Suicides and Entertainment in the American Youth: An Interrupted Time Series Approach

Dania Abuhijleh and Arthur Langlois

Department of Mathematics,

Northeastern University

Abstract

This project aims at quantifying the impact of the 2017 release of Netflix's "13 Reasons Why" on the absolute number of suicide deaths in the American youth. We will use an interrupted time series approach to quantify the impact of the series' release. This study follows on a study by Bridge et al. in 2019 which observed a significant increase in the number of monthly suicide rates following the release of the TV show. Our study aims at auditing for the aforementioned paper's results by estimating the counterfactual of the suicide death numbers in 2018 with other regressors than those used in the paper. Our analysis suggests that the release of the series was associated with an increase of about 3 suicides for 100,000 across all models, therefore our results are directionally consistent with Bridge et al's findings.

Previous work

At the first stages of the project, we were aiming at finding a way to study the determinants of suicides across the world. We initially thought of doing such using an interpretable machine learning approach using Shapley Additive Explanation values (or SHAP). While the data itself was easily accessible -such as variables related to education or income-, we were worried about omitted variable bias, that is, the potential over or under estimation of certain factors due to the absence of other determinant variables of suicides in our analysis. Omitted variable bias would easily arise in a similar analysis due to the complexity of the many determinants of suicide across the world.

As an alternative, we wanted to use the same methodology (SHAP values) to determine the determinants of suicide at the individual level. We hypothesized that we could find publicly available datasets of experiments or longitudinal studies regarding suicides. In that case, researchers who would have collected the data would be controlling for many variables, therefore reducing the bias of the included variables in our model. However, our inability to perform such analysis was due to the lack of individual level data for suicides, which was due to data privacy, as well as high costs to access the data.

We finally opted for a project aimed at assessing causal relationship rather than mere correlation (that we would have obtained with the interpretable machine learning framework), still related to suicide. We therefore came across the Bridge et al. paper and were interested to replicate their analysis with other models than what they used (Poisson and Holt-Winters).

Background

In the United States, suicide is a growing public health concern with the national age-adjused rate of suicide reported to have increased by as much as 33% between 1999 and 2017. The portrayal of various problems has rapidly changed in the past decades, and came to show the importance of responsible and accurate reporting in the news media and entertainment industry. Media depictions of suicide, mental illness, and related issues have the potential to do harm or good. Suicide contaigion, a process in which direct or indirect exposure to suicide increases the risk of subsequent suicidal behavior, has had its effects examined in some studies. Following the release of Netflix's "13 Reasons Why", a spike in suicide rates has been noticed in the American youth. We wish to analyze the methods used to quantify the impacts of this show on suicide rates across the United States, in order to determine how confident we can be in the tools used in causal inference.

Research Questions

Using aggregated state level suicide rates for the 5-24 years old population in the United States between 2000 and 2020 from the Center for Disease's Control (CDC), we focus our study on three main corollaries:

- Quantify impact of exposure to "13 Reasons Why" on suicide rates across states
- Comparative review of estimators for Interrupted Time Series (previous paper has been written with Poisson regressor, but we want to try with other methods to see how they compare).

Literature & Methodology

We will quantify the causal effect of the series release with an ITS approach. The most reliable way of estimating causal effect is through randomization, which has been widely used in medicine and social science research through randomized controlled trials. Through randomization, in-group differences are not a threat to validity as the treatment and control groups are randomly assigned, therefore allowing for the assessment of the

effect of a treatment through comparing outcomes for both control and treatment groups. However, running a randomized controlled trial (RCT) is not feasible in our situation due to the fundamental problem of causal inference: we only have access to one outcome, that is that all units are treated - we cannot observe the situation in which the unit has not been treated. As such, the statistics/econometrics community has come up with various ways of estimating causal effects without control groups, most notably with difference in difference, instrumental variables, regression discontinuity, synthetic controls and interrupted time series. Interrupted time series have been popularized by Bernal and Cummins in 2017, introducing the method to the public health and epidemiology communities.

As mentioned above, the main objective of our analysis is to analyze the comparative performance of several time series regressors, namely, ARIMA, SARIMAX and Exponential Smoothing. We provide a brief description of these mentioned methods below:

ARIMA: Autoregressive Integrated Moving Average Model, or ARIMA, models the trained data by using its previous observations as subsequent dependent variables and uses the dependency between an observation and its residual error from a moving average model applied to lagged observations. ARIMA models are set with three main parameters: p number of lag observations included in the model (or “lag order”), d number of times the raw observations are differenced (to render the data stationary), and q, the order of the moving average (the number of years on which the moving average is computed).

SARIMAX: Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX) is similar to the ARIMA model but incorporates seasonality and exogenous factors. In addition to ARIMA’s p,d,q parameters, SARIMAX is specified with another set of p,d and q parameters for the seasonality dimension of the dataset, as well as s, the periodicity of the data’s seasonal cycle.

Exponential smoothing: An exponential smoothing model is trained on the most recent observation as well as its last forecast. The last forecast populated by the model already included a part of the previous observation and a part of the previous forecast. Therefore, the previous forecast includes everything the model learned so far based on demand history.

For estimating the treatment effect, we use Interrupted Time Series. The method is useful to understand if and how an outcome has changed after an intervention was

introduced to the full population at a specific point in time, given longitudinal data of the outcome. Mathematically, a time series equation consists of four key coefficients:

$$\mathrm{Y} = \mathrm{b}_0 + \mathrm{b}_1 T + \mathrm{b}_2 D + \mathrm{b}_3 P + \mathrm{e}$$

with a Y outcome variable, T continuous variable indicating time passed from the start of the observational period, D dummy variable indicating observation collected before (=0) or after (=1) intervention, P continuous variable indicating the time that has passed since the intervention, set to 0 before the intervention. More specifically, we observe the number of suicides at time t for state j as Yjt and the potential outcome YjtN as the potential outcome without intervention (release of the series) and Yjtl the potential outcome with intervention. We therefore define the treatment effect for treated state j = 1 at year t (for t>0) as follows:

$$\tau_{1t} = Y_{jt}^{I} - Y_{jt}^{N}$$

Our treatment effect is therefore defined as the unit difference between the actual number of suicides with the estimated counterfactual (a model's prediction at time t).

Results

A standard process to follow for time series forecasting is to make the data stationary, that is, transforming the data such that its rolling average and standard deviation are 0. By removing trends and non-stationary features of the target variable, one can more accurately predict its future values. Figure 1 shows that the untransformed data is not stationary (has a mean and standard deviation that are different from 0). However, after differentiating the data and taking its natural logarithm, the data still did not reach stationarity. We computed an Augmented Dickey Fuller Test to check for the statistical significance of the stationarity, and the p-value of the AFD statistics was not statistically significant (p-value of 0.584). On the other hand, a visual inspection of the raw data and its corresponding mean and standard deviation shows that while they are positive, they are not remarkable nor significant. Additionally, ARIMA and SARIMAX models do not necessarily assume stationarity for predictions. As a result, for the remainder of the modeling exercise, we will use the untransformed data as an input.

Figure 2 displays the evolution of suicide deaths among the 4-24 years old in the United States between 2010 and 2020. The black line represents the actual line of suicide deaths, whereas the dotted lines represent the various predictions made for the post-treatment period (2017 through 2020) by ARIMA, ARIMEX and Simple Exponential Smoothing models (These models have previously been selected as they performed well in predicting the previous periods' deaths). Overall, all three models predicted a significantly lower number in the number of suicide deaths in 2017. Because every single

model is consistent in its prediction magnitude and direction, we can confirm the hypothesis. On the other hand, one can observe that the predictions are not accurate in the post-treatment period. However, as we predict additional periods, the standard error increases, so we cannot confidently draw inference from 2018, 2019 and 2020.

More specifically, Table 1 provides a more detailed analysis of the difference between the treatment and control units: the actual number of suicides in 2017 was higher than the number predicted by ARIMA and SARIMAX, which respectively predicted 3.01 and 4.3 less suicides than the actual number. A comparison with other years (2018 through 2020) show that the difference between the prediction and actual number of deaths is insignificant compared to 2017.

Figure 3 from Bridge et al. displays the mean monthly counts and rates of suicide before and after the release of *13 Reasons Why.* The months before the release of the series on March 31, 2017 they observe a significantly increasing trend in suicide rates in 10-17 year olds. Additionally, in April - the month following the release - they observe the highest monthly suicide rate over the 5-year study period, of 0.57 per 100,000 persons. They found a statistically significant effect of the release of the series on subsequent suicide rates predicted by a Holt-Winters model.

Conclusion

We used an ITS approach using three distinct modeling strategies: ARIMA, ARIMAX and Exponential Smoothing. All three methods confirmed our hypothesis that the release of the series was associated with a increase in suicide deaths in 2017 in the United States. More specifically, our models estimated that the number of suicide deaths would have been 3.01, 4.3, and 3.53 (for 100,000 individuals) without the release of the series for ARIMA, ARIMAX, and Exponential Smoothing respectively. These results are therefore directionally consistent with the baseline findings of the Bridge et al paper, which estimated the effect to be about 1.3 additional deaths per month in 2017 on average, or 15.6 over the year. However, we find the magnitude of the change to be overestimated in their analysis, which are to be attributed by their use of quasi-Poisson regression and Holt-Winters forecasting models.

Appendix

Our code and data can be found on GitHub at
https://github.com/Arthur-Langlois/capstone/tree/main

Table 1

	Treatment ARIMA	Treatment SARIMAX	Treatment SES
2017	3.084570	4.279395	3.535092
2018	1.682489	1.648183	1.893441
2019	-1.618848	-6.832198	-1.369908
2020	-0.074363	2.639562	0.168553

Figure 1

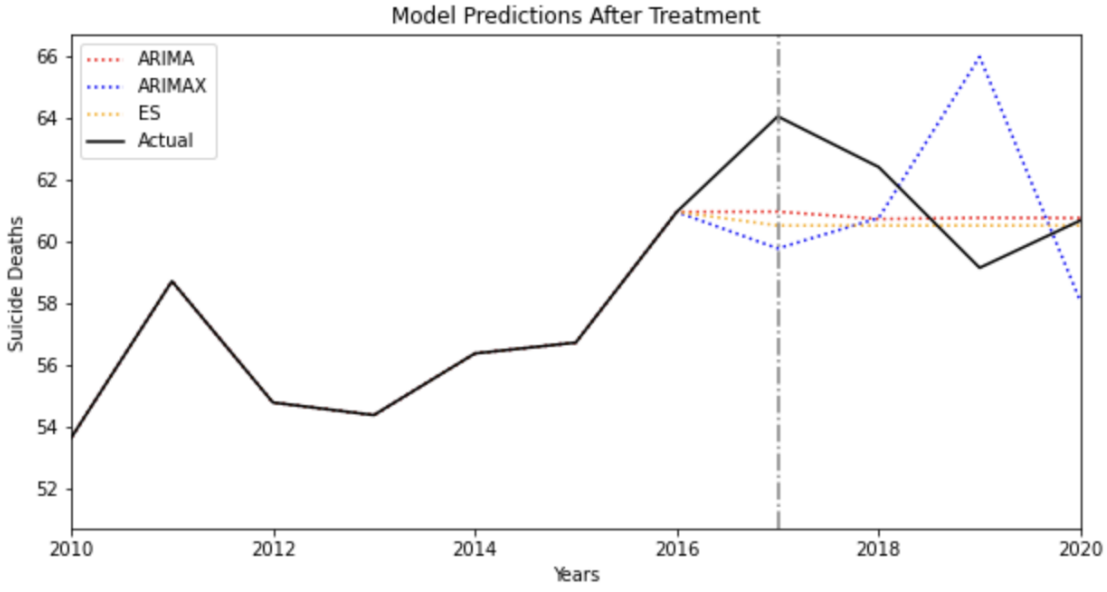


Figure 2

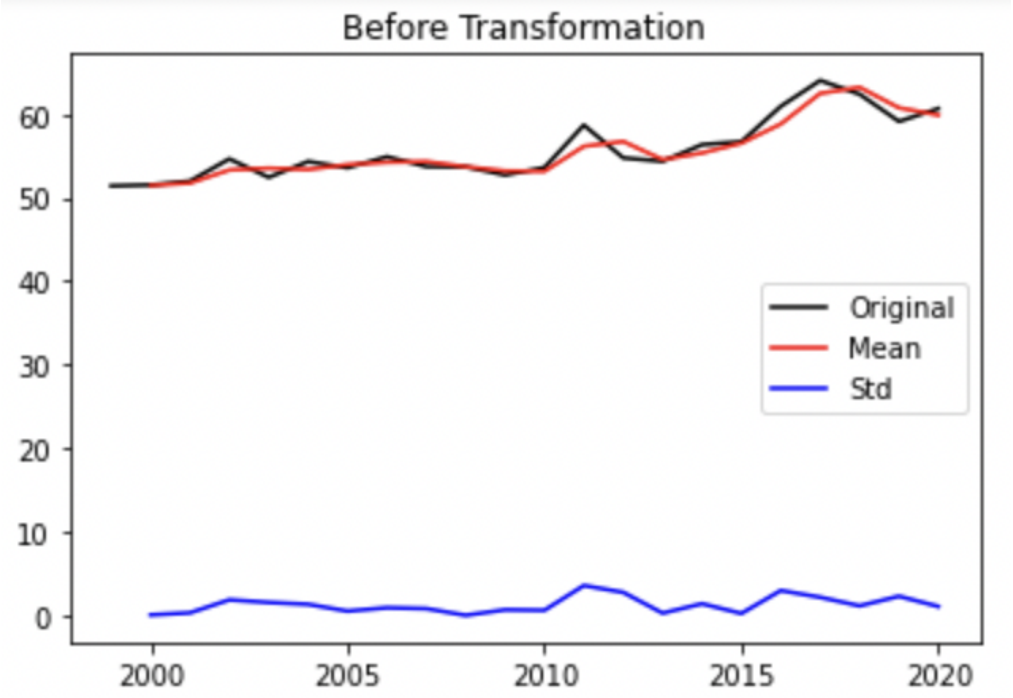
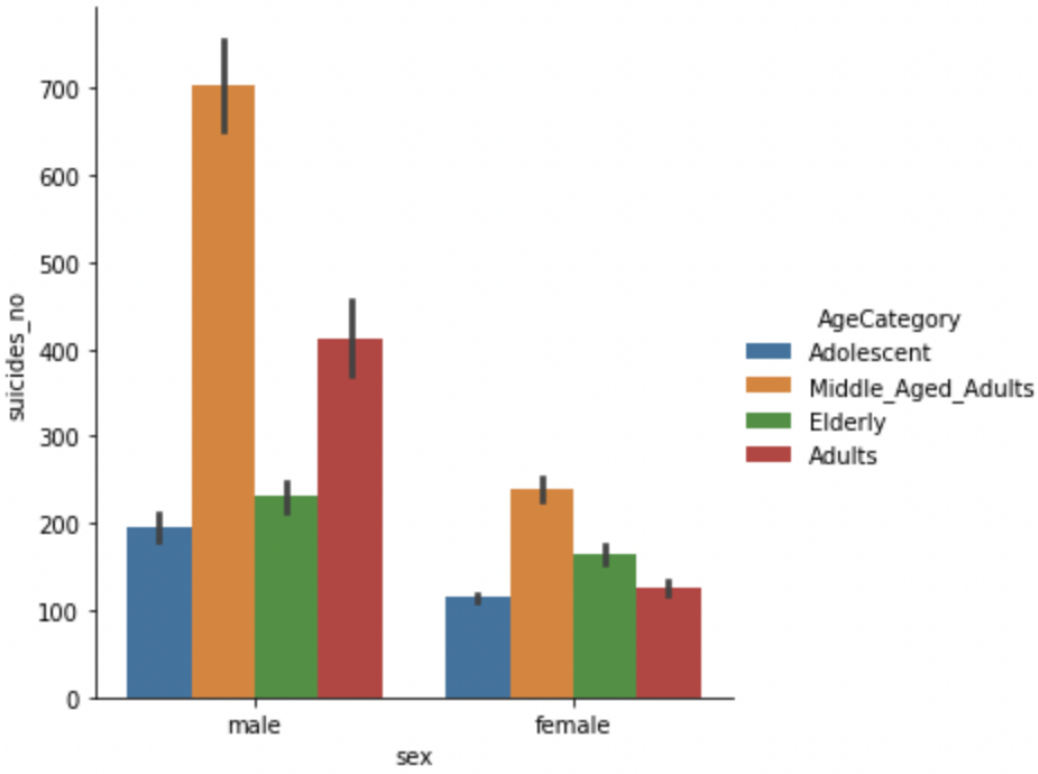


Figure 3

Characteristic	Mean Monthly Count		Mean Monthly Rate per 100,000 Persons		Pre-*13RW* Trend[a] (IRR) (95% CI)	*p*	Post-*13RW* Trend[b] (IRR) (95% CI)	*p*	Step Change[c] (IRR) (95% CI)	*p*
	Pre-*13RW*	Post-*13RW*	Pre-*13RW*	Post-*13RW*						
10–17 Years										
Overall	116.29	149.56	0.35	0.45	1.005 (1.003–1.008)	<.001	0.97 (0.95–1.00)	.057	1.29 (1.09–1.53)	.004
18–29 Years										
Overall	651.14	774.22	1.22	1.43	1.004 (1.003–1.005)	<.001	0.99 (0.98–1.01)	.443	1.07 (0.98–1.18)	.135

Figure 4



References

1. J. Asher, Thirteen Reasons Why, RazorBill/Penguin, New York, NY (2007)

2. M. Kutner, Suicide prevention expert on '13 Reasons Why': 'contagion is a real concern here' Available at: https://www.newsweek.com/netflix-13-reasons-why-suicide-contagion-experts-593367 (2017),

3. S. Stack, Media coverage as a risk factor in suicide, J Epidemiol Community Health, 57 (2003), pp. 238-240

4. A.T. Cheng, K. Hawton, C.T. Lee, T.H. Chen, The influence of media reporting of the suicide of a celebrity on suicide rates: a population-based study, Int J Epidemiol, 36 (6) (2007), pp. 1229-1234

5. T. Niederkrotenthaler, B. Till, N.D. Kapusta, M. Voracek, K. Dervic, G. Sonneck, Copycat effects after media reports on suicide: a population-based ecologic study. Soc Sci Med, 69 (2009), pp. 1085-1090

6. T. Niederkrotenthaler, M. Voracek, A. Herberth, *et al.* Role of media reports in completed and prevented suicide: Werther v. Papageno effects, Br J Psychiatry, 197 (2010), pp. 234-243

7. P. Ortiz, E. Khin Khin, Traditional and new media's influence on suicidal behavior and contagion, Behav Sci Law, 36 (2018), pp. 245-256

8. J. Pirkis, R.W. Blood, Suicide and the media. Part II: portrayal in fictional media. Crisis, 22 (2001), pp. 155-162

9. M. Gould, P. Jamieson, D. Romer. Media contagion and suicide among the young. Am Behav Sci, 46 (2003), pp. 1269-1284

10. Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. "Interrupted time series regression for the evaluation of public health interventions: a tutorial." *International journal of epidemiology* 46.1 (2017): 348-355.

11. Gardner Jr, Everette S. "Exponential smoothing: The state of the art." *Journal of forecasting* 4.1 (1985): 1-28.

12. Ho, Siu Lau, and Min Xie. "The use of ARIMA models for reliability forecasting and analysis." *Computers & industrial engineering* 35.1-2 (1998): 213-216.

13. Wangdi, Kinley, et al. "Development of temporal modeling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan." *Malaria Journal* 9.1 (2010): 1-9.