t data with monotone missingness. https://CRAN.R-project.org/package=monomvn
- Hastie, T. and Efron, B. (2013). lars: Least angle regression, lasso and forward stagewise. https://CRAN.R-project.org/package=lars
- Onorante, L. and Raftery, A.E. (2016). “Dynamic model averaging in large model spaces using dynamic Occam's window”, European Economic Review, 81: 2-14.
- Raftery, A.E., Karny, M. and Ettler, P. (2010). “Online prediction under model uncertainty via Dynamic Model Averaging: Application to a cold rolling mill”, Technometrics, 52: 52-66.

Methodology: benchmark models

- ARIMA model with dynamically changing lags (Hyndman, R.J. and Khandakar. Y. (2008). “Automatic time series forecasting: the forecast package for R”, Journal of Statistical Software, 26: 1-22.),
- the naïve forecast,
- the moving average (last 100 observations).

Methodology: forecast comparision

- First 100 observations: in-sample
- Root Mean Squared Error
- multivariate Diebold-Mariano test for the Equal Predictive Accuracy (Mariano R.S. and Preve, D. (2012). “Statistical tests for multiple forecast comparison”, Journal of Econometrics, 169: 123-130.,

Drachal, K. (2018). multMDM: Multivariate version of the Diebold-Mariano test. https://CRAN.R-project.org/package=multDM)

start from all forecasts and to reject forecasts untill in the remaining set they do not differ significantly

Methodology: MDM test

- H0: Equal Predictive Accuracy
- H0: E g(e1) = … = E g(en)
- n – numer of forecasts from different models
- g() – loss function
- e1, …, en – models' errors

Results – according to MSE

##

Models with outstanding predictive ability:

	Rank	Sc	Mean loss
DMA.1V	1	-1.2860	1.4771
LASSO	2	-0.8650	1.4300
DYN.EL.NET	4	0.8850	1.4619
LARS	3	-0.8012	1.4312
MA			1.4880

p-value: 0.1233

Number of eliminated models: 11

##

Results – according to MDM test

##

Models with outstanding predictive ability:

	Rank	Sc	Mean loss
DMA.1V	1	-1.2860	1.4771
LASSO	2	-0.8650	1.4300
DYN.EL.NET	4	0.8850	1.4619
LARS	3	-0.8012	1.4312
MA			1.4880

p-value: 0.1233

Number of eliminated models: 11

##

Conclusions

- The differences in the forecasts' accuracies between the selected methods seem to be statistically insignificant.

Thank you for your attention!

Questions: kdrachal@wne.uw.edu.pl
Details: https://CRAN.R-project.org/package=multDM