**Estimate excess COVID deaths in the US by cause of death**

Nuria Diaz-Tena1, Davit Sargsyan2, Michael Katehakis1, Javier Cabrera3, Vinguen and some of my students TBD

1Rutgers Business School, Rutgers University, NJ, USA

2Janssen Research and Development LLC, Raritan, NJ, USA

3Department of Statistics, Rutgers University, NJ, USA

**Abstract**

Causes of death incidence in America fluctuates year by year. Only deaths with an autopsy are sure to be correct, the other deaths are classified by the most probable cause without any test to prove it. Besides, healthcare conditions, economics, natural disasters, and drug overdose creates fluctuations on the American death landscape. COVID-19 has brought further uncertainties given the additional COVID test necessary to achieve the correct cause of death of the deassed Americans. The COVID deaths data don’t catch everyone whose life was shortened by the pandemic and add other people whose primary reason for dying was not COVID. This paper shows how to detect excess deaths using causes of death data. The analysis is conducted at the cause of death level to identify the over and under reported deaths by COVID or nonCOVID.

# **1. Introduction**

The COVID-19 pandemic has challenged our healthcare systems and our ability to take drastic measures to protect the population from new virus……

…..

…..

Purpose and reasons for this paper….

Definition of Excess death and multiple ways to computed using different forecasting methodologies.

Our data has few observations providing a new challenge to choose a forecasting method who can provide good forecasting estimates with these circumstances

Too many years of data will take too much weight on the old data. Recent data may not have enough observations for our forecast

1. 2013 to 2018 for training model (72 observations)
2. 2014 to 2018 for training (60 observations)
3. 2015 to 2018 for training (48 observations)

Testing set 2019

Model for our forecast

1. Exponential Smoothing
2. Sharing coefficients
3. Sinusoids
4. Seasonalities for different frequencies

Table of Diagnostic in testing set as MAE or MAD, MAPE or SSE

Plots forecasting 2019

When the best model and training set is selected. Include 2019 in the training set and forecast for COVID time.

Compare the excess deaths for each cause of death toward the COVID deaths.

New idea (3/24/23)

1. Compute excess death per state per disease with all models: 100\*(obs-pred)/
2. PCA/biplot of% excess death by disease (arrows) and state (points)