STRUCTURAL BREAK ANALYSIS IN WEEKLY DEATHS CAUSED BY COVID-19 AND OTHER DISEASE DURING THE COVID-19 PANDEMIC

XINGYU WANG

Abstract. This paper investigates the impact of the initial COVID-19 vaccine rollout on weekly death in the US with a time series structural break approach. I considered two categories: deaths directly attributed to COVID-19 and deaths caused by non- respiratory diseases during the pandemic period. To examine the impact at both the national and state levels, the paper utilizes weekly data from the US and Virginia, covering the period from March 22, 2020, to April 23, 2023. By analyzing four time series (COVID-19 weekly deaths and other weekly deaths in both the US and Virginia), the paper aims to compare and understand the structural shifts in these series. The study yields two main conclusions. Firstly, using the Chow breakpoint test with known breakpoint dates, the paper finds that the initial COVID-19 vaccine rollout resulted in a statistically significant decrease in weekly deaths attributed to COVID-19, both at the national level and in Virginia. Moreover, there is a decrease in non-COVID-19 deaths in Virginia following the vaccine rollout. Secondly, employing the Bai-Perron multiple breakpoint test, the study identifies similarities and divergences in the pandemic trends between the nation level and Virginia. While there are common patterns, such as the overall decreasing trend in weekly deaths, we also observe noticeable structural shifts that differentiate the pandemic series in Virginia from the national series.

1. Introduction

It is widely recognized that vaccination can reduce the death rate of COVID 19. Numerous studies have demonstrated the effectiveness of vaccines in reducing the severity of COVID-19 and preventing hospitalizations and deaths from a medical perspective. Since Dec. 11 2020, the Pfizer-BioNTech COVID-19 Vaccine has been available under EUA in individuals 16 years of age and older, which became the first vaccine rollout in the US.

The research interests are two-fold. The primary focus of this paper is to look at non-COVID 19 related (not influenza, pneumonia, or COVID 19 death) weekly new deaths during the pandemic and see if non-COVID related mortality has a similar behavior regarding the surging and decreasing break points due to virus spread and vaccine rollout. The second aim is to compare pandemic trend at country level and at the state level. Specifically, I picked VA as the state of interest. Thus, the paper conducts a breakpoint analysis for US and Virginia COVID-19 mortality and non-COVID-19 related mortality between the period March 2020 to March 2023. The consideration of structural changes in time series data regarding pandemic is very important because changes indicate different stages in infection and whether certain medical treatment is effective.

type	US	VA
COVID 19 weekly deaths	/	/
other weekly deaths (non COVID 19, influenza, pneumonia deaths)	/	/

Table 1. research focus

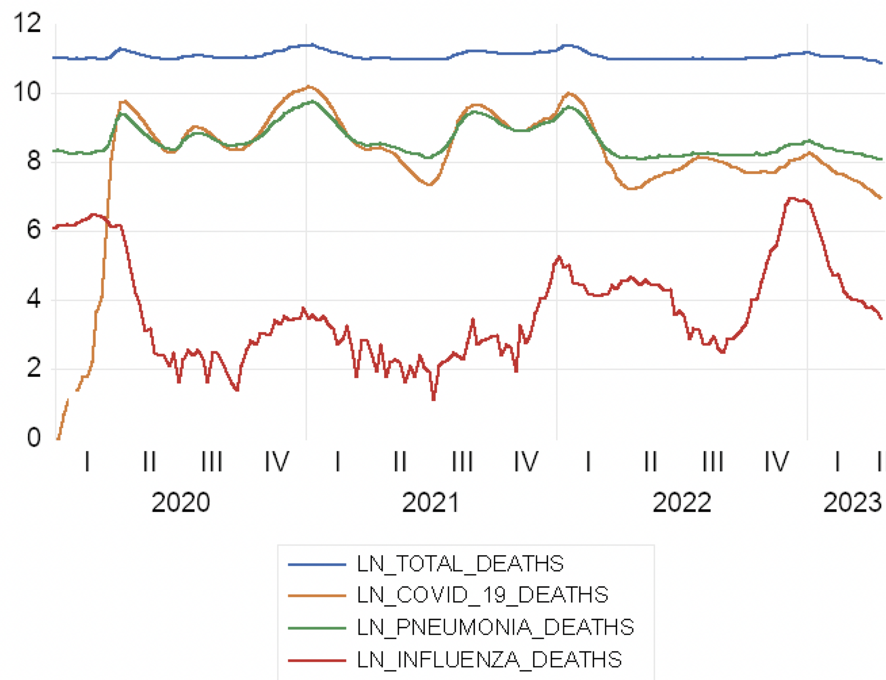

Figure 1. log-weekly death amount in the US, 12/29/2019 to 04/30/2023

2. Literature Review

In 2021 to 2022, many academic papers aim to model the disease spread and forecast future death caused by COVID-19 pandemic. A paper by Nityananda Sarkar and Kushal Banik Chowdhury in 2020 analysed the trend in COVID-19 data using a structural break approach. They used data from Brazil, India, Italy, and the UK, and were able to identify the "phase of infection" in each country. With data till May 2020, they recognized that both Brazil and India are in the increasing phase with infections rising up, while UK and Italy are in the decreasing/ containing phase. This research inspired me to think about if there is a "vaccine-effective" phase after the vaccine rollout in the US.

A 2022 paper looked at the short and long term dynamics of cause-specific mortality rates. The author found out that there is a long-run cointegrated relationship between the two most common causes of death in the US, cancer and heart disease. Although they did not discover a significant short run dependency between the two disease types, the research finds out that the two diseases are very similar in the development pattern of circulatory, respiratory, and the external mortality rates. This inspires this paper to look into the interesting question that whether the COVID-19 vaccine rollout leads to a decrease in (1) COVID-19 deaths numnber and (2) other non COVID-19 or respiratory-related deaths.

The paper *Forecasting the Trends of Covid-19 and Causal Impact of Vaccines Using Bayesian Structural time Series and ARIMA* published in 2022 disaggregated COVID-19 process into different components using bayesain strctural time-series models: a level, a local trend, seasonal impacts, and an error term. It looks at the percentage rise in the total number of cases, percentage rise in the total number of deaths. Like many other research, this research aims to forecast deaths in USA, UK and India. The research found out great variations among different countreis. They claim that using effective and quick vaccination, US and UK can reduce the number of mortality quickly; while the situation of India is more complicated, and medical treatment won't be as effective as it is in US or UK. When modeling the total number of reported daily COVID-19 cases, the paper use a ARIMA model with lags and differences determined by the AIC built in software R. There are other research concerning variations in country-wise pandemic trend, and claimed heterogeneity across countries. Figure 2 demonstrates the daily new confirmed COVID-19 deaths per million in US and India respectively. However, the problem is that different countries can use different COVID 19 testing or counting methods. Therefore, considering within country variations is worthy because data collected within the US from different states are more likely to use the same data collection methods.

FIGURE 2. daily new confirmed COVID-19 deaths per million people in red:India green: US

With many prior research taking pure mathematical models, and few works conducting an analysis with statistical modelling. Thus, it is meaningful to revisit the pandemic trend and vaccine effectiveness from a time series aspect. Some other research aims to forecast the pandemic trend with an ARIMA model without much consideration for structural changes. However, it is necessary to introduce break point analysis in this case since there are many exogenous shocks that could vary the data generate process of the pandemic series.

According to Department of Health and Human Services, the first vaccinations in the US began on December 14, 2020, and it takes approximately two weeks for vaccination to be fully effective in the body. Thus, it is reasonable to propose 12/27/2020 as a break point (decrease) in the pandemic weekly deaths amount in the US. This paper will conduct (1) a (known) break point test at 12/27/2020 for each proposed series and (2) a Bai and Perron (unknown) multiple break points test for each proposed series.

3. DATA

The data was directly collected from two sources: Centers for Disease Control and Prevention (https://www.cdc.gov/) and Our World in Data (ourworldindata.org). These two sources have information of state wise and the US DNC (daily new case) and DTC (daily total case) for COVID-19. The data for daily new death cases due to COVID-19, influenza, pneumonia is updated every week (until now, 05/2023). The number of daily deaths due to COVID-19 helps to identify the effectiveness of the vaccine treatment being used since the containment of any infection will decrease the number of daily deaths.

The data used are (i) weekly deaths due to COVID 19 and (ii) weekly deaths due to non-respiratory disease (not COVID 19, not influenza, not pneumonia) across the period March 2020 to April 2023. To compare state level and country level pandemic trend, I incorporate both US data and Virginia data. For future projects, it is worth doing an analysis for all states to observe the heterogeneity more comprehensively. While the difference in US pandemic trend and VA pandemic trend can gives us a sense of how state level pandemic trend can differ from nation level trend and provides insights for future study.

4. Econometric Model

When estimating a model, we make the key assumption that the structure of the conditional mean does not change. However, this is not always true. For instance, 2008 financial crisis leads to a change in stock market trend. Thus, it raises the motivation to use the breakpoint analysis for estimating the effect of vaccination for COVID 19 mortality rate. The steps for identifying structural breaks were referenced to *Analyzing the trend in COVID 19 data: The structural break approach*. The general steps are:

(1) Unit root test on trend and constant for each series
(2) Regression on trend and constant
(3) Single break point test for proposed break point date (12/27/2020)
(4) Sequential break tests for detecting multiple structural breaks

Considering potential structural breaks in weekly deaths due to vaccination introduction or virus variation, the model for DGP for a series can be written as:

$$y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{m} \gamma_i DU_{it} + \sum_{i=1}^{m} \delta_i DT_{it} + \epsilon_t$$

The implicit assumption is that the time series model is linear in each of the sub-periods characterised by the break points, as suggested by past papers. The parameter β_0 and β_1 are the intercept and trend coefficients irrespective of breaks; the coefficients γ_i and δ_i are the coefficients of intercept and trend for each ith break point. The paper aims to test the number (m) and date of breaks in each series to understand if COVID 19 deaths and other deaths varied due to vaccination. We use all logged data in the following analysis.

Regression and unit root test on trend and constant

Before the regression, the unit root tests are performed (Table 2). The ADF unit root test (with constant and trend, as suggested in past literature) on other deaths suggest a stochastic trend. Since forecasting is not the focus of this paper, the stochastic trend won't be discussed further. The unit root test for COVID 19 weekly death is significant, which implies a stationary series. Since as discussed in class, a deterministic trend with structural breaks can be identifies as a non-stationary series in a unit root test, we can conduct the unit root test again after identifying the structural breaks.

Sequential break tests to detect multiple structural breaks

There are several ways of determining the break point. To test the proposed vaccine effective date 12/27/2020, I used a Chow break point test on single known break date; to investigate different phase change in pandemic trend, I used the Bai-Perron multiple breakpoint test. The main idea of

type	US	VA
COVID 19 weekly deaths	p = 0.0201*	p = 0.0114*
other weekly deaths	p = 0.9978	p = 0.4847

TABLE 2. unit root test

type	US	VA
COVID 19 weekly deaths	p = 0.0009*	p = 0.0001*
other weekly deaths	p = 0.0951	p = 0.0152*

TABLE 3. single break point test for date 12/27/2020

a multiple break points test is that the true number of breaks and true time that the breaks occur generate a model with a lower sum of square error compares to other models.

5. Result

Single Chow break point tests and Bai-Perron multiple break point tests are performed for the four time series of interest.

5.1. **Chow break point test.** In general, the Chow single break point test shows that the COVID 19 weekly deaths at both the nation level and in VA has a significant structural shift after the introduction for the first vaccine. Deaths caused by other diseases shows a shift (decrease) at the VA state level but not at the nation level.

(1) COVID 19 weekly deaths: Specifically, the Chow break point test indicates a significant break in mean and trend of weekly COVID deaths case at the 1 percent for both nation level (p = 0.0009) and VA state level (p= 0.0001). For Virginia, the trend coefficient between period 3/22/2020 to 12/27/2020 is **0.0197**, while it decreases to **-0.0126** after 12/27/2020; for the US, the trend coefficient between period 3/22/2020 to 12/27/2020 is **0.013795**, while it decreases to **-0.01455** after 12/27/2020.
(2) Other weekly deaths: In general, the vaccine also tends to lead to a decrease in weekly death trend caused by COVID 19, influenza, pneumonia unrelated diseases in VA. The trend coefficient between period 3/22/2020 to 12/27/2020 is **0.0017**, while it decreases to **7.8 E-05** after 12/27/2020. However, there is no statistically significant evidence for a structural shift after 12/27/2020 at the nation level. A potential explanation is that the national data is a combination of state data, which contains variations from different states.

5.2. **Bai-Perron multiple break point test.** The multiple break point test shows the existence of structural breaks in each of the four series. The specific break point dates are presented in the appendix. However, the proposed break point (12/27/2020) is not presented in the multiple break point test results. The reason might be that the single breakpoint test assumes a specific breakpoint location and estimates the model accordingly. However, the true structural change in the data might not align precisely with the proposed date. The multiple breakpoint test, which allows for more flexibility in identifying breakpoints, may detect other dates that better capture the actual structural changes. Results for the multiple break point is shown in Figure 3. Noticeably, the test recognizes more break points in covid deaths us series. While the covid deaths in both US and VA show in general four phases: increasing, decreasing, rebounding, and "fading." Another finding is that the multiple break points for other deaths at for the US and for VA is quite similar, both have three structural breaks around the similar dates.

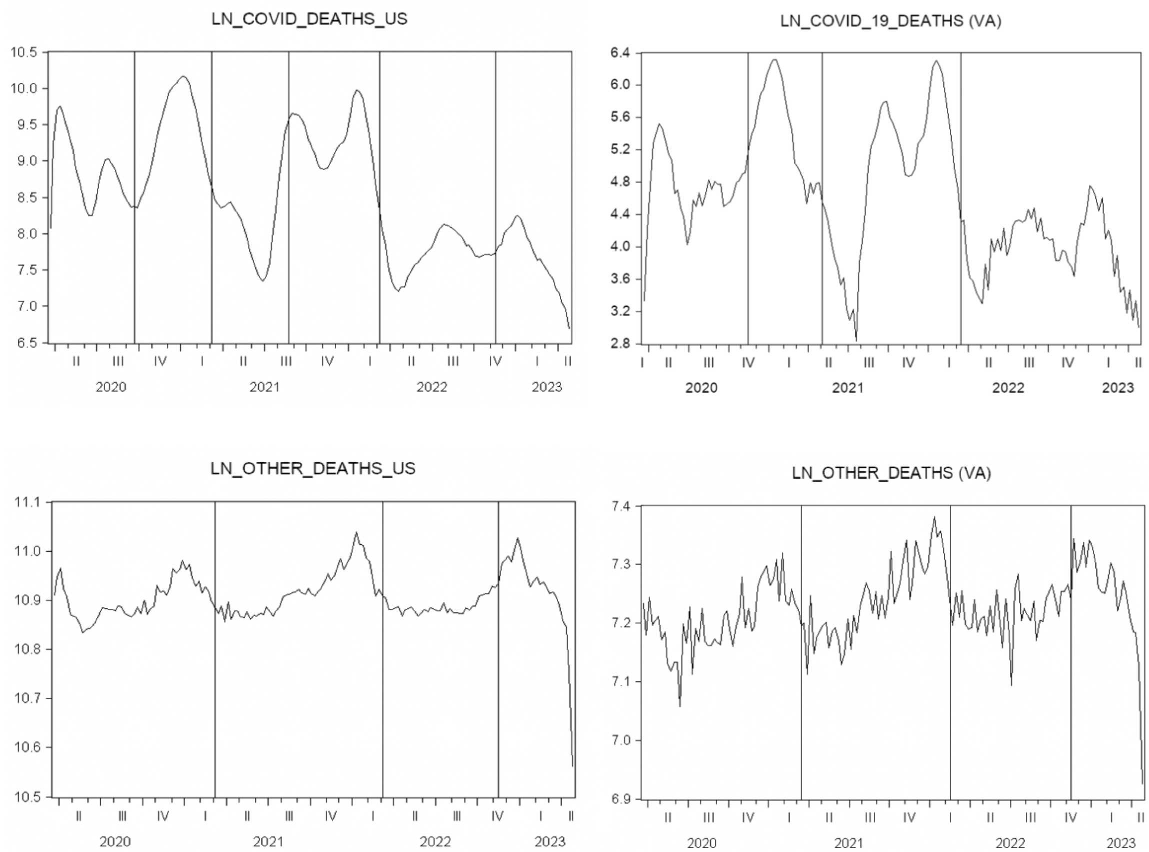

Figure 3. multiple structural breaks

6. Conclusion

The study leads to some intriguing results. Firstly, it confirms the effectiveness of vaccination by identifying a significant breakpoint occurring approximately two weeks after the initial vaccine rollout. This highlights the importance of timely immunization in reducing the spread and severity of the disease. Secondly, the research underscores the need to consider the specific state context when analyzing pandemic trends. It demonstrates that the national pandemic trend may not serve as a perfect reference for understanding the dynamics within a particular state. This emphasizes the significance of examining state-level data and policies to capture the unique factors that shape the course of the pandemic at the regional level.

This paper have several limitations. Firstly, all analysis are based on the assumption that the information obtained is correct. Secondly, only one state (Virginia) is considered. It is meaningful for future research to include more states for the break point analysis. Many research looks at variations across countries when it comes to fighting the pandemic; less authors pay attention to comparison among state-level vaccination effects. Future research should include more states to conduct a comprehensive breakpoint analysis and consider potential heterogeneity among states.

While much attention has been devoted to cross-country comparisons in combating the pandemic, there is a need to explore the variation in vaccination effects at the state level. Thirdly, the implications of the multiple break point test result should be look at more comprehensively. We specify a simple model that only contains trend and constant when testing for multiple break points. This assumption, although suggested by some prior research, may not always be appropriate. Implementing and comparing different multiple break point test methods (e.g. Global information criteria) may be useful.

As the intensity of the COVID-19 pandemic diminishes, research in this area is decreasing as well. It is crucial to seize the opportunity to learn from the pandemic and analyze it from different perspectives, including time series analysis. By exploring various approaches, we can gain a more comprehensive understanding of the pandemic's dynamics and the effectiveness of vaccination strategies among different states.

References

[1] Pierre Perron, Xiaokang Zhu (2002). Structural Breaks with Deterministic and Stochastic Trends.

[2] Colin O'hare, Youwei Li (2014). Identifying Structural Breaks in Stochastic Mortality Models.

[3] Nityananda Sarkar and Kushal Banki Chowdhury (2020). Analyzing the trend in COVID-19 data: The structural break approach

[4] Adrian E. Raftery, Jennifer L. Chunn, Patrick Gerland, Hana Sevcikova (2014). Bayesian Probabilistic Projections of Life Expectancy for All Countries. National Institutes of Health.

[5] Muhammed Navas Thorakkattle, Shazia Farhin, Athar Ali khan (2022). Forecasting the Trends of Covid-19 and Causal Impack of Vaccines Using Bayesian Structural time Series and ARIMA. Annals of Data Science (2022) 9(5): 1025-1047

type	US
COVID 19 weekly deaths	5 break points (9/20/2020, 3/07/2021, 8/22/2021, 3/06/2022, 11/13/2022)
other weekly deaths	3 break points (3/07/2021, 3/06/2022, 11/13/2022)

Table 4. Bai-Perron multiple break point test for US

type	VA
COVID 19 weekly deaths	3 break points (11/15/2020, 5/02/2021, 3/13/2022)
other weekly deaths	3 break points (3/14/2021, 2/13/2022, 11/13/2022)

Table 5. Bai-Perron multiple break point test for state VA

7. Appendix

Proposed break point (12/27/2022)

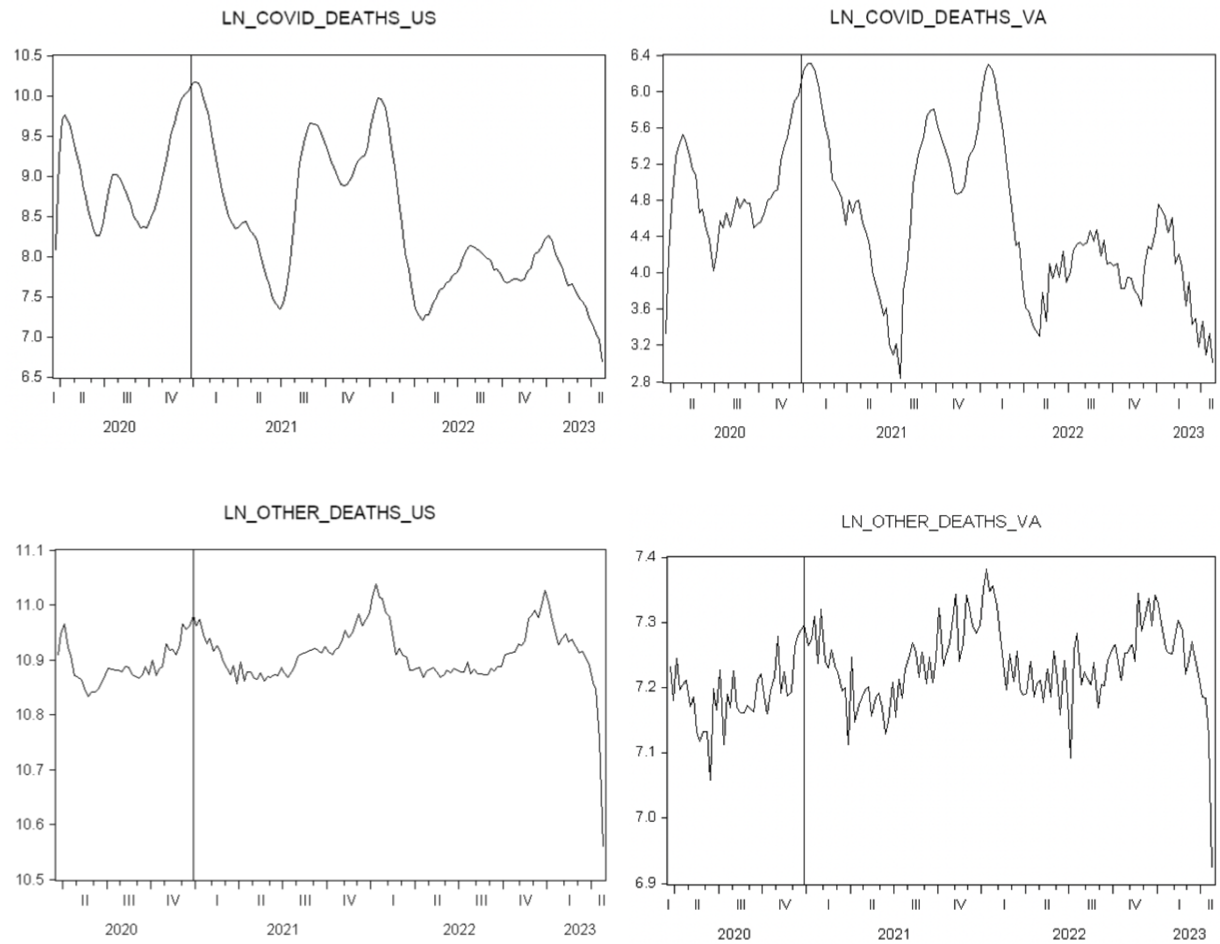

Figure 4. 12/27/2020 proposed break date, US

Dependent Variable: LN_COVID_DEATHS_VA
Method: Least Squares
Date: 05/16/23 Time: 18:31
Sample: 12/27/2020 4/23/2023
Included observations: 122

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	-0.012667	0.001959	-6.467332	0.0000
C	5.805670	0.208577	27.83466	0.0000

Sample: 3/22/2020 12/27/2020
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.019722	0.006680	2.952346	0.0053
C	4.486616	0.155232	28.90256	0.0000

Dependent Variable: LN_OTHER_DEATHS_VA
Method: Least Squares
Date: 05/16/23 Time: 18:30
Sample: 3/22/2020 12/27/2020
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.001797	0.000596	3.018152	0.0045
C	7.156103	0.013839	517.0857	0.0000

Sample: 12/27/2020 4/23/2023
Included observations: 122

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	7.80E-05	0.000162	0.482491	0.6303
C	7.229675	0.017219	419.8612	0.0000

FIGURE 5. regression before and after 12/27/2020 VA

Dependent Variable: LN_COVID_DEATHS_US
Method: Least Squares
Date: 05/16/23 Time: 18:27
Sample: 3/22/2020 12/27/2020
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.013769	0.007398	1.861077	0.0703
C	8.719457	0.171922	50.71759	0.0000

Sample: 12/27/2020 4/23/2023
Included observations: 122

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	-0.014559	0.001788	-8.144450	0.0000
C	9.808799	0.190368	51.52539	0.0000

Dependent Variable: LN_OTHER_DEATHS_US
Method: Least Squares
Date: 05/16/23 Time: 18:29
Sample: 3/22/2020 12/27/2020
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	0.001264	0.000456	2.774814	0.0084
C	10.86794	0.010585	1026.742	0.0000

Sample: 12/27/2020 4/23/2023
Included observations: 122

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	-6.61E-05	0.000143	-0.461139	0.6455
C	10.91726	0.015257	715.5564	0.0000

FIGURE 6. regression before and after 12/27/2020 US