Recession Forecasting Using Bayesian Model Averaging

MAE UCLA

Shaoyang Sun

Prof. Patrick Convery, Prof. Randall R. Rojas

June 2, 2020

Abstract

As the COVID-19 striking every aspect of the economy, there are countless workers losing jobs with living expense increasing and restricted international trade. This significant pandemic has impacted the entire economy leading us toward a recession.

When it comes to the recession forecast, in the discussion of this capstone, there are ten economic indicators included to forecast the GDP change, which are stock market index, unemployment, bond market, gold spot price, PMI, CPI, PPI, retail and net international trade, covering most of the macroeconomic features. In this case, I tried OLS linear regression, time series analysis and VAR model, Bayesian model averaging for generalized regression selection.

Compared with the model stability, predicted GDP values in the future, posterior possibilities and other evaluation criteria, the optimal choice is the generalized linear model of Gamma family with BMA method. The posterior possibility is 0.541 which is higher than other models and the prediction is consistent with the reality. In conclusion, when the indexes of SP500, PPI, CPI, and unemployment rate increase, GDP would increase at meantime and the coefficients of one-year bond rate and gold spot price are significantly negative. Therefore, it could be an efficient and favorable model for GDP predicting and recession forecast.

Keywords

Macroeconomy, Bayesian Model Averaging, VAR, Generalized Linear Model

I. Introduction

1. Background

On account of COVID-19 outbreak, there are over five million people in the world who are confirmed infected, including more than 1,600,000 American people. The quarantines and lockdowns are freezing the economy and this widespread damage turns to be an unprecedented challenge. For instance, US retail sales fall 8.7% in March, and there are nearly 36 million unemployed population in American. In this case, we hope to understand the COVID-19's hit on the macroeconomy and also give a look at the US economy recessions in the past 70 years.

2. Task Description and Challenges

The task of the capstone project is to estimate whether this strike would lead to a recession in U.S. economy. Furthermore, select the optimal model from linear regression, vector autoregression model and Bayesian models for recession forecasting in the future.

The major challenges of the task are variables selection and model evaluation. For the reason that there are plenty of macroeconomics features and which to select is a critical problem. In order to not leaving out important information and avoid multicollinearity problem, I gathered ten independent variables for prediction, which covers equity market, bond market, international trade and also demand and supply indexes.

II. Data Description

The dataset includes the quarterly macroeconomic features from the first quarter of 1950 to the first quarter of 2020. Because some features started being recorded after 1950, the dataset is imbalance panel data.

The definition of variables is explained in the chart as follows. What is worth mentioning is that, apart from GDP, all the independent variables are monthly data. In order to match the quarterly GDP data, I take the arithmetic mean of every three months as the quarterly data.

Table 1 Variables Definition

Veriables	Definition
GDP(y)	Gross Domestic Product (USD, Billions)
Recession	1: decrease of successive two quarter; 0:not recession
SP500	US stock market index
PMI	Purchasing Managers' Index(Manufacturing)
Unemployment	Seasonally adjusted unemployment rate
10YearT-noteFuture	10-Year Treasury Note Futures(%)
1-MonthBondYield	United States 1-Month Bond Yield(%)
PPI	Producer Price Index: All Commodities(Index 1982=100)
CPI	Consumer Price Index: All Commodities(Index 1982-1984=100)
XAU/USD	Spot price of gold (US Dollar)
Retail	Total Retail Trade, Index 2015=100, Monthly, Seasonally Adjusted
TradeBalance	Net trade Value (goods),Not Seasonally Adjusted (USD, Billions)

Table 1 shows the definition of variables and Table 2 is the description of the dataset. We could learn about the distribution of the data which is not normal distributed from skewness and kurtosis. Secondly, the scope of indexes and rates are in different scopes and trends.

Table 2 Data description

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GDP	1	281	6512.480	6358.948	4237.000	5703.252	5472.277	281.20	21729.120	21447.92	0.806	-0.658	379.343
Recession	2	281	0.028	0.167	0.000	0.000	0.000	0.00	1.000	1.00	5.640	29.919	0.010
SP500	3	281	611.539	741.793	180.490	471.423	219.074	17.19	3136.440	3119.25	1.399	1.192	44.252
PMI	4	281	53.139	6.890	53.200	53.347	5.733	32.23	73.800	41.57	-0.233	0.635	0.411
Unemploy	5	281	5.751	1.645	5.533	5.627	1.680	2.57	10.667	8.10	0.638	0.055	0.098
10Year Tnote	6	233	6.089	2.927	5.850	5.842	2.763	1.11	15.307	14.19	0.700	0.240	0.192
1Month bill	7	281	0.337	0.250	0.313	0.315	0.247	0.00	1.253	1.25	0.847	0.749	0.015
PPI	8	281	99.185	59.979	102.200	95.135	94.639	26.03	208.200	182.17	0.314	-1.213	3.578
CPI	9	281	162.553	125.604	94.300	150.579	76.650	38.73	424.600	385.87	0.629	-1.259	7.493
XAU/USD	10	208	542.937	449.168	383.492	489.090	225.689	35.30	1703.433	1668.13	1.116	0.021	31.144
Retail	11	241	68.824	21.572	64.678	68.515	29.299	33.68	107.858	74.18	0.167	-1.285	1.390
Trade Balance	12	241	-2.303	2.618	-1.091	-2.002	1.661	-7.93	0.072	8.01	-0.781	-1.017	0.169

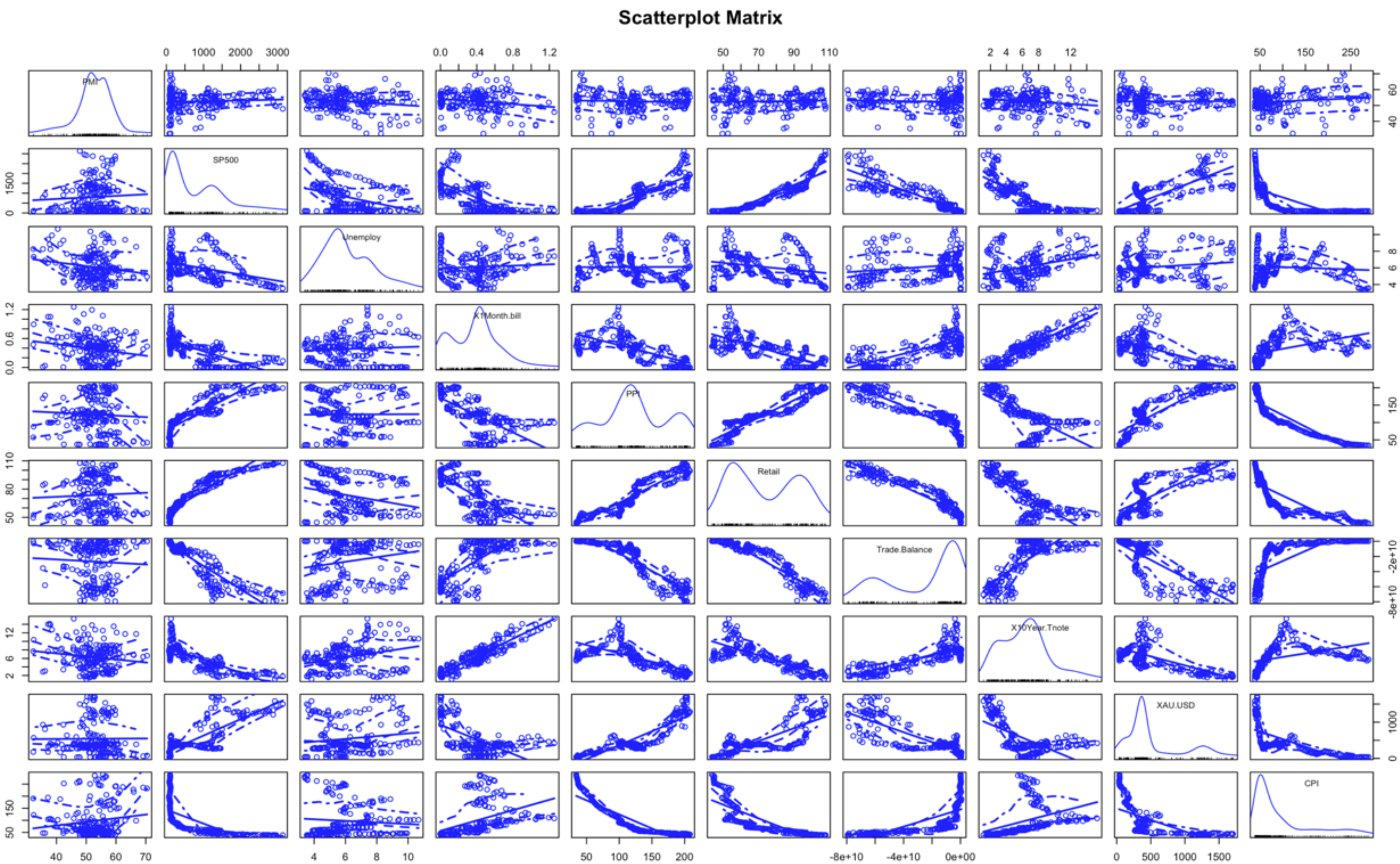

Figure 1 Scatterplot matrix

The Figure 1 visualizes the distribution of 10 independent variables and Figure 2 shows the specific value of correlation, in order to check the collinearity. In Figure 2, the darker the color is, the higher correlation is. As a result, most of the independent variables suffer the multicollinearity problem except for PMI and unemployment rate, like the spot price of gold would move to the same direction with stock market, PPI, retail sales and the opposite direction with bond market, CPI, and trade balance.

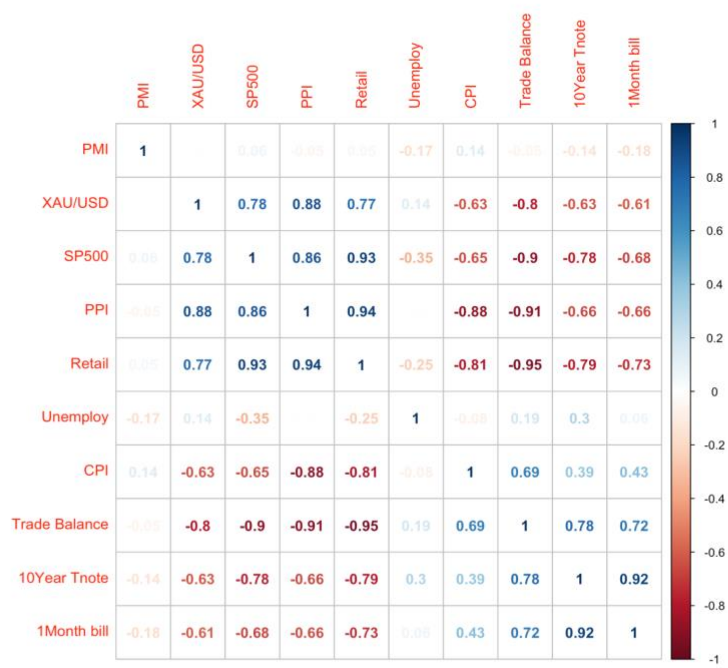

Figure 2 Correlation matrix

Figure 3 shows the importance of independent variables using random forest algorithm, implying that the stock price, SP500, would take priority to consider when it comes to economy prediction. The specific rank of importance is calculated by sequence of the split in each node, and the value of importance is calculated by Gini index change or impurity decline of each random forest tree formed by this feature. Also, the variables importance is similar so I would take them all in account to the regression as follows.

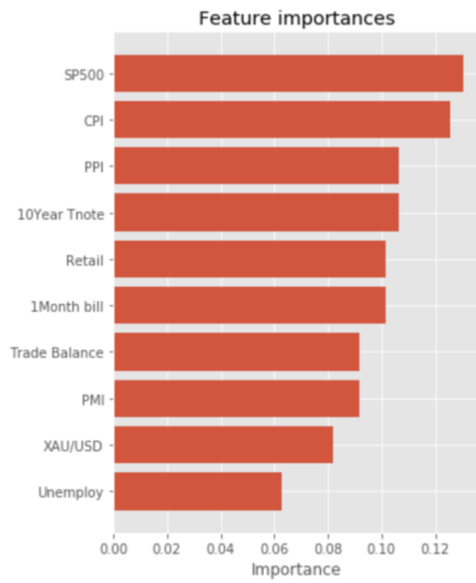

Figure 3 Feature importance

III. Linear Regression

Before running the linear regression, use the method of Mallow's CP to get the subset regressions for model selection. From Figure 4 we could find that the model with seven independent variables has the lowest mean squared prediction error and the variables are PMI, SP500, 1-month bond rate, PPI, retail trade, international trade balance and CPI according to the order in Mallows CP showed below.

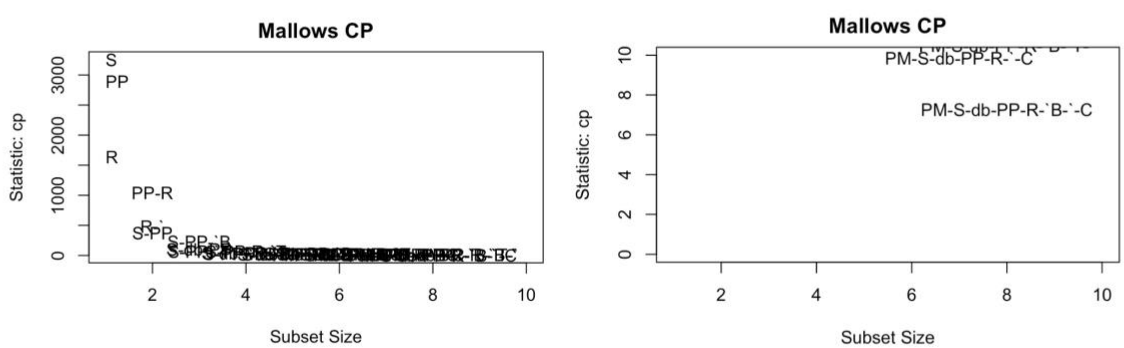

Figure 4 Model selection

In order to build the reliable model, use 5-fold cross validation method. The mean squares of predicted errors are similar in each fold, so the stability of the model is confirmed.

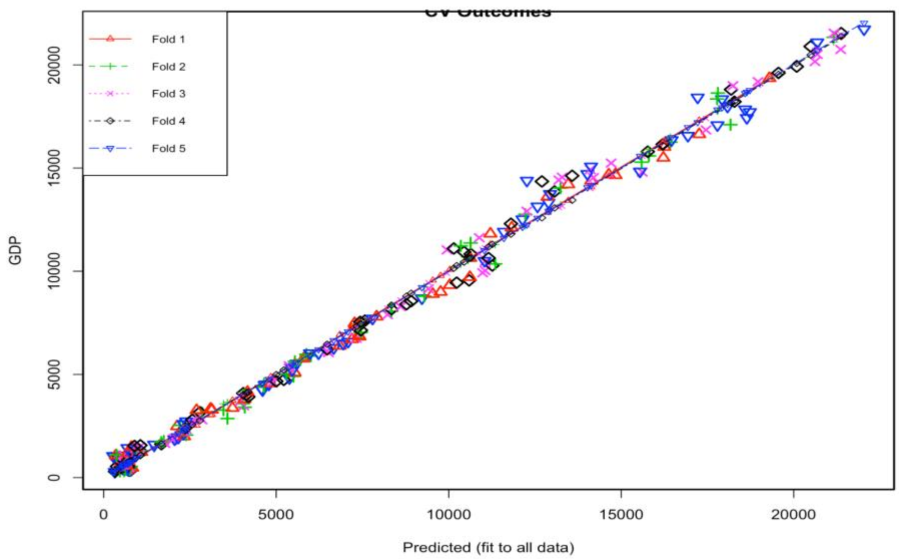

Figure 5 Cross validation outcome

Table 3 Regression table below shows the optimal linear regression model. When PMI index, 1-month bill rate and net trade values increase, GDP would decrease and GDP would increase with the increase of stock price, CPI, PPI indexes, and retail trade. The model explains 99.6% of the variance of GDP and the p values of the coefficients are nearly zero implying very strong statistical significance.

Table 3 Regression table

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-4623.0	965.297	-4.79	0.000
PMI	-17.5	4.793	-3.65	0.000
SP500	2.6	0.135	19.26	0.000
`1Month bill`	-1675.3	183.329	-9.14	0.000
CPI	7.4	1.623	4.56	0.000
PPI	58.6	2.780	21.06	0.000
Retail	61.5	10.506	5.85	0.000
`Trade Balance`	-84.6	50.621	-1.67	0.096

Figure 6 below is the GDP prediction based on OLS model and it shows that there would be decrease in the next quarter, which predicted the reality correctly. The x-axis is the time index. For the reason that in the first quarter of 2020 there are signal of recession which are increasing unemployment rate, CPI and PPI and stock index decreasing, the prediction for next quarter would be GDP decrease based on the latest monthly data of April, 2020.

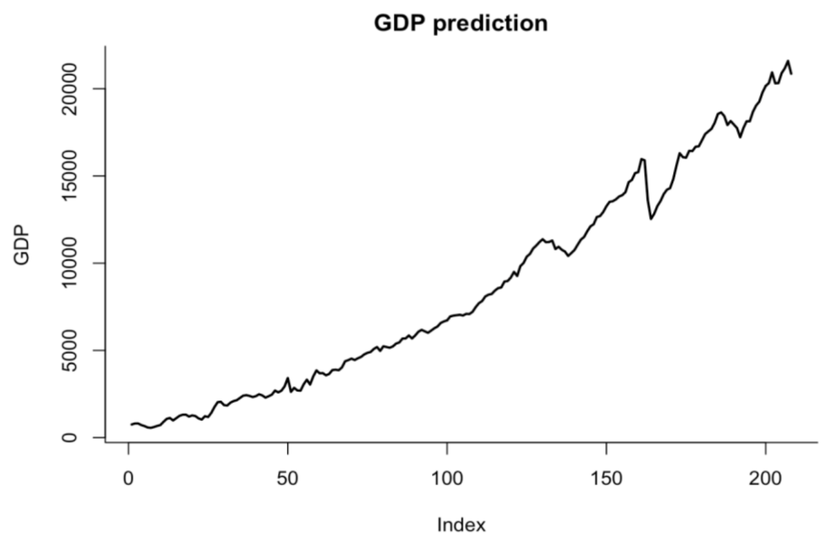

Figure 6 Prediction

Table 4 VIF test illustrates the VIF values of variables, most of which are larger than 10. Also, with the finding of high correlation between independent variables, there is multicollinearity problem in the OLS model for sure. With this problem, the standard error of the coefficients and model would increase and it's hard to differentiate the individual effects of each explanatory variable. Also, the significance test of parameters loses efficiency. Then the efficiency and reliability of the whole estimation would decrease significantly.

Table 4 VIF test

	VIF
PMI	1.17
SP500	13.80
`1Month bill`	2.90
CPI	34.33
PPI	33.14
Retail	67.42
`Trade Balance`	23.06

IV. Vector Autoregression Model

Vector autoregression (VAR), as the generalization of autoregression model, is the stochastic process model which could be used to capture the linear interdependencies among multiple time series. Under this case, we could predict GDP based on lagged values of 6 independent variables with the consideration of model dimensions.

To begin with, we should check the stability of all the time series before building up the VAR model. However, it could be found that the obvious trends and seasonality from the plot of logged data. Therefore, take the first difference of the variables in Figure 7.

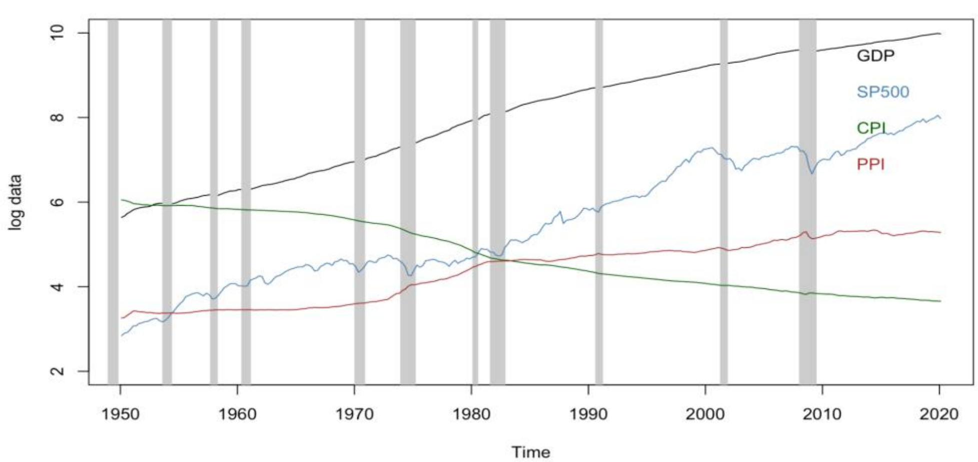

Figure 7 Time series data

In addition, test the stationarity of all the variables using the unit root test with AIC rule. If the series could reject the null hypothesis of unit root test under 5% confidence level, use granger test to get the granger causality of dependent variables. As a result, I take the first difference for GDP, SP500, CPI, PPI, 1-month bond rate and then the data as Figure 8 showed is stable enough for VAR model.

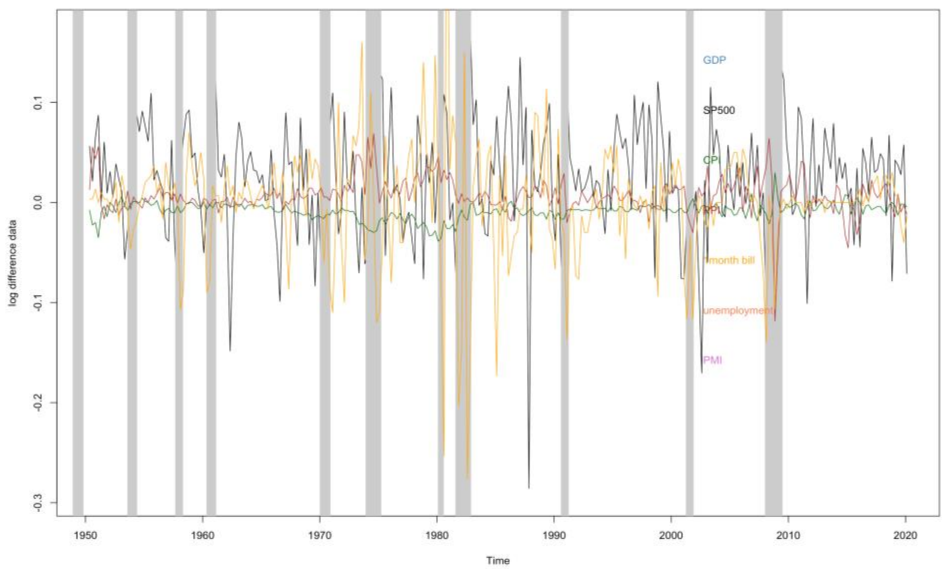

Figure 8 Stable series

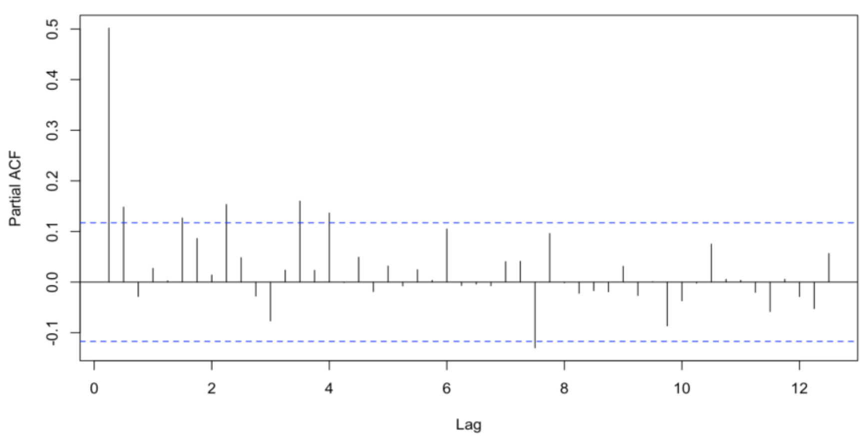

Figure 9 GDP Pacf

The partial autocorrelation function Figure 9 above illustrates that lag four is the optimal choice for AR model. And the choice is confirmed by the VAR selection function based on AIC and FPE values which are Akaike rule and final prediction error. Therefore, select the lag four for VAR model. The result shows that, the first lag of SP500 and PMI would significantly increase GDP, and the second lag of 1-month bond rate and PMI, the fourth lag of unemployment rate would significantly decrease GDP. The model explains 42% of the variance of the GDP. The coefficients chart is showed in the Table 5.

Table 5 Coefficients of VAR model

	Estimate	Std. Error	t value	Pr(>|t|)
dgdp.l1	0.074	0.085	0.866	0.387
dsp.l1	0.023	0.009	2.630	0.009
dbill.l1	-0.004	0.010	-0.389	0.698
dcpi.l1	-0.217	0.150	-1.449	0.149
dppi.l1	-0.053	0.053	-1.003	0.317
unempts.l1	0.006	0.013	0.457	0.648
pmits.l1	0.043	0.009	4.532	0.000
dgdp.l2	0.128	0.085	1.509	0.133
dsp.l2	0.008	0.009	0.849	0.397
dbill.l2	-0.017	0.010	-1.724	0.086
dcpi.l2	-0.095	0.149	-0.638	0.524
dppi.l2	-0.014	0.054	-0.262	0.793
unempts.l2	0.006	0.022	0.258	0.797
pmits.l2	-0.022	0.012	-1.900	0.059
dgdp.l3	0.044	0.081	0.543	0.587
dsp.l3	0.005	0.009	0.607	0.544
dbill.l3	0.004	0.010	0.431	0.667
dcpi.l3	-0.209	0.147	-1.425	0.155
dppi.l3	-0.072	0.054	-1.326	0.186
unempts.l3	0.019	0.022	0.890	0.374
pmits.l3	0.017	0.012	1.393	0.165
dgdp.l4	0.083	0.075	1.112	0.267
dsp.l4	0.004	0.009	0.472	0.637
dbill.l4	-0.012	0.010	-1.230	0.220
dcpi.l4	-0.104	0.155	-0.674	0.501
dppi.l4	-0.004	0.054	-0.068	0.946
unempts.l4	-0.030	0.013	-2.283	0.023
pmits.l4	-0.003	0.009	-0.383	0.702
const	-0.130	0.040	-3.246	0.001

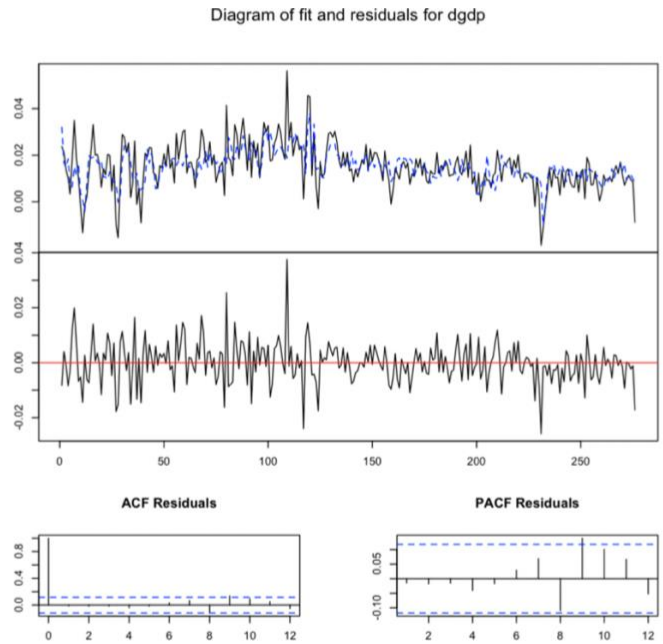

Figure 10 Diagram of fit and residuals for GDP

The Figure 10 above shows the outcome of the VAR model. In the top graphs, the x-axis is the time series and y-axis is the numerical value of differenced log GDP data and residuals of the fit. The fitted value, in blue dashed line, is very close to original data and residuals are in the shape of white noise. Moreover, the acf and pacf plots of residuals also illustrate that the VAR model is appropriate to explain the existing data.

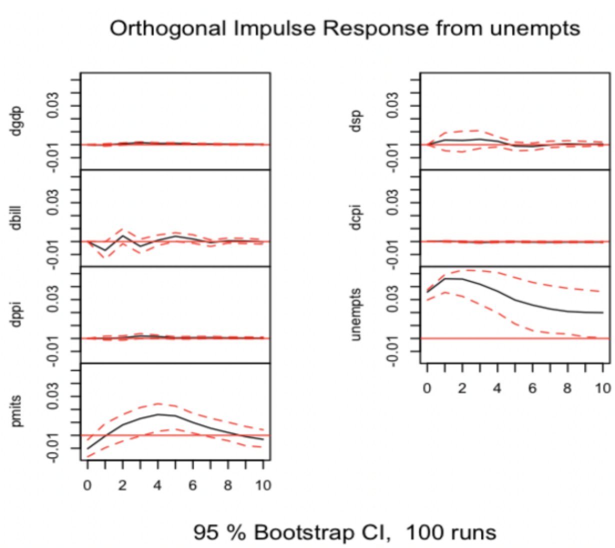

Figure 11 Impulse Response test

Also, I check the stability of the model by impulse test showed above in Figure 11. It shows the reactions of all variables when there's a unemployment rate decrease and GDP would increase after four quarters.

Table 6 Prediction and confidence interval

fcst	lower	upper
0.009	-0.006	0.025
0.005	-0.012	0.022
0.011	-0.007	0.029
0.010	-0.009	0.029
0.011	-0.008	0.030
0.011	-0.008	0.030
0.012	-0.007	0.031
0.012	-0.007	0.031
0.012	-0.007	0.032
0.013	-0.007	0.032

The Also, I check the stability of the model by impulse test showed above in Figure 11. It shows the reactions of all variables when there's a unemployment rate decrease and GDP would increase after four quarters.

Table 6 above shows the forecasted values of differenced log GDP in the next 10 quarters and the 95% confidence interval. It could be found that all the confidence intervals cross zero. Therefore, although the values are positive, the forecast of VAR model is unsatisfied.

V. Bayesian Model Averaging

1. Bayesian theory

Bayesian Model Averaging is a crucial method when facing considerable model uncertainty, variables selection and goal of prediction. It is kind of unsupervised machine learning model and parametric estimation of frequentist method. The fundamental idea is to find the parameters (θ) of the underlying probability function p(θ|yn) from data (y).[1]

[1] Clyde, M. (1999). Bayesian model averaging and model search strategies.

$$p(\theta|y^n) = \frac{p(y^n|\theta)p(\theta)}{\int p(y^n|\theta)p(\theta)d\theta} = \frac{\mathcal{L}(\theta)p(\theta)}{\int \mathcal{L}(\theta)p(\theta)d\theta}$$

The function above shows the basic Bayesian estimation theory and we only need to consider $\mathcal{L}(\theta)p(\theta)$ because the denominator is the normalizing constant.

2. BMA Method

In Bayesian model averaging (BMA), we calculate a weighted combination of the posterior distributions from the different models based on Bayesian theorem, instead of one single model. Specifically, it uses the cross validation to determine the optimal weights and the procedure is as follows.[2] Starting with noninformative priors pj(θj) and take subsets from data (y) which is the training set used to update the priors by Bayesian theorem. When it comes to choosing competing models, the criterion used here is BIC[3] or the Schwarz criterion that choosing the model maximizes:

$$log\mathcal{L}_j(\hat{\theta}_j) - (d_j/2)logn$$

In conclusion, using Bayesian theorem we could get several candidate models and instead of choosing one BMA method averages over the models and provide precise prediction.[4]

3. Data Preparation

Separate the dataset and use 80% of the original data as training set and 20% as testing set so as to evaluate the model. Also, the BMA method is one of the machine learning models, so we should normalize all the variables. In the case of generalized linear model, transform the data into normal distribution in the Gaussian family, and transform the data into percentage form in the Gamma family.

4. Results

4.1 BMA Generalized Linear Model

4.1.1 Gaussian Family

For the reason that the dependent variable, GDP is continuous and univariate, try the method of gaussian family first.

2 Wasserman, L. (2000). Bayesian model selection and model averaging.
3 Claeskens, G., & Hjort, N. L. (2008). Model selection and model averaging.
4 Ching, J., & Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging.

The Table 7 below shows the output of the BMA model with training data, and there are six choices from Bayesian methods. We choose the model1 with six independent variables, where PMI and 1-month bond rate have negative effect on GDP and SP500, PPI, CPI and retail have positive effect on GDP. The optimal model is selected based on BIC rule with lowest BIC value and largest posterior probability.

Table 7 BMA selection①

	p!=0	EV	SD	model 1	model 2	model 3	model 4	model 5
Intercept	100	-0.000111	0.00541	-4.82e-04	6.78e-04	8.22e-04	-1.19e-03	-5.22e-04
SP500.x	100.0	0.341319	0.02736	3.24e-01	3.70e-01	3.67e-01	3.26e-01	3.39e-01
PMI.x	46.0	-0.007319	0.00898	-1.67e-02	.	.	.	.
Unemploy.x	0.0	0.000000	0.00000	.	.	.	.	.
X10Year.Tnote.x	36.1	-0.025348	0.03613	.	-6.77e-02	-6.87e-02	.	.
X1Month.bill.x	97.7	-0.062026	0.01925	-7.22e-02	-4.72e-02	-4.50e-02	-6.79e-02	-7.37e-02
PPI.x	100.0	0.506621	0.05718	5.50e-01	4.66e-01	4.49e-01	5.56e-01	5.19e-01
CPI.x	63.1	0.070897	0.05912	1.24e-01	.	.	1.17e-01	8.56e-02
XAU.USD.x	17.0	0.006141	0.01649	.	.	2.26e-02	.	.
Retail.x	63.9	0.126050	0.09997	2.07e-01	.	.	1.97e-01	1.40e-01
Trade.Balance.x	45.6	-0.047422	0.05495	.	-1.12e-01	-1.15e-01	.	-5.28e-02
nVar				6	5	6	5	6
BIC				-6.56e+02	-6.54e+02	-6.53e+02	-6.53e+02	-6.52e+02
post prob				0.414	0.175	0.086	0.085	0.056

Then evaluate it by using the same model with testing data. As a result, the BIC increases from -656 to -93.6 and posteriori probability decreases from 0.414 to 0.188. The optimal choice also changes with different quantity of independent variables.

In the Figure 12 below,[5] it shows the model choices with orange is the color to use when the variable estimate is positive, blue is the color to use when the variable estimate is negative, and white is the color to use when the variable is not included in the model.[6] In the validation part, the optimal is

5 McAllester, D. A. (1999, July). PAC-Bayesian model averaging.

6 Fletcher, D., & Turek, D. (2012). Model-averaged profile likelihood intervals.

the model1, which is not satisfactory may because the dataset doesn't fit Gaussian distribution.

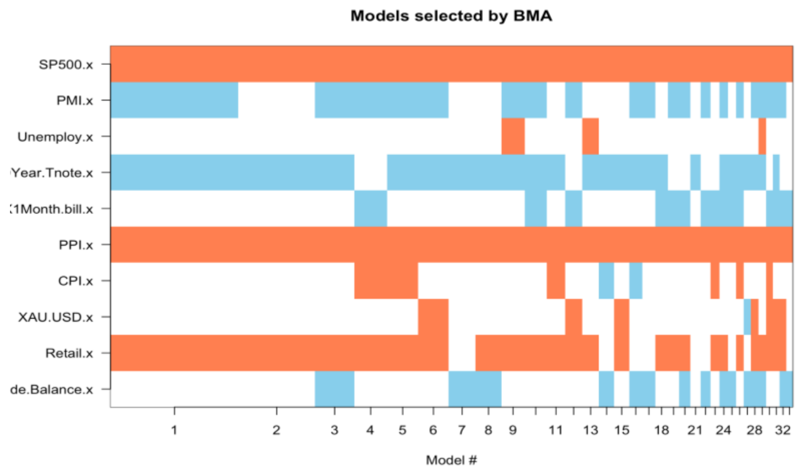

Figure 12Outcome visualization

4.1.2 Gamma Family

Then try another generalized linear model of Gamma family with BMA method. Table 8 below illustrates the BMA outcomes and the posterior probability of the first choice with lowest BIC is about 0.541, which is higher than before. The coefficients of SP500, unemployment rate, PPI and CPI are positive and meantime CPI has the largest impact on GDP change. And when 10-year treasury bond rate and gold spot price increase one unit, GDP would decrease 7.15 and 2.95 respectively.

Table 8 BMA selection②

	p!=0	EV	SD	model 1	model 2	model 3	model 4	model 5
Intercept	100	-4.74200	0.8860	-4.761	-5.599	-3.951	-4.781	-4.366
SP500.x	95.5	2.13497	0.8997	2.469	2.070	0.880	2.615	2.504
PMI.x	4.2	-0.00101	0.0767	.	.	.	.	.
Unemploy.x	83.8	1.03514	0.6120	1.163	1.416	.	1.492	1.230
X10Year.Tnote.x	100.0	-7.11435	1.0612	-7.155	-6.502	-7.061	-8.310	-7.055
X1Month.bill.x	7.7	0.07775	0.3684	.	.	.	1.005	.
PPI.x	100.0	6.33342	1.1285	6.564	5.548	6.360	6.575	6.020
CPI.x	100.0	62.38525	2.9024	62.520	63.460	60.830	62.781	62.370
XAU.USD.x	100.0	-2.71005	0.7608	-2.947	-2.586	-1.883	-3.138	-2.802
Retail.x	17.1	0.30982	0.8194	.	1.840	.	.	.
Trade.Balance.x	5.1	-0.01417	0.1177	.	.	.	.	-0.279
nVar				6	7	5	7	7
BIC				-529.515	-526.605	-526.463	-525.623	-524.783
post prob				0.541	0.126	0.118	0.077	0.051

Then use the testing data for model evaluation. The validation outcome of Gamma model is better than before, with similar optimal choice and there are six independent variables, the same as the training set did. The Figure 13 below is the validation output.

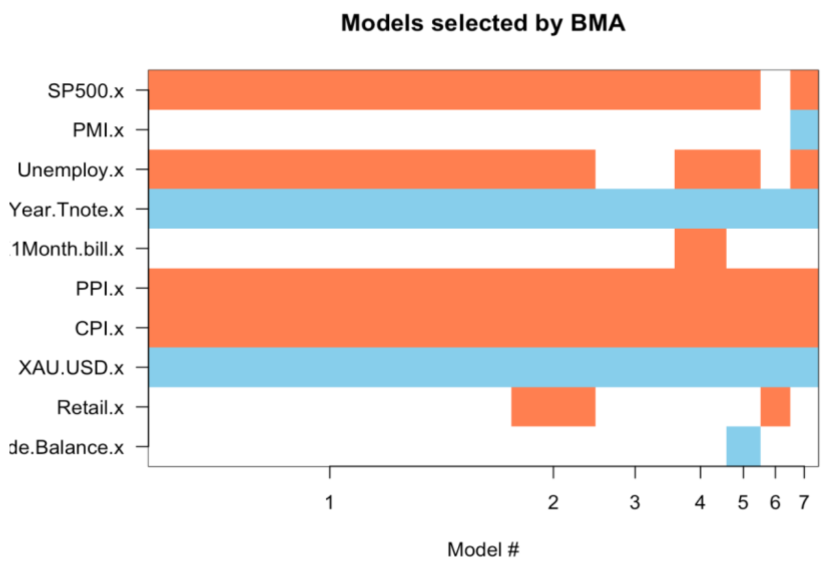

Figure 13 Outcome visualization

Therefore, we could use the model for forecast. The Figure 14 below shows the prediction of GDP using generalized linear model with Gamma family. The black line shows the original normalized data and the red line shows the fitted values. The red line after 2020 is the prediction of next 4 quarter and blue lines are the sketch of 95% confidence interval. It could be found that there would be an obvious economy recession in 2020, which fits the reality correctly.

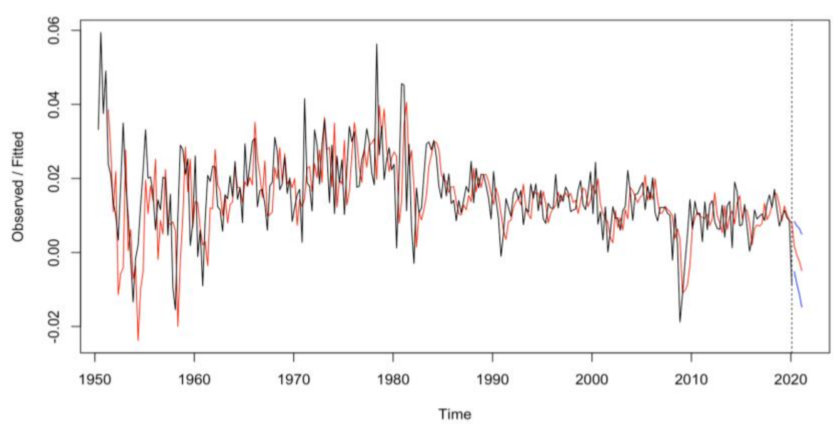

Figure 14 GDP Forecast

4.2 BMA Linear Regression

The main idea of this method is to select the most efficient model with Bayesian model averaging and the linear regression is a simple way to fit the data.

Table 9 BMA selection ③

	p!=0	EV	SD	model 1	model 2	model 3	model 4	model 5
Intercept	100.0	-5065.802	1332.895	-5496.189	-5136.318	-4190.951	-3555.782	-5961.465
SP500	100.0	2.513	0.172	2.516	2.432	2.598	2.521	2.563
PMI	100.0	-26.689	6.505	-25.959	-28.084	-25.900	-28.230	-26.494
Unemploy	5.9	2.203	12.911	.	.	.	.	37.101
X10Year.Tnote	3.8	0.107	10.017	.	.	.	.	.
X1Month.bill	100.0	-1753.259	250.418	-1737.237	-1719.463	-1831.689	-1829.396	-1644.790
PPI	100.0	58.797	7.132	62.942	53.793	59.721	49.055	60.826
CPI	97.2	9.541	2.686	10.736	8.825	8.696	6.218	10.972
XAU.USD	32.1	0.181	0.313	.	0.520	.	0.573	.
Retail	100.0	73.383	13.457	72.371	83.800	59.048	69.198	78.238
Trade.Balance	18.3	-20.800	55.217	.	.	-88.451	-104.635	.
nVar				6	7	7	8	7
r2				0.994	0.995	0.994	0.995	0.994
BIC				-830.928	-829.431	-827.614	-826.774	-826.712
post prob				0.489	0.231	0.093	0.061	0.059

Table 9 BMA selection ③above illustrates the outcomes of BMA method applied in linear regression. It could be found that the model explains more than 99% of the GDP variance and the posterior probability is 0.489, a little lower than generalized linear model. The coefficients of the first choice are SP500, PMI, 1-month bond rate, PPI, CPI, retail, showed in the plot below, which are exactly the linear model of the first part using Mallow's CP method before. Therefore, the evaluation and forecast would be the same as before.

VI. Conclusion and Future Work

According to the analysis of three types of method, it could be found that the most efficient model is the generalized linear model of Gamma family with BMA method. Firstly, the OLS linear regression suffers from severe multicollinearity problem. Even though the model could explain nearly the total variation of GDP change, the outcome and prediction are suspicious based on the variables selected. Secondly, using the technique of time series analysis, we could find the granger causality between GDP and the SP500, PMI, unemployment rate and 1-month bond rate and their lagged terms after making sure all time series are stable. However, the model is less stable than before with even worse prediction because the confidence interval of forecast is too large to determine the GDP change in the next few quarters.

Thirdly, try the Bayesian model averaging for generalized regression selections with normalized data because using the machine learning methods. After comparing with the model stability, BIC, posterior possibilities, and the contrast of training set and testing set, select the generalized linear model of Gamma family as the final choice. The posterior possibility is 0.541 and the prediction is consistent with the reality. The difference between training score and testing score is the least compared with the other two models. In conclusion, the indexes of SP500, PPI, CPI, and unemployment rate, one-year bond rate and gold spot price could significantly explain the GDP change. In conclusion, using Bayesian model averaging method, we could build an efficient and favorable model for GDP predicting and recession forecast.

While the data with 281 rows is too limited to use machine learning method, and 10 independent variables are still not enough for macroeconomy forecast. In the future, there should be larger scale data and more comprehensive independent variables for the more precise prediction. Secondly, the forecasts in the discussion before are all in forms of nominal GDP values instead of the status of recession. Next, it may be helpful to try logistic regression and the model evaluation could be more intuitional by confusion matrix. Thirdly, using Monte Carlo method for updating, selecting and Bayesian models averaging in order to build a more efficient and stable model.