1. **Introduction**

An accurate state-wise assessment of excess mortality during the COVID-19 pandemic is necessary to understand effectiveness of health policies in reducing mortality. However, the numbers of reported COVID-19 deaths represent only a partial count of the total death toll from the COVID-19 pandemic. The number of excess COVID-19 pandemic deaths is defined as the difference between the number of deaths during the pandemic and the number of deaths that would have been expected if the pandemic had not happened. The expected deaths without COVID-19 are forecasted using historical data (i.e., death counts before the pandemic). Gaps exist between the reported and excess deaths related to COVID-19 pandemic. This study examined how the differences between the reported and the predicted COVID-19-related deaths numbers changed over time as new testing and treatment procedures as well as new policies were implemented.

Three different types of models were employed in this study to forecast the expected number of deaths without COVID-19: (1) logarithmic trend with monthly dummy variables and autoregressive terms, (2) exponential smoothing, and (3) ARIMA models. The residuals from each of these models were utilized to fit an autoregressive tensor model for causes of death other than COVID-19. Information related to changes in one cause of death could influence and enhance the forecast for another cause of death. For instance, an increase in diabetes-related deaths in the last three months could affect heart disease deaths, potentially leading to either a decrease in deaths due to heart disease (competing risk between diseases) or a slight in deaths from heart disease (due to the same reason diabetes deaths have increased). The models were based on only a few parameters since the amount of data available was limited. Only 50 months of data was used for training the models. The models were run on the training set (January 2015 until February 2019) and evaluated on a hold-out period before COVID-19 (March 2019 to February 2020). Once the best model was selected, it was rerun using data from January 2015 to February 2020, including the previous hold-out period to include the latest data in the model before the forecast.

1. **Methods** **and Results**
   1. **Data**

All deaths data by month and state were downloaded from the CDC1-2 mortality dataset from January 2015 to September 2023. Note that, since there is a 6 months lag for the assignment of non-natural causes of death (unintentional accidents and self-harm), the data for the last 6 months is not considered reliable for this cause of death. There were 50 observations for each state (i.e., for the 50 states and the District of Columbia) and 165 causes of death. The data was used to build forecasting models for the number of deaths from 14 main causes of death (not including COVID-19) and the combination of the other 151 causes of death for the pre-Covid period from January 2015 to February 2019.

Population data by state and year were collected from the Census3 for the period from 2015 to 2022; the 2022 population was used in 2023 as the 2023 data was not available.

* 1. **Excess mortality definition and metric to forecast.**

Excess mortality during the Covid-19 pandemic covering the time period from March 2019 to September 2023 was defined as the difference between the number of actual deaths and the number of forecasted deaths calculated as if the pandemic had not occurred:

Excess deaths = Actual deaths – Forecasted deaths without COVID (1)

In order to compare across the states, the death rates were normalized to the state populations::

Crude death rate by month = 100,000\* (Monthly deaths)/Population (2)

Additionally, the crude (monthly) death rate was adjusted for the number of days in each month to eliminate month length effect:

Daily crude rate = Monthly crude rate / Number of days in the month (3)

Figure 1 compared the monthly death counts in the US from January 2015 to May 2023 to the monthly crude rates. The seasonality from both metrics was the same, but the crude rate did not increase as much as the number of deaths.

Figure 2 compared the daily crude rate in the US from January 2015 to May 2023 to the monthly crude rate. The trend were the same with both metrics but the seasonality effect was smoother using the daily crude rate than the monthly crude rate.

The number of excess deaths was estimated using the dependency between excess deaths and daily crude rate as expressed in equation (4).

Excess deaths = Observed deaths – Forecasted deaths without COVID

= Observed deaths - Forecasted daily crude rate \* Number of days in a month \*

Population/100,000