MGT 6203 Group Project Final Report

Team #: 89

Team Members:

1. Bill Hariaczyi; gtID 903454213
2. Igal Seagal; gtID 903541024

Abstract

Milton Friedman said “Inflation is always and everywhere, a monetary phenomenon. It's always and everywhere, a result of too much money, of a more rapid increase in the quantity of money than an output.“ The purpose of this paper is to explore the relationship between inflation in the United States and US monetary policy. The paper analyses the theoretical links of inflation (CPI) with M2 money supply, M2 velocity of money, Federal debt, and Federal Funds rate.

Inflation is a primary driver of economic activity impacting spending, investment, and economic growth. Adverse effects of inflation include: The opportunity cost of holding money increases, hoarding of goods for fear of increasing prices, erosion of purchasing power, and generalized wealth destruction

Modern Monetary Theory (MMT) is an economic theory that suggests that the government could simply create more money without consequence. The US monetary policy being enacted in recent history seems to be adhering to a MMT like approach, while not explicitly labeled as such the fact remains that 40% of all US dollars ever printed were created in just the last few years. Supporters of modern monetary theory (MMT) acknowledge that inflation is theoretically a possible outcome from printing of fiat currency and debt creation, they say it is highly unlikely and can be fought with policy decisions in the future if required.

Monetarism is an economic school of thought which states that the supply of money in an economy is the primary driver of economic growth. Monetarism is closely associated with economist Milton Friedman, who argued, based on the quantity theory of money, that the government should keep the money supply steady, expanding it slightly each year to allow for the natural growth of the economy due to the inflationary effects that can be brought about by the excessive expansion of the money supply.

Multiple linear regression and random forest regression models both showed significant CPI predictive power from three factors - M2 money supply, M2 velocity and Federal Funds rate.

Introduction

Is it necessary for governments to adhere to a balanced budget? According to modern monetary theory (MMT) governments in control of a fiat currency can and should print as much money that is needed to fund entitlements, pay for government expenditures, and build infrastructure with little regard for potential inflationary consequences. MMT states that deficit spending is what builds people’s savings and promotes employment and that the only limit that the government has when it comes to spending is the availability of real resources. MMT suggests that in the unlikely event inflation occurs, it can be controlled by reducing government spending and raising taxes.

Monetarism on the other hand proposes the idea that the total amount of money in circulation in an economy determines the rate of economic growth and that money supply is key to economic stability. Expansionary monetary policy increases the money supply and can result in stimulating the economy along with increasing inflationary pressure, conversely contractionary monetary policy decreases the money supply in the economy, driving down asset prices, and helping to combat inflation. The quantity theory of money is central to the concept of monetarism. The basic equation for the quantity theory is called The Fisher Equation. The fisher equation states the following:

$$(M)(V)=(P)(T)$$

Where M= Money Supply, V = Velocity of money, P= price level, T=Volume of transactions

Our initial hypothesis is that money supply (M) and the velocity of money(V) will be significant variables in predicting inflation(P).

We will primarily use Federal Reserve Economic Data (FRED) to perform our analysis, FRED is an online database consisting of hundreds of thousands of economic data time series from scores of national, international, public, and private sources. FRED is created and maintained by the Research Department at the Federal Reserve Bank of St. Louis.

We will Investigate the impact of M2 money supply on generalized inflation (CPI) to determine if changes in CPI can be predicted from M2 money supply changes. Other inflationary factors to investigate include velocity of money, public debt, and federal funds rate. We will use multiple regression and random forest regression to explore whether any of these factors are significant.

Data

Reports from the FRED database were combined into one csv file (Table 1), CPI and M2 supply is reported monthly while the remaining data sets are reported on a quarterly basis (all datasets were joined by date). The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services. M2 supply is a measure of the money supply that includes cash, checking deposits, and easily-convertible near money. The velocity of money is calculated by dividing a country's gross domestic product by the total supply of money, it is a measurement of the rate at which money is exchanged in an economy. The Federal Funds Rate is the target interest rate set by the FOMC, this is the rate at which commercial banks borrow and lend their excess reserves to each other overnight. Raising the rate makes it more expensive to borrow disincentivizing economic activity and should thus negatively impact inflation.

	Date	CPI	M2_Supply	M2V	Fed_Debt	Fed_Funds_Rate
1	1966-01-01	13.451	462.0	1.713	320999	4.559
2	1966-04-01	13.619	469.3	1.712	316097	4.914
3	1966-07-01	13.691	470.9	1.733	324748	5.411
4	1966-10-01	13.860	475.7	1.744	329319	5.560
5	1967-01-01	13.881	481.6	1.739	330947	4.819
6	1967-04-01	13.965	492.1	1.708	322893	3.992

Table 1 – Joined Dataset

Model Analysis

1. Multiple Linear Regression

Data Exploration

The entire data set will first be explored before splitting into training and test sets for model analysis. Factors were checked for multicollinearity, Federal Debt and M2 money supply show a correlation (Figure1)

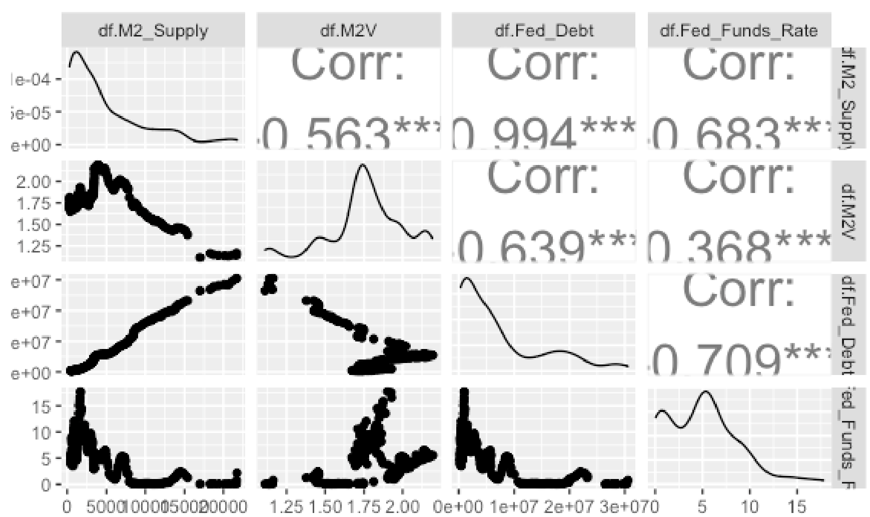

Figure 1 – Correlation Matrix

Further checking via Variance Inflation Factors (VIF) confirm multicollinearity, M2 money supply and fed debt have VIF values greater than five.

```
M2_Supply            M2V       Fed_Debt Fed_Funds_Rate
92.765715       1.787982     102.365082       2.179247
```

Removing federal debt results in remaining explanatory variables showing low correlation.

```
M2_Supply            M2V Fed_Funds_Rate
 2.730648       1.659553       1.926874
```

Multiple regression analysis on the full data set shows a good fit (Adj R2 = 0.95) and all factors are significant as well as the overall model (F-test). Residuals vs Fitted plot does not indicate presence of heteroskedasticity but there is a potential autocorrelation in error terms. If there is a pattern in the error terms then the *estimated* standard errors will underestimate the *true* standard errors. P values along with confidence and prediction intervals will be narrower than they should be resulting in overconfidence in the capability of the model.

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   -8.310e+01  5.142e+00 -16.162  < 2e-16 ***
M2_Supply      7.138e-03  1.484e-04  48.103  < 2e-16 ***
M2V            6.108e+01  2.474e+00  24.688  < 2e-16 ***
Fed_Funds_Rate -6.235e-01  1.611e-01  -3.871 0.000142 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.671 on 222 degrees of freedom
  (525 observations deleted due to missingness)
Multiple R-squared:  0.9546,	Adjusted R-squared:  0.9539
F-statistic:  1554 on 3 and 222 DF,  p-value: < 2.2e-16
```

Figure 2 – Regression Summary for full dataset

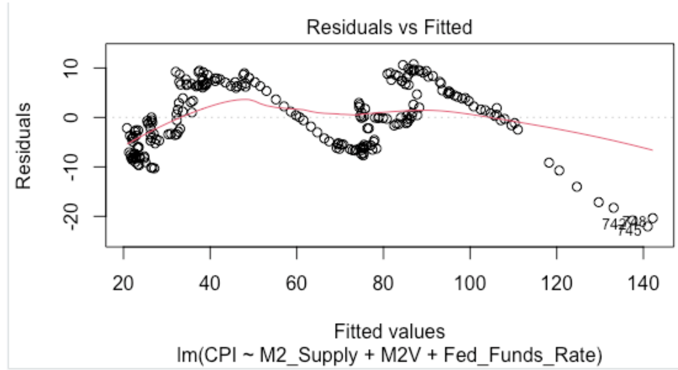

Figure 3 – Residuals vs Fitted for full dataset

The Durbin-Watson test is used to detect autocorrelations in a linear model

```
lag Autocorrelation D-W Statistic p-value
  1       0.9560736    0.03847042       0
Alternative hypothesis: rho != 0
```

The value of autocorrelation can range from -1 to 1, where the range 0 to 1 represents positive autocorrelation and a range of -1 to 0 represents negative autocorrelation. The results show a positive autocorrelation value of 0.956. The DW statistic is 0.0384 and the p-value is significant; DW value < 1.5 indicates positive autocorrelation, whereas DW > 2.5 indicates negative autocorrelation. If the DW value is between 1.5 and 2.5 there is no autocorrelation. The overall results of the the Durbin-Watson test indicate positive serial correlation.

There are a couple of ways to deal with autocorrelation including changing the model to obtain non-autocorrelated errors or expanding the confidence intervals around the regression coefficients to account for the violation of the model assumption of non-autocorrelated errors.

We will apply heteroskedasticity and autocorrelation (HAC) robust standard errors correction, this will result in wider confidence intervals applied to the original model (figure 4)

```
t test of coefficients:

                 Estimate  Std. Error t value  Pr(>|t|)
(Intercept)    -8.3100e+01  2.4133e+01 -3.4434 0.0006865 ***
M2_Supply       7.1381e-03  7.2908e-04  9.7904 < 2.2e-16 ***
M2V             6.1083e+01  1.0854e+01  5.6278 5.475e-08 ***
Fed_Funds_Rate -6.2351e-01  6.6034e-01 -0.9442 0.3460791
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 4 – HAC robust standard errors test

Fed Funds Rate coefficient is no longer statistically significant, we will rerun the regression with M2 supply and M2 velocity and compare to robust standard errors (Figure 5). Coefficients remain significant with robust standard errors applied (Figure 6).

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.672e+01  4.510e+00  -21.45   <2e-16 ***
M2_Supply    7.686e-03  1.056e-04   72.78   <2e-16 ***
M2V          6.506e+01  2.350e+00   27.69   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.793 on 247 degrees of freedom
  (522 observations deleted due to missingness)
Multiple R-squared:  0.9574,	Adjusted R-squared:  0.9571
F-statistic:  2777 on 2 and 247 DF,  p-value: < 2.2e-16
```

Figure 5 –Regression summary for full dataset

```
t test of coefficients:

               Estimate  Std. Error t value  Pr(>|t|)
(Intercept) -9.6725e+01  1.7391e+01 -5.5617 6.928e-08 ***
M2_Supply    7.6862e-03  4.4989e-04 17.0845 < 2.2e-16 ***
M2V          6.5061e+01  8.8228e+00  7.3741 2.496e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 6 – HAC robust standard errors test

Model Fit

Data will be split into 80% training and 20% test. The training set will be fit to two multiple linear regression models; Model 1 will evaluate three significant factors and Model 2 will have Fed Funds Rate removed based on robust standard errors criteria. Tests will be performed to check if residuals have non-linear patterns and are normally distributed, as well as if the standardized residuals are spread equally along the range of fitted values. Cook's distance will be used to identify influential points and the Durbin-Watson test will be used to detect autocorrelations.

Name	Y	X1	X2	X3
LR Model1	CPI	M2_Supply	M2V	Fed_Funds_Rate

```
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -8.034e+01  5.514e+00 -14.570  < 2e-16 ***
M2_Supply      7.255e-03  1.616e-04  44.900  < 2e-16 ***
M2V            5.916e+01  2.667e+00  22.182  < 2e-16 ***
Fed_Funds_Rate -5.345e-01  1.717e-01  -3.113  0.00216 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.444 on 177 degrees of freedom
Multiple R-squared:  0.9562,	Adjusted R-squared:  0.9555
F-statistic:  1288 on 3 and 177 DF,  p-value: < 2.2e-16
```

Figure 7 – Regression summary for training dataset

Name	Y	X1	X2	X3
LR Model2	CPI	M2_Supply	M2V	

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.572e+01  5.362e+00  -15.99   <2e-16 ***
M2_Supply    7.567e-03  1.298e-04   58.28   <2e-16 ***
M2V          5.977e+01  2.724e+00   21.94   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.599 on 178 degrees of freedom
Multiple R-squared:  0.9538,	Adjusted R-squared:  0.9533
F-statistic:  1837 on 2 and 178 DF,  p-value: < 2.2e-16
```

Figure 8 – Regression summary for training dataset

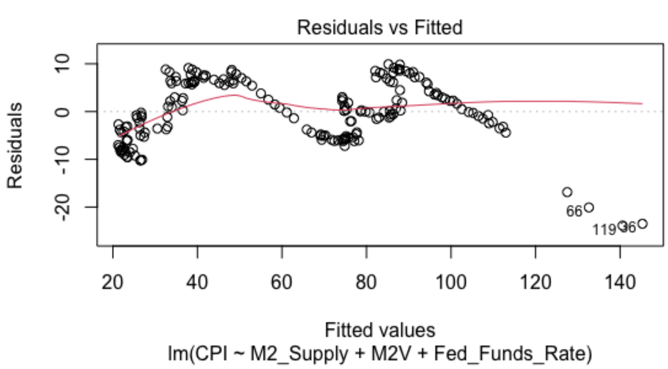

Figure 9 – Residuals vs Fitted for training dataset

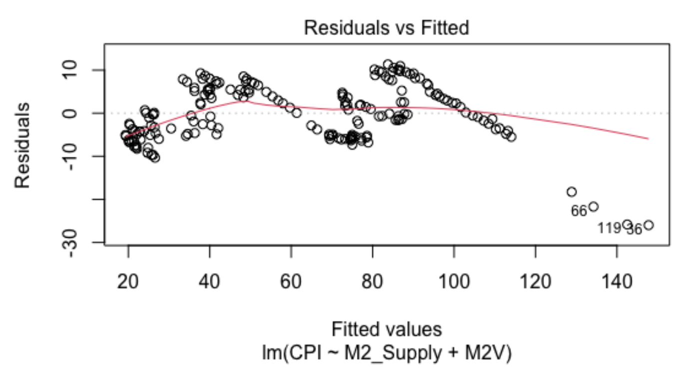

Figure 10 – Residuals vs Fitted for training dataset

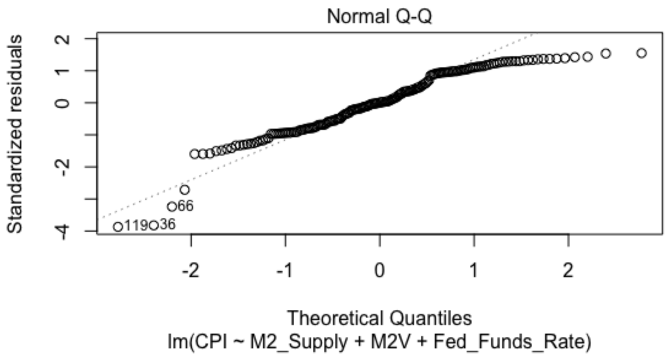

Figure 11 – Q-Q Plot for training dataset

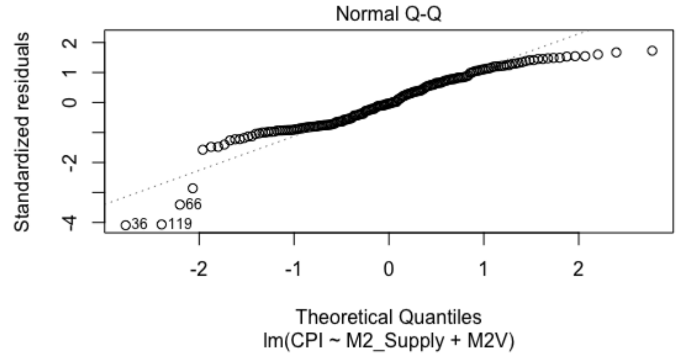

Figure 12 – Q-Q Plot for training dataset

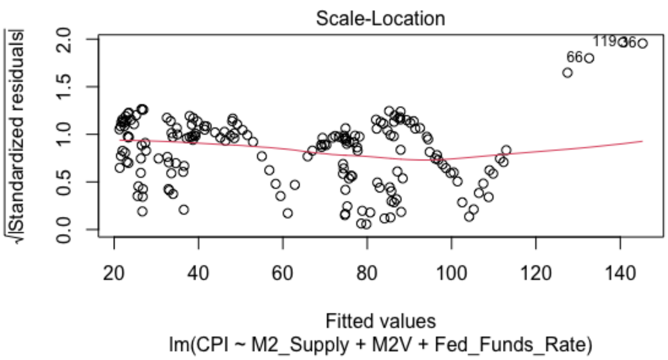

Figure 13 – Std Residuals vs Fitted for training dataset

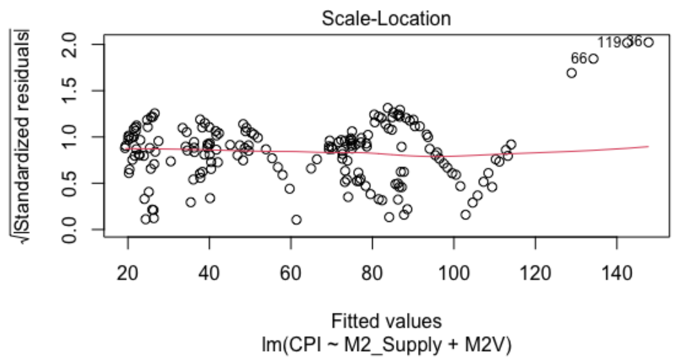

Figure 14 – Std Residuals vs Fitted for training dataset

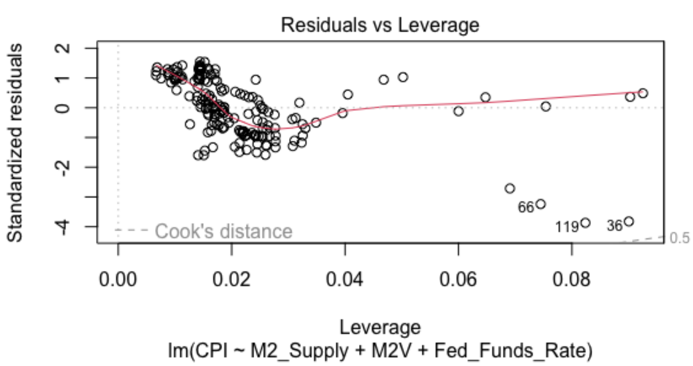

Figure 15 – Cook's Distance for training dataset

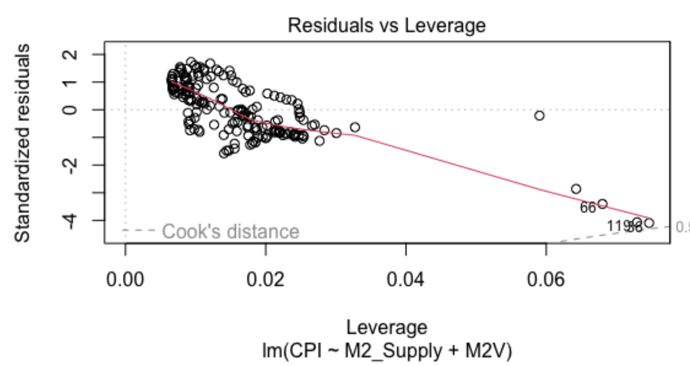

Figure 16 – Cook's Distance for training dataset

```
lag Autocorrelation D-W Statistic p-value
  1     -0.08792102      2.171856   0.226
```

Figure 17 – Durbin-Watson test for training dataset

```
lag Autocorrelation D-W Statistic p-value
  1     -0.09424986      2.184646     0.2
```

Figure 18 – Durbin-Watson test for training dataset

For both models multiple regression analysis shows a good fit (Adj R2 ~ 0.95) and all factors are significant as well as the overall model (F-test). Increasing the money supply and velocity of money results in increased inflation (CPI). As expected, increasing the fed funds rate has a negative effect on inflation. Residuals vs Fitted plot does not indicate presence of heteroskedasticity but there is autocorrelation in the error terms. Unlike the full dataset, the sample training dataset fails the Durbin-Watson test for autocorrelation, presumably this is due to the random sampling of the training dataset. Q-Q plot shows some deviation from normality at the upper and lower tail, there is no linear trend in the Scale-Location plot, and no high leverage points identified with Cook's Distance statistic.

2. Random Forest Regression

The same training and test datasets utilized for linear regression were used to evaluate two random forest regression models. The number of trees (ntree) and the number of variables tried at each split (mtry) will be varied to minimize the RMSE. To determine the optimal mtry the train() function was utilized from the caret library. This function sets up a grid of tuning parameters for a number of classification and regression routines, fits each model and calculates a resampling based performance measure. Optimal mtry was determine by searching along a grid and utilizing 5-fold cross validation. The number of trees was determined by running the model several times and examining the plot showing error vs number of trees.

Name	Y	X1	X2	X3
RF Model1	CPI	M2_Supply	M2V	Fed_Funds_Rate

	mtry	RMSE	Rsquared
1	1	2.837	0.992
2	2	1.484	0.998
3	3	1.092	0.999

Table 2 – Model 1 mtry results

Name	Y	X1	X2	X3
RF Model2	CPI	M2_Supply	M2V	

mtry	RMSE	Rsquared
1	2.249	0.995
2	0.968	0.999

Table 3 – Model 2 mtry results

```
 randomForest(formula = CPI ~ M2_Supply + M2V + Fed_Funds_Rate,
 = TRUE, mtry = 3, na.action = na.omit)
               Type of random forest: regression
                     Number of trees: 125
No. of variables tried at each split: 3

          Mean of squared residuals: 1.06348
                    % Var explained: 99.89
```

Figure 19 – Random Forest summary for training dataset

```
 randomForest(formula = CPI ~ M2_Supply + M2V, dat
 2, na.action = na.omit)
               Type of random forest: regression
                     Number of trees: 100
No. of variables tried at each split: 2

          Mean of squared residuals: 0.9200329
                    % Var explained: 99.9
```

Figure 20 – Random Forest summary for training dataset

	Overall
M2_Supply	75.461049
M2V	10.579857
Fed_Funds_Rate	9.767887

Table 4 – Model 1 Variable Importance

	Overall
M2_Supply	55.144849
M2V	8.688513

Table 5 – Model 2 Variable Importance

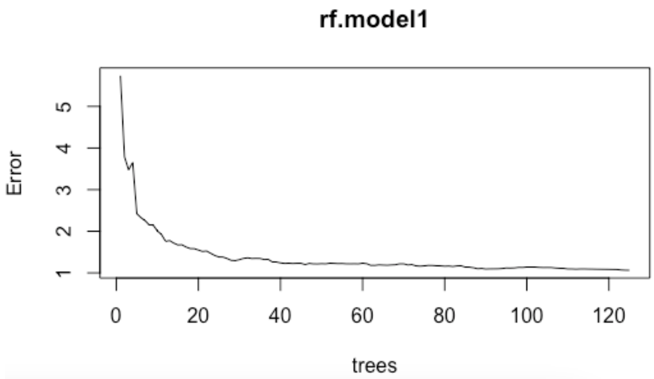

Figure 21 –Error vs #trees for training dataset

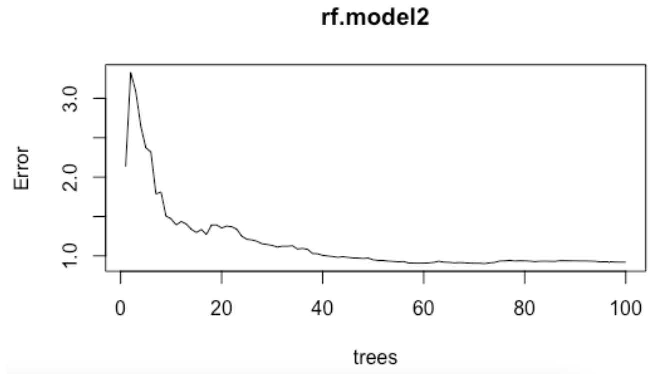

Figure 22 –Error vs #trees for training dataset

For both models the random forest regression analysis shows low mean of squared residuals (~1%) and high percent variance explained (~99%). M2 supply was ranked as the most important variable in both cases as well. RF model 1 and RF Model 2 optimal mtry results were three and two, respectively. The number of trees selected for model 1 (ntree = 125) was slightly higher than model 2 (ntree = 100).

Results

Model predictions were generated utilizing the randomly sampled test dataset (20%). Model RMSE was calculated, and plots were created for each of the four models. The linear regression models (Figure 23, Figure 24) show a strong linear relationship (actual vs theoretical) with a weak sinusoidal pattern superimposed. The sinusoidal pattern is likely related to the autocorrelation present in the dataset, the predicted data points in the peak and trough areas lie outside the 95% confidence intervals. Both of the random forest models (Figure 25, Figure 26) show a better fit compared to the linear regression models.

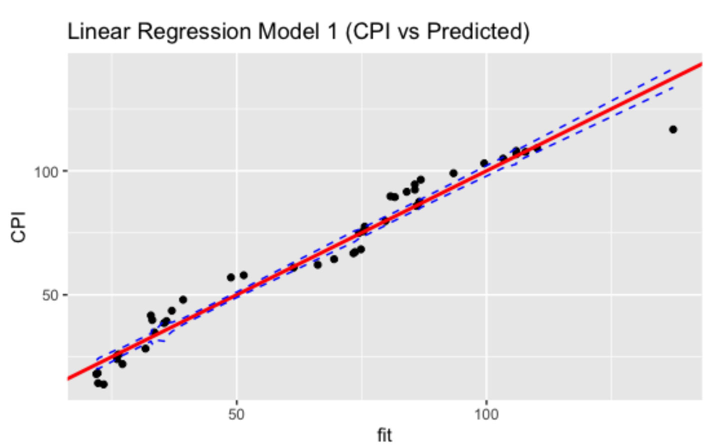

Figure 23 – Linear Regression Model 1 (3 predictors)

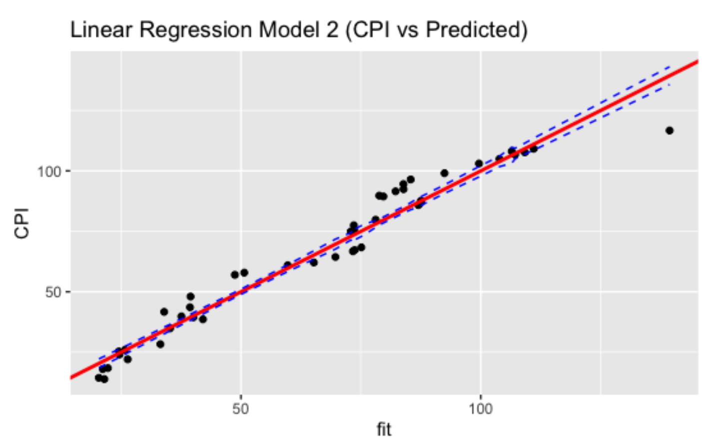

Figure 24 – Linear Regression Model 2 (2 predictors)

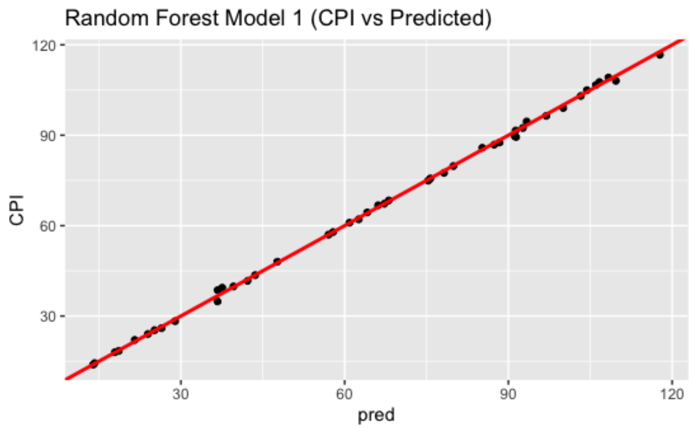

Figure 25 – Random Forest Regression Model 1 (3 predictors)

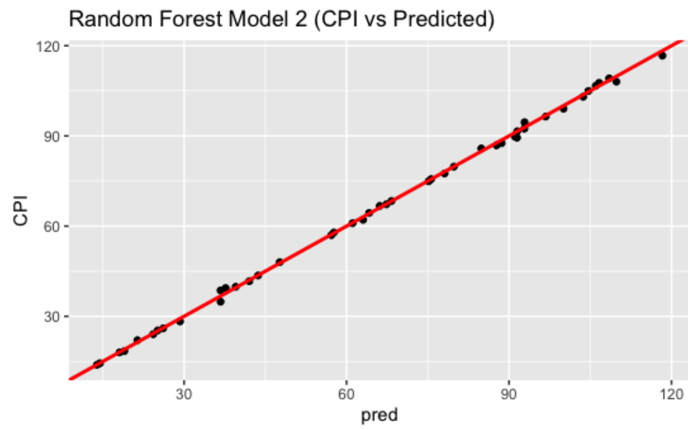

Figure 26 – Random Forest Regression Model 2 (2 predictors)

Model	Model_RMSE	Predictor1	Predictor2	Predictor3
LR Model1	6.088	M2_Supply	M2V	Fed_Funds_Rate
LR Model2	6.389	M2_Supply	M2V	
RF Model1	0.808	M2_Supply	M2V	Fed_Funds_Rate
RF Model2	0.885	M2_Supply	M2V	

Table 5 – Model RMSE

Table 5 summarizes the RMSE results for all four models. The random forest models have significantly lower RMSE compared to the linear regression models. The best performing model (RF Model 1) is the random forest model containing all three predictors.

Conclusion

Multicollinearity, positive autocorrelation, and significant factors have been identified through exploratory analysis. Both linear regression and random forest regression models appear to be good predictors of inflation (CPI) given M2 money supply, M2 velocity of money, and Federal Funds rate. The coefficients in the linear regression models indicate that increasing M2 money supply and M2 velocity will increase inflation while increasing the Federal Funds Rate will decrease inflation. The random forest regression models outperformed the linear regression models, they appear to be more capable of handling autocorrelation. It seems Milton Friedman was right.

Works cited

European Journal of Multidisciplinary Studies

Theories of Money Supply: The Relationship of Money Supply in a Period of Time T-1 and Inflation in Period T- Empirical Evidence from Albania (PhD Cand. Sorina Koti, Tomi Bixho)