Forecasting the Next Recession

Nikhil Gupta

Stuart Miller

Motivation

The Need	Definition	Proposed Solution
We would like to forecast the next recession.	Two consecutive quarters of decline in GDP is considered a working definition of a recession.*	We will forecast the GDP for the next 2 quarters to see if the models can predict the next recession.

* https://www.investopedia.com/terms/r/recession.asp

Dataset

Response Variable

- Response Variable: % change in GDP (USA)
- Quarterly observations from 1982 to 2019
  - 151 observations

Exogenous Variables

- An additional 20 exogenous variables were also collected.
- Contained economic indicators related to the labor market, monitory policy, consumer related data, business environment, stock data, exchange rates and several macro-economic factors

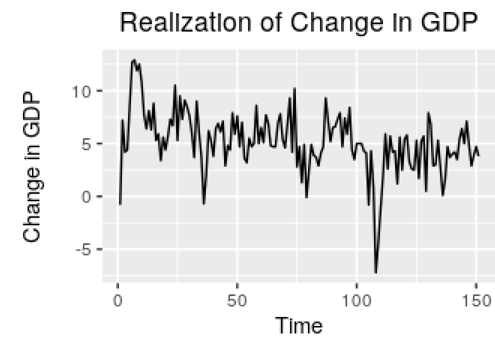
Data was collected from the Federal Reserve Bank of St. Louis Economic Data (FRED) https://fred.stlouisfed.org/

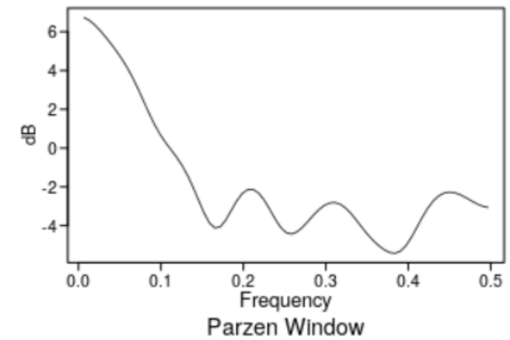
Univariate EDA: Checking Stationarity

Condition 1 & 2: Constant Mean & Variance

- Constant Mean
  - The mean does not appear to change over time. There does not appear to be evidence of a deterministic signal or oscillatory process.

- Constant Variance
  - The realization does not appear to show sufficient evidence of non-constant variance.
  - However, we only have one realization so this is difficult to assess.

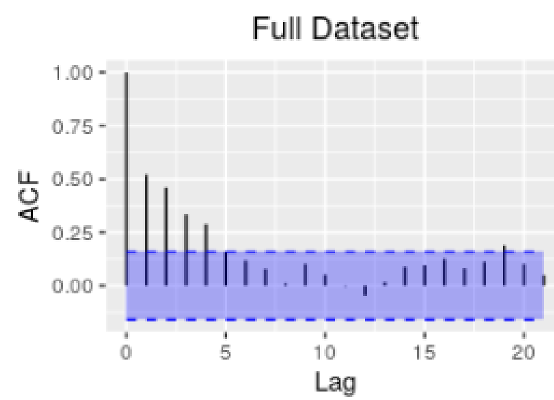


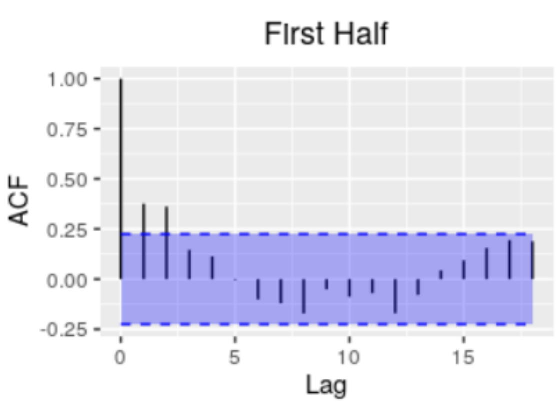


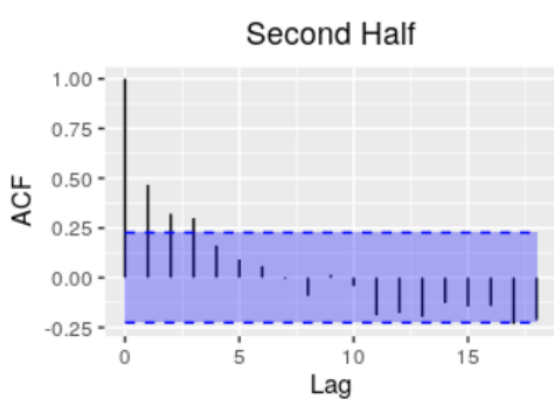
Univariate EDA: Checking Stationarity

Condition 3: The correlation of X_{t_1} and X_{t_2} depends only on $t_2 - t_1$

- The significant ACFs of the first and second half of the realization exhibit similar. Additionally, they appear to exhibit similar characteristics as the full data set.







Univariate Modeling: Model ID

An ARMA(2, 0) was selected by both AIC and BIC.

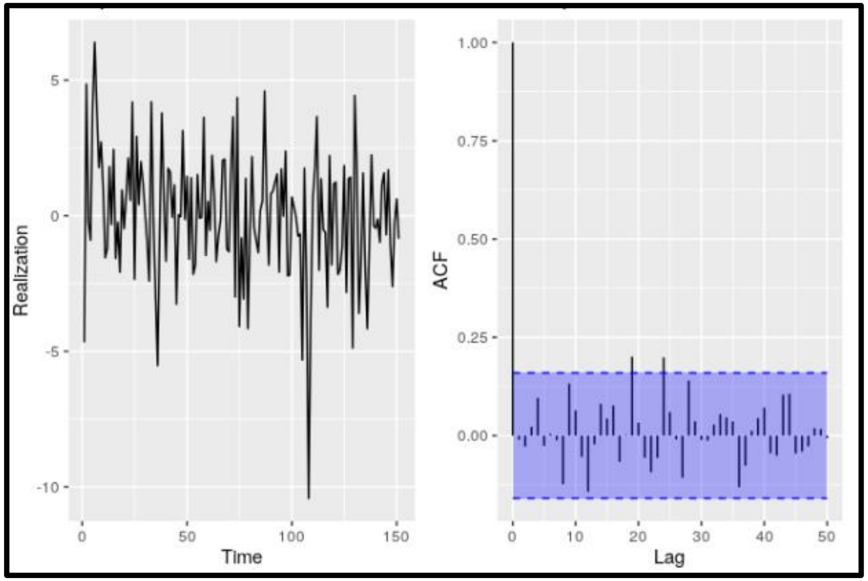
p	q	AIC
2	0	1.765702
1	1	1.772199
1	2	1.776732
2	1	1.778418
3	0	1.778594

p	q	BIC
2	0	1.825648
1	1	1.832145
1	0	1.854565
1	2	1.85666
2	1	1.858346

ARMA(2,0) Factored Model

$$(1 - 0.7391B)(1 + 0.3472B)(X_t - 5.149) = a_t$$
$$with\ \hat{\sigma}_a^2 = 5.618$$

Univariate Modeling: White Noise Evaluation



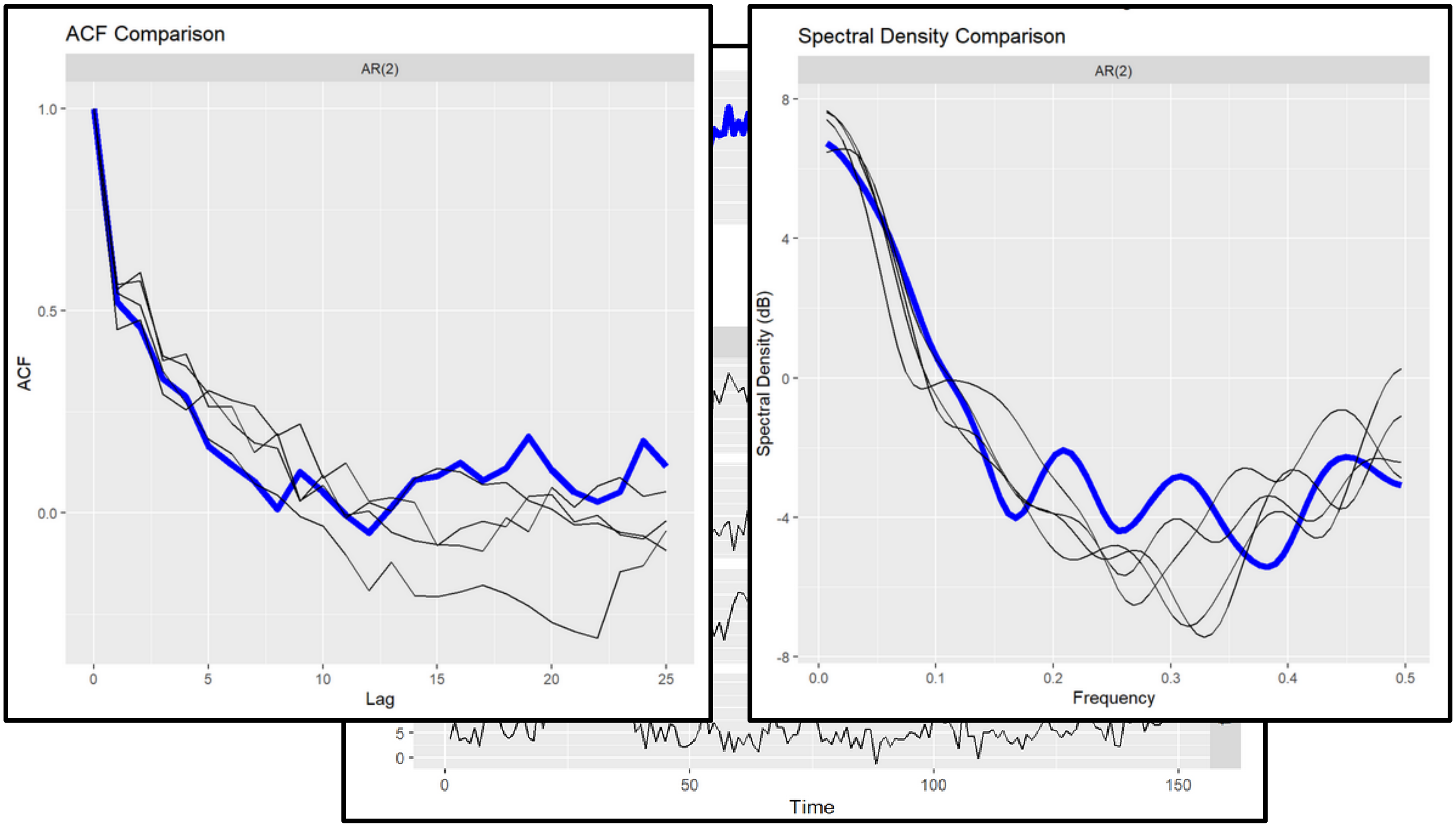
test	K	chi.square	df	pval	Decision
Ljung-Box test	24	32.08723	24	0.1248443	FTR NULL
Ljung-Box test	48	53.20827	48	0.2806241	FTR NULL

The time plot of the model residuals appear to be generally consistent with white noise.

Only two of the autocorrelations of the residuals appear to be marginally significant. This is not unusual with at 95% confidence level.

The Ljung-Box test fails to reject the null hypothesis at K = 24 and K = 48.

Univariate Modeling: Simulated Realizations

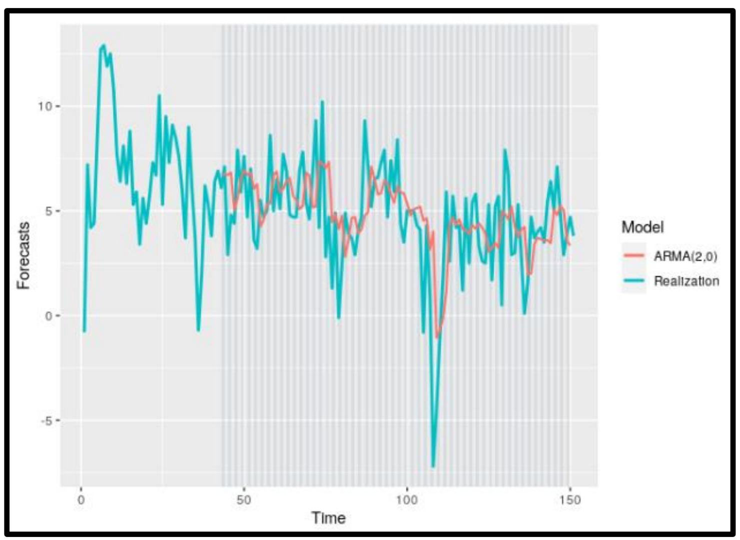


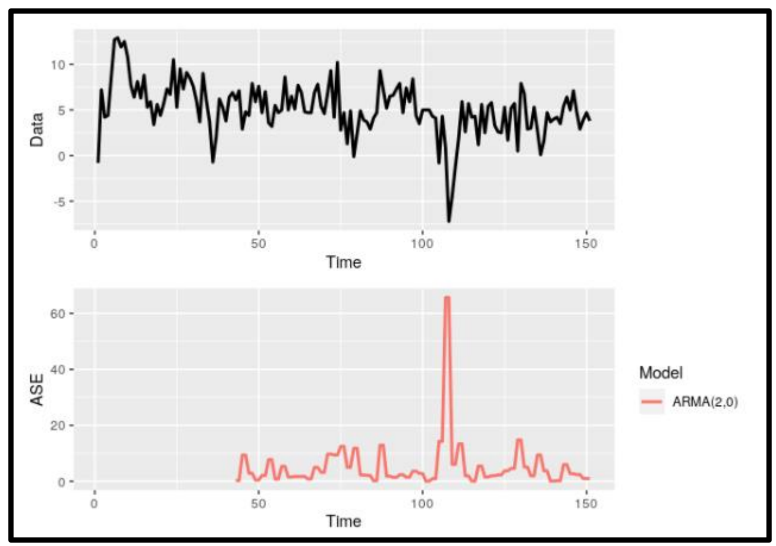
Univariate Modeling: Performance

Batch Size = 50 observations (Use 4 years of data to predict the next 2 quarters)

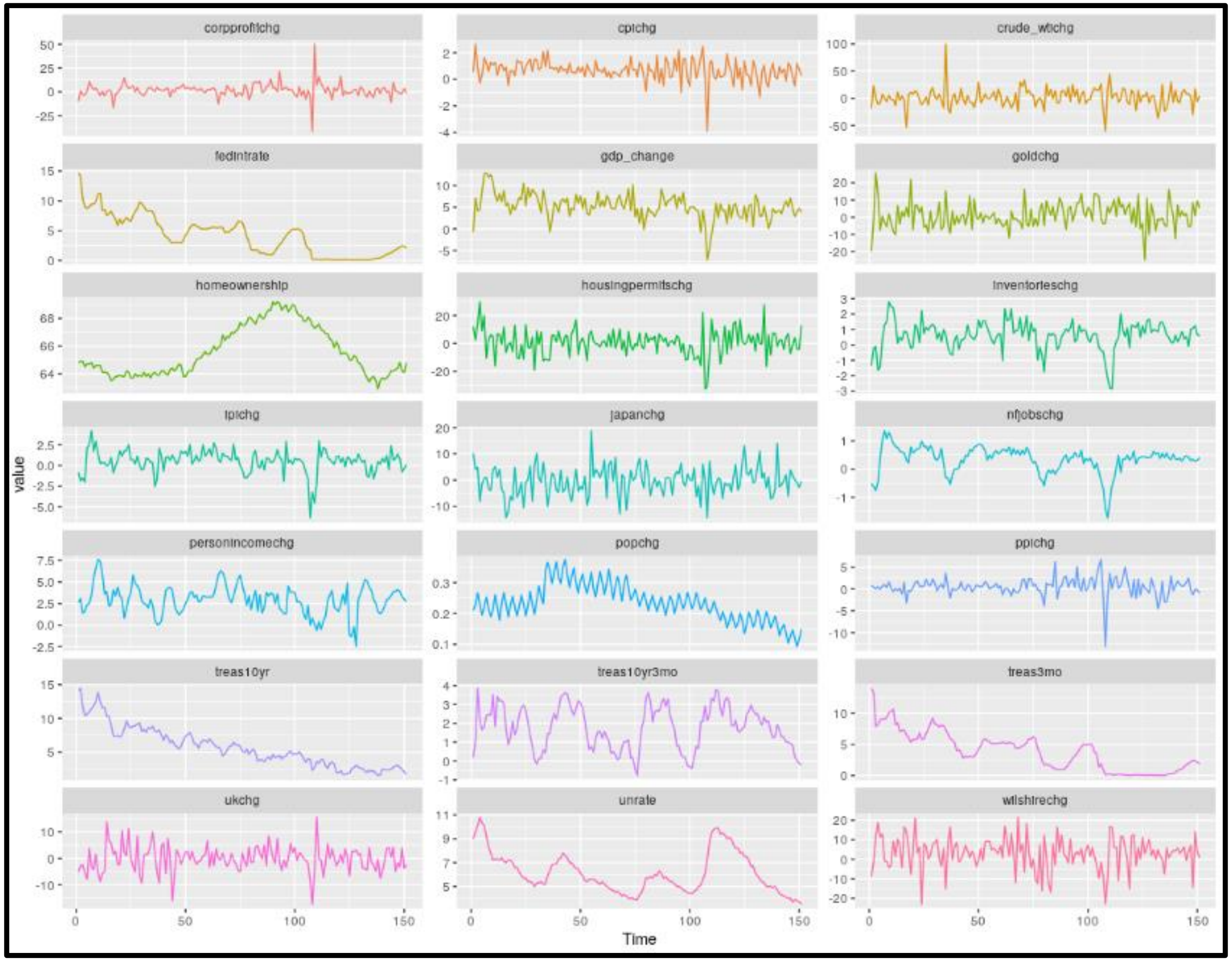
As expected, the ARMA model appears to capture the movement of the realization. However, it does not capture the sharp changes in the realization.

Generally, the model appears forecast ASE less than 15 over the sliding window. The primary error occurs at the large change in step 100.





Multivariate EDA: Realizations



Multivariate EDA: Cross Correlation Analysis

Exogeneous Variable Cross-correlations

variable	max_ccf_index	max_ccf_value
nfjobschg	0	0.6823141
ipichg	0	0.5907157
inventorieschg	0	0.4852442
treas10yr	0	0.4677121
treas3mo	0	0.4453255
fedintrate	0	0.4162709
personincomechg	0	0.3711154
homeownership	-12	-0.2841890
cpichg	0	0.2797316
unrate	-5	0.2668851
housingpermitschg	-1	0.2664965
wilshirechg	-3	0.2598483
ppichg	0	0.2579584
treas10yr3mo	-9	0.2085462
corpprofitchg	0	0.1764240
goldchg	-8	-0.1509366
crude_wtichg	0	0.1366070
popchg	0	0.1327717
japanchg	-9	-0.1071158
ukchg	-1	0.0953229

Observations:

- Several variables show strong cross-correlation with "GDP change".
- Most of the strongly cross corelated exogenous variables show maximum cross correlation at lag = 0.

NOTE: We only considered negative lags in this evaluation since we would not have access to future values while building the models

VAR Modeling: Process

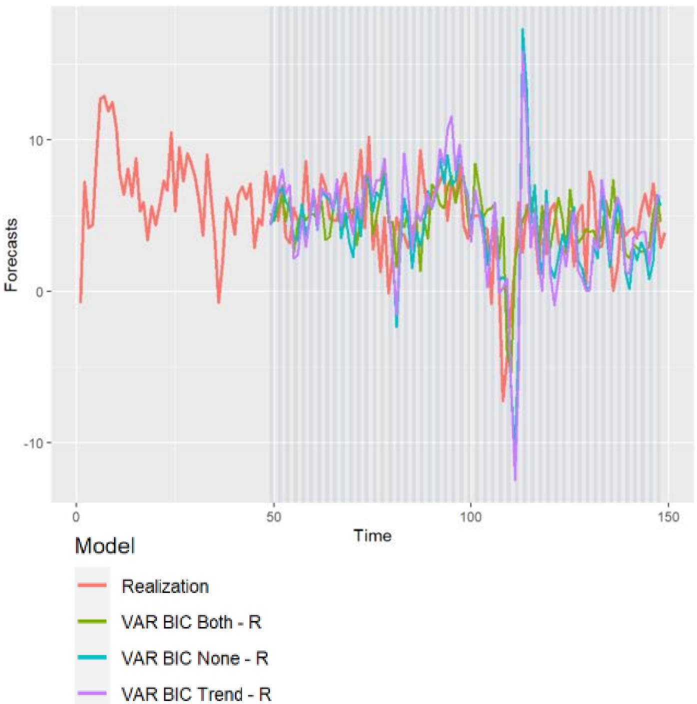
Need for variable selection to reduce overfitting

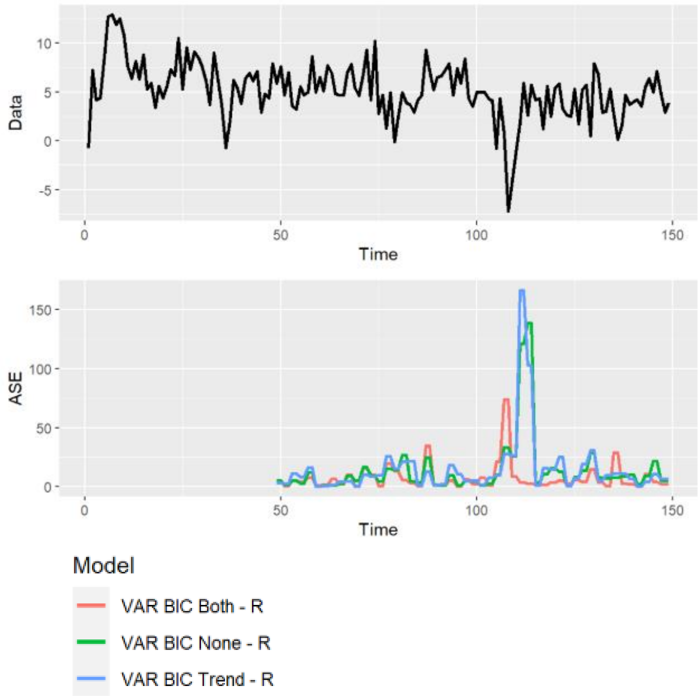
1. Use VARselect and BIC to select the maximum lag to consider for various trend types
2. Fit the model with selected lag from VARselect
3. Remove insignificant elements
    - Remove variables that are not significant at any lag
    - Reduce the maximum lag to the maximum significant found in the fit.
4. Select trend type based on ASE performance

Model	VARSelect p	Sig. Lags	Significant Variables
VAR BIC Both	6	3	gdp_change, nfjobschg, cpichg, ppichg
VAR BIC None	6	6	nfjobschg, corpprofitchg
VAR BIC Trend	6	6	nfjobschg, corpprofitchg

VAR Modeling: Model ID

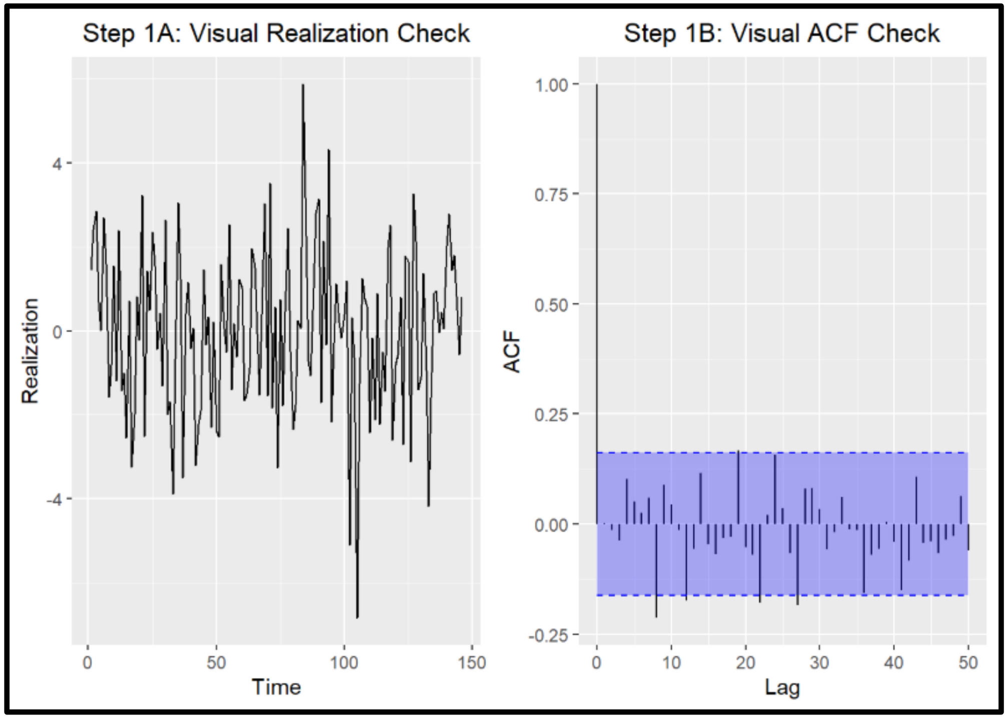
Batch Size = 50 observations (Use 4 years of data to predict the next 2 quarters)





VAR Modeling : White Noise Evaluation

“VAR BIC Both – R” Model



The time plot of the model residuals appear to be generally consistent with white noise.

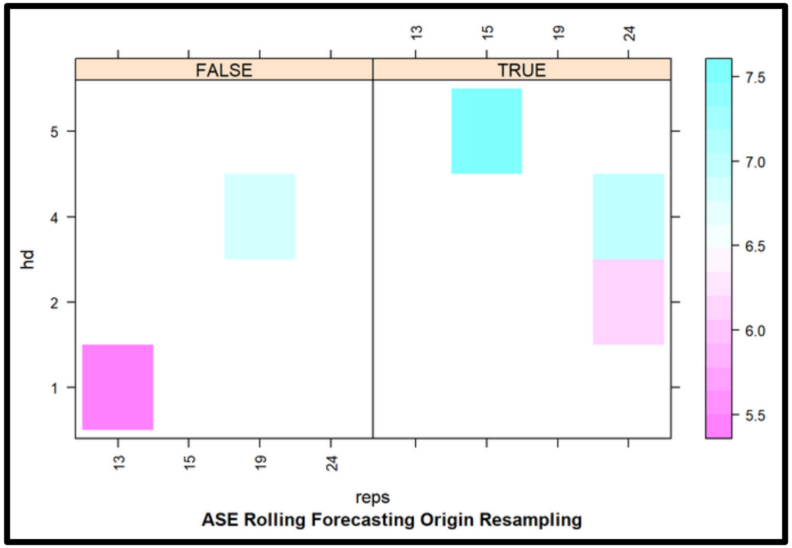
A few autocorrelations are marginally significant, but this is with in the 95% confidence level.

MLP Modeling: Grid Search

There are a large number of variables and hyperparameters that could affect the performance of an MLP model.

We used a random search to find a good set of hyperparameters

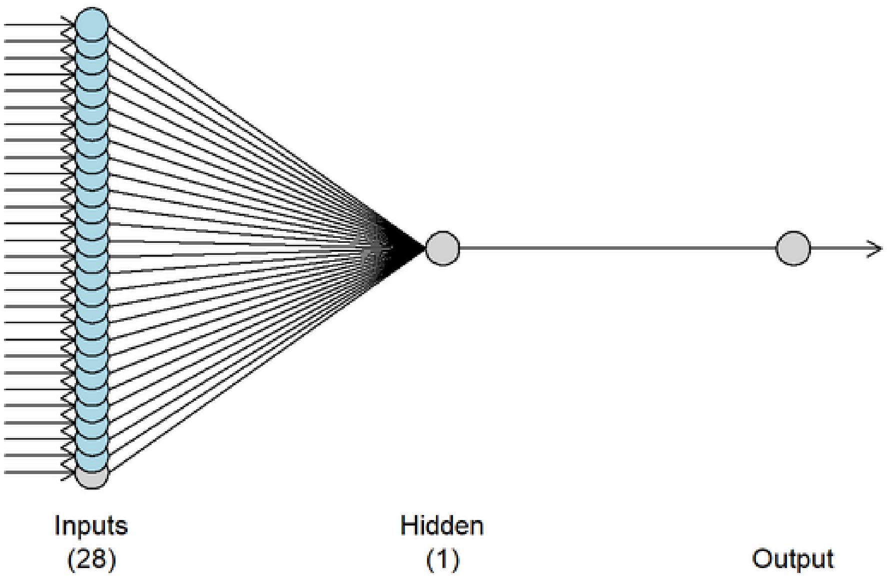
Batch Size = 50 observations (Use 4 years of data to predict the next 2 quarters)



Best Hyperparameters
- Hidden Layers: 1
- Repetitions: 13
- Use Seasonality: False

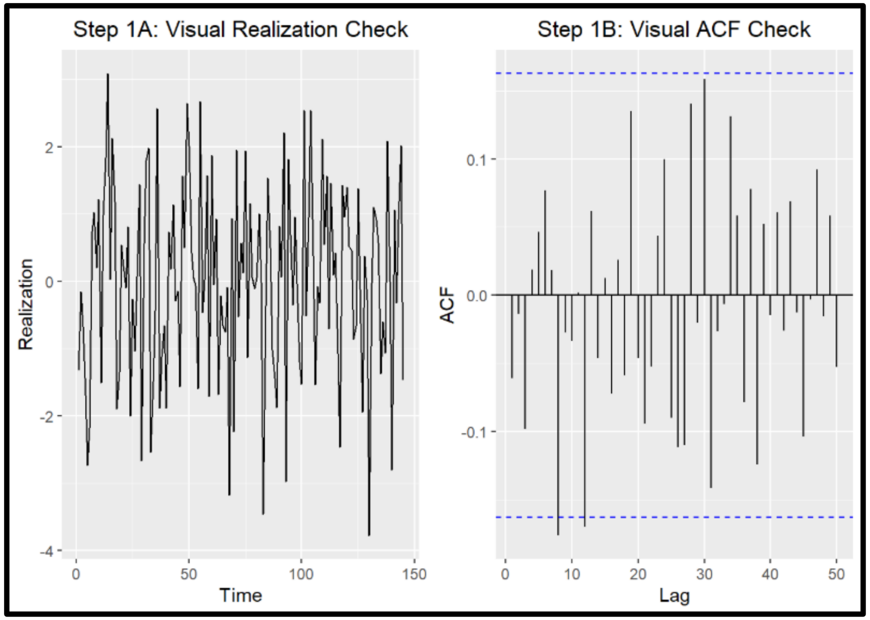
MLP Modeling: Final Model

```
## MLP fit with 1 hidden node and 13 repetitions.
## Univariate lags: (3)
## 17 regressors included.
## - Regressor 1 lags: (3)
## - Regressor 2 lags: (1,2,3)
## - Regressor 3 lags: (2)
## - Regressor 4 lags: (3)
## - Regressor 5 lags: (2)
## - Regressor 6 lags: (2)
## - Regressor 7 lags: (1,3)
## - Regressor 8 lags: (2)
## - Regressor 9 lags: (1,3,4)
## - Regressor 10 lags: (1,3)
## - Regressor 11 lags: (4)
## - Regressor 12 lags: (1,3)
## - Regressor 13 lags: (2)
## - Regressor 14 lags: (3,4)
## - Regressor 15 lags: (1,3,4)
## - Regressor 16 lags: (1)
## - Regressor 17 lags: (1)
## Forecast combined using the median operator.
## MSE: 2.0414.
```



1 Univariate Lag + 27 Exogenous Variable Lags

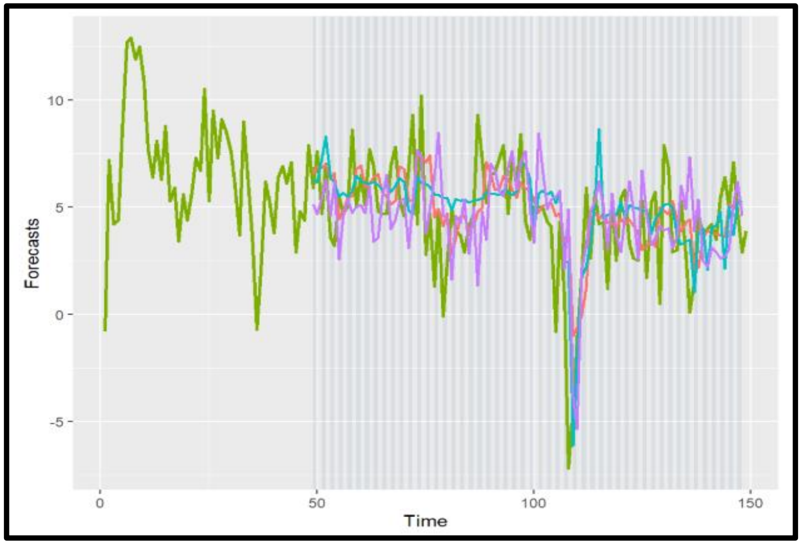
VAR Modeling : White Noise Evaluation

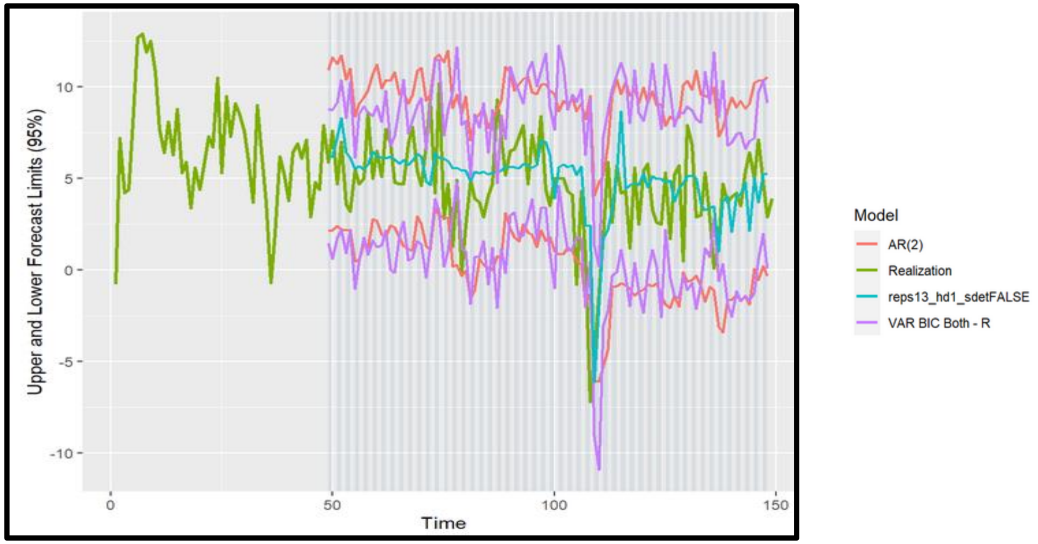


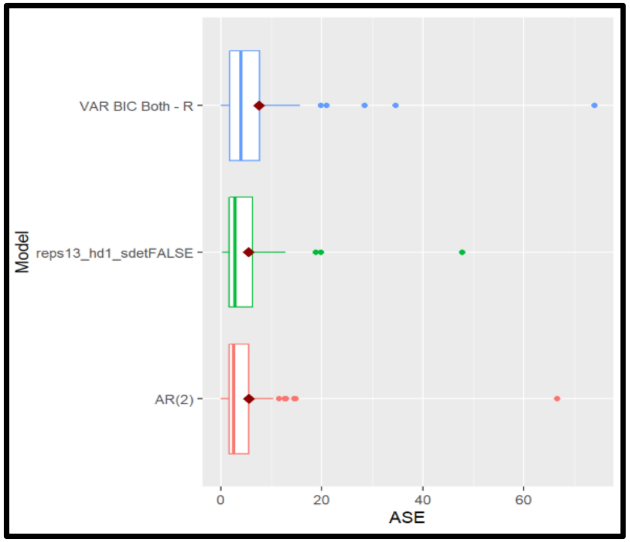
The time plot of the model residuals appear to be generally consistent with white noise.

A few autocorrelations are marginally significant, but this is with in the 95% confidence level.

Base Models: Comparison







AR(2) shows the best performance

- Lowest mean ASE
- Tightest distribution of rolling window ASEs

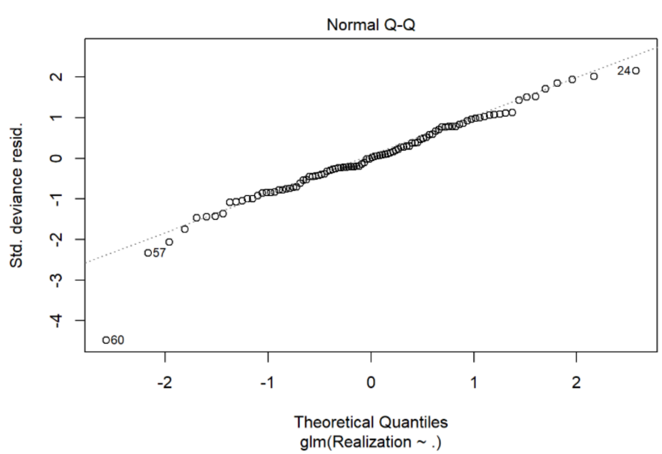
All models show a large ASE value, occurring at the steep dip in the realization

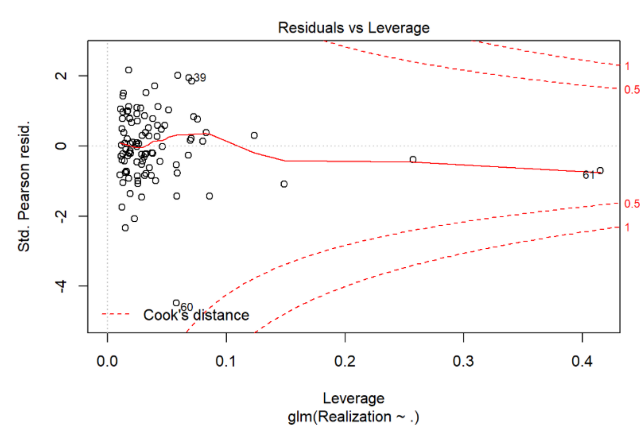
Modeling: Ensemble

Three types of ensemble models were created

- Mean of the forecasts of the base models
- Median of the forecasts of the base models
- Linear regression combining the forecasts of the base models

Coefficient	Estimate	Std. Error	Pr(>|t|)
(Intercept)	0.41118	0.74790	0.58374
AR(2)	0.31278	0.24443	0.20376
VAR BIC Both - R	-0.09548	0.15598	0.54190
reps13_hd1_sdetFALSE	0.60804	0.22717	0.00875*





Ensemble Modeling: Performance

Base Models and Ensemble Models were used to make predictions on the 2 holdout (test) observations

Observations:

GLM ensemble provides lowest ASE

Median is slightly better on the first data point.

Median forecast lower values for second test data point.

Model	Test ASE
AR(2)	0.2069
VAR	0.8537
MLP	0.8493
Median	0.4128
Mean	0.1937
GLM	0.1859

	Base Model Forecasts		
GDP Change	AR(2)	VAR	MLP
3.9	4.08603	3.766524	4.852671
3.8	4.415752	2.500141	2.910566

	Ensemble Model Forecasts		
GDP Change	Median	Mean	GLM
3.9	4.08603	4.235075	4.280197
3.8	2.910566	3.275486	3.323373

Conclusion

- GDP data was very noisy and overall, models are not able to capture variance in this data.
- The univariate model AR(2) performs better than VAR and MLP models.
- Ensembles appear to improve forecasts, but further analysis should be performed.
- Addition of other exogenous variables with even stronger cross correlations may improve performance of multivariate models.

Reproducible Research: Code for the complete analysis is available on GitHub

Youtube Video:

References:

- Jim Chappelow, *Recession*, Investopedia. Accessed March 6, 2020. https://www.investopedia.com/terms/r/recession.asp
- U.S. Bureau of Economic Analysis, Gross Domestic Product [A191RP1Q027SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A191RP1Q027SBEA, March 6, 2020.
- U.S. Bureau of Labor Statistics, All Employees, Total Nonfarm [PAYEMS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PAYEMS, March 6, 2020.
- Board of Governors of the Federal Reserve System (US), 10-Year Treasury Constant Maturity Rate [DGS10], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DGS10, March 6, 2020.
- Board of Governors of the Federal Reserve System (US), Effective Federal Funds Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDFUNDS, March 6, 2020.
- U.S. Bureau of Economic Analysis, Real Disposable Personal Income [A067RO1Q156NBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A067RO1Q156NBEA, March 6, 2020.
- U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCNS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCNS, March 6, 2020.
- U.S. Bureau of Economic Analysis, Population [POPTHM], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/POPTHM, March 6, 2020.
- U.S. Bureau of Economic Analysis, Corporate Profits After Tax (without IVA and CCAdj) [CP], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CP, March 6, 2020.
- Federal Reserve Bank of St. Louis, Spot Crude Oil Price: West Texas Intermediate (WTI) [WTISPLC], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/WTISPLC, March 6, 2020.
- ICE Benchmark Administration Limited (IBA), Gold Fixing Price 10:30 A.M. (London time) in London Bullion Market, based in U.S. Dollars [GOLDAMGBD228NLBM], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GOLDAMGBD228NLBM, March 6, 2020.
- Board of Governors of the Federal Reserve System (US), Japan / U.S. Foreign Exchange Rate [EXJPUS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/EXJPUS, March 6, 2020.
- Board of Governors of the Federal Reserve System (US), U.S. / U.K. Foreign Exchange Rate [EXUSUK], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/EXUSUK, March 6, 2020.