

Discover the Influence of Economic Variables on CPI

Ruiming Min<sup>1</sup>, Zicheng Wang<sup>2</sup>, Qi Yang<sup>3</sup>, & Shumeng Zhang<sup>4</sup>

<sup>2</sup> UIUC

Author Note

Enter author note here.

The authors made the following contributions. Ruiming Min: Conceptualization, Data curation, Resources, Formal Analysis, Methodology, Writing - Original Draft Preparation, Writing - Review & Editing; Zicheng Wang: role 3, role 4; Qi Yang: role 1, role 2; Shumeng Zhang: role 3, role 4.

## Abstract

One or two sentences providing a **basic introduction** to the the problem being addressed by this study. One sentence summarizing the main result. Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge. One or two sentences to put the results into a more **general context**. Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

## Discover the Influence of Economic Variables on CPI

**Introduction**

A brief introduction/motivation for the problem at hand, relevant details about the data, additional relevant scientific information, and what is to be addressed.

Citation example:

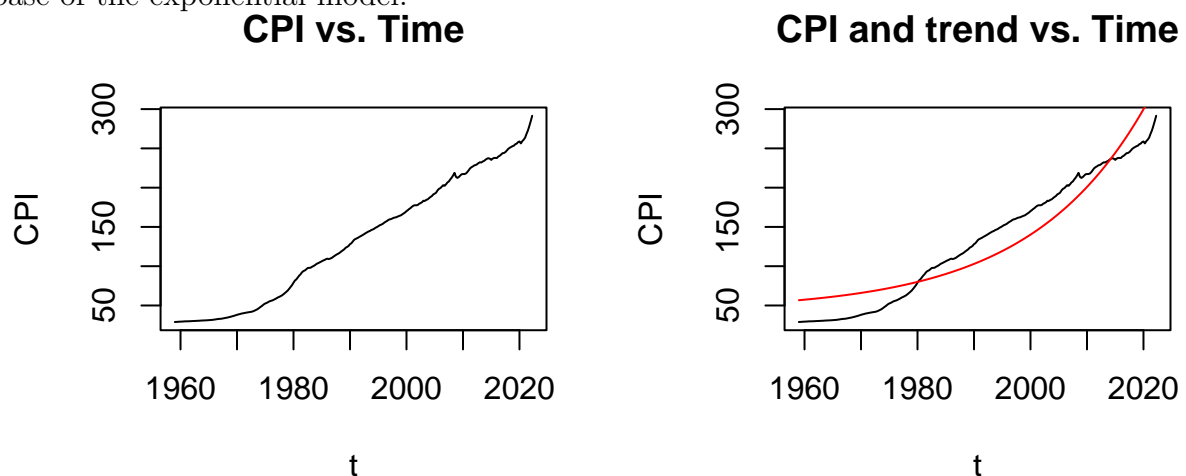
citation within parenthesis [tsa4]

**Methods**

To address the problem at hand, we used the following three steps to analyze the data: 1. detrending, 2. lag analysis, 3. adding the economic variables.

**Detrending**

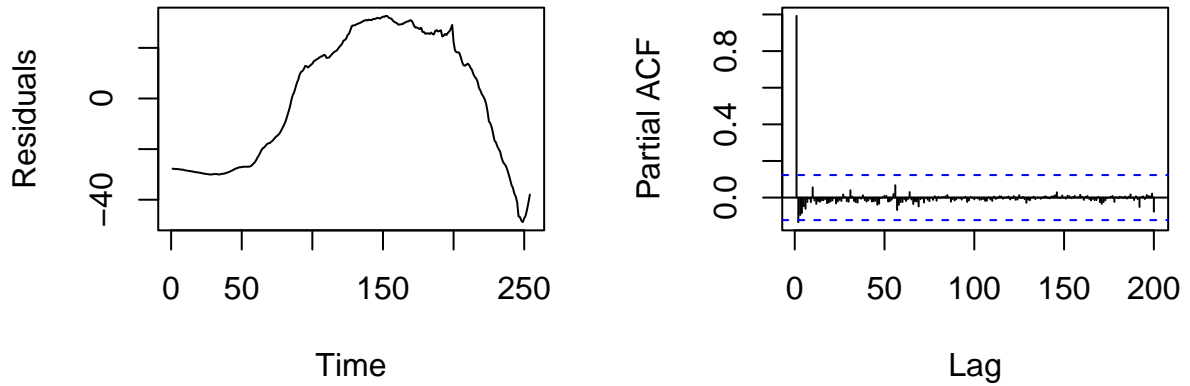
Observating the original data, the CPI has a significantly exponential trend. So we first use a exponential model to fit the data. Moreover, according to the nature of the CPI, which is highly related to inflation rate, and the average of inflation rate, we choose 1.05 as the base of the exponential model.

**Lag Analysis**

After detrending, we use PACF to figure out whether or not the data has lags. From the plot below, the data has one lag. Therefore, the CPI may be a  $AR(1)$  process and we

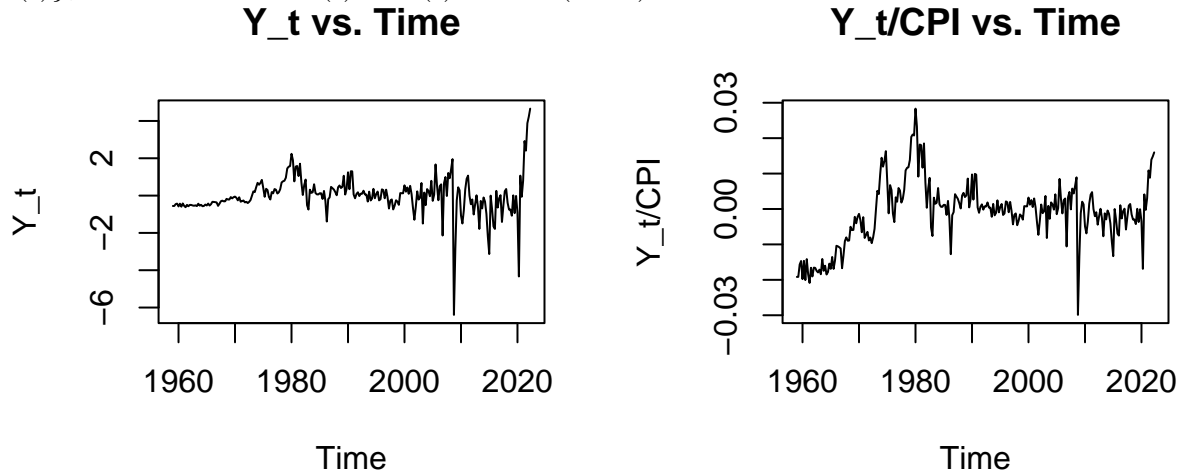
improve our model by adding the first order lag of CPI to the model.

### Residuals of detrending model vs. -PACF of Residuals of detrending m



### Adding the economic variables

After the detrending and adding the first order lag of CPI, the time series of CPI,  $\{CPI(t)\}$ , has become  $CPI(t) = T(t) + CPI(t - 1) + Y_t$ .



The plot above shows that  $Y_t$  still does not behave like a white noise. Additionally, the significant depressions at 2008 and 2020 and the significant peak at 1979 reflected the influence of the economic crisis [gross2019iran] [williams2010uncontrolled] [forbes\_2019\_strange\_new\_world]. So we add some economic variables to the model to regress the final model. Also, we add the sine functions of time to the model to fit the seasonal effect of the business cycle.

According to the nature of CPI [blanchard2004macroeconomics],

$CPI(t) = \frac{\sum_{i \in \mathcal{P}} P_{i,t} Q_i}{\sum_{i \in \mathcal{P}} P_{i,0} Q_i}$ , where  $\mathcal{P}$  is the set of all goods and services,  $P_{i,t}$  is the price of good or service  $i$  at time  $t$ , and  $Q_i$  is the quantity of good or service  $i$  at the base period. Since the Wage-Setting Relation,  $W = \mathcal{A}P^e F(u, z)$ , where  $W$  is the nominal wage,  $\mathcal{A}$  is the price mark-up,  $F(u, z)$  is the function of unemployment rate and other variables,  $u$  is the unemployment rate, and  $z$  is the output(GDP), and the Price-Setting Relation,  $P = (1 + m)\frac{W}{\mathcal{A}}$ , where  $m$  is the desired mark-up, it is concluded as

$$CPI(t) = \frac{\sum_{i \in \mathcal{P}} (1 + m_i) \frac{W_i}{\mathcal{A}} Q_i}{\sum_{i \in \mathcal{P}} P_{i,0} Q_i} = \frac{\sum_{i \in \mathcal{P}} (1 + m_i) P_{i,t}^e F(u, z) Q_i}{\sum_{i \in \mathcal{P}} P_{i,0} Q_i}$$

.

To simplify the formular, we assume  $m_i = m \forall i$  and  $P_{i,t}^e = P_{i,t-1}$ , we get  $CPI(t) = CPI(t-1)(1 + m)F(u, z)$ . By the assumption of  $m$  is small and  $F(u, z) = 1 - \alpha u + \beta z$ , we get  $CPI(t) = CPI(t-1)(1 + m)(1 - \alpha u + \beta z) \cong CPI(t-1)(1 + m - \alpha u - \beta z)$ . We could conclude  $Y_t = CPI(t) - CPI(t-1) = CPI(t-1)(m - \alpha u + \beta z)$ . For easy-computation, we assume  $CPI(t-1)$  is a constant in the cross term, we get  $Y_t = M - \mathcal{A}u + \mathcal{B}z$ . Therefore we add the unemployment rate, the real consumption, real grovement spending, and the real investment to the model.

To summarize the model, we get the final model as follows:

$$\begin{aligned} CPI(t) = & a_0 + a_1 \cdot 1.05^t + a_2 CPI(t-1) && \text{(trend and lags)} \\ & + a_3 \sin\left(\frac{2\pi(t-1959)}{40}\right) + a_4 \sin\left(\frac{2\pi(t-1980)}{24}\right) + a_5 \sin\left(\frac{2\pi(t-1980)}{48}\right) && \text{(business cycle)} \\ & + b_1 C_t + b_2 u_t + b_3 G_t + b_4 I_t + W_t && \text{(economic variables)} \end{aligned}$$

where  $W_t$  is a white noise.

## Results

### Fitting Results

The fitted result of the model is shown below. The plot shows that the model fits the data well.

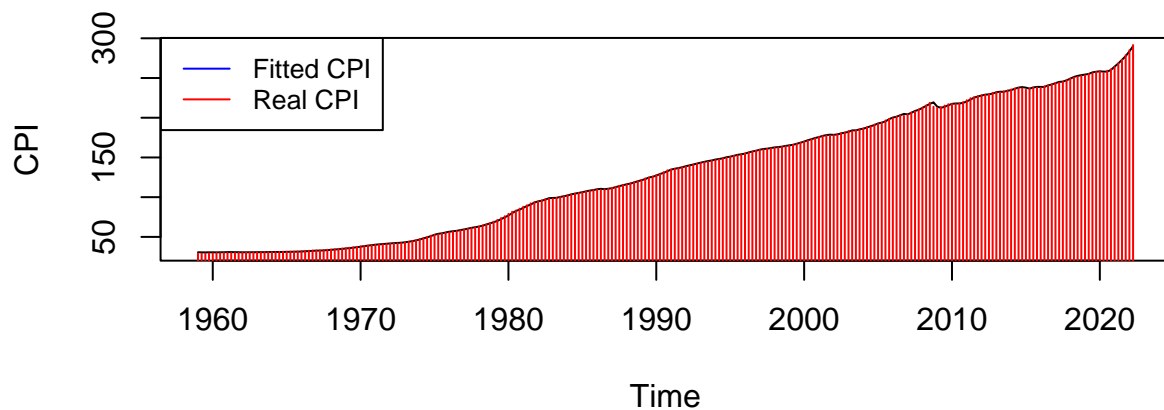
Coefficients	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.159	4.662	4.632	5.90 $\times 10^{-6}$
I(1.05 <sup>t</sup> )	$-8.779 \times 10^{-42}$	1.245	-7.050	1.83 $\times 10^{-11}$
sin_1	-3.881	6.540	-5.934	1.01 $\times 10^{-8}$
sin_2	-1.028	1.593	-6.455	5.80 $\times 10^{-10}$
sin_3	-3.760	9.484	-3.965	9.66 $\times 10^{-5}$
CPI_lag1	$9.735 \times 10^{-1}$	1.351	72.033	<2 $\times 10^{-16}$
unemp_rate	$1.847 \times 10^{-1}$	5.406	3.416	0.000 745
grove_exp	$6.250 \times 10^{-4}$	2.178	2.870	0.004 470
invest	$7.797 \times 10^{-4}$	3.737	2.087	0.037 966
consump_real	$3.423 \times 10^{-3}$	3.337	10.259	<2 $\times 10^{-16}$

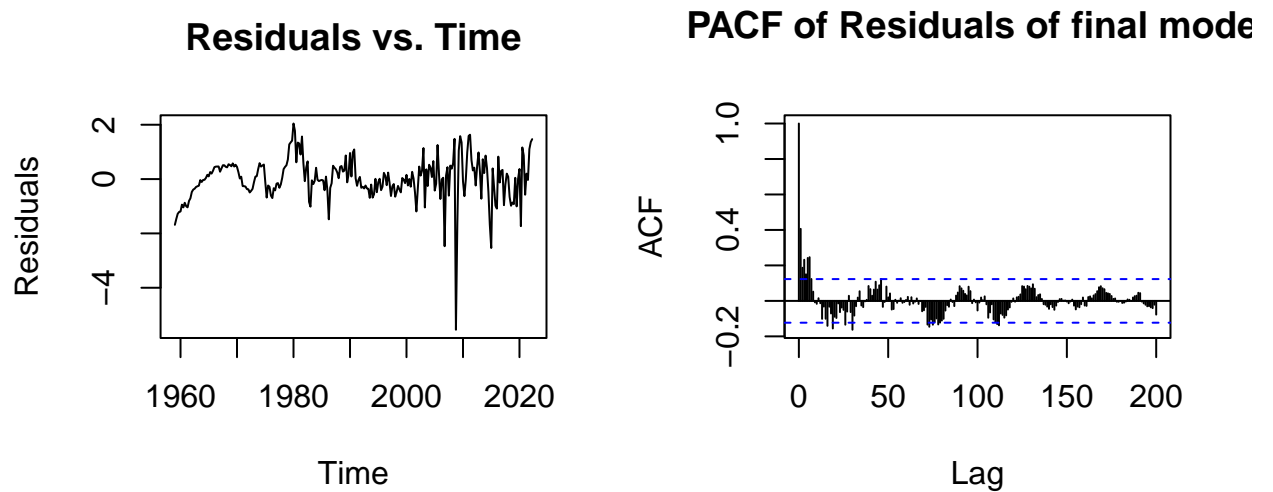
Residual standard error: 0.8011 on 244 degrees of freedom

Multiple R-squared: 0.9999, Adjusted R-squared: 0.9999

F-statistic: 2.688e+05 on 9 and 244 DF, p-value: < 2.2e-16

### Fitted CPI & real CPI vs. Time

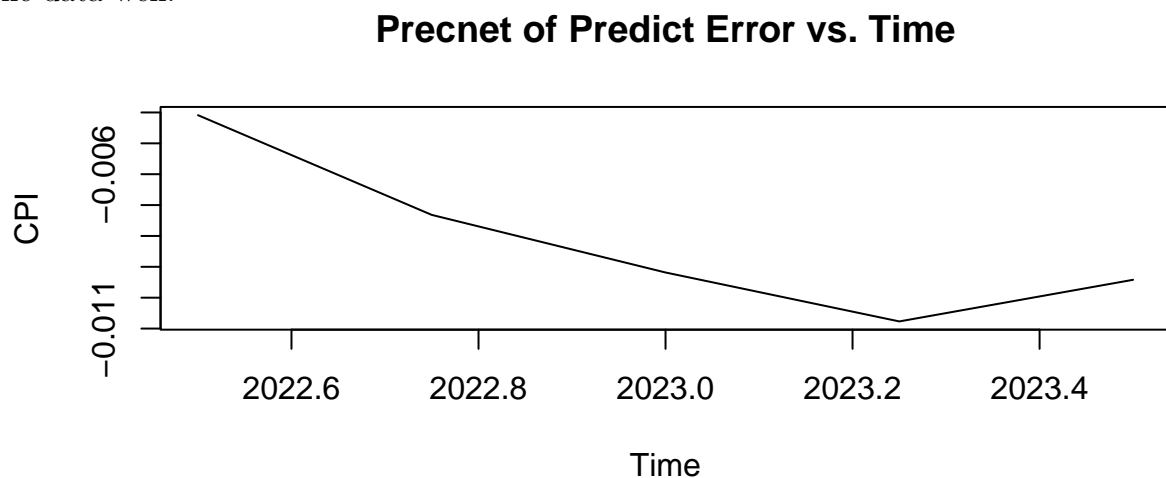




We can find although there are still a depression at 2010 but the residuals of the final model seem like a white noise at other point. And the ACF plot shows that most of the correlation of residuals are within the confident interval.

### Forecasting Results

The forecasting result of the model is shown below. The plot shows that the model fits the data well.



And the mean square error of the model is  $7.185918 \times 10^{-5}$

**rest.** A presentation of the results of your analysis. Interpretations should include a discussion of statistical versus practical import of the results.

### **Discussion**

A synopsis of your findings and any limitations your study may suffer from.



## References

## Appendix (Optional)

Any R codes or less important R outputs that you wanted to keep- can go in here.