# A Core Problem with Forecasting Inflation: Evidence for Adding Food and Beverage Prices to Core CPI

**John Macke**
**Pujan Thakrar**
**Mark Vandre**

**ECON 403B: Applied Econometrics**
**Final Project**
**March 11th, 2019**

**Introduction**

While the consumer prices index (CPI) is considered an important measure of inflation in the United States, economists typically look at core inflation. This is due to food and beverage prices as well as energy prices being extremely volatile and hard to predict. Michael Owyang, of the Federal Reserve Board of St. Louis, explained that food and beverage prices were extremely dependent on weather historically, while energy prices were dependent on OPEC. Although energy prices remain fairly volatile, food price volatility has declined. According to the Federal Reserve Board of St. Louis, food prices have become much more stable since the 1990s. Furthemore, economists from the Federal Reserve Board of St. Louis state that food prices are the best predictor for future inflation.

Given that food and beverage prices have become more stable and are considered to be a strong predictor of future inflation, we are interested in forecasting inflation by including food and beverage prices. We would like to determine if we can get a better or comparative forecast to those of the Survey of Professional Forecasters in the past. More specifically, will adding food and beverage CPI back into core CPI provide a more accurate forecast of full CPI inflation? A better forecast in inflation will help central banks and governments gain a more accurate view of the future economy so that they can direct policy more efficiently.

**Data**

In order to answer our question, we required several data series. First, we required monthly, historical CPI data for all 8 components of CPI. These components are food and beverages, housing, apparel, transportation, medical care, education and communication, recreation, and other goods and services. We also required full CPI and core CPI historical data. We retrieved all the CPI data from the Federal Reserve Economic Data page (FRED). In addition, we retrieved historical data for the individual CPI component weights. This data was reported annually from the Bureau of Labor Statistics (BLS) and dates back to 1987. Finally, for comparison purposes, we retrieve historical, quarterly CPI inflation forecasts from the Survey of Professional Forecasters.

For the purposes of our report, we convert all CPI data to monthly CPI inflation data by taking the year-over-year percentage change of each observation and subset data beginning January 1994 and ending December 2018. We do this for two reasons. The first is due to limited data availability for both education and communication CPI and medical CPI. The second reason was because food and beverage prices stabilized in the mid-90s. Finally, we only require the professional forecasts for the last year and therefore we retrieve this data from 2018Q1 to 2018Q4.

**Methodology**

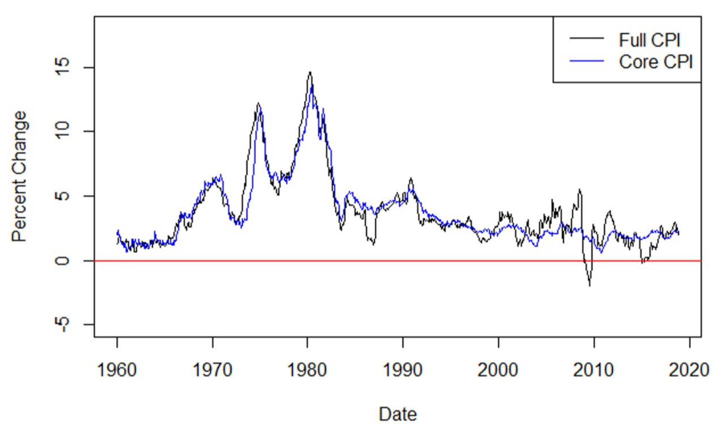
The consumer price index has always been one of the main measures of inflation for the United States (Board of Governors). According to the Bureau of Labor Statistics, “CPI is a measure of the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services”. This market basket is carefully composed of the 8 components shown in Table 1. Typically, FRED will aggregate the components using the set of weights determined by the BLS from their consumer expenditure surveys that occur every three years. Therefore the weights applied tend to be lagged by three years.

Table 1: Components and Weights

| Component | Weights |
|---|---|
| Food and Beverages | 14.314 |
| Housing | 42.202 |
| Apparel | 2.959 |
| Transportation | 16.348 |
| Medical Care | 8.682 |
| Recreation | 5.694 |
| Education and Communication | 6.596 |
| Other Goods and Services | 3.204 |

The primary difference between core CPI inflation and full CPI inflation is that core CPI inflation excludes food and beverage prices, while full CPI inflation includes food and beverages prices. The reason food and beverages have been excluded from core CPI inflation in the past is that this component has been volatile over time relative to the other components. Figure 1 below illustrates the difference between full and core CPI inflation.

**Figure 1: CPI Full vs CPI Core**

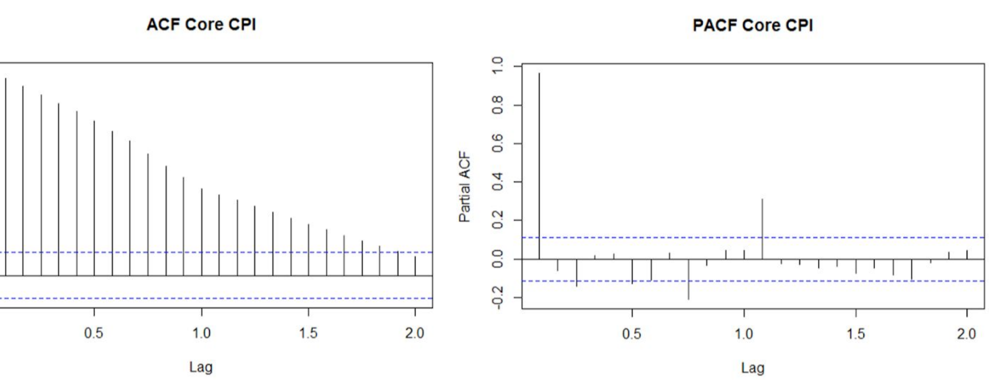


As we can see from the above plot, core CPI inflation is fairly in line with full CPI inflation until the mid-90s where the two series begin to diverge. At this point, core CPI inflation tends to remain consistent over time while full CPI inflation is much more volatile. This change in volatility may be due to a change in the underlying behavior of the components or as a result of core CPI inflation not capturing an important component, such as food and beverage prices.

We begin by forecasting the full CPI inflation rate using core CPI inflation to get a baseline estimate of the forecasting accuracy of core CPI inflation. Before forecasting core CPI inflation we split our data into a training and testing sample. The training sample consists of data from January 1994 to December 2017 and our testing sample consists of data from January 2018 to December 2018.

In order to choose a model, we begin by looking at the ACF and PACF of core CPI inflation. We notice in Figure 2 that the ACF slowly decays to zero while the PACF has significant spikes at lags 2 and 12. This indicates that the series is an autoregressive process (AR).

Figure 2:




We test the significance of the spike at lag 12 using the Box-Pierce test and find that it is statistically significant. Because we have monthly data, we believe the spike at lag 12 may be the result of a seasonal component within the series. To check for this we test the significance of the spikes at lag 24 and 36 and get the following results.

Box-Pierce test

data: core.ts

X-squared = 1956.9, df = 12, p-value < 2.2e-16

Box-Pierce test

data: core.ts

X-squared = 2200.7, df = 24, p-value < 2.2e-16

Box-Pierce test

data: core.ts

X-squared = 2205.5, df = 36, p-value < 2.2e-16

The significant spikes in the PACF at lags 12, 24, and 36 indicate that core CPI inflation has some seasonality that must be incorporated into the model. We examine a decomposition of the series, given by Figure 3, in order to get a better idea of the trend and seasonal components.

**Figure 3: Core CPI Inflation Decomposition**

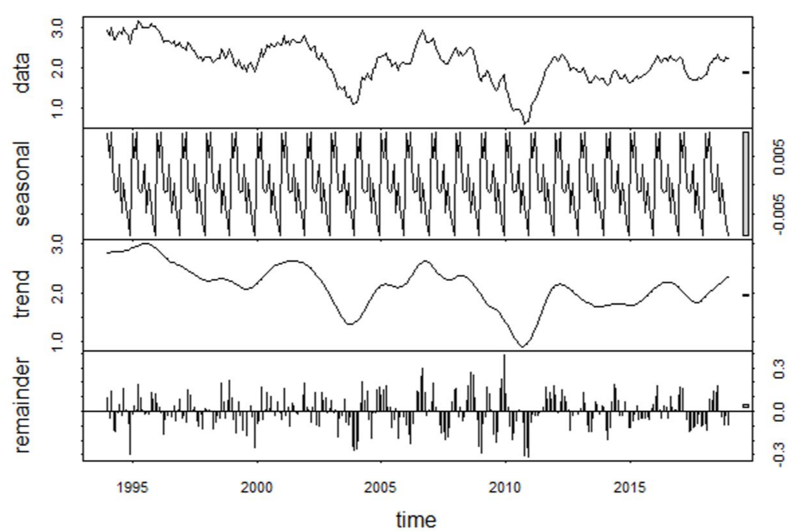



Figure 3 confirms that seasonality is present. However, we can see that the scale of the seasonal component is small relative to the overall series. Based on the time period of inflation data that we have, it appears that there is a slight decreasing trend in core CPI inflation from 1994 to present. Using this information we fit an AR(2) s-AR(1) with a drift to core CPI and get the following results for the residuals.

**Figure 4: Residuals**

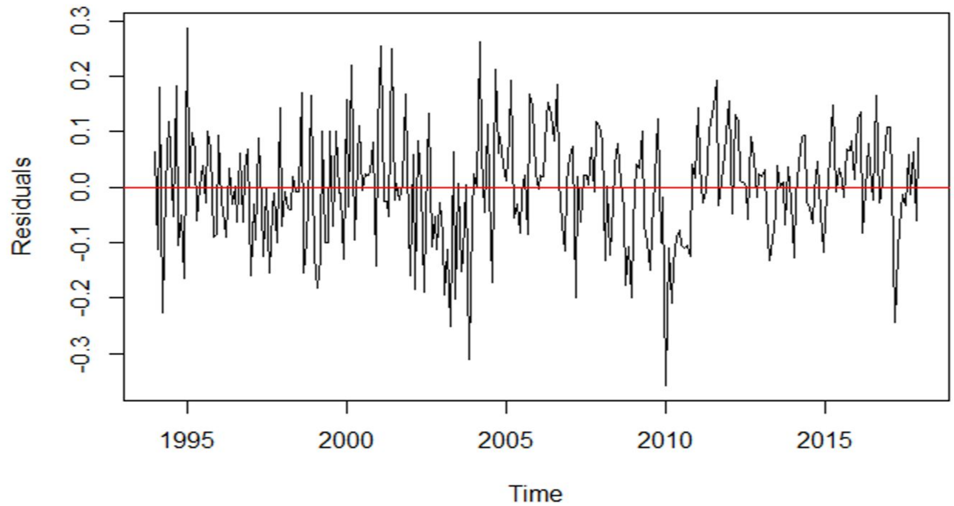
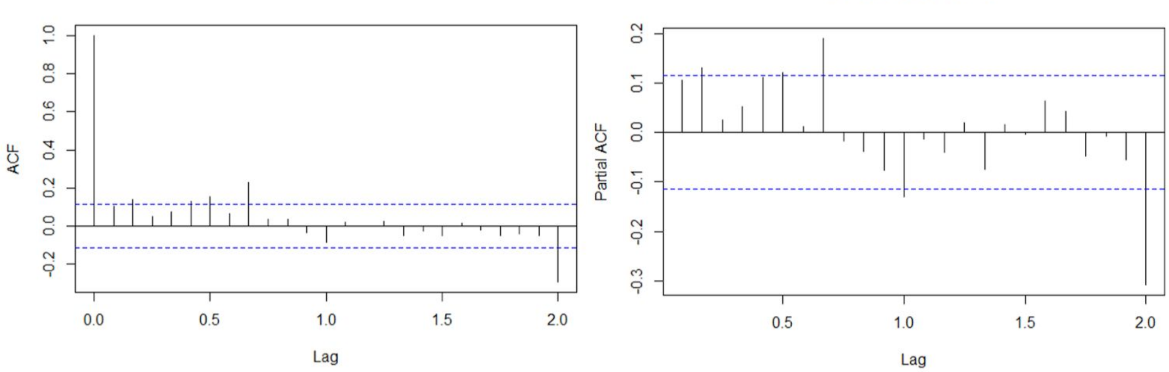
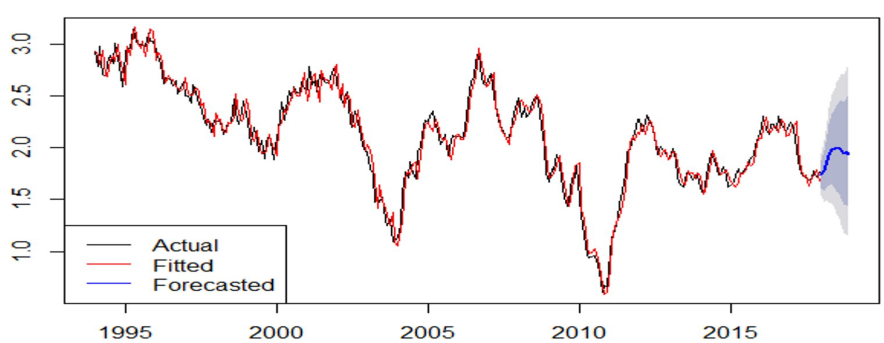



**Figure 5:**




From the ACF and PACF of the residuals we can see that they are still not white noise and therefore we first difference the data. After doing so, we fit our model to the training sample of core CPI inflation and forecast 12 months ahead.

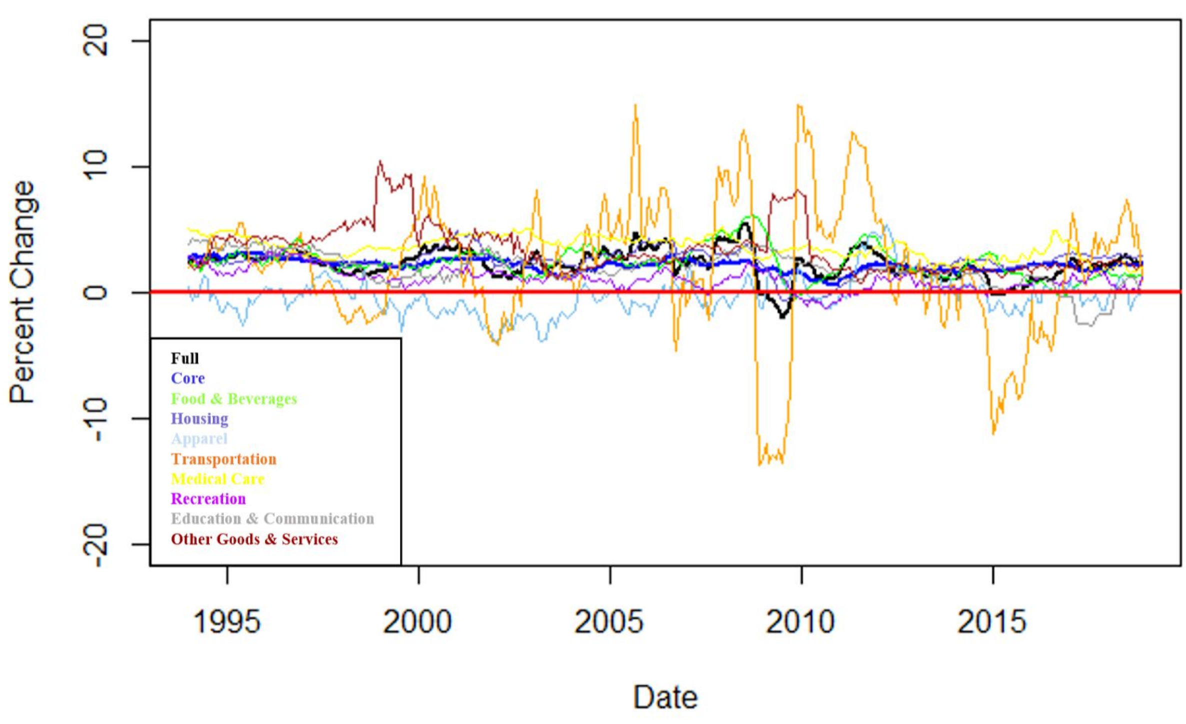
**Figure 6: Forecasts from ARIMA(2,1,0)(1,0,0)[12] with drift**



After forecasting 12 months ahead, we compute the RMSE of our forecasted values with the actual full CPI inflation values for the year 2018 and get an RMSE of 0.57. Given that our data is in terms of percent change, for interpretability, it would be useful to use the accuracy measure MAPE because it is represented in terms of the percentage error. In this case the MAPE is 0.20. For comparison, we fit a model using the auto arima function in R, which suggests an ARIMA(2,1,1)(1,0,2) with a drift which gives us an RMSE of 0.58 or a MAPE of 0.22. Given this information, we find that our fitted model performed better over the testing sample. While an RMSE of 0.57 is not bad, we would like to improve on this by constructing a new CPI inflation metric that includes food and beverage prices in the core CPI inflation. We will use the RMSE of 0.57 and MAPE of 0.20 as baselines for future models.

In order to do this we take a closer look at the components that make up full CPI inflation shown in Figure 7.

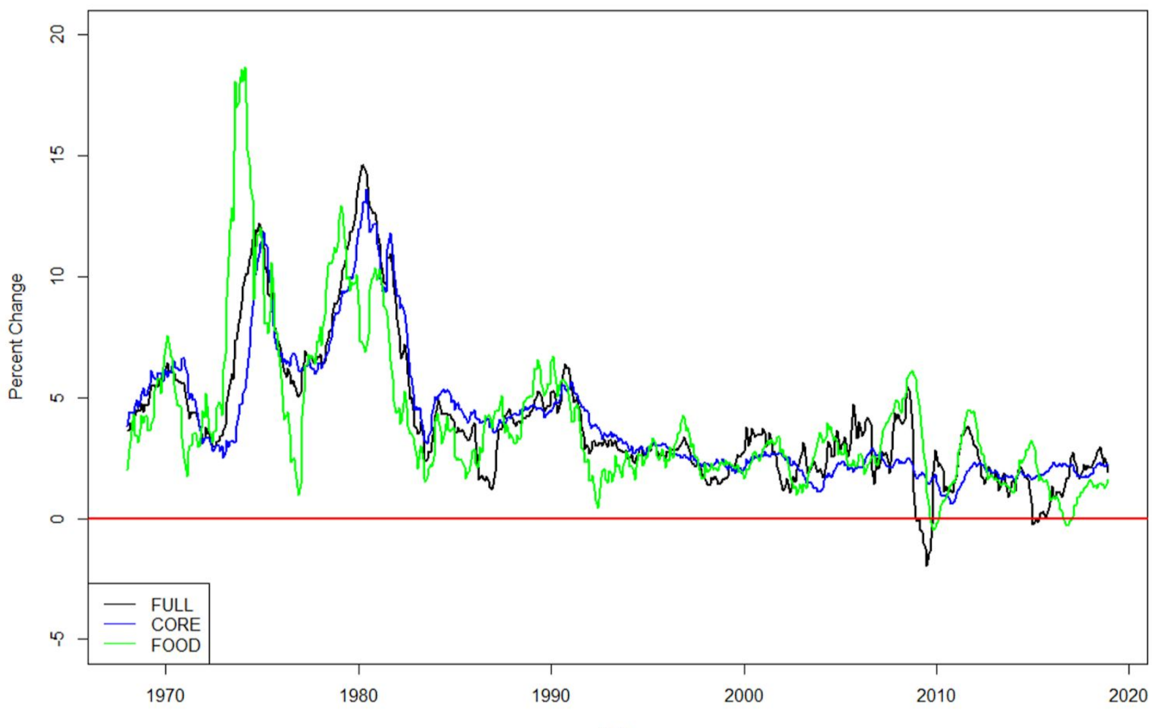
**Figure 7: CPI Indices**




From figure 7, we can see that transportation and other goods and services are more volatile than the other components. Apparel tends to be lower than all the other components but has a relatively consistent level of CPI inflation over time. The other components seem to be relatively stable in terms of volatility. This raises the question of why food and beverage prices are not included when forecasting full CPI inflation.

To get a better idea, full and core CPI inflation are plotted in figure 8 along with food and beverage CPI inflation. Historically, we see that core CPI inflation seemed to be a reasonable metric for forecasting full CPI inflation. However, since the mid-90s, it seems that food and beverage CPI inflation follows the true value of full CPI inflation more closely. Therefore, it it would be worth considering for future forecasts.

**Figure 7: CPI Indices**




In order to understand how each component relates to full CPI inflation, we compose the following RMSE table (Table 2), which is constructed by looking at the RMSE between each component's CPI inflation and full CPI inflation. The results are ordered below from best indicators to the worst indicators based on the RMSE.

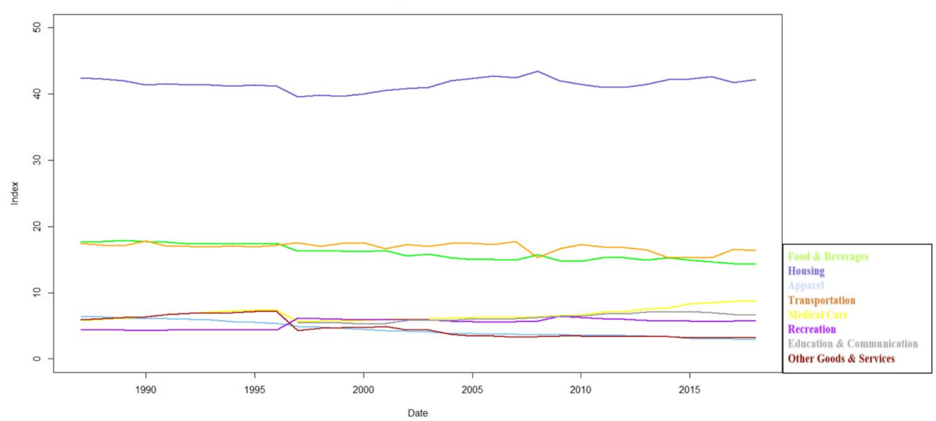
Table 2: RMSE Table

| CPI Component | RMSE |
|---|---|
| Housing | 1.04 |
| Food and Beverages | 1.28 |
| Education and Communication | 1.62 |
| Recreation | 1.69 |
| Medical Care | 1.76 |
| Other Goods and Services | 2.51 |
| Apparel | 3.11 |
| Transportation | 4.33 |

As shown in the table above, the Food and Beverages component is one of the best indicators for predicting full CPI inflation. We can also see that transportation is the worst performing indicator of full CPI inflation. This may be a result of many goods and services within the transportation component being reliant on energy prices, such as gas, which are highly volatile.

We use BLS data for the weights of each component to construct our own core CPI inflation metric that includes food and beverage prices. Before constructing our new metric, we examine how the weights of each component have changed over time. In the plot below, we see that the component weights have remained fairly consistent over time. While some of the component weights have increased slightly since 1987, the difference is not appreciable. Given this information, we use the most recent weights (December 2018) in constructing a new core CPI.

**Figure 8:** CPI Weights Over Time



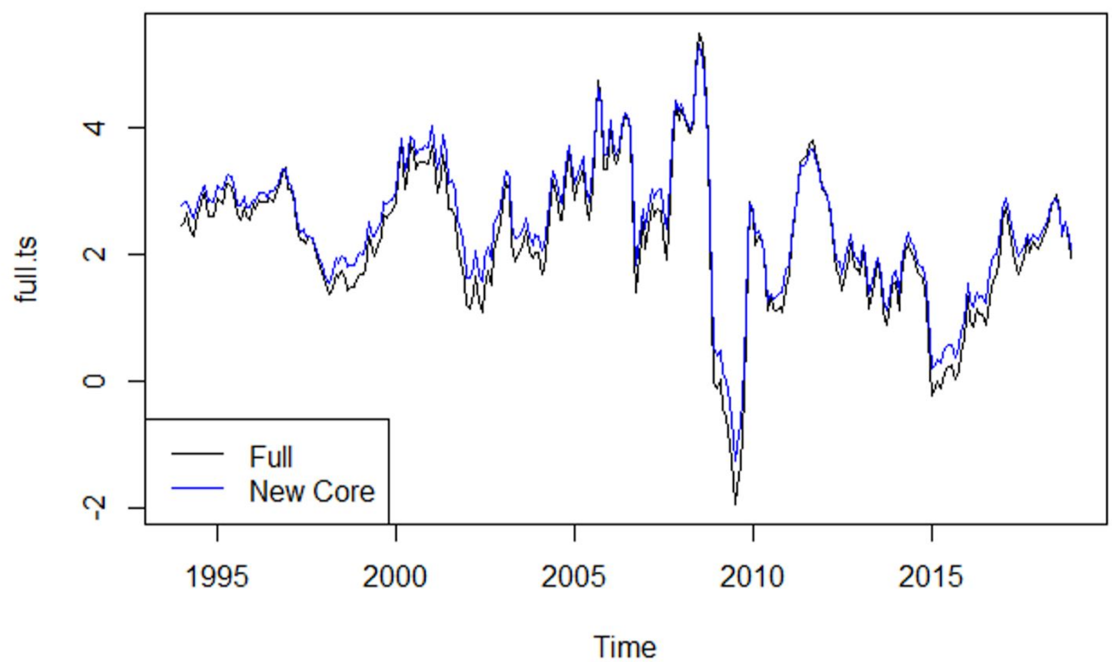

To construct our new CPI metric we use the following equations, where $CPI_{i,t}$ is the value of component i in period t and $w_{i,t}$ is the weight associated with component i. The first equation gives us the CPI at time t, given by the CPI of each component at time t multiplied by its weight. The second equation is used to calculate the growth rate as percentage change from a year ago. We create a function in R to perform these calculations for us.

$$CPI_t = \sum_{i=1}^{N} w_{i,t} CPI_{i,t}$$

$$\pi_t = \frac{CPI_t - CPI_{t-12}}{CPI_{t-12}}$$

We then apply our function to all the individual components of full CPI and their respective weights. As we can see in Figure 9, our "new core" is approximately equally to full CPI inflation as expected. However, there exists a small degree of error due to rounding and/or differences in computation.

**Figure 9: CPI Full Inflation & CPI New Core Inflation**




Now we fit a model to our "new core" metric. Again, we split the data into training and testing similar to the way we did before with core CPI inflation. We start by examining the ACF and PACF of the series and notice that the ACF decays to zero and the PACF has significant spikes at lags 1 and 12. This indicates an AR process, which is also evident from the persistence in the series. Additionally, we  examine the trend and seasonal components in figure 11.

**Figure 10:**

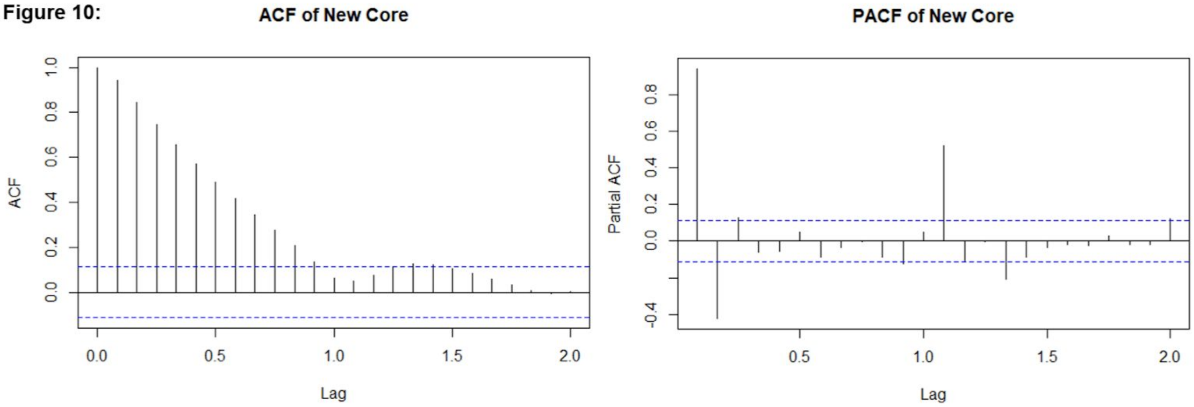
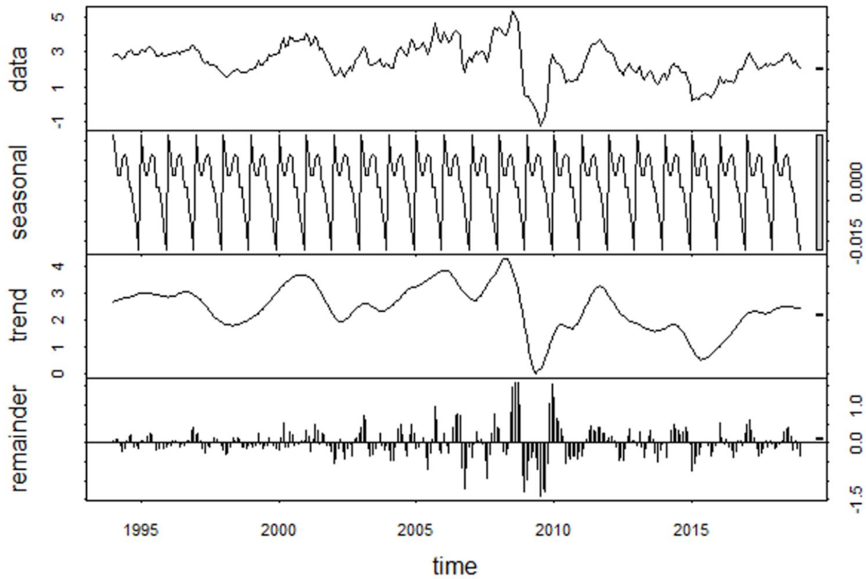
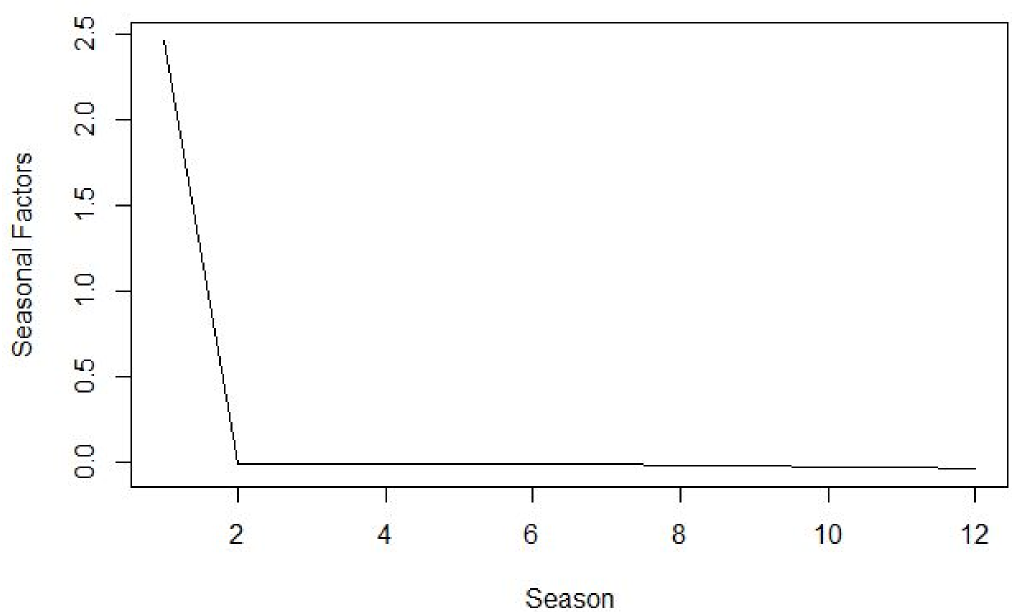


**Figure 11: New Core CPI Inflation Decomposition**




From Figure 11, we see that there appears to be a slight decreasing trend in our "new core" CPI inflation from 1994 to present. Again, seasonality is present even though the scale of the seasonality component is fairly small relative to the overall series. Initially, we consider fitting a AR(1) s-AR(1) with a drift to this series, but after regressing the series on the seasonal component, we find that the coefficients are all close to zero and insignificant. This is represented in Figure 11.5.

**Figure 11.5: Plot of Seasonal Factors**



Given that the seasonal component was negligible, we chose to fit and AR(1) with a drift and get the following. Figures 12 and 13 below show that the model we fit to the data does not fully reduce the residuals to white noise.

**Figure 12: Residuals**

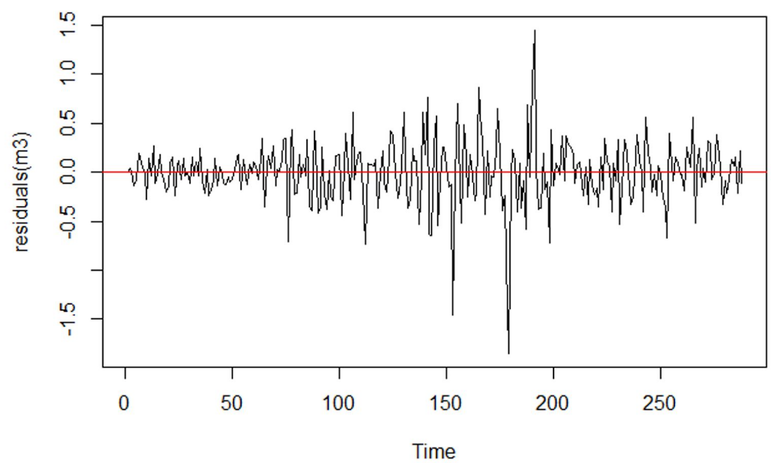
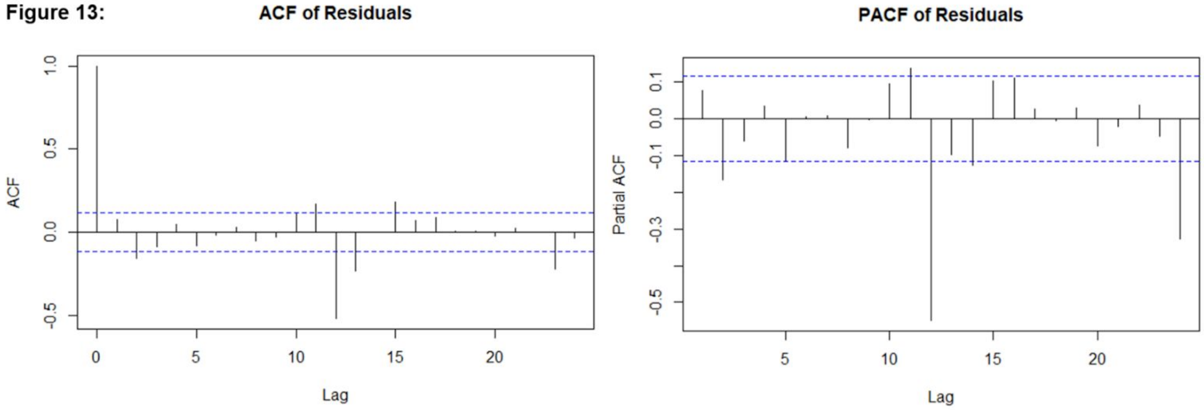
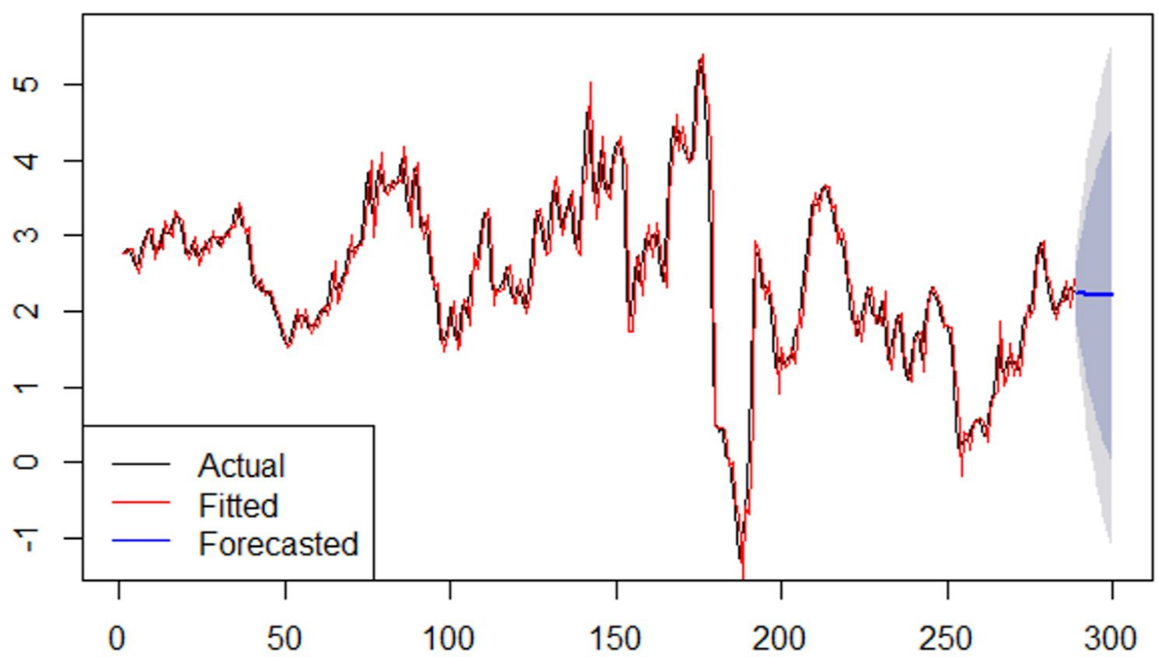



**Figure 13:**

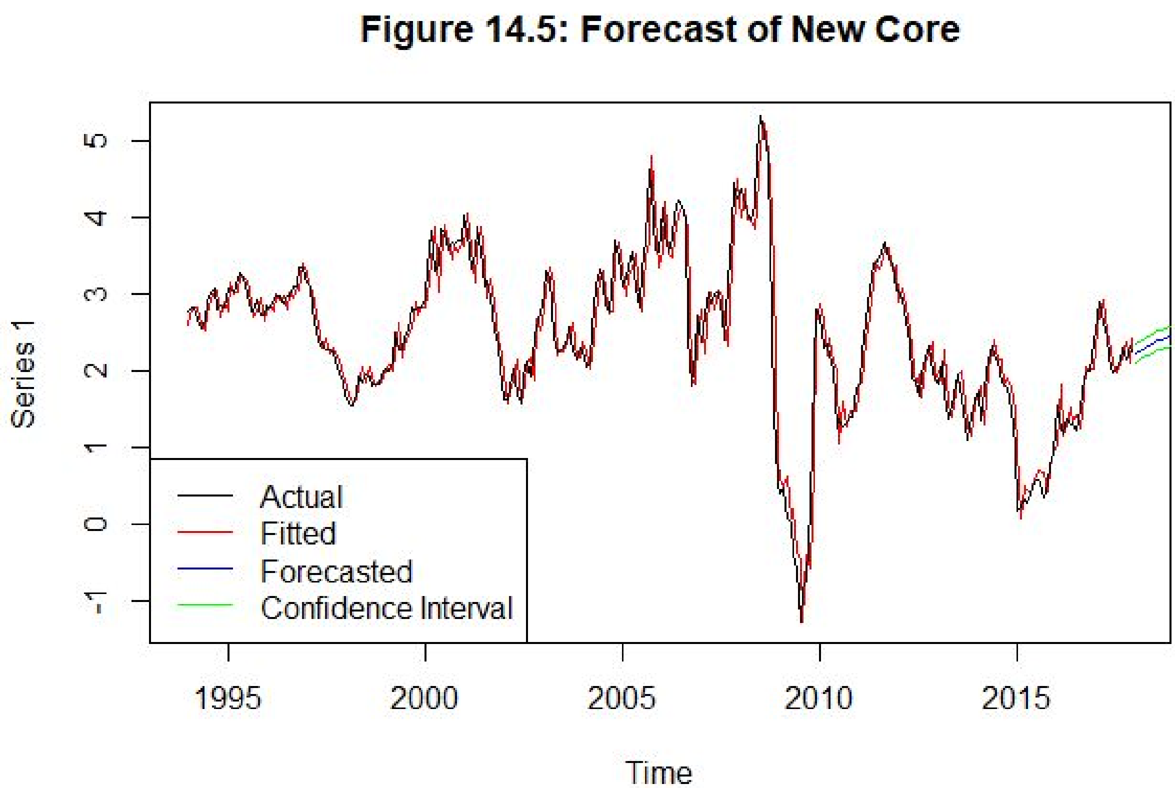



We first difference the data and fit an ARIMA(1,1,0) to the training sample of our “new core” CPI inflation and forecast 12 months ahead, which can be seen in Figure 14. We compare the forecasted values to full CPI inflation for January 2018 to December 2018 and calculate an RMSE of 0.37 and a MAPE of 0.11.

**Figure 14: Forecast of New Core**




As a comparison we use the auto arima function in R, which suggests an ARIMA(1,1,2) and gives an RMSE of 0.38 and a MAPE 0.11. However, we decide to square the residuals to see if we can apply a GARCH model to improve the forecast. After doing this, we notice that the ACF and PACF of the squared residuals have some dynamics that we can incorporate. We fit a sGARCH(1,1)ARIMA(1,0,1), which yields an RMSE of 0.31 and a MAPE of 0.10. This can be seen in Figure 14.5 below. Comparing this to our baseline, we observe that our "new core" CPI inflation is better at forecasting the full CPI inflation rate than the original core CPI inflation. This indicates that including food and beverage prices with core CPI to forecast full CPI inflation would improve the accuracy of the forecast.



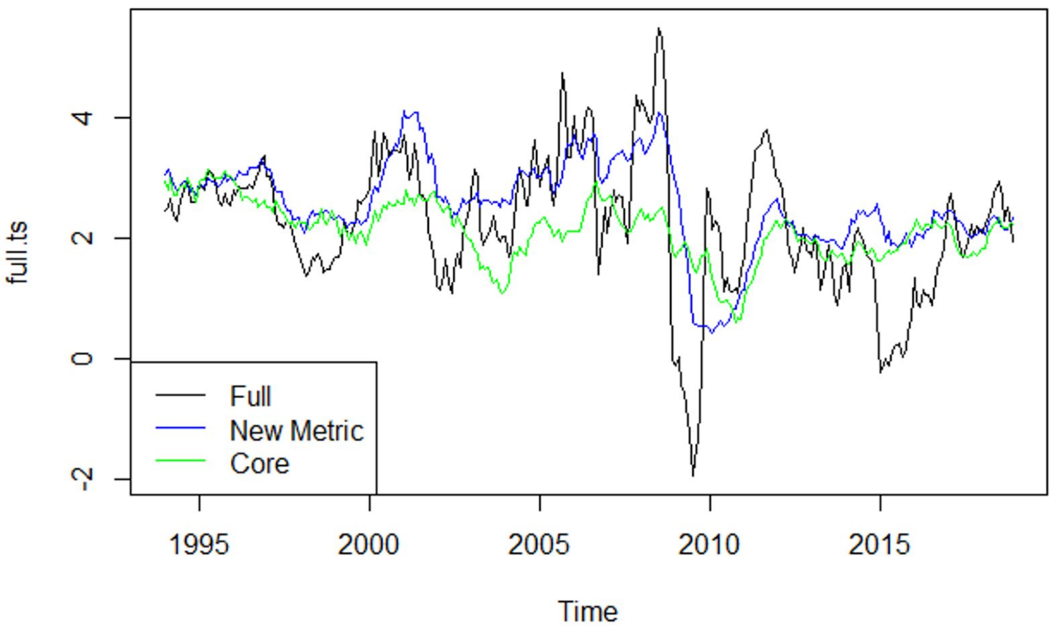

**Additional Analysis**

Given that the forecasts from our "new core" CPI inflation  are better than the forecasts from the original core CPI inflation, we want to determine if it is also better than forecasts from the Survey of Professional Forecasters.

To do this, we convert our "new core" CPI inflation and full CPI inflation observations to quarterly data by taking the mean of the forecasts in each quarter (3 months). We do this to compare our forecasts with the professional forecasts that are reported in quarterly frequency. The professional forecasts give an RMSE of 0.29 and a MAPE of 0.09, which beats our model by a percentage error of 0.01.

Since we were very close to the professional forecasts with our model, we want to use different criteria to build a new core CPI inflation that we will refer to as "new CPI metric". "New CPI metric" consists of the 5 components that had the lowest RMSE as shown in Table 2. These components are: housing, food & beverages, medical care, recreation, and education & communication.This encompasses approximately 77% of full CPI. We compare our "new CPI metric" to both full CPI inflation and core CPI inflation in Figure 15.

**Figure 15: Full CPI Inflation & New CPI Metric & Core CPI Inflation**




We examine the ACF and PACF of our "new CPI metric" and notice that the ACF decays to zero and the PACF has significant spikes at lags 1 and 12. Again, this indicates an AR process, which can be seen in the persistence of the series in Figure 15. Additionally, we examine the trend and seasonal components of the series in Figure 17 and see a small seasonal component and a slightly decreasing trend. Similarly to before, this seasonal component is negligible and therefore we do not include it.

**Figure 16:**

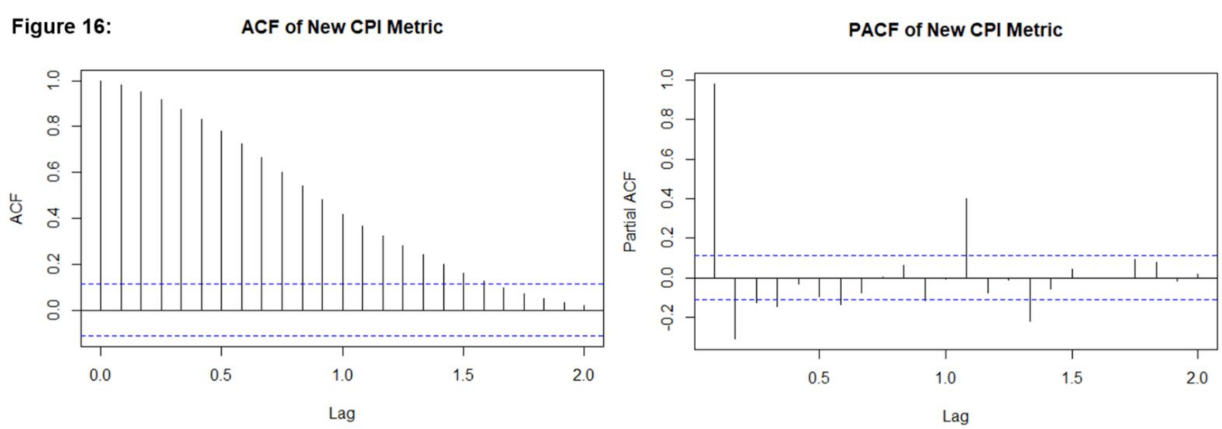
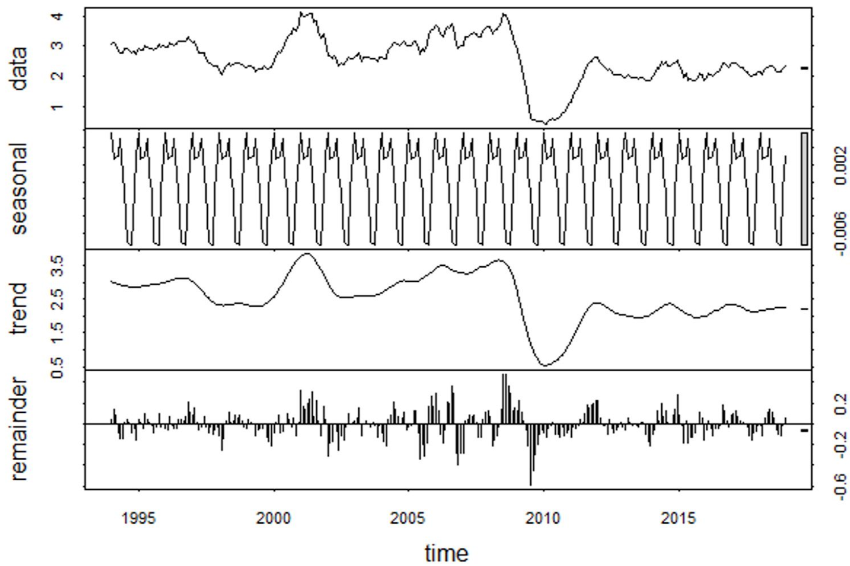


**Figure 17: New CPI Metric Decomposition**




We fit an sGARCH(1,1)ARIMA(1,0,1) to the training set and compare the forecasted values to the actual inflation for 2018 and get an RMSE of 0.36 and a MAPE of 0.11. The forecast itself can be seen in Figure 19. The “new CPI metric” that we created, using the five components that predict full CPI inflation the best, performs worse than the “new core” CPI inflation measure we created initially. This is surprising as we would have expected “new CPI metric” to have more predictive power as it only includes the components that predict full CPI inflation the best individually.

**Figure 18: Squared Residuals**

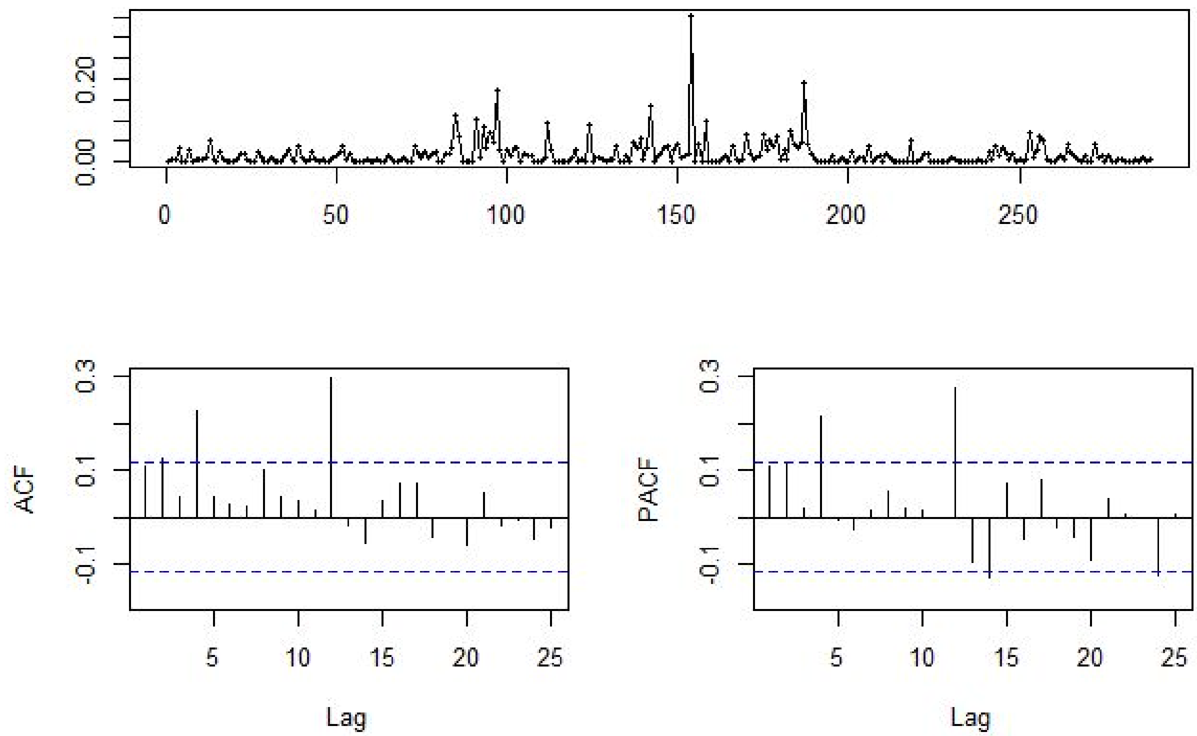
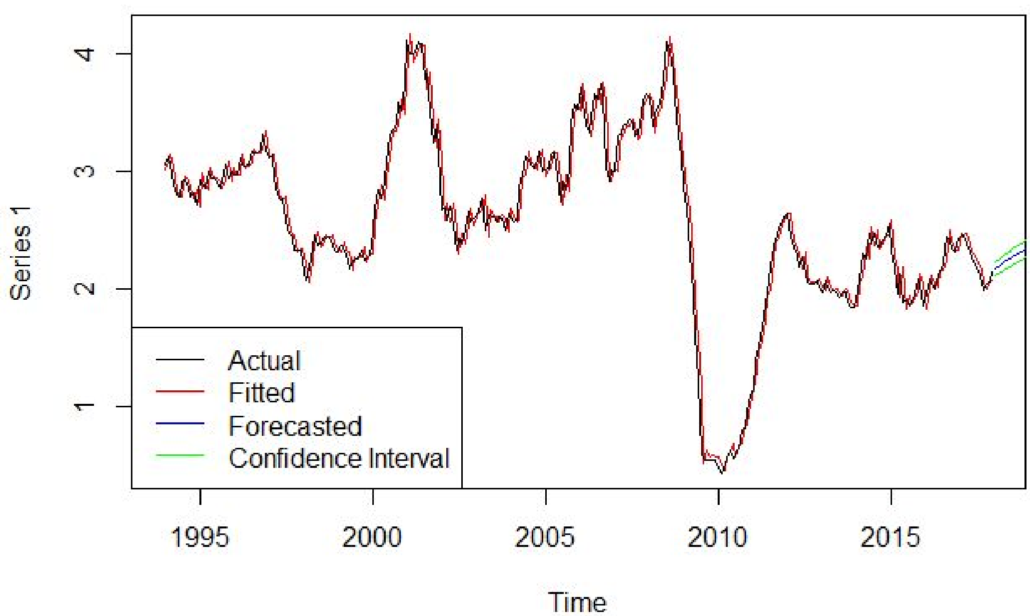



**Figure 19: Forecast RMSE Based Metric**

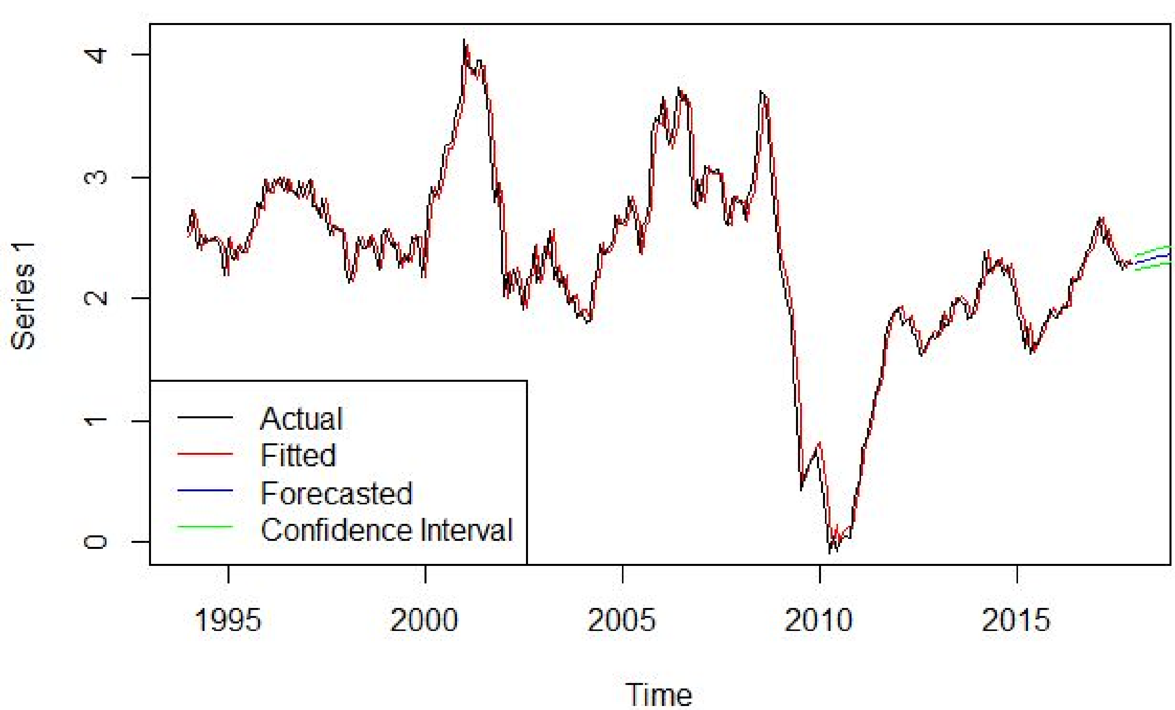


Next, we try to develop another metric using the top five components based on MAPE as seen in Table 3 below. This includes the following components: recreation, apparel, education & communication, housing, and other goods & services. This contains approximately 60% of full CPI. Similarly to the RMSE based metric, we fit an sGARCH(1,1)ARIMA(1,0,1) and get an RMSE of 0.32 and a MAPE of 0.10. The forecast is provided in Figure 20. Again, this is worse than our original forecast that included all CPI components.

Table 3: MAPE Table

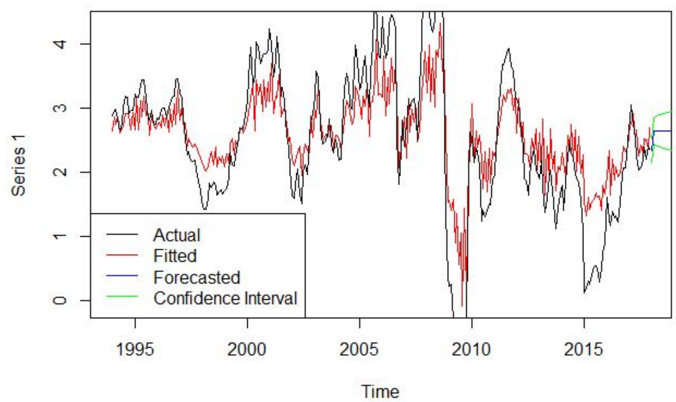
| CPI Component | MAPE |
|---|---|
| Recreation | 1.85 |
| Apparel | 2.69 |
| Education and Communication | 3.02 |
| Housing | 3.38 |
| Other Goods and Services | 4.26 |
| Medical Care | 4.68 |
| Food and Beverages | 5.01 |
| Transportation | 16.17 |

**Figure 20: Forecast MAPE Based Metric**



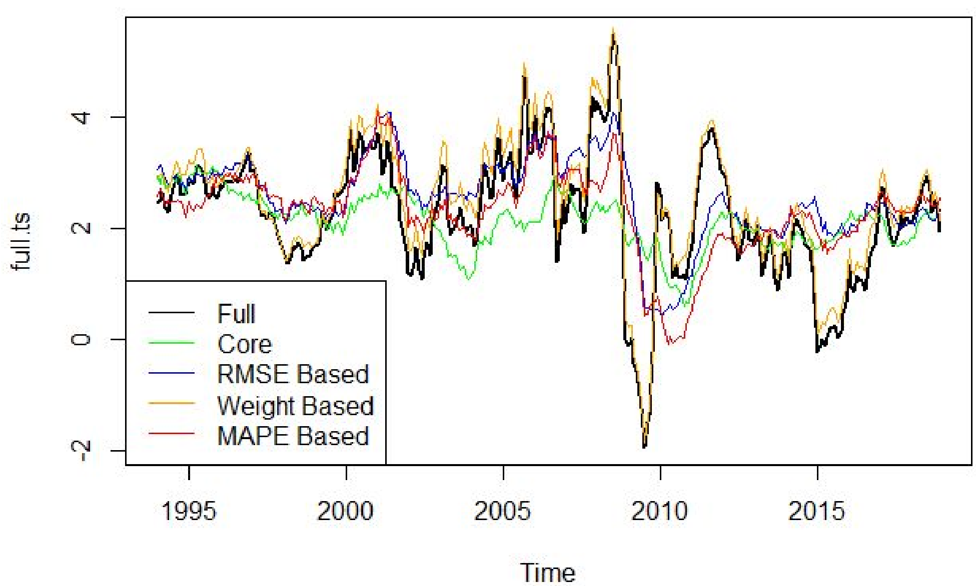

One thing we need to consider about previous metrics that we built is that they did not include components that consisted of a large proportion of full CPI inflation. This is likely influencing the forecast as it is not a representative basket. Therefore, we decide to build a new metric containing the top 5 components based on weight. More specifically, we look at components that contribute the most towards CPI. This includes food & beverages, housing, transportation, medical care, and education & communication. We fit an sGARCH(1,1)ARIMA(0,0,1) and get an RMSE of 0.34 and a MAPE of 0.13. This is seen in Figure 21 below.

**Figure 21: Forecast Weight Based Metric**




Visually, we can compare our different metrics in Figure 22.

**Figure 22: Full, Core, & New CPI Metrics**



Finally, we forecast using both a Holt Winters and an ETS model which gives the same accuracy measures. We get an RMSE of 0.32 and a MAPE of 0.11. We compare our metrics in Table 4.

Table 4: Model Comparison Table

| Model type | Model fit | RMSE | MAPE | Weight |
|---|---|---|---|---|
| Core CPI | ARIMA(2,1,0)(1,0,0) with drift | 0.57 | 0.20 | 86% |
| SPF | NA | 0.29 | 0.09 | 86% |
| New Core CPI (all components of full CPI) | sGARCH(1,1)ARIMA(1,0,1) | 0.31 | 0.10 | 100% |
| Top 5 components based on RMSE | sGARCH(1,1)ARIMA(1,0,1) | 0.36 | 0.11 | 77% |
| Top 5 components based on MAPE | sGARCH(1,1)ARIMA(1,0,1) | 0.32 | 0.10 | 60% |
| Top 5 components based on weight | sGARCH(1,1)ARIMA(0,0,1) | 0.34 | 0.13 | 88% |
| Additional | ETS and Holt Winters | 0.32 | 0.11 | 100% |

**Future Work**

In the future, we would like to apply different weights on each of the components to see if that would improve the overall accuracy of the forecast. Additionally, we would like to try implementing a model with time varying weights on these components to see if this improves the forecast accuracy. Furthermore, we would try different forecasting models such as a VAR model to incorporate other variables that may be impacting inflation. Finally, it would be interesting to see if a combination of forecasts from different models would provide a better forecast overall.

**Conclusion**

In conclusion, we found that including food and beverage prices to core CPI improves the forecast accuracy for full CPI inflation. This shows that price volatility of food and beverages has decreased significantly over time and is now considered to be a good predictor of full CPI inflation. As a result, central banks and forecasters may be able to improve their inflation forecasts by incorporating food and beverage prices into their models. This has implications for monetary policy as a more accurate forecast will allow for better informed decisions.

**REFERENCES**


Board of Governors of the Federal Reserve System. (n.d.). What is inflation and how does the Federal Reserve evaluate changes in the rate of inflation? Retrieved from https://www.federalreserve.gov/faqs/economy_14419.htm.

Dolmas, J., & Wynne, M. (2008). Measuring Core Inflation: Notes from a 2007 Dallas Fed Conference. *Federal Reserve Bank of Dallas, Staff Papers,2008*(4). Retrieved from https://www.dallasfed.org/~/media/documents/research/staff/staff0802.pdf.

Gavin, W. T., & Mandal, R. J. (2017, November 24). Predicting Inflation: Food For Thought. Retrieved from https://www.stlouisfed.org/publications/regional-economist/january-2002/predicting-inflation-food-for-thought

Federal Reserve Bank of Philadelphia. (n.d.). Historical Data Files for the Survey of Professional Forecasters. Retrieved from https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/data-files/cpi

Owyang, M. (2016, May 31). Should Core Inflation Be Measured Differently? Retrieved from https://www.stlouisfed.org/on-the-economy/2015/august/core-inflation-measure-food-pcepi

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCSL, March 3, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items Less Food and Energy [CPILFESL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPILFESL, March 3, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Apparel [CPIAPPSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAPPSL, March 4, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Education and Communication [CPIEDUSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIEDUSL, March 4, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Food and Beverages [CPIFABSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIFABSL, March 3, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Housing [CPIHOSNS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIHOSNS, March 4, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Medical Care [CPIMEDSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIMEDSL, March 4, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Other Goods and Services [CPIOGSSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIOGSSL, March 4, 2019.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Recreation [CPIRECSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIRECSL, March 4, 2019.

U.S. Bureau of Labor Statistics. (n.d.). [Table 1 (2015-2016 Weights). Relative importance of components in the Consumer Price Indexes: U.S. city average, December 2018]. Raw data.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: Transportation [CPITRNSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPITRNSL, March 4, 2019.

U.S. Bureau of Labor Statistics. (n.d.). Consumer Price Index Frequently Asked Questions. Retrieved from https://www.bls.gov/cpi/questions-and-answers.htm#Question_2