The Effect of Daylight Saving on Driver Fatalities:
A Machine Learning Approach

Jacob Gudmundsen
Alex Johnson
Emily Maynes
Jordan Miles
Larissa Nowjack

Econ 484
April 12, 2021

Abstract

Our analysis focuses on identifying the causal effect of daylight saving time on fatal traffic accidents. With accident data from the Fatality Analysis Reporting System and US government census data, we use post double-selection LASSO to flexibly control for potentially confounding covariates. Using the variables selected by LASSO, we utilize a synthetic control model on state-level data to estimate the causal effects if Arizona were to implement daylight saving time. Our results show that Arizona would experience approximately .25 more accidents on the Monday following the DST time change, but the point estimate is not statistically significant.

I. Introduction

Mitigating the risk of driver fatalities through policies and new technology is of interest to both private and public sectors. Approximately 79% of traffic fatalities are attributed to risk factors including speeding and driving under the influence according to the US transportation National Highway Traffic Safety Administration (NHTSA). Other identified risk factors for traffic fatalities include behavioral changes and impaired driver vigilance, but the remaining 21% of accidents report no underlying causes. Previous studies explore the effect of daylight saving time (DST) clock shifts on traffic incidents (Allen and Varughese; Büschenfelde, Meyer, and Lohse; Cummings and Lambe; Hicks et al.) and point to sleep deprivation as the primary mechanism.

The existing literature provides mixed conclusions about the causal impact of DST on traffic accidents. One study found no statistically significant impact on the effect of DST clock shifts on driving (Cummings and Lambe) while another study found that traffic accidents increased significantly in the week following DST (Hicks et al.) Using the machine learning assisted technique post double-selection LASSO to control for high dimensional covariates, we seek to identify a causal effect of DST on fatal traffic accidents in Arizona. We compare the outcomes in Arizona which does not observe DST to the outcomes from the remaining 49 areas using a synthetic control model.

II. Data

We obtained fatal accident data from the U.S. National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS). This dataset contains all accidents in the United States that led to a fatality within thirty days of the accident. Estimates are based on the "Person" series of data for 49 states and Washington, DC, for years 2015 – 2019. Additional data controlling for state-level demographic information was obtained through the United States Census Bureau for the same years. This includes state-level data on the total population size and different age populations (including adolescent and elderly age groups, etc.), as well as gender ratios and measures of the citizenship status of individuals in each state and the number of individuals living at or below the poverty level. Using a report from the National Institute on Alcohol Abuse and Alcoholism, we also included measures of alcohol consumption by state. Additionally, we created a web scraping program to gather average temperature data (in °F) in each of the states in our sample over the relevant years. We examine the average number of fatal traffic accidents that occurred on the Mondays four weeks before the DST event and the Monday afterwards by collapsing the FARS data over the five years examined. Summary statistics on the sample of covariates later selected through PDS are included in Table 1.

An initial comparison of the average fatal accidents the Monday after to the four Mondays before the treatment, or a "naive" simple difference, is illustrated in Figure 1. The states included are those later identified by the synthetic control as comparable to Arizona. At first glance we observe more fatal accidents on average in the week after DST in both Arizona and comparable states with the latter showing a greater change in fatal accidents. However, we will examine the marginal difference in projected fatal accidents between observed and synthetic Arizona and test the significance of this difference.

III. Method

After gathering all of the state level data and features for our model, we restricted the data to five dates of interest: the four Mondays before DST and the Monday afterwards. We isolated accident data for the dates of interest by the unique number of accidents per day in each state in each year. We had data for five years, then collapsed them into one dataset, using an average across all years over the index of the days of interest. We then merged datasets to include all relevant covariates. Given more computational power, interacting the features with each other could have generated more covariates and potentially increased accuracy.

We used post double selection (PDS) on our covariates, which applies machine learning to econometric processes to aid in selecting a subset of covariates from the overall set. Specifically, PDS uses LASSO regression involving an X matrix of characteristics (whose columns are the relevant covariates), an outcome vector Y, and a treatment vector D. One advantage of this technique is that it can optimally force estimated coefficients completely to zero, resulting in a subset of covariates chosen from a much larger dataset. How many of the ß coefficients are pushed to zero depends on the level of a penalizing coefficient, α. In order for LASSO to optimally choose which covariates to keep, we must have sparsity, which assumes that only a relatively small amount of all of our covariates actually affect our outcome of interest. We note that in most practical cases, sparsity is a strong assumption to make, but approximate sparsity may be sufficient for LASSO to return correct and effective results. In the context of our study, we work specifically with data that is approximately sparse, and, as such, we preliminarily conclude that the LASSO method will correctly return a subset of the most important covariates.

The PDS method uses LASSO twice: first, it uses a LASSO regression of the treatment on all of the covariates collected. Second, PDS requires another LASSO regression of the outcome on all of the covariates collected, like so:

$$(1)\ LASSO \quad - \quad D \quad x_1 \ldots . x_p$$
$$(2)\ LASSO \quad - \quad Y \quad x_1 \ldots . x_p$$

where D is the treatment, Y is the outcome, and each x_i is a distinct covariate. After both regressions, the LASSO method will have optimally forced several ß coefficients to zero. Using the union of (1) and (2), $\{x_i\}D \cup \{x_i\}Y$, we fit a synthetic control model of Y on the treatment D and these covariates to estimate causal effects.

We use k-fold cross-validation at both stages of PDS to choose the optimal penalty coefficient α for LASSO. Cross-validation involves splitting a dataset into k folds, then fitting the model on all the folds except for the kth fold for a value of α. Once the model is fit, it is then tested on the left out fold. We then calculate the error rate, which, in most cases, is the mean squared error (MSE) of the left out fold. After repeating this process k times, we average all the MSE results. We repeat this process for many different values of α and pick the optimal alpha resulting in the highest test set accuracy. In this study, we implemented the sklearn package GridSearchCV in Python to iteratively select the optimal α values, which are .1 for equation (1) and .0001 for (2).

Our research design then implements a synthetic control method. Synthetic control methods combine elements from both matching and difference-in-difference techniques to evaluate intervention effects in a treated group. Synthetic control constructs a "synthetic" group using non-negative weights of observable characteristics from the control group, in this case, the subset of covariates resulting from PDS, and matches them to the treated group where they cumulatively sum to one. In our case, the treated unit is the state of Arizona, while the rest of the areas in our dataset form the control group used to create synthetic Arizona.

IV. Results

After running PDS on our original 69 covariates, we have 9 significant covariates to use in synthetic control. This corresponds to a test set accuracy of 0 for the LASSO regression of D on X and -4.137 for the LASSO regression of Y on X. Though this accuracy is low, we are not concerned because prediction is peripheral to our analysis. Rather, we use PDS to optimally select significant variables for our causal model.

The synthetic Arizona is a weighted average of 26 states (see Table 2 for the weighting breakdowns.) Observed outcomes for Arizona are compared with those from the synthetic Arizona in Table 3 and the outcomes for both across the dates examined are illustrated in Figure 2. Our results show that if Arizona were to adopt DST there would be an increase in .25 fatal accidents on the Monday following DST. Pretreatment trends of Arizona versus synthetic Arizona follow each other closely which strengthens the validity of this result, especially in the week prior to DST. However, in determining the statistical significance of our synthetic model and following the typical in-space placebo constructions, we calculate that Arizona is the 34th most extreme of all the tests, as illustrated in Figure 3. This implies that Arizona possesses a p-value of 34/50 (or .68), which strongly suggests that we fail to reject the null hypothesis.

V. Conclusion

Past research gives mixed results on the effect of DST on fatal accidents the Monday after the time change takes place. We add to the current literature by providing a causal analysis of the effects in Arizona, one of two states that does not practice DST. After using post double-selection LASSO to flexibly control for covariates, we use synthetic control to model how fatal accidents would change in Arizona if it implemented DST. The results show that accidents would go up by .25 on average, but the estimate is not statistically significant.

Our research provides evidence that the lack of DST in Arizona does not likely affect the number of fatal car crashes there. Though outside the scope of our study, it seems that DST is not a large contributor to fatal accidents in other states. Whatever trouble individuals may have with "springing forward," it does not seem enough to increase crashes. However, we recognize several shortcomings with our research design that may have prevented detection of a significant or larger effect. First, we were limited in the state-wide covariates we were able to obtain. More robust data on weather, miles driven, driving culture and other behaviors, etc., would enrich the quality of our model, allowing us to better control for differences among states. Additionally, due to limited computational power, we were unable to interact and otherwise transform our base covariates, limiting the fit of our model. We were also unable to perform synthetic control on all other states and carry out a comparable analysis to that of the synthetic Arizona. These are fruitful concerns for future research with better data and higher computing power.

Works Cited

Hicks, Gregory J., et al. “Fatal Alcohol-Related Traffic Crashes Increase Subsequent to Changes to and from Daylight Savings Time.” Perceptual and Motor Skills, Volume 86, Issue 3, June 1998, pp. 879–882, doi:10.2466/pms.1998.86.3.879.

Kaplan, Jacob. Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends 1977-2018. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-01-16. https://doi.org/10.3886/E105583V5

Lambe, Mats and Cummings, Peter, “The Shift to and from Daylight Savings Time and Motor Vehicle Crashes, Accident Analysis & Prevention, Volume 32, Issue 4, 2000, Pages 609-611, ISSN 0001-4575, https://doi.org/10.1016/S0001-4575(99)00088-3.

Meyer, K., M. Büschenfelde, and Lohse AW. "Daylight savings time and traffic accidents." *The New England Journal of Medicine* (1996).

National Center for Statistics and Analysis. (2020, December). Fatality Analysis Reporting System (FARS) analytical user’s manual, 1975-2019 (Report No. DOT HS 813 023). National Highway Traffic Safety Administration.

Varughese, Jason and Allen, Richard P, Fatal accidents following changes in daylight savings time: the American experience, Sleep Medicine, Volume 2, Issue 1, 2001,Pages 31-36, ISSN 1389-9457, https://doi.org/10.1016/S1389-9457(00)00032-0.

Appendix

Table 1

	Date Index	Accidents in Week 4	Average Accidents	Population	Avg consumption of beers (per capita)
count	250	250	250	250	250
mean	3	0.4014	1.8043	6465514.02	267.8281
std	1.4170	1.1298	1.6133	7256675.935	52.9177
min	1	0	0	581024	146.3703
25%	2	0	1	1817305	231.1111
50%	3	0	1.5	4556707	262.5185
75%	4	0	2	7404107	299.2592
max	5	8.4	8.4	39283497	435.5555

Figure 1

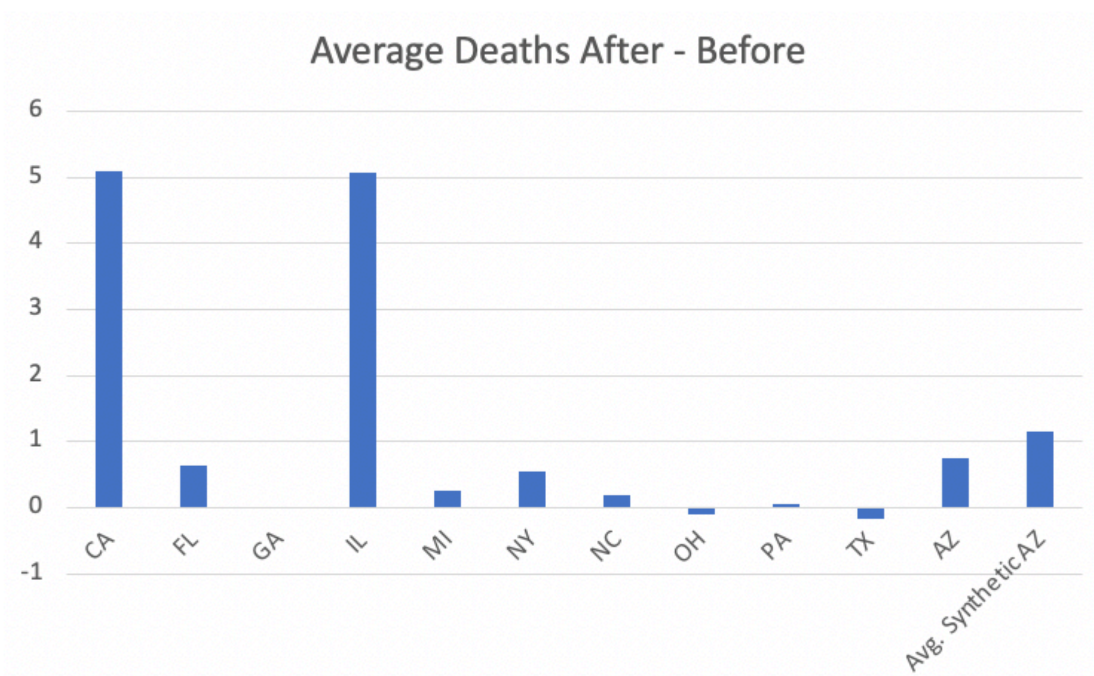

Table 2

Synthetic Control Weighting			
State	*Weight*	*State*	*Weight*
Alabama	0.0226	**New Jersey**	0.0425
California	0.0745	**New York**	0.0674
Colorado	0.0109	**North Carolina**	0.0496
Florida	0.0616	**Ohio**	0.0564
Georgia	0.0510	**Oklahoma**	0.0156
Illinois	0.0552	**Pennsylvania**	0.0561
Indiana	0.0332	**South Carolina**	0.0181
Kentucky	0.0165	**Tennessee**	0.0333
Maryland	0.0283	**Texas**	0.0698
Massachusetts	0.0290	**Utah**	0.0175
Michigan	0.0492	**Virginia**	0.0438
Minnesota	0.0129	**Washington**	0.0380
Missouri	0.0211	**Wisconsin**	0.0104

Table 3

Variable	**Arizona**	**Synthetic Arizona**	**WMAPE**	**Importance**
Accidents in Week 4	0.60	0.59	0.42	0.14
Average Accidents	1.65	1.95	0.65	0.14
Population	7050299.00	13729152.34	7529807.02	0.15
Avg consumption of beers (per capita)	266.07	245.89	32.25	0.14
Avg consumption of wine (per capita)	78.39	81.78	20.85	0.14
Avg shots of liquor (per capita)	163.50	166.30	22.98	0.14
Temperature (squared interaction term)	3636.09	3088.20	868.19	0.15

Figure 2

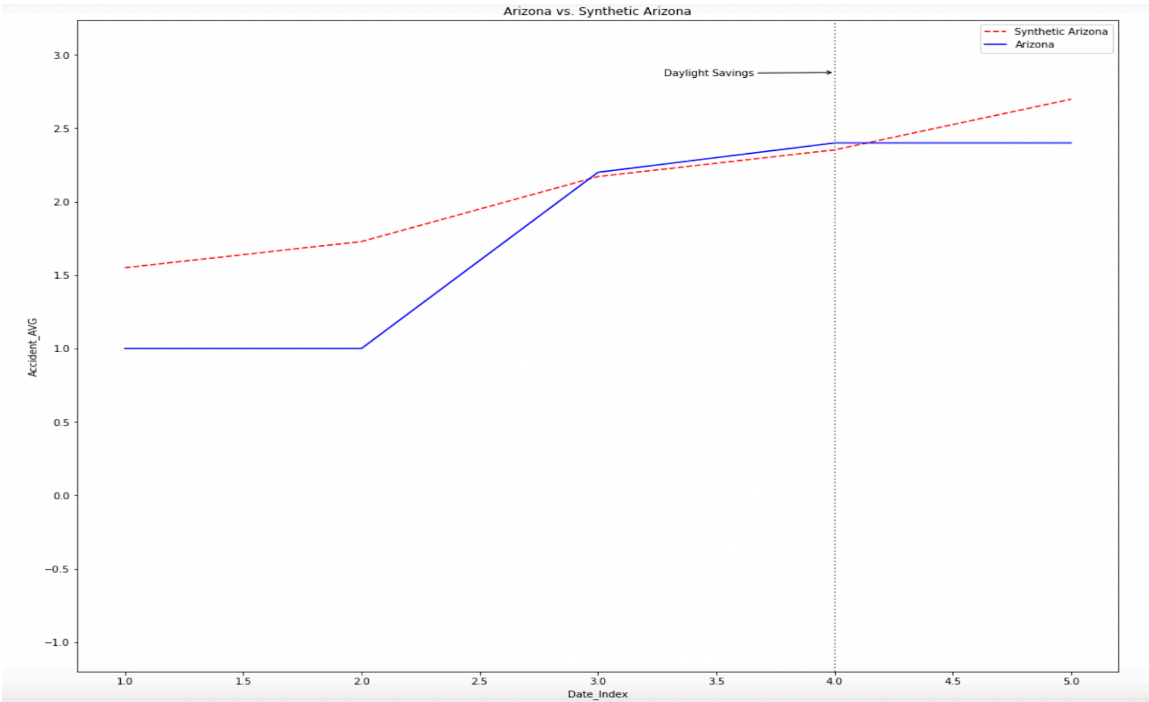

Figure 3