

The Impact of COVID-19 on the US Consumer Price Index (CPI)

Hanyang Shang *

The Harker School, Santa Clara, 95050, United States

* Corresponding author: 25hernys@students.harker.org

Abstract. This research investigates the profound impact of irregular commodity price changes during unusual circumstances, particularly focusing on the COVID-19 pandemic. The research stems from the recognition that global events have the potential to trigger irregular and drastic changes in commodity prices. This research holds significant implications for both consumers and policymakers in navigating the complexities of global crises. The research employs the Consumer Price Index (CPI) as a crucial indicator to chart drastic changes in consumer product prices from January 2019 to September 2023. To establish a baseline, CPI data from January 2010 to December 2018 is utilized, and an Autoregressive Integrated Moving Average (ARIMA) model is employed to forecast hypothetical CPI for the target period. The comparison between forecasted and observed CPI data reveals the vast influence of the COVID-19 pandemic on consumer prices. This research explores various subcategories of CPI, emphasizing sectors such as Food and Beverage, Transportation, Housing, and Medical Care. The research found that COVID-19 leads to delays and lagging effects in certain sectors reflecting the overall increase in consumer product prices. The Medical Care CPI also exhibited the most complexity, reflecting the immediate and direct impact of the pandemic on healthcare costs.

Keywords: CPI, ARIMA, COVID-19.

1. Introduction

Consumer Price Index (CPI) reflects the impact of commodity price changes on the daily lives of residents. Researchers continue to explore ways to improve the methodologies used in calculating and measuring the CPI. This includes considerations for the basket of goods, weighting mechanisms, and adjustments for quality changes. Research in this area explores how consumer behavior influences CPI. This includes studies on the impact of changes in consumer preferences, purchasing habits, and decision-making on CPI trends. Normally, CPI data time series can be expected to reflect seasonal changes in economic conditions. However, irregular and drastic commodity price changes can take place in unusual circumstances, such as during and after the COVID-19 pandemic, when extreme automobile sales prices and auto-service costs spiked right after the COVID-19 pandemic in the US. This study aims to use CPI to chart such drastic changes under unusual circumstances, so consumers can be better prepared when similar events happen. This is important as people are now undergoing several wars globally, which may impact the consumer market globally as natural disasters like COVID-19.

Recent development in time series data analysis methods includes using Machine Learning where the model is first trained on a dataset to fit the model parameters. The model is then applied to the target dataset for analysis such as forecasting [1]. This study utilizes a similar approach to the CPI time series data to focus on the target period of January 2019 - September 2023, when COVID-19 and its aftermath impact consumer product prices the most. To establish a baseline for comparative study, CPI data from January 2010 - December 2018 is used as a training dataset to fit the model and as the basis for projected CPI that would be expected for January 2019 - September 2023. This basis period is selected as the markets have largely recovered from the 2008 global financial crisis and been relatively stable. There are various economic models for CPI forecasting, such as VAR, FAVAR, Bayesian VAR models, and Error Correction models [2]. As ARIMA is identified as one of the most widely used models for CPI forecast [3, 4], it is used in this study to forecast the hypothetical CPI for the target period. The “forecasted” hypothetical CPI is then compared with the observed CPI data. Although the SARIMA model can account for regular seasonal data fluctuations, it is not expected to

account for drastic changes in time series data, such as during COVID-19 natural disasters, or wars. This study aims to demonstrate the impact on the CPI of such events.

2. Methodology

2.1. Data Selection

The U.S. CPI data is downloaded from the U.S. Bureau of Labor Statistics website <https://www.bls.gov/cpi/>. The CPI data selected is All Urban Consumers (Current Series) (Consumer Price Index - CPI). The monthly data uses the base of 100 from 1980. The data provides a continuous monthly time series from the starting month and no additional preprocessing is required.

2.2. Model Selection and Fitting

The ARIMA model is selected for forecasting hypothetical CPI values. The `auto.arima()` function in R is used to automatically find the best model fitting. The method is first to fit the ARIMA model using the data from January 2010 to December 2018 for model training. The best-fitting ARIMA model is then used for “forecasting” the hypothetical CPI as if it would be expected without COVID-19 impact.

2.3. Comparative Study of CPI Subcategories

The U.S. CPI time series is divided into several subcategories, including food, housing, transportation, healthcare, education communication, etc. A comparative study is carried out between the subcategories of food, transportation, housing, and healthcare, which may most prominently reflect COVID-19 impact.

3. Data Processing and Analysis

For this study, CPI data will be selected from All Items, and subcategories in Food and Beverage, Transportation, Housing, and Medical Care. These subcategories represent the most fundamental aspects of resident expenditure during the COVID-19 pandemic. For each subcategory, the model training time series consists of monthly CPI data from January 2010 - December 2018, which is used for ARIMA forecasting. The observed time series consists of monthly CPI data from January 2010 - September 2023. The two-time series are compared in a chart to visualize the impact of COVID-19.

3.1. Analysis for All Items CPI

Use `auto.arima()` in R to obtain the ARIMA model for the All Items CPI time series from January 2019 - December 2018. The best-fitting model is ARIMA (0, 1, 1). The ARIMA forecast uses prediction intervals of 99.5% and 85.5%.

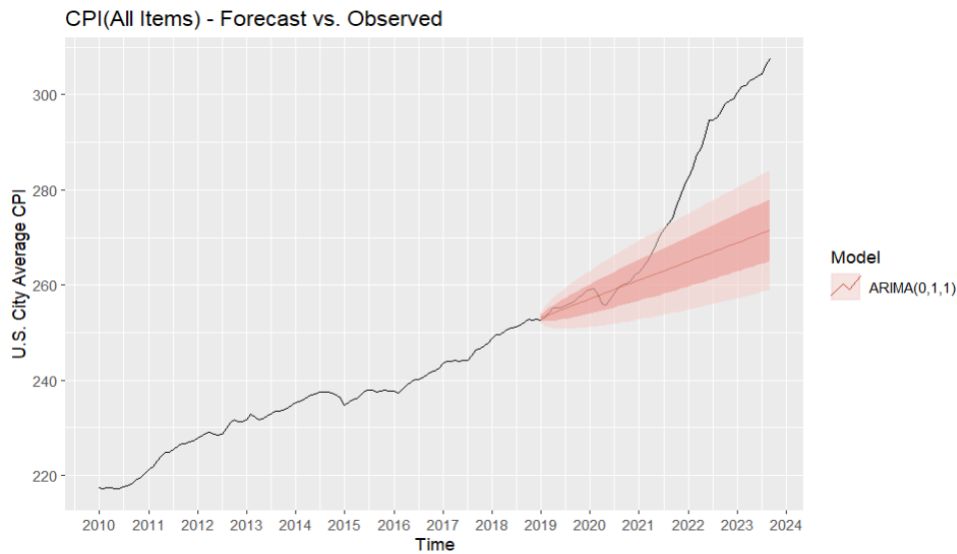


Figure 1. Forecast and observed for all items CPI (Photo/Picture credit: Original).

In Fig. 1, the trajectory of the observed All Item CPI shown in Fig. 1 can be divided into three segments. In the first segment (January 2019 - January 2020), the observed CPI largely traces the prediction with a slightly upward trend. The impact of COVID-19 isn't prominent for all consumer products combined. Some products may have experienced price increases due to higher demand and reduced production. In contrast, other products may have experienced price decreases due to a lowered level of demand under the pandemic situation.

In the second segment (January 2020 - June 2020), the All-Item CPI even trends downward, which may be explained by further economic slowdown and reduced demand and circulation of consumer products. However, this reduced level of demand appears to be temporary, as prices start to spike after June 2020. This CPI decrease can benefit consumers because of lowered prices of goods and services but can hamper regions with moderate economics, especially for producers [5].

In the third segment (June 2020 - September 2023), the All-Item CPI shows an accelerated upward trend. This can be explained as a result of increased consumer demand, while production and distribution are still slow as the pandemic continues. It is not until the end of 2022 that the accelerated upward trend starts to abate. As it's nearing the end of the pandemic, the economy starts to stabilize, and some sectors are aggressively compensating for business losses during the pandemic, such as travel, transportation, etc. The third segment demonstrates a significant upward trend of CPI which must be treated with caution by monetary and fiscal authorities for tight monetary and fiscal policy measures to avoid run-away inflation, as a current high level of inflation can often foster anticipation of more severe inflation [6].

3.2. Analysis of Food and Beverage CPI

Use `auto.arima()` in R to obtain the ARIMA model for the Food and Beverage CPI time series from January 2019 - December 2018. The best-fitting model is ARIMA (0, 2, 1). The ARIMA forecast uses prediction intervals of 99.5% and 85.5%.

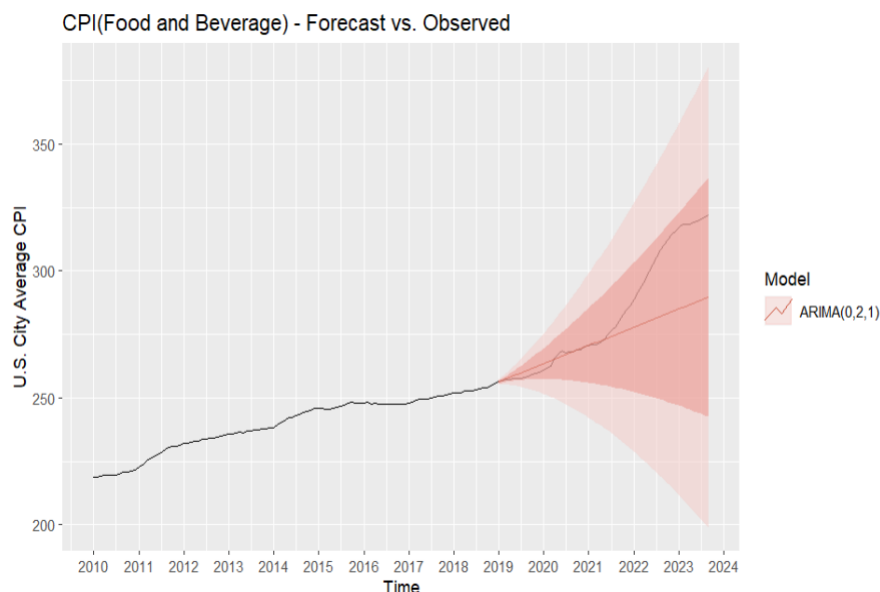


Figure 2. Forecast and observed for food and beverage CPI (Photo/Picture credit: Original).

In Fig. 2, the trajectory of the observed Food and Beverage CPI as shown in Fig. 2 can also be divided into three segments. In the first segment (January 2019 - June 2022), the observed CPI largely traces the prediction yet at a slightly lower level, with brief fluctuation during the second quarter of 2020. Food and beverage consumption was mostly impacted by the pandemic lockdown when many restaurants had to shut down or cut off in-person services.

In the second segment (June 2022 - June 2021), the Food and Beverage CPI very closely traces the ARIMA forecast line. This can be from the combined factor that while the restaurant business is still slow due to the lockdown policy, the steadily increasing demand from other food supply channels such as online shopping, and group shopping may offset the soft demand, and thus generate the CPI line aligns with the forecasted trend.

In the third segment (June 2021 - September 2023), like the observed All Item CPI, the Food and Beverage CPI shows an accelerated upward trend. The shape of this CPI trend mirrors that of the All-Items CPI very well, though mostly within the ARIMA forecast range. As the most fundamental component in consumer consumption, the Food and Beverage CPI mirrors the All-Item CPI closely to reflect the general fluctuation in economic conditions during the pandemic.

3.3. Analysis for Transportation CPI

Use `auto.arima()` in R to obtain the ARIMA model for the Transportation CPI time series from January 2019 - December 2018. The best-fitting model is ARIMA (0, 1, 1). The ARIMA forecast uses prediction intervals of 99.5% and 85.5%.

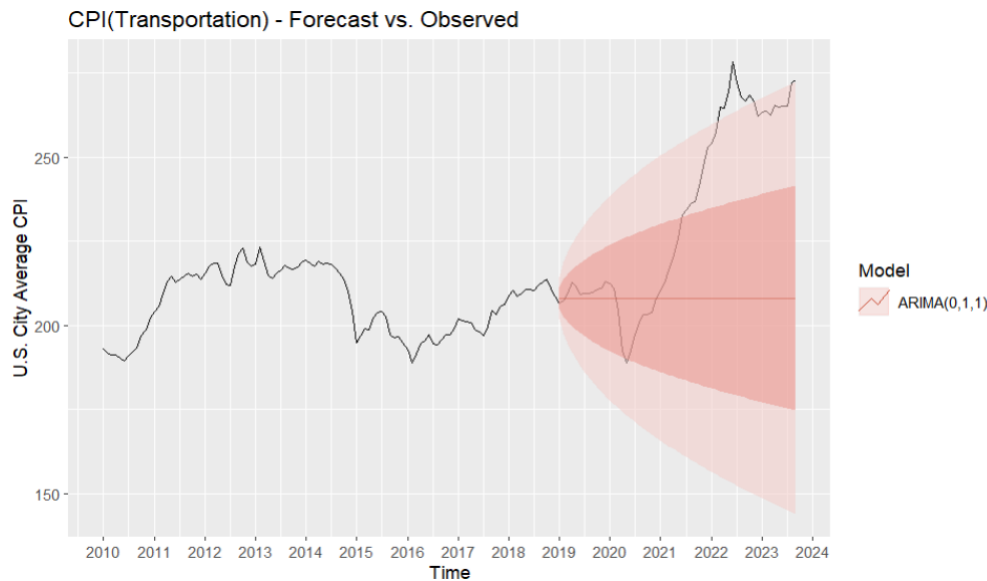


Figure 3. Forecast and observed for transportation CPI (Photo/Picture credit: Original).

The Transportation CPI as shown in Fig. 3 differs from other subcategory CPIs in that it shows more fluctuation historically, with a less apparent upward trend due to inflation. In the comparison chart, the trajectory of the observed Transportation CPI can be divided into three segments. However, the timeframes for each segment differ from those of Food and Beverage CPI. The transportation sector has also been affected by the lockdown policies for public transportation. Furthermore, private transportation is also reduced which isn't as hard a demand as food and beverage. Lots of workers switch to remote work to further reduce the demand for transportation and related products. In the US, the aviation and transport sectors account for more than half of the oil demand, and mobility constraints quickly translated into decreased oil consumption [7]. This is reflected in the sharp drop in Transportation CPI in the first half 2020.

In the second segment (June 2020 - June 2022), the observed CPI has been steadily climbing out of the AMIRA forecast range. This is very similar to the upward trend of Food and Beverage CPI, which reflects the economic recovery and increase in demand for transportation. However, due to policy constraints, especially those involving foreign parties, the transportation sector itself has undertaken a much slower recovery. It has been observed that the cost of ocean cargo liners has increased multiple times due to shortages, while the domestic markets for consumer goods are recovering much faster and pumping up the demand.

In the third segment (June 2022 - September 2023), the Transportation CPI starts to flatten out and show fluctuation more similar to historical patterns. As the policy starts to loosen up, airline and ocean transportation prices are beginning to return to pre-pandemic levels, though still elevated in comparison with historical price levels. The Transportation CPI is most subject to policy and regulations, especially in air and ocean transportation. The CPI curve corresponds well with the transportation policy changes during the pandemic.

3.4. Analysis of Housing CPI

Use `auto.arima()` in R to obtain the ARIMA model for the Housing CPI time series from January 2019 – December 2018. The best-fitting model is ARIMA (1, 2, 1). The ARIMA forecast uses prediction intervals of 99.5% and 85.5%.

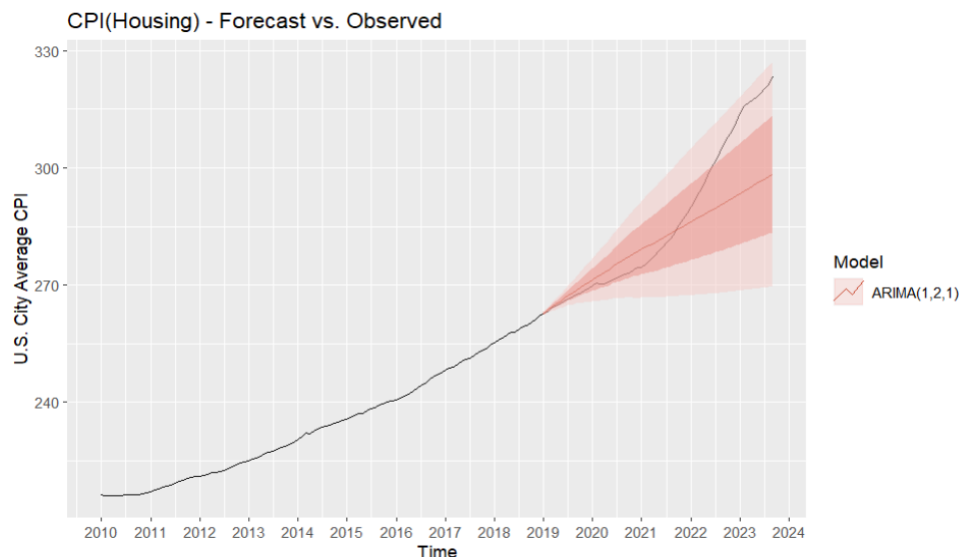


Figure 4. Forecast and observed for housing CPI (Photo/Picture credit: Original).

The general shape of the Housing CPI curve during the pandemic as shown in Fig. 4 mirrors that of the All-Items CPI as shown in Fig 1, which can be used as a proxy of the general economic condition. The Housing CPI shows two distinctive segments in the first segment of January 2019 - January 2021. The Housing CPI steadily traces the forecasted curve but at a lower level in comparison. It means the housing market was soft during this period, but due to the limited liquidity of housing properties, the CPI curve remains relatively stable without violent fluctuations.

In the second segment of January 2021 - September 2023, the Housing CPI picks up a steady upward trend outpacing the forecasted CPI. This upward trend coincides with the general economic recovery as demonstrated in the All-Items CPI curve. However, the Housing CPI responds less violently and the entire observed CPI curve resides within the ARIMA forecasted range.

3.5. Analysis for Medical Care CPI

Use `auto.arima()` in R to obtain the ARIMA model for the Medical Care CPI time series from January 2019 - December 2018. The best-fitting model is ARIMA (2, 1, 0). The ARIMA forecast uses prediction intervals of 99.5% and 85.5%.

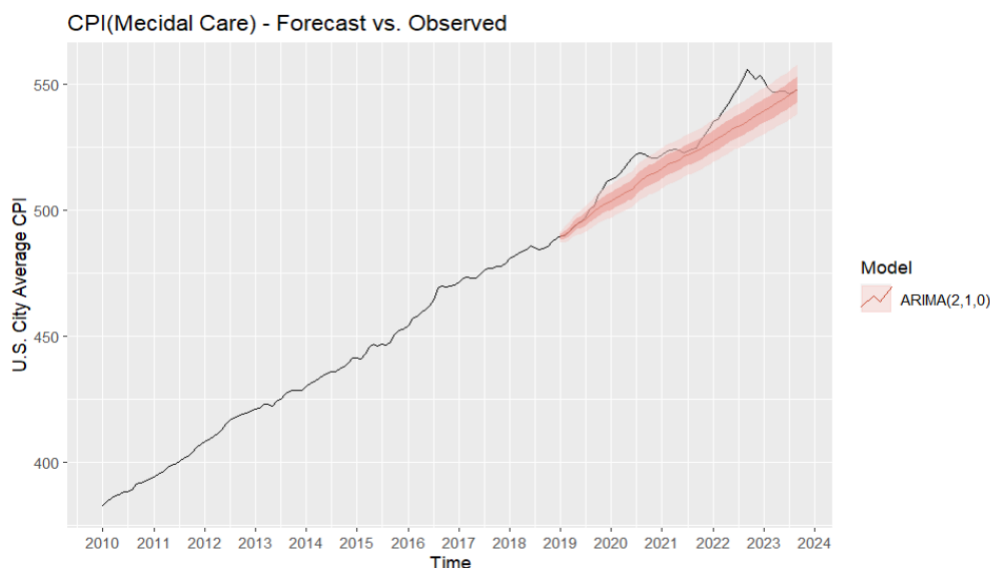


Figure 5. Forecast and observed for medical care CPI (Photo/Picture credit: Original).

The observed Medical Care CPI curve demonstrates a distinctly different shape compared with the CPI curves of all other subcategories studied so far. As the CPI time series has a self-driven internal

dynamic mechanism, this seemingly erratic curve as shown in Fig. 5 reflects the complex economic impact on the sector most closely and intricately related to the COVID-19 pandemic [8]. This sector is also the most heavily regulated and the consumer market condition is subject to many other complex factors. For example, the shortage of clinical care facilities, personnel, and supplies has distorted the consumer price of medical care products to a great extent at the beginning of the pandemic. Several waves of changes in governmental policy and the availability of vaccines and other medical supplies further distorted the market to make it difficult to identify any clear pattern, especially when compared with the forecast. It is worth breaking down the Medical Care CPI for a more detailed study at finer granular levels using the CPI data from the Bureau of Labor Statistics.

It's worth noting that all other subcategories of CPI have demonstrated a downward trend or lower-level curve compared with the forecast, which can be explained as a general or sectional slowdown at the beginning of the pandemic, the observed Medical Care CPI curve shows an elevated level above the forecast from the beginning. This can be attributed to the fact that the COVID-19 pandemic hits the medical care sector most directly and thus inflates consumer prices more directly and immediately.

4. Conclusion

This paper selects CPI monthly data of January 2009 - September 2023 for study and uses the period of January 2009 - January 2019 as a training dataset to fit ARIMA models for All Items CPI and several subcategory CPI respectively. Then the ARIMA model forecasts for the target period of January 2019 – September 2023 is compared with observed All Items CPI and subcategory CPI data to demonstrate the impact of the COVID-19 pandemic. The CPI comparison charts are analyzed by segments of curve trend, and interpreted within the U.S. economic contexts of that period to identify the economic factors contributing to the difference between the forecast and observed CPI data.

If people use the All-Items CPI as the general indicator, it demonstrates the vast impact of the COVID-19 pandemic on consumer product prices overall from January 2019 to September 2023. Several subcategory CPIs, such as Food and Beverage, Transportation, and Housing more closely follow the pattern of the All-Items CPI. It's interesting to observe the lagging effect of the pandemic impact on certain sectors, such as housing. Some sectors are more subjective to policy influence, such as transportation. The Medical Care CPI shows the most complexity because this sector is the most directly and immediately affected by the pandemic, and it is also subject to the many changes in public health policy during the pandemic. The pandemic impact on Medical Care CPI warrants further study at more granular levels to break down the CPI data into sub-sectors.

Prospects include studies that expand the scope to include international comparisons. Collaborative efforts among countries or regions could provide a more comprehensive understanding of how global events impact consumer prices on a broader scale, such as in the study of variations of crude oil prices at international levels. With advancements in data analytics and real-time data processing, exploring the development of real-time monitoring systems for CPI is another goal. Also, the development of big data and artificial intelligence will provide a more accurate method of CPI prediction. This would enable a more immediate response to unexpected events and faster dissemination of information to the public and policymakers. After monitoring the CPI data, understanding how to communicate CPI data and its implications to the public effectively is crucial. Future research could explore communication strategies that help consumers comprehend the significance of CPI changes and take appropriate actions.

References

- [1] Sarangi P K. A Literature Review on Machine Learning Applications in Financial Forecasting. *Journal of Technology Management for Growing Economies*, 2020, 11 (1): 23 - 27.
- [2] Shapovalenko N. A Suite of Models for CPI Forecasting. *Visnyk of the National Bank of Ukraine*, 2021 (252): 4 - 36.

- [3] Du Y, Cai Y, Chen M, et al. A novel divide-and-conquer model for CPI prediction using ARIMA, Gray Model and BPNN. *Procedia Computer Science*, 2014, 31: 842 - 851.
- [4] Ghazo A. Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy. *International Journal of Financial Research*, 2021, 12 (3): 70.
- [5] Imron M, Utami W D, Khaulasari H, et al. ARIMA Model of Outlier Detection for Forecasting of Consumer Price Index (CPI). *Jurnal Ilmu Matematika dan Terapan*, 2022, 16 (4): 1259 - 1270.
- [6] Nyoni T. ARIMA modeling and forecasting of Consumer Price Index (CPI) in Germany. 2019.
- [7] Nyga-Łukaszewska H, Aruga K. Energy prices and COVID-immunity: The case of crude oil and natural gas prices in the US and Japan. *Energies*, 2020, 13 (23): 6300.
- [8] Xiao Z. CPI Prediction Based on ARIMA Model. 7th International Conference on Education, Management, Information and Computer Science (ICEMC 2017). Atlantis Press, 2016: 135 - 138.