003 - Review of literature

Karel Räpp, Henry Enno Turu

28 11 2020

1 Review of literature

1.1 Link between financial markets & real economy

Firstly, to dig deeper into various variables that help to predict future macroeconomic conditions, it is essential to understand what is the mechanism that potentially enables using the financial markets for predicting (movements in) the real economy. The base of it being one of the foundations of macroeconomic theory - Fischer theory, according to which the nominal interest rate is equal to the real interest rate and inflation. Hence, in the case of high interest rates and deflation, the outcome should be a recessionary period. Thus, the current stock/bond/commodity market prices include expectations about future performance & earnings - consequently on the aggregate level, the expectations can be generalized to the whole economy.

1.2 Volatility as a predictor of GDP

As the studies about first moments of financial variables have become fairly obsolete, researchers have started focusing on another level - e.g. (implied) volatilities. Fornari & Mele (2019) use the countercyclicality of financial volatility to construct a prediction model and conclude that stock volatility is a significant variable in predicting business cycles. The conclusion is even stronger when combining volatility with term spread - in such a case, their proposed model would have predicted at least 3 of the last recessions. Ferrara et al. (2014) continue the same path - they mix daily financial volatility with monthly industrial production and achieve significant results in predicting GDP; nevertheless their results are limited to the timeframe of Great Recession in 2008-2009.

Similarly to stock prices, option prices reflect future expectations. Taking this into account, David & Veronesi (2014) continue to discover implications of at-the-money (ATM) implied volatility (IV). They find that a positive shock to stock and bonds ATM IV is followed by a decline in future rates and probability of deflation. Thus, there exists a positive relationship between IV shock and possibility of recession.

Building upon the aforementioned various research, Cremers et al. (2017) have analyzed over 20 years of data to find out whether Treasury yield implied volatility can be used to predict different macroeconomic and financial measures such as growth & volatility of GDP, industrial production, employment. Using daily at-the-money option data from 1990 May until 2016 November for different treasury bonds and bills, they daily calculated (treasury) implied volatilities (YIV) using Black's model, which in essence is an adjusted model of the famous Black & Scholes model to value options on future contracts (Black, 1976) - see equation (1).

$$c = e^{-rT} + [FN(d_1) - KN(d_2)] \tag{1}$$

where

$$d_1 = \frac{ln(F/K) + 0.5\sigma_t^2]}{\sqrt{\sigma T}} \qquad\qquad d_2 = d_1 - \sqrt{\sigma T} \tag{2}$$

The options are chosen on the basis of exercise price being closest to the price of the underlying bond future - i.e. then its closest to at-the-money. This is done because Ederington & Lee (1993) & (1996) argue that contracts that are closest to at-the-money possess a strong link between spot and future markets. Therefore, these options can be treated as they are options on the bond spot market itself. Furthermore, they argue that those options tend to be the most liquid ones.

After obtaining the daily time series of implied volatility, the authors simply average the daily time series to obtain monthly data, which is regressed with different macroeconomic & control variables.

$$\sum_{j=1}^{j=H} log(1 + GDP_{i,t+j}) = \alpha_H + \beta_H \sigma_{IV,t}^{INT} + Controls + \varepsilon_{t+H} \tag{3}$$

The authors possess the data for different maturities (1,5,10,30 years). The research concluded that specifically a 5-year Treasury note significantly predicts most of the aforementioned elements, even after controlling for many other predictors such as term spread, stock returns, stock market implied volatility. This is so as Brandt et al. (2007) find that the price discovery tends to mostly happen in contracts with maturity of 5-year, both for cash and future markets.

1.3 Forecasting asymmetries

In the fairly recent wave of research of forecasting, the problems and limitations regarding full-sample forecasting have become more prevalent. More specifically, it has been found that the business cycles have a statistically significant effect on the model's predictive ability. Furthermore, through using state-dependent models, the effects of business cycle asymmetries can be evaluated on the forecasting performance of the model.

For example, in the Handbook of Economics, Chauvet & Potter (2013) evaluate the accuracy of different models with regards to the performance during recessionary periods and expansionary periods using the classifications for US recessions and expansions from National Bureau of Economic Research (NBER). They conclude that for all different models tested, the GDP growth is significantly harder to forecast during recessions when compared to expansionary periods. Furthermore, based on the results they state that although the forecasting ability of some of the models is relatively good during expansions, most of them fail during recessions.

Therefore, there is reason to believe that the model introduced by Cremers & Fleckenstein & Gandhi might not offer significant information about the future during calm days. Arising from this, we replicate their models while accounting for the effects of different economic phases and see whether it still offers a robust outcome.

1.4 Advanced methods for macroeconomic forecasting

Based on the recent research on forecasting, there is significant evidence on the poor performance of simple autoregressive models when forecasting macroeconomic real activity. In an IMF working paper "Deus ex Machina? A Framework for Macro Forecasting with Machine Learning" published in February 2020, Bolhuis and Rayner bring out the following key shortcomings of a simple OLS-based forecasting model:

- Collinearity
- Dimensionality
- Predictor relevance
- Non-linearity

To combat some of these shortcomings, Siliverstovs & Wochner (2020) assessed the forecasting performance of a dynamic factor model (DFM) compared to a simple autoregressive model. Furthermore, they also included the state-dependent subsamples to their research. The main finding is that there is a significant performance improvement in forecasting capability during the recessionary periods when using the dynamic factor model. Thus, it further illustrates the need for accounting the state-dependency and the importance of using more advanced models (especially during the recessionary periods) over simple autoregressive models to account for the shortcomings of simple OLS-based models.

However, as brought out by Bolhuis & Rayner (2020), while the DFM-s account for collinearity and dimensional shortcomings, they are still unable to account for the predictor relevance and non-linearity. This in turn increases the root mean square forecast errors (RMSFE) and thus, makes the predictions less accurate.

A solution for this can be found through implementing machine learning (ML) models which have been increasingly taken into use also in applied economics research. Researches have been drawn more and more towards the ML methods in forecasting mainly due to its ability to take into account nonlinearity and its emphasis on out-of-sample forecasting to avoid overfitting. Furthermore, the ML methods have been found to have better performance with regards to the forecasting accuracy and robustness (Carrasco & Rossi, 2016)

1.5 Random Forest

In our research we focus on a random forest based method as it requires little tuning. For example, one of the main problems regarding the models, overfitting, has shown to have little effect on the RF model (Díaz-Uriarte & Alvarez de Andrés, 2006).

Random forest itself is a supervised learning method i.e. it predicts outputs based on existing input and output pairs. The concept of random forest in essence is relatively simple.

It builds upon the decision trees i.e. random forest consists of multiple standalone decision trees. Figure 1.1 below depicts a single decision tree. The example here uses variables and makes a simple logical decision (yes/no) based on whether the variable's value is bigger or smaller than the comparison variable. Through these kinds of logical iterations the decision tree arrives at the end value. Furthermore, the random forest then is an average of the individual decision trees.

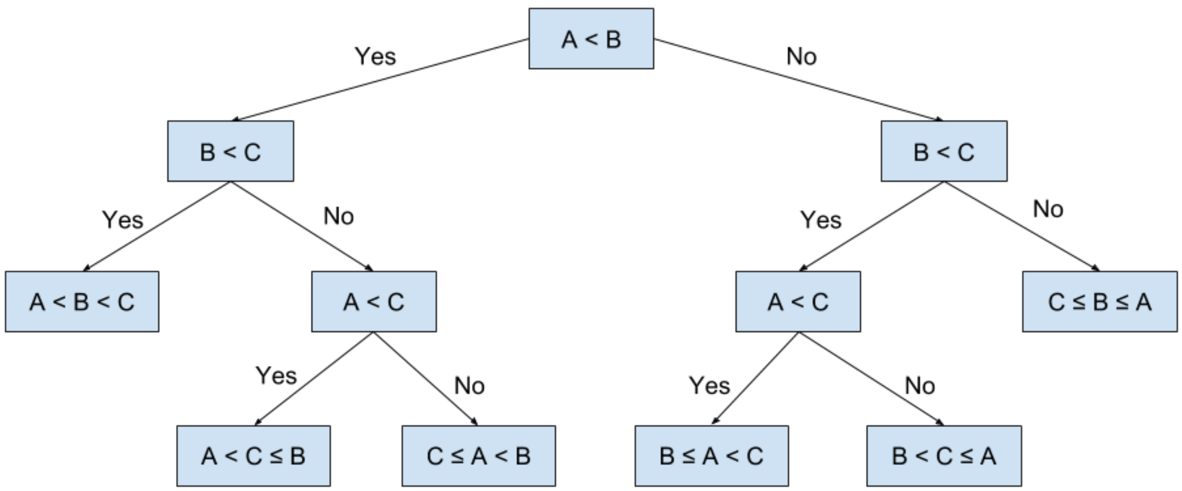

Figure 1.1: Example of a decision tree. Figure extracted from Niculaescu, O. (2018).

To take it together, our paper builds upon Fleckenstein et al while improving the methodology in 2 ways:

1. Addressing the topic of forecast model's state dependency
2. Introducing machine learning method in search of more accurate forecasting performance

Proceedingly, our hypotheses are the following:

Hypothesis 1. The treasury bond implied volatility (YIV) is a significant predictor of future macroeconomic real activity.

Hypothesis 2. Due to the business-cycle related asymmetries, the model based on the full-sample forecasts is inefficient in predicting during turbulent time periods.

Hypothesis 3. Using the RF model, it is possible to significantly reduce the RMFSE of the forecast thanks to its ability to take into account collinearity, predictor relevance and non-linearity while also maintaining the interpretability.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, *3*(1-2), 167–179. https://doi.org/10.1016/0304-405X(76)90024-6

Bolhuis, M. A., & Rayner, B. (2020). *Deus ex Machina? A Framework for Macro Forecasting with Machine Learning, WP/20/45, February 2020*.

Brandt, M. W., Kavajecz, K. A., & Underwood, S. E. (2007). Price discovery in the treasury futures market. *The Journal of Futures Markets*, *27*(11), 1021–1051. https://doi.org/10.1002/fut

Carrasco, M., & Rossi, B. (2016). In-Sample Inference and Forecasting in Misspecified Factor Models. *Journal of Business and Economic Statistics*, *34*(3), 313–338. https://doi.org/10.1080/07350015.2016.1186029

Chauvet, M., & Potter, S. (2013). Forecasting output. *Handbook of Economic Forecasting*, *2*, 141–194. https://doi.org/10.1016/B978-0-444-53683-9.00003-7

Cremers, M., Fleckenstein, M., & Gandhi, P. (2017). Treasury Yield Implied Volatility and Real Activity. *Journal of Financial Economics (JFE), Forthcoming*. https://doi.org/10.2139/ssrn.3006473

David, A., & Veronesi, P. (2014). Investors' and Central Bank's uncertainty embedded in index options. *Review of Financial Studies*, *27*(6), 1661–1716. https://doi.org/10.1093/rfs/hhu024

Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*, 1–13. https://doi.org/10.1186/1471-2105-7-3

Ederington, L. H., & Lee, J. H. (1993). *How Markets Process Information: News Releases and Volatility* (Vol. 48, pp. 1161–1191). https://doi.org/10.1111/j.1540-6261.1993.tb04750.x

Ederington, L., & Lee, J. (1996). The Creation and Resolution of Market Uncertainty: The Impact of Information Releases on Implied Volatility. *The Journal of Financial and Quantitative Analysis*, *31*(4), 513–539. http://journals.cambridge.org/abstract%7B/_%7DS0022109000023784

Ferrara, L., Marsilli, C., & Ortega, J. P. (2014). Forecasting growth during the Great Recession: Is financial volatility the missing ingredient? *Economic Modelling*, *36*, 44–50. https://doi.org/10.1016/j.econmod.2013.08.042

Fornari, F., & Mele, A. (2019). Financial volatility and real economic activity. *Journal of Financial Management*, *1*(2), 155–196. https://doi.org/10.4324/9780429456572

Siliverstovs, B., & Wochner, D. (2020). State-Dependent Evaluation of Predictive Ability †. *Journal of Forecasting*. https://doi.org/10.1002/for.2715