State-dependent forecasting of macroeconomic real activity

Research Design Description

Authors:

Karel Räppo

Henry Enno Turu

Supervisor:

Boriss Siliverstovs

Riga

2020

Contents

1 Introduction

Forecasting the real economy has been a widely researched topic as it is crucial for effective policy making - nevertheless, there are many options in doing so, be it equity, fixed income, or commodity markets.

Stock & Watson (2003) reviewed over 93 working papers & articles and documented numerous variables (interest rates, term spreads, returns, exchange rates, etc.) that have been used for forecasting macroeconomy. Nevertheless, there is no strong agreement which can do so consistently, as per the conclusion made: ”Some asset prices have been useful predictors of inflation and/or output growth in some countries in some time periods” (Stock & Watson, 2003, p. 822). Furthermore, even less variables are there that can predict macroeconomic real activity through longer time periods (i.e. time periods of over a year).

However, in a work by Cremers et al. (2017), they tackle the topic of predicting the macroeconomic real activity through the treasury implied volatility (YIV) that is derived from options on treasury bond futures of different maturity periods. Furthermore, they focus on predicting the periods of 12-36 months ahead and find that the YIV is able to predict the GDP growth successfully during these periods. Furthermore, they show that YIV is able to predict also consistently industrial production, consumption, and employment.

Thus, the first part of our research aims to replicate the research by Cremers et al. (2017) - predicting GDP growth using YIV.

In the second part of the research we build upon the research of Cremers et al. (2017) in two ways.

Firstly, to account for the state-dependency of forecast models, as shown to be relevant by Chauvet & Potter (2013) and Siliverstovs & Wochner (2020), we conduct separate regression analysis for both recessionary periods and expansionary periods (classification of the US business cycles obtain from NBER).

Secondly, we propose using machine learning models to account for the shortcomings of the simple OLS-based models. More specifically, the inability of simple OLS-based models to account for collinearity, dimensionality, predictor relevance and non-linearity (Bolhuis & Rayner, 2020). To elaborate, we apply the framework of a random forest (RF) model tailored specifically for macroeconomic forecasting (MRF) as it is able to tackle the aforementioned problems - the MRF model is shown to significantly decrease the root mean square forecasting errors (RMSFE) while also being interpretable.

Thus, our research question consists of three sub-questions:

1. **Can options on Treasury bond futures effectively predict the USA's macroeconomic and financial activity?**
2. **If so, is this conclusion robust even after taking into account different business cycles.**
3. **Do Machine Learning methods such as Macroeconomics Random Forest (MRF) improve the forecast accuracy?**

2 Review of literature

2.1 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.

2.2 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 calculated daily (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

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

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.

2.3 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.

2.4 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)

2.5 Random Forest

In the research, we focus on a random forest thanks to its ability to account for complex nonlinearities. 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.

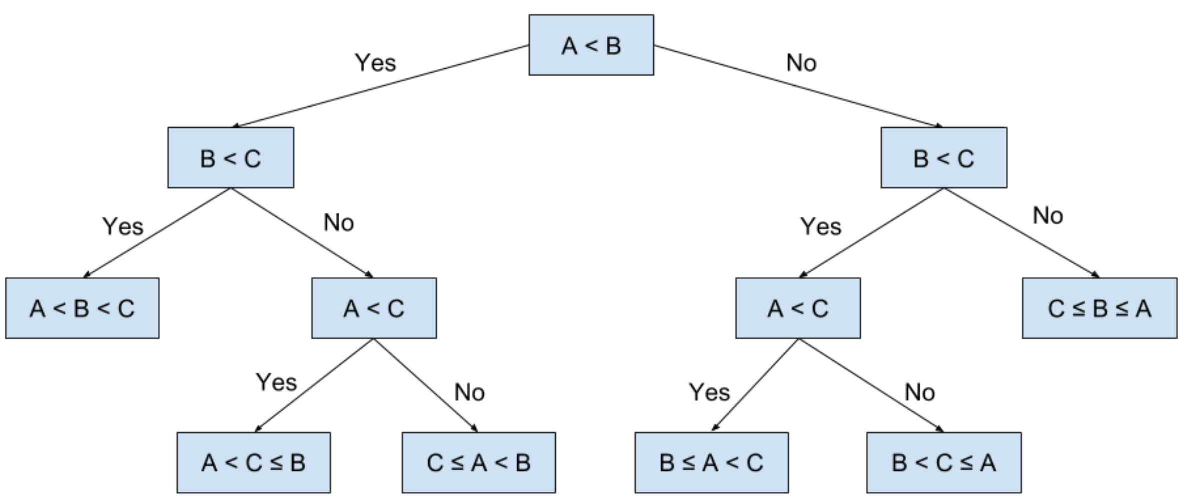

Figure 2.1: Example of a decision tree. Figure extracted from Niculaescu, O. (2018).

Figure 2.1 above 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.

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.

3 Methodology

3.1 Linear regressions

The first part of our methodology consists of replicating the methodology conducted by Cremers et al. (2017). In other words, we test whether the treasury implied volatility can be used to predict the future macroeconomic real activity.

To quantify the predictability of macroeconomic real activity using the treasury implied volatility, we run an ordinary least squares (OLS) regression. In the first regression, we take the 5-year YIV and use it to predict the logarithmic values of year-on-year growth rate of the real GDP. In the equation, H is equal to the periods predicted - e.g. if H=4, it means that we are taking the rolling overlapping average of GDP growth over the 4 quarters.

$$\sum_{j=1}^{j=H} log(1 + GDP_{i,t+j})/H = \alpha_H + \beta_H \sigma_{IV,t}^{INT} + \varepsilon_{t+H} \tag{4}$$

In addition, by using an autoregressive model, an additional regression with the lagged values of macroeconomic variables is run to see whether the predictive ability will improve. The number of lagged variables included is decided by the Akaike information criterion (AIC).

We then proceed to include different control variables to the regression model such as term spreads, credit spreads, stock returns, stock market implied volatility.

$$\sum_{j=1}^{j=H} log(1 + GDP_{i,t+j})/H = \alpha_H + \beta_H \sigma_{IV,t}^{INT} + Controls + \varepsilon_{t+H} \tag{5}$$

Next, to determine the direction of the causality, we run a Vector Autoregressive model (VAR) Granger Causality test on every variable to make sure that e.g. YIV indeed granger-causes movements in GDP.

3.2 Adjusting for different economic cycles

In the second part of our research we build upon the research conducted by Cremers et. al. We test whether the model holds when accounting also for the possible business cycle performance asymmetries as suggested by Siliverstovs & Wochner (2020). We perform separate regressions on the subsamples i.e. forecasting performance during expansionary and recessionary periods. More specifically, we compare the root-mean-square-forecasting-error (RMSFE) of the predictive model during recession and expansion with full sample forecast to find out whether full sample forecast's relative RMSFE forecasts are robust during the expansionary and recessionary subsamples (in other words, we check whether a distributional shift occurs during the recessionary and/or expansionary periods).

3.3 Macroeconomic Random Forest (MRF)

As Machine Learning methods successfully tackle nonlinearities that cannot be accounted for in simple linear models, the forecasting gains tend to be the highest during times of high economic uncertainty (Coulombe et al., 2020). As our research includes many different financial crises, we expect to find forecasting gains by utilizing Random Forest whose benefits were elaborated in review of literature.

However, in our research, we use an algorithm developed by Coulombe (2020) named Macroeconomic Random Forest. His proposed model further develops off-the-shelf random forest with regards to two areas:

- **Statistical efficiency** - due to the nature of random forest approximation, achieving smooth linear relationships takes many splits which in case of a smaller time series can result in the loss of degrees of freedom which in turn increases the error terms. MRF, however, includes a linear component which can capture these less complex relationships with less splits, saving the degrees of freedom for more complicated sequences.

- **Interpretation** - random forest model itself is often considered as a black box model in terms of interpretability, and external interpreter algorithms are used to understand it. However, MRF tackles this problem through adding the linear part - after splitting the initial parameter into many pieces of the whole forest, it allows to directly interpret the coefficients of the underlying surrogate models[1].

The general model (see Equation (6)) proposed by Coulombe (2020). The main difference between MRF and RF is that the latter is a restricted model ,where Xt=1.

$$y_t = X_t \beta_t + \varepsilon_t \tag{6}$$

$$\beta_t = F(S_t) \tag{7}$$

The author has provided us with an R-package[2], which dramatically simplifies our research process and makes it more feasible for a Bachelor level thesis.

[1] The full derivation & optimization can be found in Coulombe (2020)

[2] Macroeconomic Random Forest R package can be downloaded form https://philippegouletcoulombe.com/code

4 Data

For the regression analysis, we (have) extract(ed) the following data:

- Quarterly 5-year Treasury Implied Volatility
- Dependent Variables:
  1. Quarterly GDP year-on-year growth rate
  2. Monthly industrial production
  3. Monthly non-farm payroll
  4. Monthly consumption
- Control Variables
  5. US treasury interest rates to construct term spreads
  6. credit spreads
  7. stock returns
  8. stock market implied volatility
  9. Housing Starts

As mentioned in the review of literature, YIV is constructed using Black model & deriving implied volatility through option prices, time-to-maturity, etc. The data regarding options on treasury bond futures, however, is collected by CME and available only through accessing their database.

Due to the data being behind the paywall, we had to resort to other measures to access the data. Firstly, we have contacted prof. Cremers, Gandhi & Fleckenstein, who are ready to share their data with us if the CME group gives their consent. Nevertheless, we cannot rely solely on this possibility, and thus found an alternative.

As researchers typically do not typically post underlying data with their research, various plot digitizers have seen an exponential increase in use. Drevon et al. (2017) researched intercoder reliability, during which over 3500 data points were extracted with WebPlotDigitizer from 36 different graphs. Nevertheless, they controlled the validity of the results and concluded that there was a near perfect correlation (r=0.989 with p-value <0.01) between extracted and actual data. Nevertheless, the limitations mentioned highlight coders previous experience with plot-digitizing tools.

Furthermore, Burda et al. (2017) also highlight that systematic reviewers often tend to have data constraints which is why plot digitizers are of a great help. They estimated data using WebPlotDigitizer and conclude that the extraction done by different coders was consistent; nevertheless, in the case of continuous data (compared to event data), the distribution varied more. Whatsoever, the intreclass coefficient for both types of plots was over 95%.

We also used the WebPlotDigitzer in our research and as validity test extracted GDP from the same graph as YIV time series & plotted it with actuals - see the graph below.

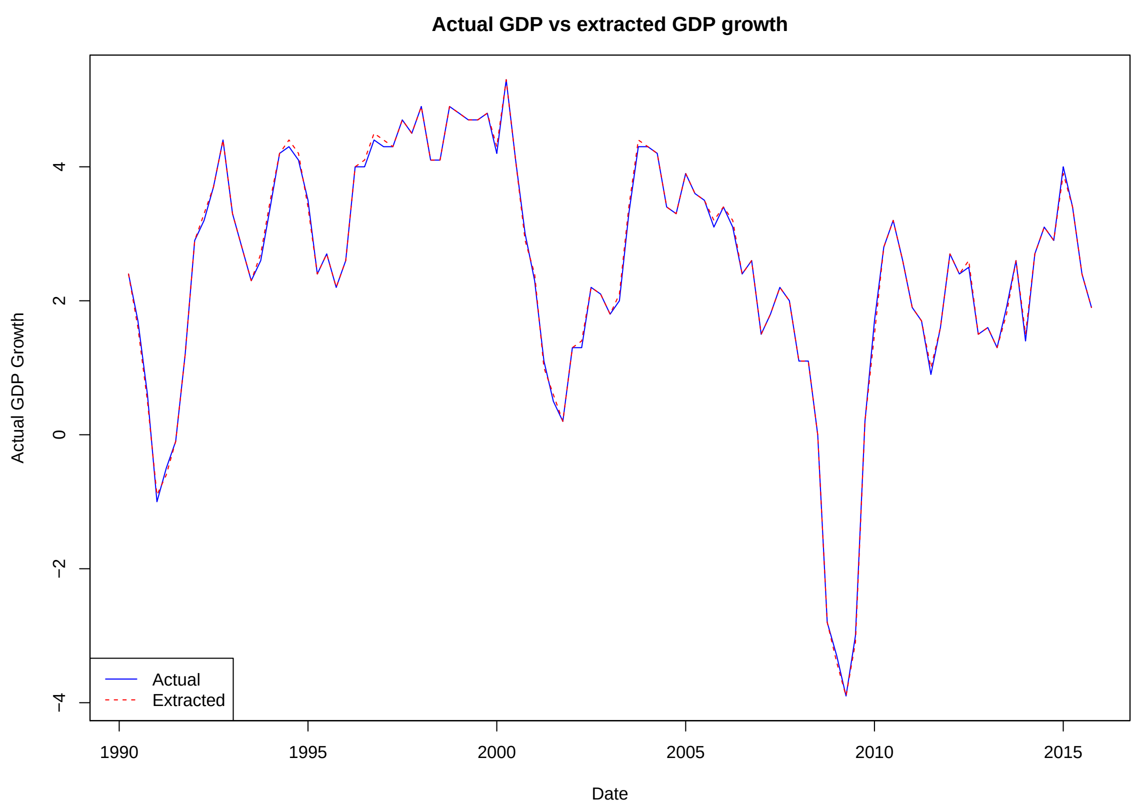

Figure 4.1: Actual GDP growth vs extracted GDP growth in % using WebPlotDigitizer

The data regarding GDP year-on-year growth rate was extracted on a quarterly basis from Archival FRED (2020). Furthermore, the vintage 2019-12-20 was extracted as it coincided the best with the data used in the research paper by Cremers et al. (2017). Other dependent variables (industrial production, non-farm payroll and consumption) & housing data was extracted from the same database but on a monthly basis. The data for risk free rates that will be used in constructing term spreads was extracted from the dataset of The Federal Reserve Board working paper. (Gurkaynak et al., 2006)

5 Appendices

5.1 Appendix A

Notes: This table includes summary statistics for main variables used in our research. Statistics include mean, standard deviation,, min, 1st quartile, median, 3rd quartile, max & number of valid data points. In Panel A, different YIV data is summarized. In Panel B, we have listed the main dependent variables which are used for predictions. GDP denotes the year-on-year growth rate(quarterly data), CON denotes YOY consumption growth(monthly data), EMP describes YOY growth rate for non-farm payroll and lastly IND stands for Industrial production YOY growth (monthly data). In Panel C, different control variables are listed: SVEN1F01 - 1 year treasury bond par yield.

Table 5.1: Summary Statistics

	Mean	Std.Dev	Min	Q1	Median	Q3	Max	N.Valid
Panel A: YIV								
YIV	3.34	1.31	1.39	2.60	3.00	3.62	9.21	103
Panel B: Dependent Variables								
GDP	2.50	1.78	-3.92	1.71	2.61	3.98	5.30	103
CON	4.88	1.95	-3.03	3.92	5.11	6.22	9.02	312
EMP	1.07	1.67	-5.00	0.20	1.60	2.20	3.50	312
IND	2.00	4.05	-15.33	1.19	2.74	4.16	8.54	312
Panel C: Control Variables								
SVEN1F01	3.89	2.40	0.21	1.38	4.35	5.88	9.29	6486
VIX	19.83	7.64	10.82	14.20	17.76	23.54	62.64	312
HOUSNG	1.01	18.36	-54.80	-7.25	2.80	12.70	50.00	312

Note:
Additional control variables will be added upon construction. Furthermore, currently the frequency of the datasets differs for different variables but this will be addressed in the research process.

5.2 Appendix B.

Notes: This table includes regression using GDP & YIV. Controls will be added during research process. The equation for the regression is the following:

$$\sum_{j=1}^{j=H} log(1 + GDP_{i,t+j})/H = \alpha_H + \beta_H \sigma_{IV,t}^{INT} + Controls + \varepsilon_{t+H} \tag{8}$$

Table 5.2: Regression output

	H12	H18	H24
Panel A: YIV			
R-Squared	21.98	21.41	23.73
Intercept	4.38	4.23	4.19
Beta	-0.57	-0.52	-0.50
t-statistic.(Intercept)	11.17	11.51	12.43
t-statistic.YIV	-5.25	-5.11	-5.41

Note:

Preliminary regression output. Initial data needs to be re-checked.

5.3 Appendix C

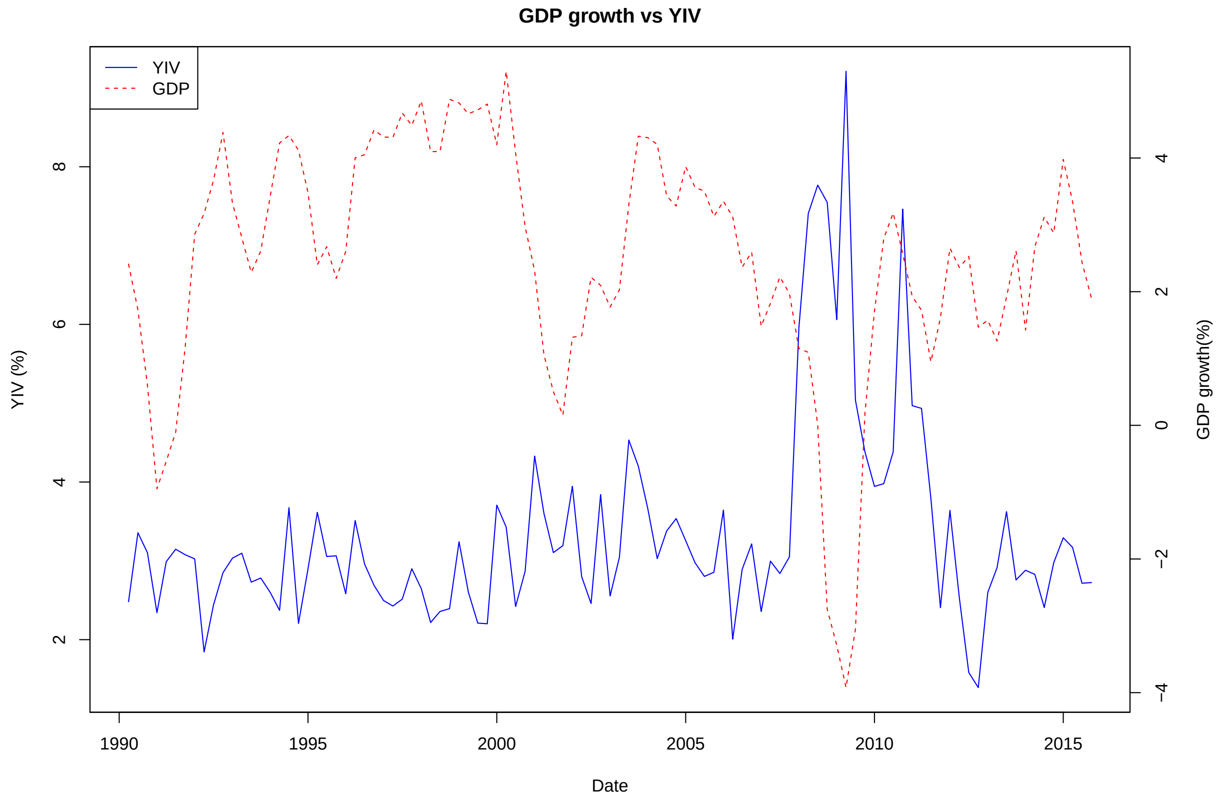

Figure 5.1: GDP Growth(%) vs 5-year Treasury Implied Volatility

6 References

Archival FRED. (2020). *Dataset. (Vintage 2020-20-12)*. https://alfred.stlouisfed.org/

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

Burda, B. U., O'Connor, E. A., Webber, E. M., Redmond, N., & Perdue, L. A. (2017). Estimating data from figures with a Web-based program: Considerations for a systematic review. *Research Synthesis Methods*, *8*(3), 258–262. https://doi.org/10.1002/jrsm.1232

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

Coulombe, P. G. (2020). *The Macroeconomy as a Random Forest* [PhD thesis]. https://doi.org/10.2139/ssrn.3633110

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

Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data. *Behavior Modification*, *41*(2), 323–339. https://doi.org/10.1177/0145445516673998

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

Gurkaynak, R. S., Sack, B. P., & Wright, J. H. (2006). *The U.S. Treasury Yield Curve: 1961 to the Present*. https://doi.org/10.2139/ssrn.920183

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

Stock, J. H., & Watson, M. W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, *XLI*, 788–829.