Causal Forests in Inflation Studies: Evaluating the Accuracy of Inflation Theories

Daniel R. Hogan
Quinnipiac University
Data Science Capstone, DS-480
Spring 2025

Abstract

After COVID-19, inflation has been a primary focus of many central bankers around the world, due to the lasting economic effects from our response to the pandemic. This has brought many economists to revisit the true impact of traditional inflation theories, as well as new theories. In order to measure the 'effectiveness' of each theory I employed a causal forest to identify the causal relationship of three specific inflation theories, Monetarism, New Keynesian, and the Fiscal Theory of the Price Level. By using a range of heterogeneous and confounding effect controls, I isolate the causal treatment of unemployment, percent change in money in circulation and percent change in government debt over gross domestic product, representing New Keynesian theory, Monetarism and Fiscal theory of the Price Level respectively. By comparing the treatment effect from each model, I gauge the effectiveness of each theory in practice. Unemployment had the largest average treatment effect, at -.14 on inflation. Monetarism was the second most supported inflation theory, as it had a treatment effect of .12. The fiscal theory had no treatment effect. I draw the conclusion that continuing to base inflation theories around the movement of the phillips curve and New Keynesian theory remains the most viable option for both central bankers and private firms.

Introduction

After a record 13 years of continual growth in the developed world, Covid-19 changed the economic trajectory worldwide. Specifically, over the past five years, almost every economy worldwide has been subject to price increases, inflation, due to supply chain issues and increased government spending (Alessandria et al). This has led to a significant decline in consumer sentiment, or the satisfaction of an everyday consumer. Normally, consumer sentiment and economic growth move hand-in-hand, but there has been a discrepancy since the pandemic (Harris, et al). This creates an issue for Central banks and governments, as the sentiment can provide insight into how the economy and inflation may act in the upcoming months. This prompted economists to revisit inflation theories as central banks around the world used a wide range of methods to try and ease inflation with various results.

A foundational aspect of economics revolves around understanding inflation and throughout the past 125 years, the field has developed numerous theories regarding the causation of inflation. In the early 1920's John Maynard Keynes developed the foundations for the theory that inflation and unemployment rates moved in opposite directions of each other (Phillips, 283). Today, many economists subscribe to the 'New-Keynesian' school of thought, built off the ideas of Keynes. 40 years later, Milton Friedman challenged this idea, proving that printing money is the primary driver of inflation. At the time, it was hard to disprove, but over the decades, the causal relationship between inflation and M2, or money in circulation, has decayed. Finally, a relatively new theory is the Fiscal Theory of the Price Level (FTPL). Developed by John Cochrane, this theory states that inflation is driven by the government's ability to pay off their debts and consumer expectations of how long the status quo will remain (Binder and Kamdar,

1301). However, due to the natural complexities of the economy, these theoretical models often only hold in certain economic conditions, or in the case of the FTPL, are impossible to implement in the real world due to the need for future data.

A causal forest is a new technique that uses the same tree structure as a random forest, but the causal forest focuses on extracting the causal relationship between the treatment, T, and the dependent variable, Y. Additionally, there are controls for both heterogeneity and exogenous effects, allowing the user to isolate the treatment. By averaging the treatment effect (ATE) of each theory on inflation, it will be possible to extract how effective each theory is to compare across one another and against their respective theories. With the Global Macroeconomic Database (Müller, Xu, Lehbib, and Chen (2025), I can truly test all three theories using ten variables measuring various aspects of the economy, spanning 70 years totaling to 2338 observations. To achieve this I will be implementing a causal forest to extract the true causal relationship of each inflation theory.

Methods*:*

Dataset:

The database I used was the Global Macroeconomic Database (GMD) (Müller, Xu, Lehbib, and Chen (2025)). It is one of the largest databases for macroeconomic data, containing information on various economic indicators for countries all around the world.The data is collected by hand through going through historical records for nations, combining both online and paper sources. Values that appeared missing or untruthful (due to political manipulation) were calculated using values from the year prior and year following. From this set, I have picked out 10 variables going back to 1972 totalling to 575 observations. Due to larger periods of

missing data, many countries included in this subset do not have a series of continuous observations, rather the dataset is a collective of all viable observations from the dataset that do not contain missing observations. Additionally, I used data on consumer sentiment from the Federal Reserve Economics Database (Federal Reserve Bank of St. Louis).

To identify the effectiveness of each theory, I will be utilizing a causal forest through Python (Version 3.13.2) (Python Software Foundation), specifically the scikit-learn package (Pedregosa, F. et al. (2011)). To further enhance my model, I employed Kmeans to establish clusters within the dataset. Figure 1 shows the elbow plot used to determine the optimal number of clusters to use in the dataset.

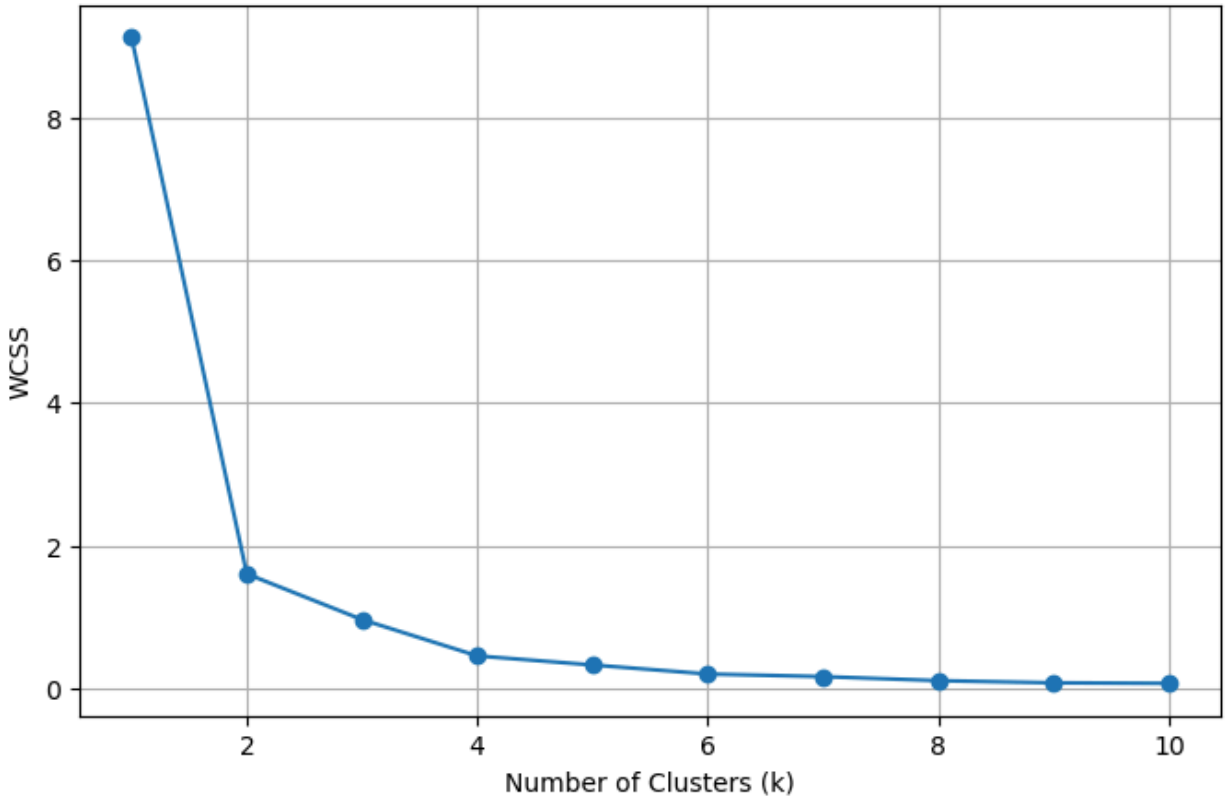

Figure 1. The elbow plot shows the fit of each cluster. This is measured within the cluster sum of squares (WCSS). The more clusters there are, the closer each observation is to each other within the specific cluster.

I decided to opt for four clusters to ensure each cluster contains similar features without completely losing out on variation within each cluster.

The causal forest model will allow me to find the average treatment effect of each model. Due to the limitations of the model, I selected only one variable to 'represent' each inflation theory. For monetarism, I chose the percent change in M2 as the treatment effect, as classic literature states that this is the primary driver of inflation. It is derived from the quantity theory of money equation, MV = PY (Lucas, Jr, 249). In the equation: M is money in circulation, V is velocity of money, P is the price level and Y is real Gross Domestic Product (GDP). From *Monetary Neutrality* by Rober E. Lucas, Jr. we hold both V and Y constant, so all changes come from M and P. The resulting percent change in price level, P, is inflation. Thus, if we move M up or down, the price level will move in a corresponding manner, such that for every per cent increase in M2, there is a per cent increase in inflation.

I used the percent change in government debt year over year to measure the treatment of the Fiscal Theory of the Price Level (FTPL). For this model, I assumed an inverse relation between the year over year percent change in national debt and market confidence that the debts will be paid. For both the FTPL and Monetary theory, using percent changes allows us to equally compare effects across countries despite drastic differences in various economic measurements.

Finally, I used unemployment as the treatment variable in the model measuring the New Keynesian theory. The literature relies on an inverse relation between unemployment and inflation, better known as the Phillips curve, thus it is reasonable to use unemployment as the treatment variable (Phillips, 283).

In isolating the treatment effect, I used central bank rate, long-term interest rate as well as the two treatment variables that were not being directly tested in that model for confounding

controls, or variables that affect both the treatment and dependent variable. I assumed that all treatment variables had a significant effect on inflation, thus treated them as confounding variables. Additionally, the tie between long term interest rates, central bank rates and inflation have long been established, thus were included. The heterogeneity controls included consumer confidence, imports as a percent of GDP , the real GDP of each country and the K Means classification. Heterogeneity controls allow the model to get an accurate representation of economic health each year and account for the variation across time. Furthermore, I imposed a two year lag on inflation, as inflationary policy often takes time to have a noticeable effect (Havranek et al).

Results:

Table 1 provides the average inflation statistics for each of the bins created by the K-nearest neighbors algorithm. Figure. 1-3 show scatterplots of each treatment variable against inflation. In figure 1, there is a concentration of points in the bottom half of the graph, and tapers with no distinct directions, while figure 2 shows a positive correlation.

Table 1: The descriptive statistics of inflation for the K-means groupings, including the cluster number labels and the amount of observations.

kmeans_cluster	count	mean	std	min	max
0	172.0	2.60	1.97	-0.69	10.21
1	138.0	2.15	1.47	-0.83	9.029
2	111.0	7.74	3.41	1.25	15.16
3	78.0	1.632	1.47	-1.35	5.39

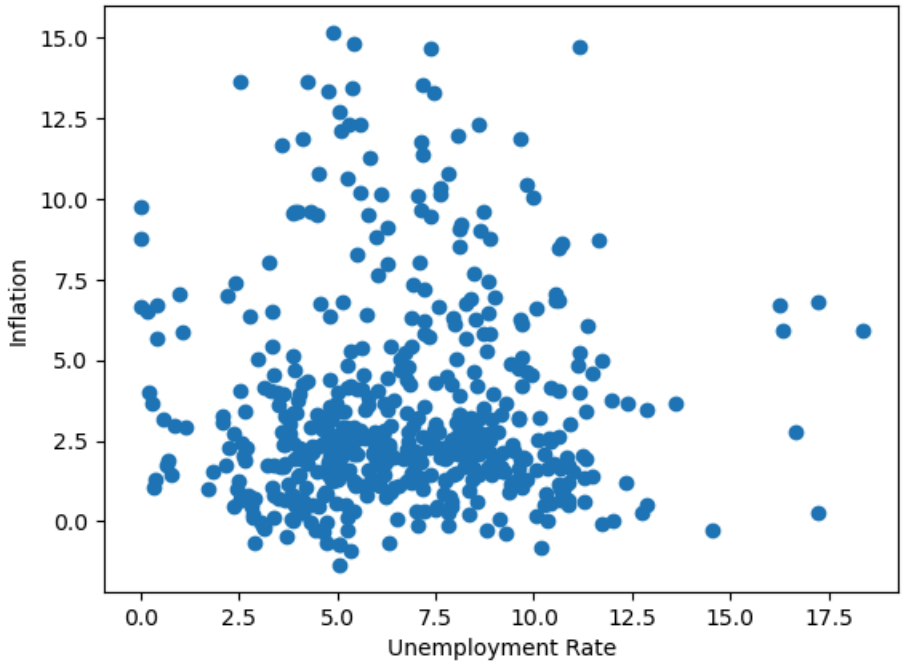

Figure 2: The scatterplot for unemployment vs. inflation is representative of the New Keynesian theory.

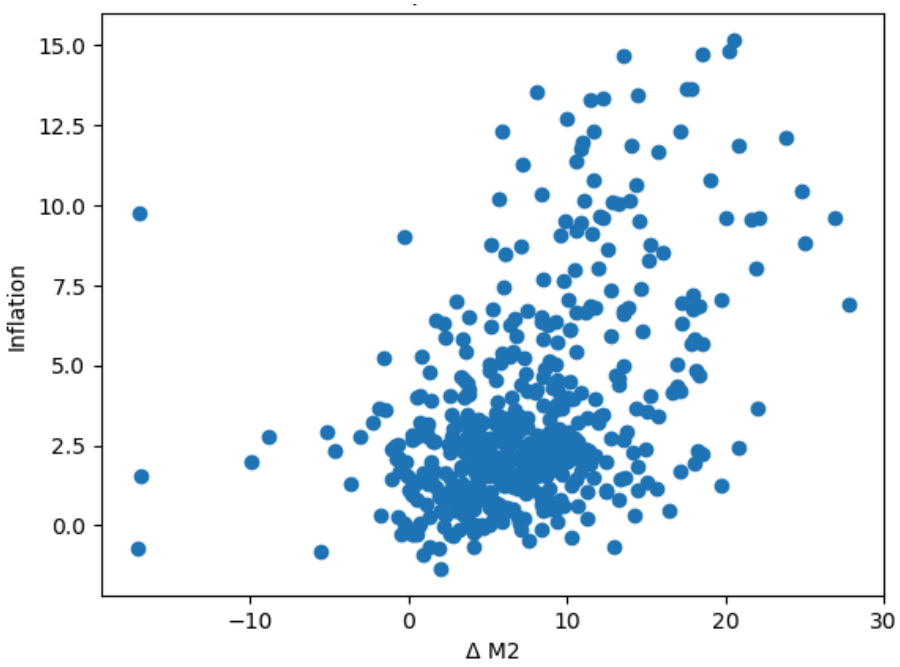

Figure 3: The scatterplot for Monetarism shows a grouping of observations between 0-10% change in M2 and 0 and 5% inflation.

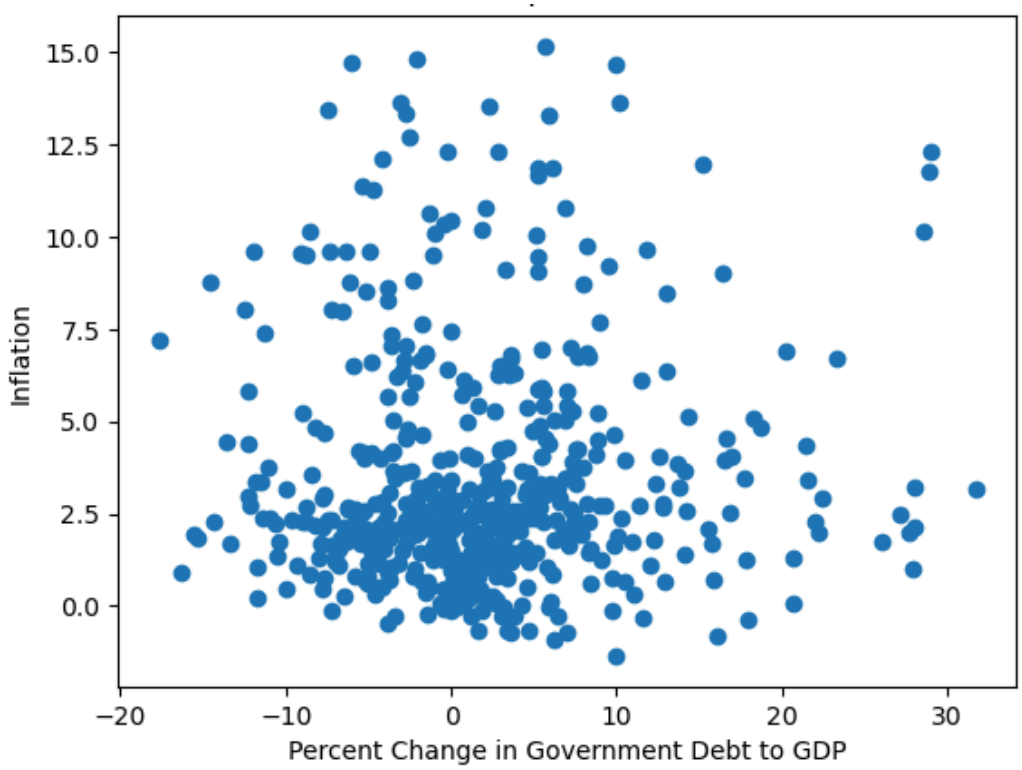

Figure 4: The scatterplot for the Fiscal Theory of the Price Level shows a concentration of observations between -10% and 10% change in Government Debt / GDP at low inflation levels.

As shown in Figure 5, each of the treatment variables, unemployment, percent change in government debt, and percent change in monetary supply show a causal relation that supports their respective theory. Specifically, for every unit increase in the treatment variables, there is a -.14, .0 and .12 change in inflation respectively.

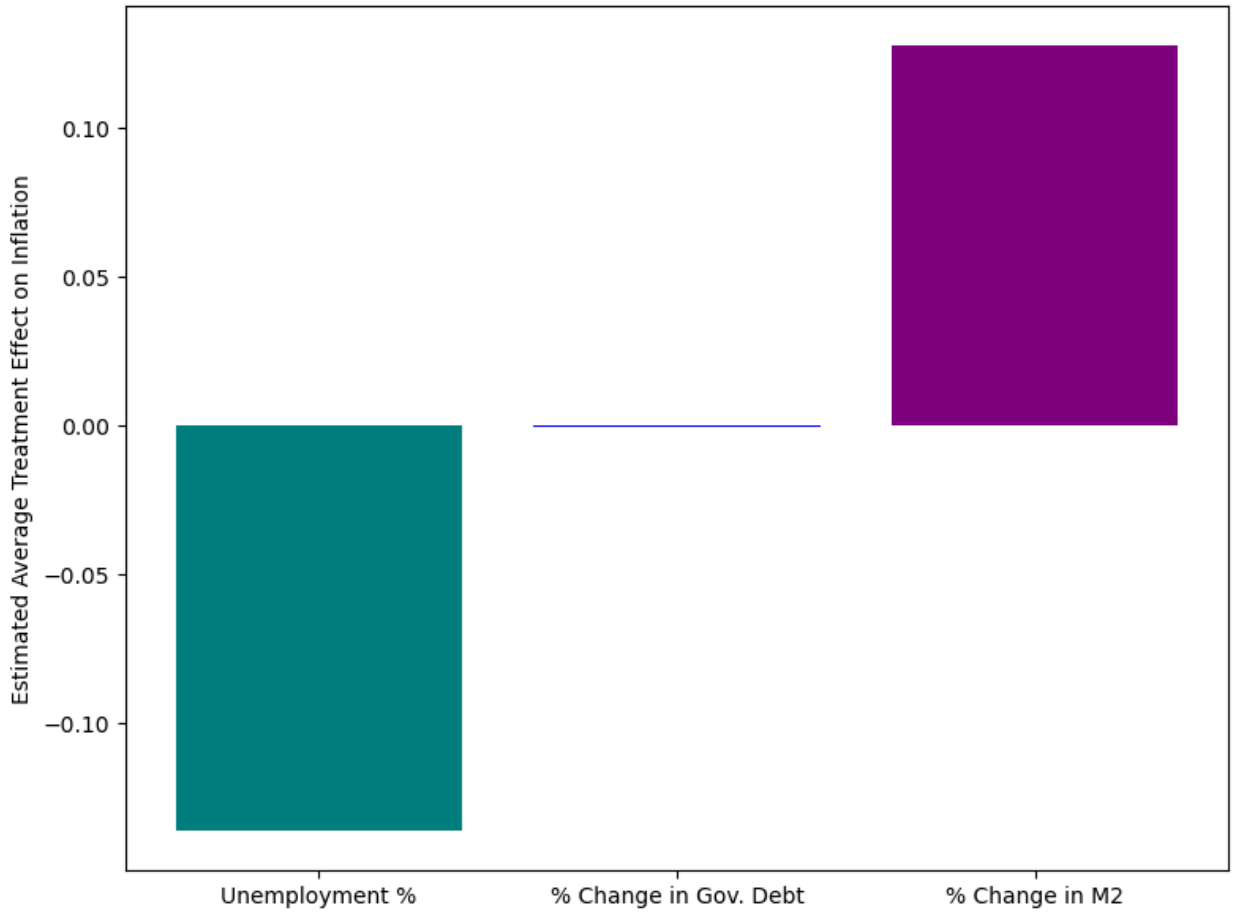

Figure 5: The bar graph shows the treatment effect of the primary variable on inflation for every unit increase in the primary variable. From left to right, the effect of New Keynesian, Fiscal Theory of the Price Level and Monetarism on inflation.

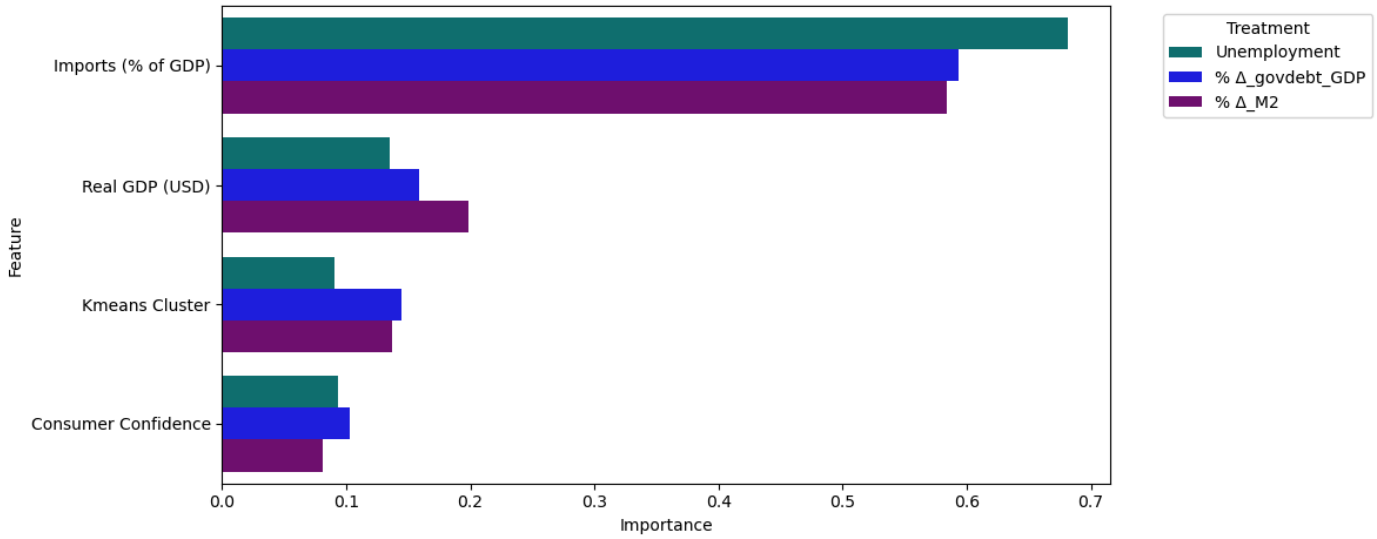

Figure 6: The bar chart shows the weight, or 'importance' of each heterogeneity control in the model. Importance is the measurement of how much effect each control variable had on the treatment variable in determining the magnitude of the effect on inflation.

Each variable shown in Figure 6 has its own relationship with the corresponding treatment effect, but does not have a direct relation with inflation. For Figures 7-9, the distributions of each model's calculated treatment effect is shown. The resulting distribution is a bimodal distribution, where there is a concentration on each side, giving the graphs two 'peaks'.

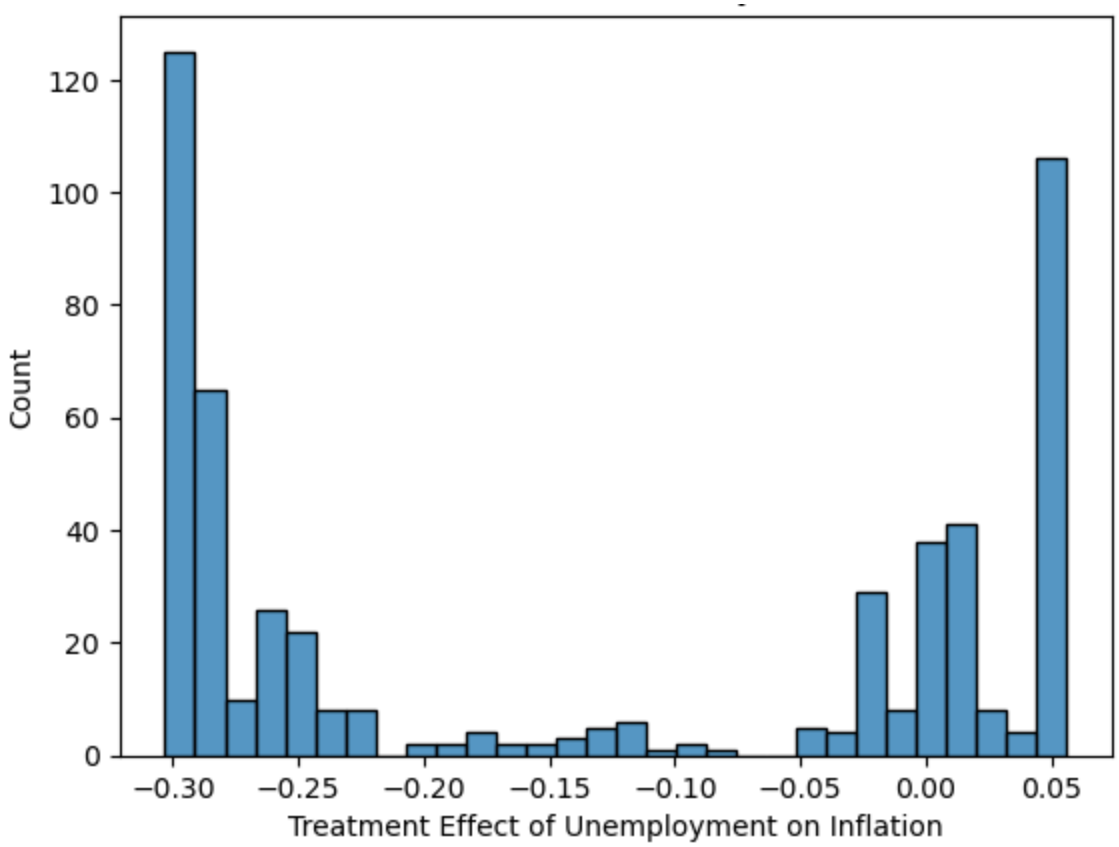

Figure 7: The histogram shows the distribution of results from each individual causal tree on the relation between unemployment and inflation. The mean of the distribution sits around -.14.

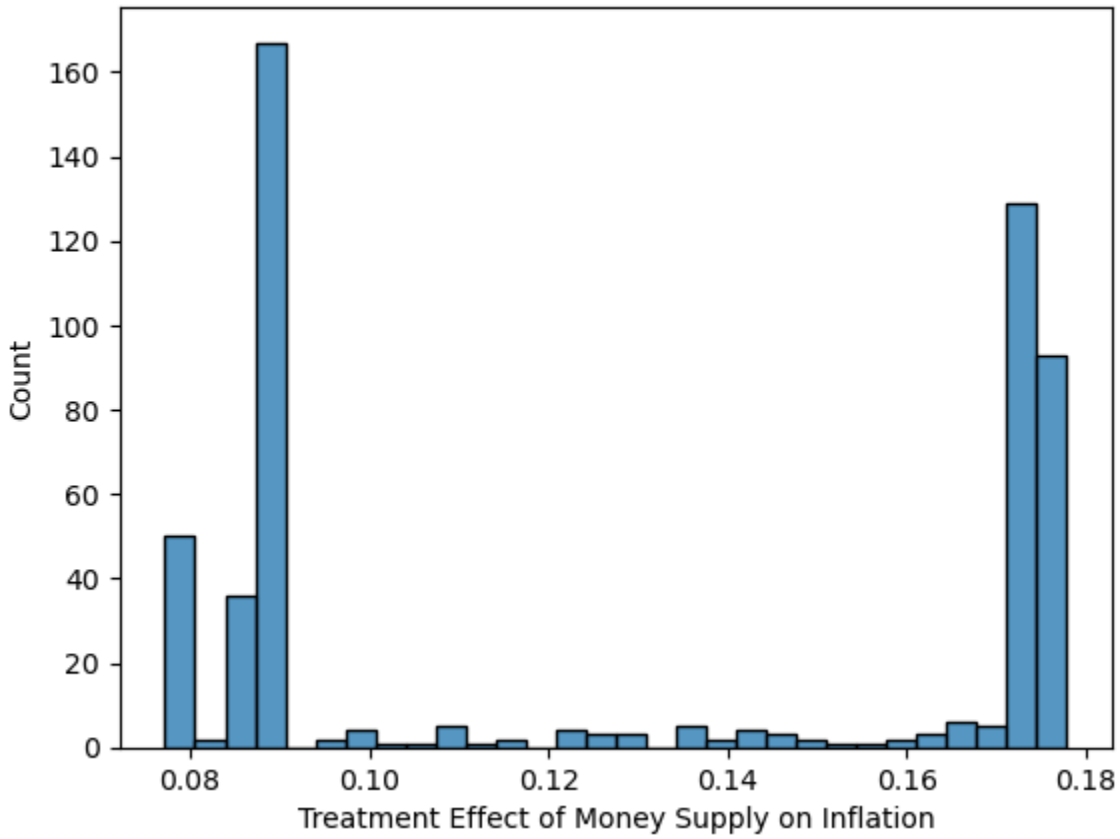

Figure 8: The histogram shows the distribution of results from each individual causal tree on the relation between percent change in M2 and inflation. The mean of the distribution sits around .12.

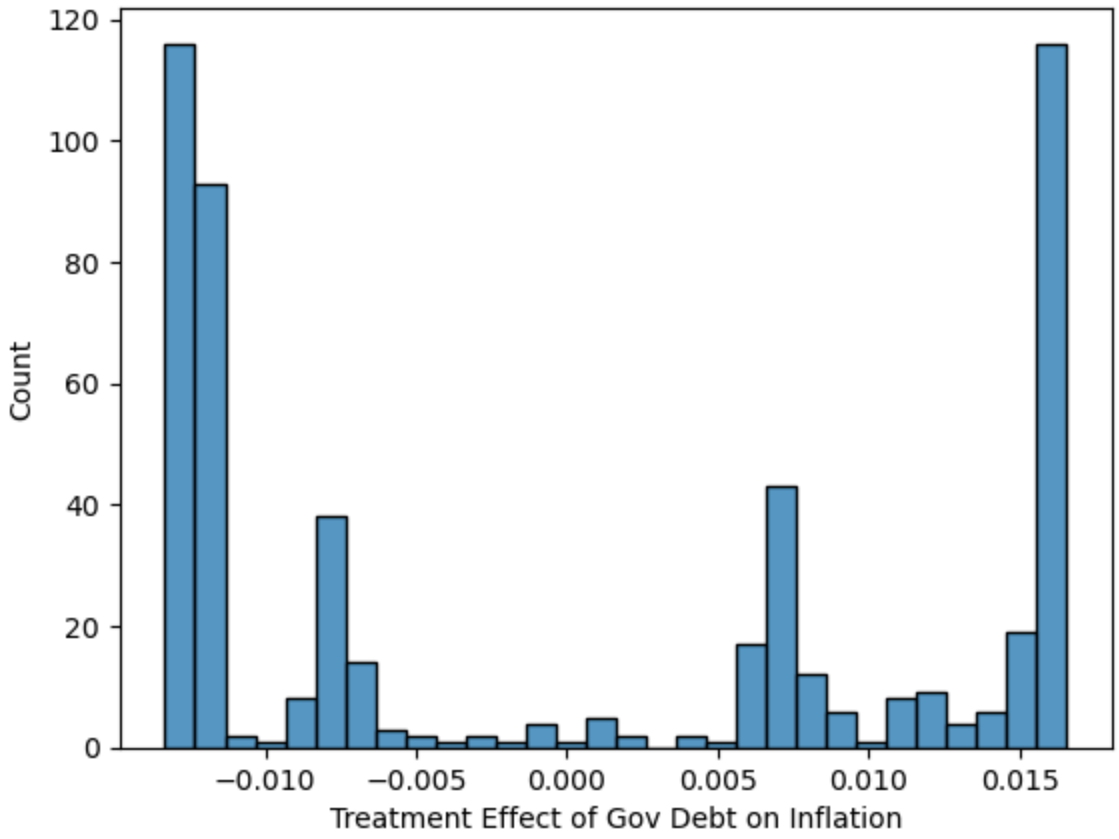

Figure 9: The histogram shows the distribution of results from each individual causal tree on the relation between percent change of government debt and inflation. The mean of the distribution sits at 0.

Table 2 shows the Variance Inflation Factor (VIF) for every variable used across all three models. Figure 10 shows the exact correlation between each possible pair of variables. Both the chart and table include the heterogeneity and confounding effect controls. A high variance inflation score is caused by multicollinearity between variables.

Table 2: The table lists the VIF with each variable's corresponding value. A value over ten is considered to be troublesome. The average VIF is 5.21.

Variable	VIF	Variable	VIF
Long Term Interest Rate	**13.34**	**Consumer Confidence**	**2.19**
Central Bank Rate	**7.48**	**Percent Change in M2**	**3.49**
Inflation	**6.27**	**Imports / GDP**	**3.36**
Unemployment	**5.28**	**Real GDP in USD**	**1.33**
Percent Change in Government Debt / GDP	**1.14**		

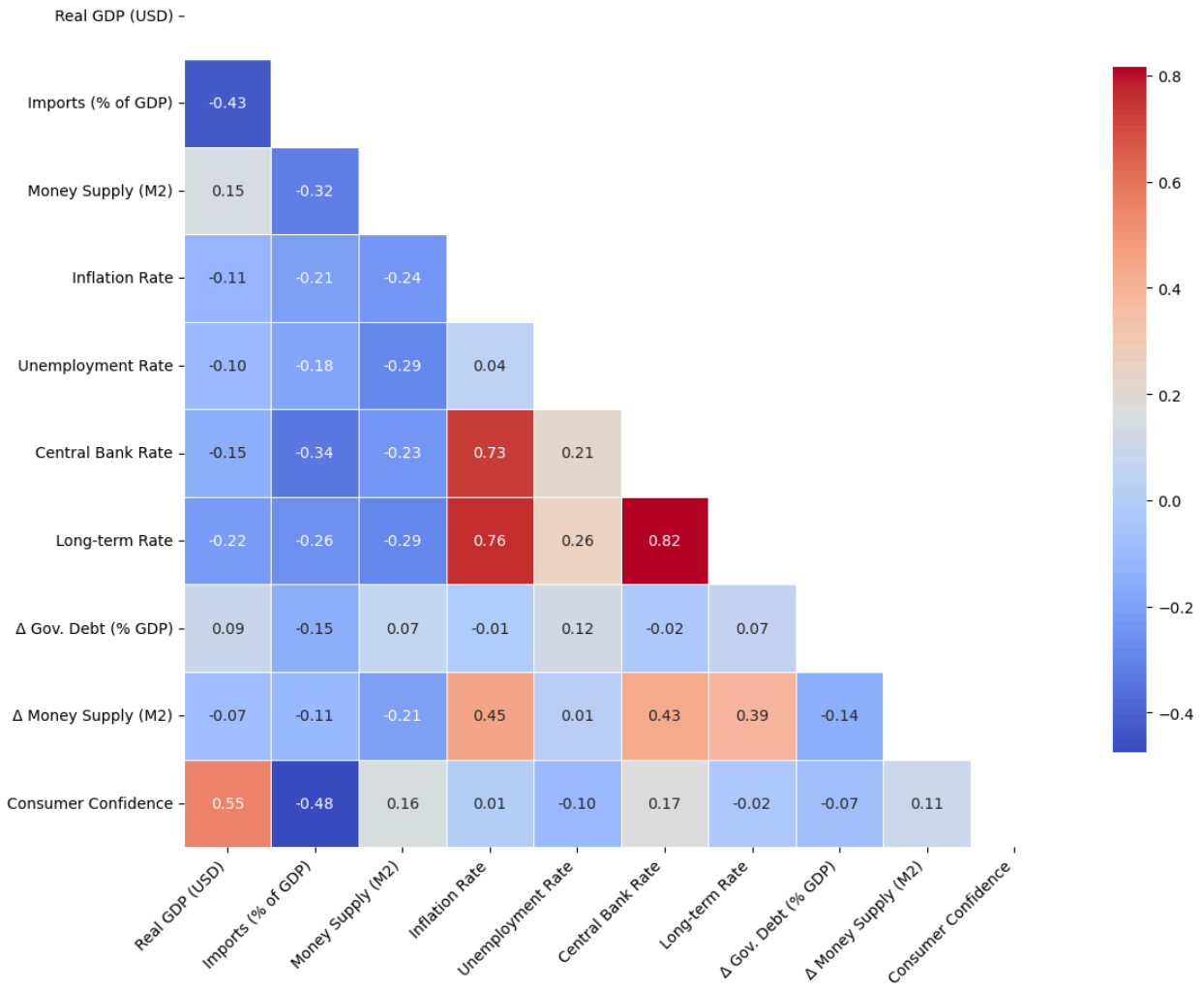

Figure 10: The chart shows a heat map of the correlation between all variables used. The higher the absolute value of the number, the more correlated the two variables are. Negative numbers

are interpreted as an inverse correlation while positive numbers are interpreted as a positive correlation.

Discussion

The causal forest estimating the Fiscal Theory of the Price Level effect shows little support for the theory, the magnitude of the model's estimation is almost negligible due to the near-zero estimation. Furthermore, the feature importance highlights the impact of imports and real GDP, over consumer confidence. With consumer confidence being a key part of the FTPL, the ideal model would have placed a larger emphasis on this feature. I believe a poor representation of the theory led to the model producing no real findings.

The causal relationship found for unemployment and inflation lines up more with the econometrics interpretation. While econometric literature suggests a causal relation of -.3, which is to say that for every unit increase in unemployment, there is a direct decrease of .3 in the inflation rate. The causal forest model finds a causal relation of -.14, suggesting that the phillips curve has flattened out even more over the past decade. However, as seen in Figure 7, the bimodal distribution shows that unemployment has frequently had a causal effect over -.3 on inflation. Thus, permitting certain conditions we know the model lines up with econometric literature.

Despite the flattening of the phillips curve over the past few years, (Hazell, et al.) the model is trained on over 50 years of data, which should yield a more impactful relationship. However, the feature importance does highlight real GDP as a primary driver in determining the effectiveness of unemployment on inflation, which is loosely related to how New Keyensian believes the theory to function.

Certainly this model yields more information, establishing a true causal relationship between unemployment and inflation, but there are still significant gaps within the model to address. But the model does not give us any reason to doubt that the current central bank policies should find alternative theories.

The model built to test Monetarism had an ATE of .12. There is a weak, but present causal relationship between the treatment variable and inflation. Additionally there is an even distribution of variation among the controls, shown in Figure 6, leading me to believe that the model is capturing a lot of the real-world movement and causality.

But, the low value of .12, raises questions on the effectiveness of the model. There are numerous sustained instances of inflation where economists are certain it was induced by an increase in money in circulation, particularly in cases of hyperinflation. However, the feature importance is spread out, which I believe provides some level of validity to the model as we can be more certain that it is capturing more of the variance throughout the data. Since the economy is composed of millions of different actors, having only one or two variables accounting for the variance does not make sense on even an intuitive level.

Finally, in table 2, the variable of Long Term Interest Rate has a reported VIF of 13.34, which normally is of concern. However, as noted in figure 10, there is a correlation of .83 with the Central Bank's rate and a correlation of .78 with inflation. It is then reasonable to assume that a portion of the high VIF score comes from these two variables; sharing correlation is not a concern with either as both interest rates are used to control for confounding effects, and since the long term interest rate is a control, correlation with inflation will improve the model. Furthermore, I am not concerned with what each confounding control is controlling, as that is not the aim of the model, and only that combined all variables will reduce the bias within the model.

Thus, this high VIF score is not a concern to the performance of the models.

The next steps for my research will be to start building models for predicting inflation based on each theory. Future research should plan to address the issue of 'oversimplification' for each theory. Reducing each theory to a linear relation of a single variable to inflation is not realistic for a practical application. As seen in Figures 2-4, inflation is often fairly low, but we need a true understanding of inflation in order to address the inflationary, or deflationary periods to sustain economic growth. To do this better controls need to be established for each model. Instead of testing three treatment variables on similar models, each model should be made with more specific grouping of control variables and additional hyperparameters. Being limited to one treatment variable on each model resulted in crude simplifications which certainly limited the ability to properly address the primary research question.

Furthermore, the research can pursue more situational treatments. Here I would test the effectiveness on various subgroups, divided based on economic indicators as the K-means clustering. In Table 1 we can see that cluster 2 is primarily composed of higher inflation rates while cluster 3 contains the lower end of values. The feature importance provided by each model will also provide information into how effective each inflation theory is, given certain parameters.

The last step from there would be work on a predictive model to predict a region's economic indicators. By predicting the economic movement of the key parameters, central banks can then prescribe the best policy based on the forecasts. This would allow the central banks to tailor the policy to an unprecedented level, maximizing their output while minimizing the detrimental effects from economic downturn.

Work Cited

Alessandria, George, et al. The Aggregate Effects of Global and Local Supply Chain Disruptions: 2020–2022. Policy Research Working Paper 10303, East Asia and the Pacific Region, Office of the Chief Economist, World Bank, Feb. 2023. https://documents1.worldbank.org/curated/en/099358202132331977/pdf/IDU04fa983470a55e045cf086ea01ed3050976da.pdf. Accessed 2 May 2025.

Binder, Carola, and Rupal Kamdar. "Expected and Realized Inflation in Historical Perspective." Journal of Economic Perspectives, vol. 36, no. 3, Summer 2022, pp. 131–156. American Economic Association. JSTOR, https://www.jstor.org/stable/27151257. Accessed 2 May 2025.

Harris, Elad, et al. "The Paradox Between the Macroeconomy and Household Sentiment." Economic Studies at Brookings, Brookings Institution, 8 Jan. 2025, https://www.brookings.edu/wp-content/uploads/2024/10/20250108_ES_HarrisEtAl_Sentiment.pdf. Accessed 2 May 2025.

Havranek, Tomas, and Marek Rusnak. "Transmission Lags of Monetary Policy: A Meta-Analysis." International Journal of Central Banking, vol. 9, no. 4, Dec. 2013, 38 pp., https://www.ijcb.org/journal/ijcb13q4a2.pdf. Accessed 2 May 2025.

Hazell, Jonathon, Juan Herreno, Emi Nakamura, and Jón Steinsson. "The Slope of the Phillips Curve: Evidence from U.S. States." The Quarterly Journal of Economics, vol. 137, no. 3, 2022, pp. 1299–1344

Müller, Andreas, Jiajun Xu, Karsten Lehbib, and Wenjie Chen. Global Macroeconomic Database (GMD). 2025

Phillips, A. W. “The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957.” Economica, vol. 25, no. 100, 1958, pp. 283–99. JSTOR, https://doi.org/10.2307/2550759. Accessed 2 May 2025.

Robert E. Lucas, Jr. – Prize Lecture. NobelPrize.org. Nobel Prize Outreach 2025. Mon. 5 May 2025. <https://www.nobelprize.org/prizes/economic-sciences/1995/lucas/lecture/>

Federal Reserve Bank of St. Louis. Consumer Confidence Indicators, Multiple Countries. FRED, https://fred.stlouisfed.org/. Accessed 5 May 2025.

Wager, Stefan, and Susan Athey. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” arXiv, 14 Oct. 2015, arXiv:1510.04342v4 [stat.ME], https://arxiv.org/pdf/1510.04342. Accessed 2 May 2025.

Link to Github:

https://github.com/DrHogan-R/DS_Capstone