

What is Driving U.S Inflation Higher?

Date: April 2022

Team 1: Team Inflate

Members: Dylan Jin, Srila Maiti, Maria Manna, Joseph Roberts

Github Repo: https://github.com/UC-Berkeley-I-School/Project2_Jin_Maiti_Manna_Roberts

Overview & Research Question

As of February 2022, we have seen inflation rising globally. Specifically, within the United States, the Consumer Price Index (CPI) inflation rate hit 7.9%⁽¹⁾, an inflation release we have not seen since January 1982. Rising inflation is problematic for society because as prices rise, wages typically do not keep up. This causes a squeeze on everyday people and their spending. If persistently high inflation continues, the probability of entering a recession is greatly increased⁽¹²⁾.

Our goal in this project is to dive into the CPI data published by the U.S. Bureau of Labor Statistics (BLS) and answer the question, **what is driving U.S. inflation higher?**

Describing the Data

Our main data source is published by the BLS. The BLS measures labor market activity, working conditions, prices changes, and productivity in the U.S economy⁽⁵⁾. A key measure the BLS produces is the Consumer Price Index (CPI). The CPI represents the prices paid by U.S. consumers for a basket of goods and services. Each month the BLS records roughly 80,000 prices, representing a sample of all goods and services⁽⁶⁾. This data is then weighted and aggregated into an aggregate price index.

Data Assumptions and Processing

In this section we discuss accessing the data, processing the data, and its structure.

We accessed data from the BLS through the data repository because of its ease of access programmatically and because it has all of the relevant inflation data we need.

Data Decisions and Assumptions

- The dataset was already curated and needed limited transformations.
- Specific transformations added were calculations of the year-over-year and month-over-month inflation rates and converting our year-month information to a date-time object.
- We chose to only incorporate the monthly data provided by our dataset, and excluded aggregated time measures like annual and semi-annual data.
- There were trailing white spaces in some fields that we removed. The detailed process we used is below under Cleansing and Processing.
- We pre-processed the data to facilitate easier access for all team members.

Cleansing and Processing

We broke the process into two sections. In the first section, we cleaned and aggregated all of the data into a master Pandas DataFrame and then saved the object as a .pickle file. In the second section, we pulled the series that we were interested in, into simple files for the team to easily access. For a detailed description of the data cleansing process steps, see the Appendix.

Challenges

The main issue we had with the BLS data was understanding where to start. The data starts at a simple high-level aggregated price series and then goes very granular into the baskets. We chose to start with the most recently published weights to gain a deeper understanding of what baskets and levels we would be interested in for our analysis. You can see our reconciled weights here: [BLS Index Weights](#). In this file, you will also see our data mapping to the series we used in our analysis.

Some additional challenges we encountered outside of the challenges noted above in the Data Decisions and Assumptions section are listed below:

- The data in the basket subcategories did not always have the same date ranges. For example, one of the three sub-items for food at home had data that started in 1969 while the other two sub-items had data that started decades earlier. This caused some problems drilling down into the subcategories.
- The sampling rate was not consistent. Most ranges contained monthly data, some bi-monthly, some bi-monthly at different intervals, others seasonal.
- The BLS data we utilized was not seasonally adjusted, but many complementary data sources did have seasonally adjusted data. This made it challenging to find data that could be easily used with our BLS data.

Analysis

The CPI All-Items is a top-level index that consists of roughly 80,000 items. We start our analysis by exploring the top-level index first, and providing our findings. Next, we dive into the main 4 sub-categories broken down by food, energy, commodities less food and energy, and services less energy and services. In the below table we see the 4 subcategories' weights summing to 100%, representing the All-Items Index reported by BLS.

Index	Index Weight (%)
All Items	100
Food	13
Energy	8
Commodities less food and energy	22
Services less energy services	57

In Figure 1, we see historical inflation rates overlaid with periods of recession (in gray), which can at times be explained by major events such as world wars and changes in government policy. For example, because of the events after Covid-19, we have seen inflation increase to levels that we have not seen since 1982. There are many themes today that mirror the Great Inflation Period from 1970 to 1980s ⁽³⁾.

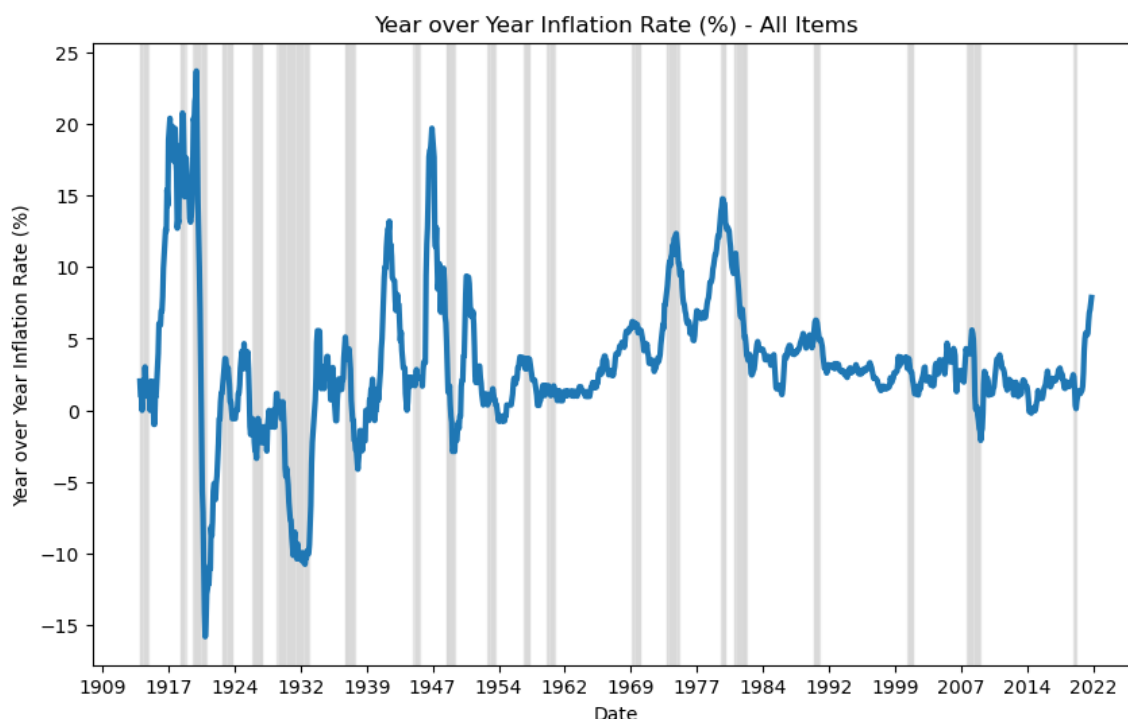


Figure 1: Year over Year Inflation - All Items with Recession Overlay

Next, we will dive into each of the 4 underlying categories to understand what is causing the higher inflation.

Food (Including Food at Home and Food Away from Home)

The first category we will review is food, which accounts for 13.4% of the main index. Food is broken up into two main subcategories – food at home and food away from home. This section will explore food as a possible contributor to the recent rise in inflation.

Index	Index Weight (%)
All Items	100
Food	13.4
Food at Home	8.2
Food Away from Home	5.2

From 1914 to 1950, food was fairly volatile. Since then and with very minor exceptions, food has been generally stable, hovering between a 0% and 0.1% change in inflation. Of the two subcategories of food, food at home has greater historical variability. A potential reason for this is that the foodservice industry is able to purchase larger quantities at economies of scale, helping to stabilize rapid peaks and dips in food inflation. This is demonstrated in Figure 2 below.

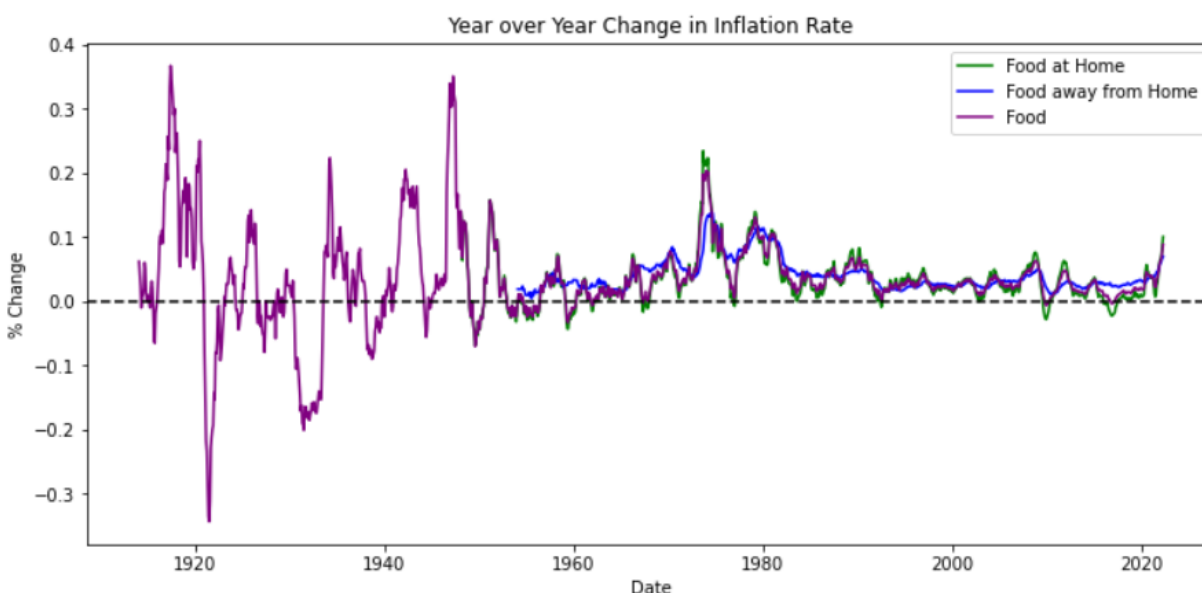


Figure 2: Year over Year Inflation - Food and Food Subcategories

When comparing food to the main index, the basic trends align with each other as shown in Figure 3. This reinforces the theory that food is a contributing factor to the rise and fall of inflation.

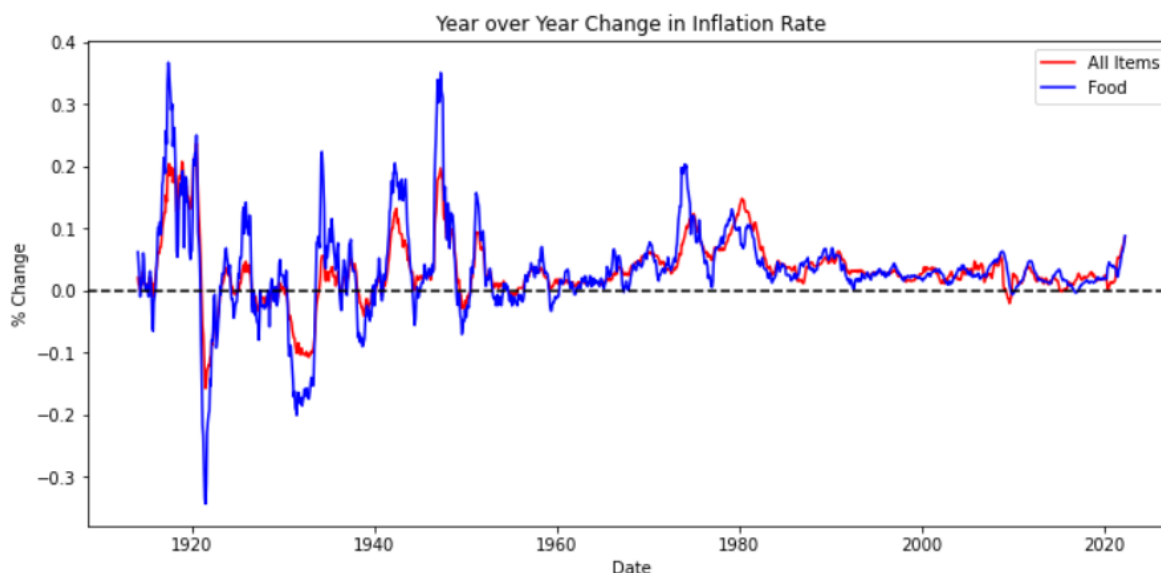


Figure 3: Year over Year Inflation - All Items vs. Food

Energy (Energy Commodities and Energy Services)

Energy has a significant impact on our daily lives. Energy prices impact the cost of heating our homes, filling the gas tanks of our cars, and virtually everything we consume, including our groceries, the clothes we wear, and the electronic devices we use.

Energy Commodities:

The primary energy commodities are crude oil, gasoline, heating oil, coal, natural gas, electricity, ethanol, and uranium ⁽⁹⁾.

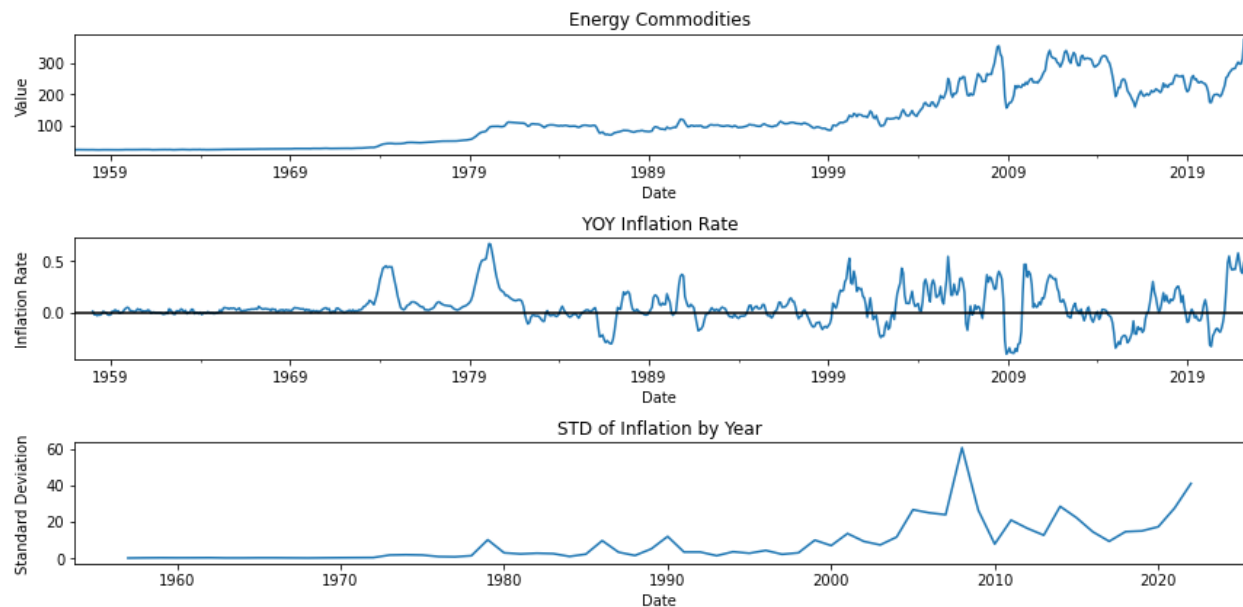


Figure 4: Value, Inflation Rate, and Standard Deviation - Energy Commodities

In Figure 4, there was a noticeable spike in inflation in the mid-1970s due to the OPEC oil embargo against the US in 1973. We also see an inflation spike in the early 1980s due to decreased domestic production and “Black Sunday.” There was a noticeable dip in 1986 from the crash of oil prices. We can see increasing volatility in energy commodities since 2000 from the standard deviation chart. More recently, we have seen the largest drop in inflation due to the great recession of 2008. Presently, in 2022, we see inflation has increased since the initial dip from reduced travel at the start of the Covid-19 pandemic.

Energy Services:

Energy services play a vital role in the production, transformation, transportation, distribution, and sale of energy.

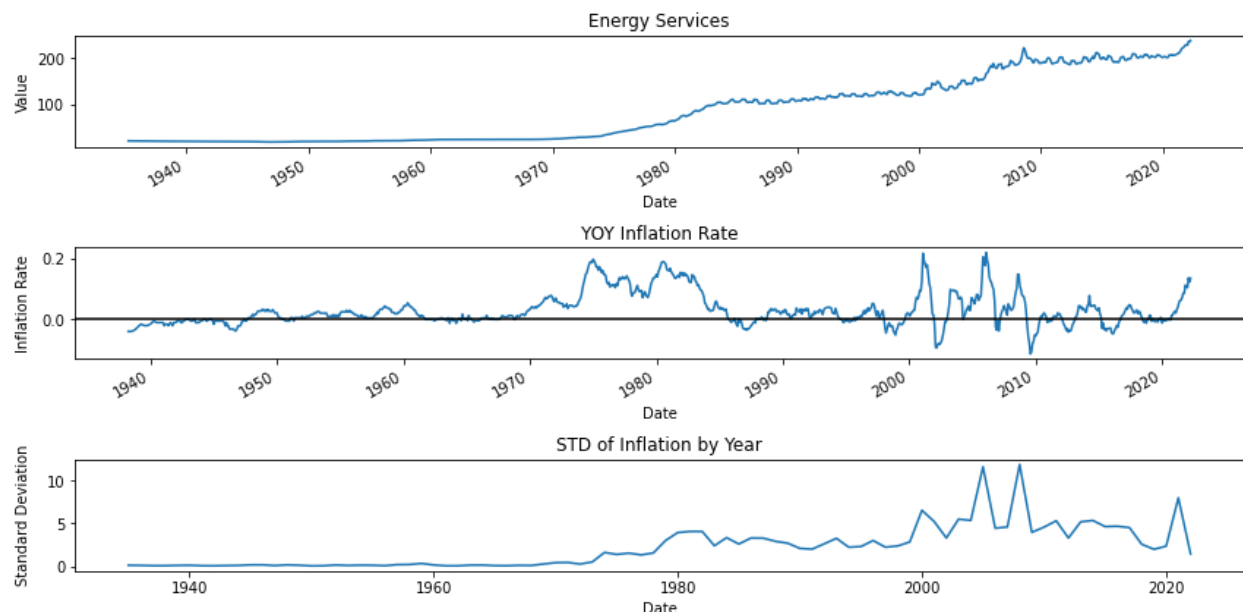


Figure 5: Value, Inflation Rate, and Standard Deviation - Energy Services

In Figure 5, we can see a slight increase in the fluctuation of inflation in energy services since the 1973 OPEC embargo. The trend toward the privatization and liberalization of national energy markets since the 1990s has increased the participation of service providers from the private sector ⁽¹⁰⁾. Clean energy technologies and

innovation have led to a relatively stable and steady increase in energy services, relatively unaffected by the volatile changes in energy commodities. Much of that is likely due to energy service providers expanding beyond petroleum and the continued push for clean and renewable energy.

Energy Commodities and Services vs All Items:

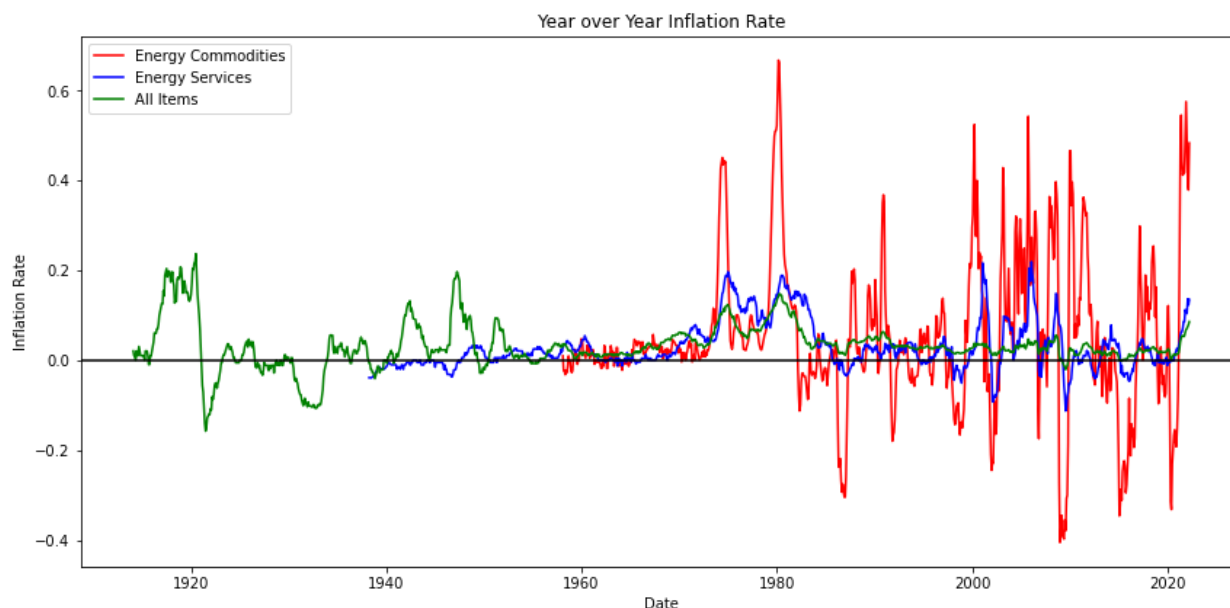


Figure 6: Inflation Rate - All Items, Energy Commodities, and Energy Services

In Figure 6, we can see while the inflation of Energy Services closely matches All Items, the inflation rate of Energy Commodities has greatly diverged since the mid-1970s, often several-fold the general inflation rate.

Commodities Less Food and Energy Commodities

In this section we will dive into commodities less food and energy (commodities).

Index	Index Weight (%)
All Items	100
Commodities less food and energy	21.8
Transportation commodities less motor fuel	8.8
Household Furnishings and Supplies	4.0
Apparel	2.5
Other Items	6.5

While historically we see commodities typically being lower than the headline inflation, Figure 7 shows commodities being a large contributor to the recent rise in headline inflation.

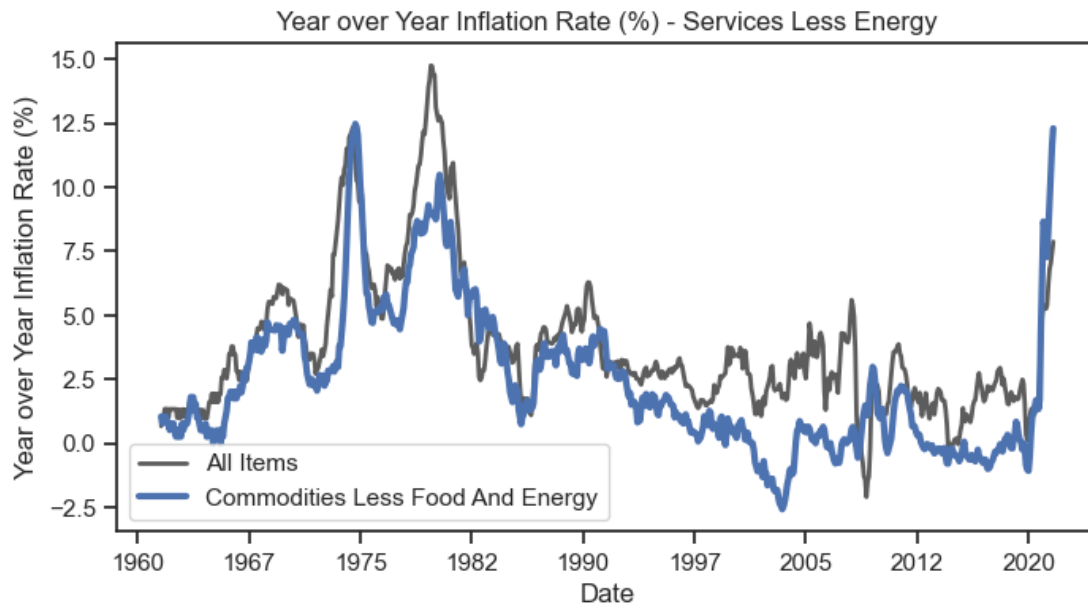


Figure 7: Inflation Rate - All Items Vs Commodities Less Food and Energy

In the below table we can see the latest inflation release for commodities was significantly greater than its historical average, comparing 12.3% versus 2.3%. While each subcategory is a contributor to this, the most significant contribution comes from transportation, with the latest release at 23.9% versus its historical average of 2.1%.

	all_items	commodities_less_food_and_energy	transporation	household_furnishings	apparel
last_value_inflation	7.9	12.3	23.9	10.3	6.6
mean	3.2	2.3	2.1	-0.1	1.9
std	5.0	2.8	5.2	2.3	6.4
min	-15.8	-2.6	-1.8	-3.1	-39.6
median	2.6	1.5	0.4	-0.9	1.4
max	23.7	12.5	23.9	10.3	82.0

Services less Energy Services

We will first look at the services less energy services subcategory, which represents 57% of the main index. Next, we will explore the top 5 underlying indices by weight: shelter, medical care, transportation, education and communication, and recreation.

Index	Index Weight (%)
All Items	100
Services less energy services	57
Shelter	32.8
Medical care services	7.0
Transportation services	5.6

Education and communication services	5.5
Recreation services	3.2
Other Items	2.8

Focusing only on the recent rise in inflation in Figure 8, we see that inflation from services has not risen as high as the main inflation rate. This tells us that inflation from services is not the main contributor to the high inflation that is being reported.

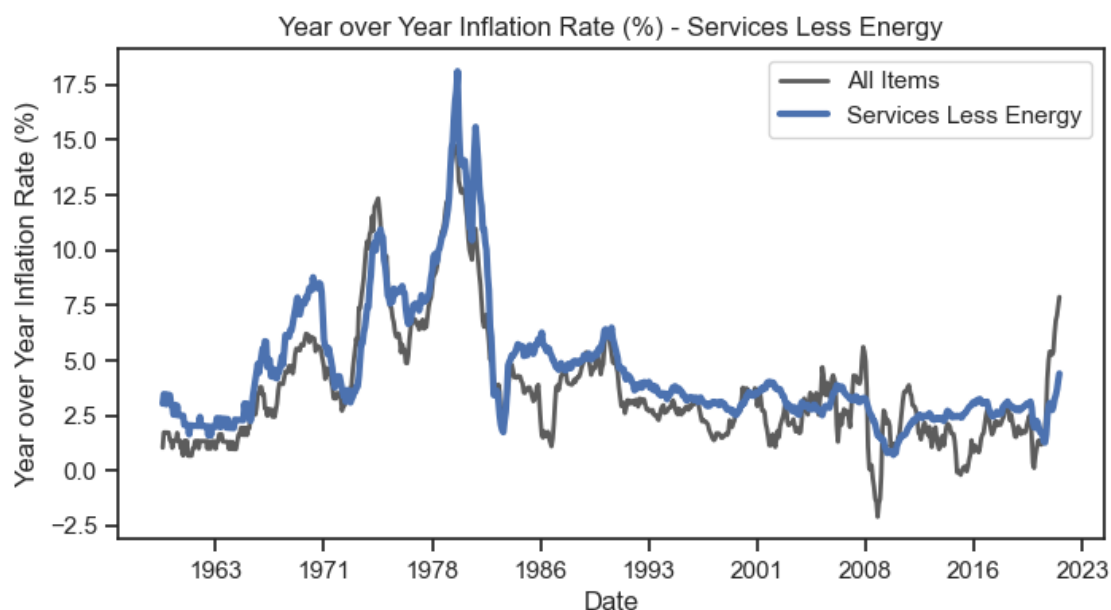


Figure 8: Comparing Services Less Energy Services to Main Inflation Rate

In the below table, you can focus on the latest inflation release for each category and compare it to its historical distribution. Focusing on services less energy we see the last print being close to its historical average, providing further evidence that services are not the leading cause of headline increases in inflation.

	all_items	services_less_energy	shelter	medical_care_services	transportation_services	education_communication_services	recreation_services
last_value_inflation	7.9	4.4	4.7	2.4	6.6	1.7	5.1
mean	3.2	4.5	4.1	6.1	4.8	1.3	2.4
std	5.0	2.8	3.2	3.6	5.0	1.2	0.9
min	-15.8	0.7	-0.7	0.8	-8.7	-2.6	0.3
median	2.6	3.5	3.3	5.0	3.6	1.6	2.3
max	23.7	18.1	20.9	24.2	38.2	2.9	5.1

Diving further into the table and into the subcategories we see most of the categories are above their historical average except for Medical Care Services. Medical Care Services is significantly below its average, 2.4% versus 6.1%.

During the great inflation period, 1970 looked to be driven by services moving higher, and the second wave in 1974 seemed to be across the board. Compared to now we are seeing significantly higher inflation from non-services items.

Conclusion

The recent increase in headline inflation has been driven mostly by commodity prices, and not services. We see increases in services overall but this seems normal for each of those categories. Within commodities, we see significant increases in transportation, which includes motor vehicles, parts and equipment, and other related items ⁽¹³⁾. We also see significant price increases in energy commodity prices, we suspect this to be driven by demand coming out of the Covid-19 pandemic, and the Russia and Ukraine conflict.

One area of future research is to dive deeper into what is causing higher commodity prices. A possible target of interest is globalization and its effects.

References

1. <https://www.bls.gov/news.release/cpi.nr0.htm>
2. <https://www.bls.gov/cpi/>
3. <https://www.federalreservehistory.org/essays/great-inflation>
4. <https://www.bls.gov/bls/history/home.htm>
5. <https://www.bls.gov/bls/faqs.htm>
6. <https://www.bls.gov/cpi/questions-and-answers.htm#:~:text=We%20record%20the%20prices%20of,for%20goods%20and%20services%20purchased.>
7. <https://fred.stlouisfed.org/>
8. <https://www.federalreserve.gov/monetarypolicy/monetary-policy-what-are-its-goals-how-does-it-work.htm#:~:text=The%20Federal%20Reserve%20Act%20mandates,for%20monetary%20policy%20is%20commonly>
9. <https://commodity.com/energy/>
10. https://www.wto.org/english/tratop_e/serv_e/energy_e/energy_e.htm
11. <https://www.worldometers.info/oil/oil-production-by-country/>
12. <https://www.nytimes.com/2022/03/17/business/federal-reserve-inflation-recession.html>
13. <https://www.bts.gov/components-consumer-price-index-transportation>

Appendix A

Cleansing and Processing

Section 1: Aggregating Data

The notebook for this section is located here: [1. process_timeseries_data.ipynb](#)

Step 1. Append all time-series files on top of each other, into a large DataFrame. For each of the files, remove the empty spaces from each item in the series_id column and clean up the column names.

Step 2. Bring in each of the Meta Information files to join to all time-series DataFrame. For each of the files, remove the empty spaces from each item in the series_id column and clean up the column names.

Step 3. Join each of the Meta Information files with the all time-series file, creating a master DataFrame called “all_items_enhanced”.

Step 4. Final cleaning. We remove unnecessary periods by filtering the DataFrame to only include the months of the year. Next, we create a DateTime column and normalize the monthly data to month-end periods. We then set a multi-index to series_id and date, for easy access later.

Step 5. Save the DataFrame as a .pickle and .csv file in the datastore folder.

Section 2: Individual Timeseries

The notebook for this section is located here: [2. process_single_timeseries_data.ipynb](#)

Step 1. We bring in the “all_items_enhanced.pickle” DataFrame into this notebook.

Step 2. Build a dictionary of series_ids associated with the file name we will save off.

Step 3. Loop through each time-series in the dictionary. Add 2 columns, a year over year and a month over month inflation rate calculated as a percentage change on CPI a price index.

Step 4. Save off the DataFrame as a .pickle for easy access.

Section 3: Example Data Access

The notebook for this section is located here: [3. Example_getting_data_and_plotting.ipynb](#)

We added an additional example notebook to help users understand how to access the data to perform their own analysis.