An Analysis of Vehicle Collision Rates & Daylight Savings Time

Adeline Southard

Abstract

This project discusses the following question: does the Spring transition to daylight savings time have an impact on fatal car collisions in the United States? To investigate this, the FARS (Fatality Analysis Reporting System) data in the United States from 1975-2022 is used and regression discontinuity methods are employed to answer this question. Possible confounders are identified, discussed, and accounted for including, most notably, holidays, day-of-week, and day-of-year variation. A comparison to Arizona, which does not observe daylight savings time, is also included. This analysis finds a significant increase in fatal vehicle collisions after Spring transition to daylight savings time, while the data from Arizona alone is inconclusive.

Crash Data

To investigate the possible impact of daylight savings time on car collisions in the United States, data first needed to be sourced. Data was retrieved from the FARS (Fatality Analysis Reporting System) from the NHTSA (National Highway Traffic Safety Administration) database (here). The data spans the years 1975 to 2022, and can be indexed by state the crash occured in. This data was collated into one file for use.

To start, *all* of the data was plotted to visualize the number of crashes over time for the entire United States from 1975 to 2022.

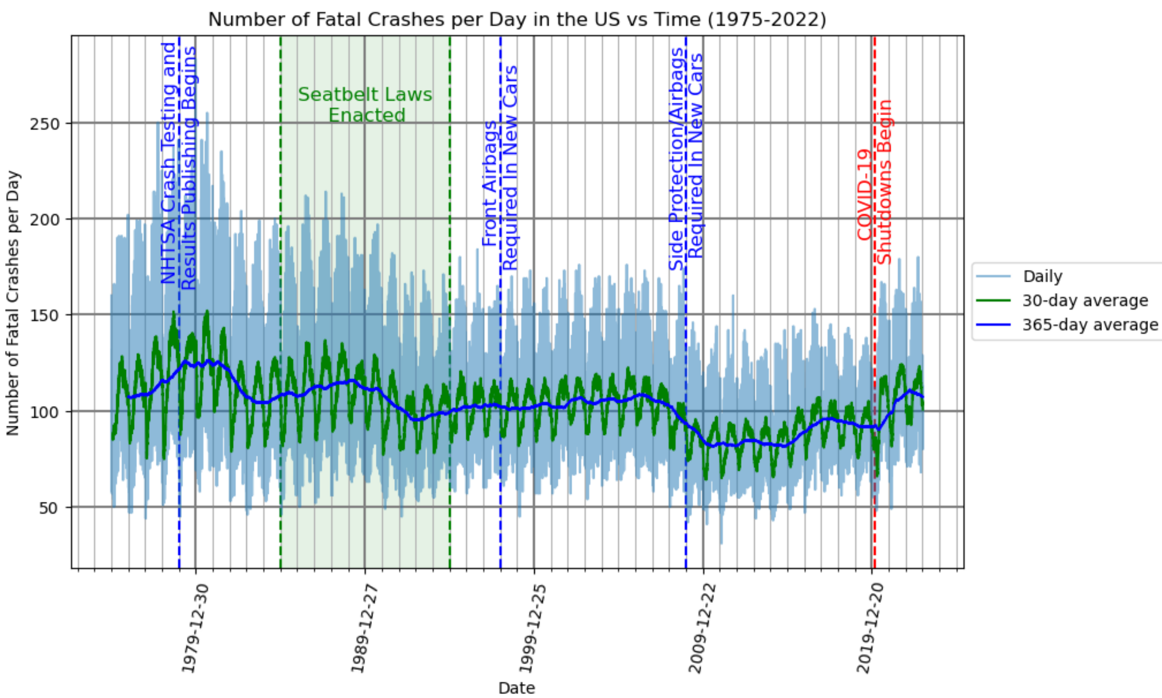

Figure 1. Visualization of all crashes from the FARS data set from 1975 to 2022. Notable events that occurred over time are marked, and the averages for 30-day periods and 360-day periods are also plotted.

Note the effect of the COVID-19 pandemic on the number of fatal vehicle collisions. This will come into play later when discussing confounders in the data.

Accounting for Confounders

Holidays

After cleaning the FARS data and considering initial data exploration, the question arose of what confounders might be taking effect on the data when considering the shift to daylight savings time. The most obvious are Holidays. Holidays and other events like the Super Bowl are associated with social gatherings and subsequent loss of sleep (late night parties) and/or increased alcohol consumption and other drug use. It's fair to posit, then, that holidays could have an effect on the crash data, and if a holiday falls near the daylight savings time change there could be confounding in any analysis that is performed on the data.

For the purposes of this analysis, the following holidays were marked: New Years Day (January 1st), Super Bowl Sunday (first Sunday in February), Mardi Gras (the day before Ash Wednesday), Valentine's Day (February 14th), St. Patrick's Day (March 17th), 420 (April 20th), Cinco de Mayo (May 5th), Memorial Day (last Monday in May), Indepdence Day (July 4th), Labor Day (first Monday in September), Halloween (October 31st), Blackout Wednesday (the Wednesday before Thanksgiving), Thanksgiving (the fourth Thursday in November), Christmas Eve (December 24th), Christmas Day (December 25th), and New Year's Eve (December 31st).

A potentially significant holiday for this analysis is St. Patrick's Day, which is associated with a significant increase in alcohol consumption. This is particularly important since St. Patrick's Day is in March and is close to the spring transition to daylight savings time, often falling in the week immediately following the spring transition. The omission of considering St. Patrick's day may lead to an overestimation of the effect of the spring transition on the number of fatal crashes.

Days that happen to fall on holidays are removed from the data set to account for any potential confounding from them on fatal vehicle collisions.

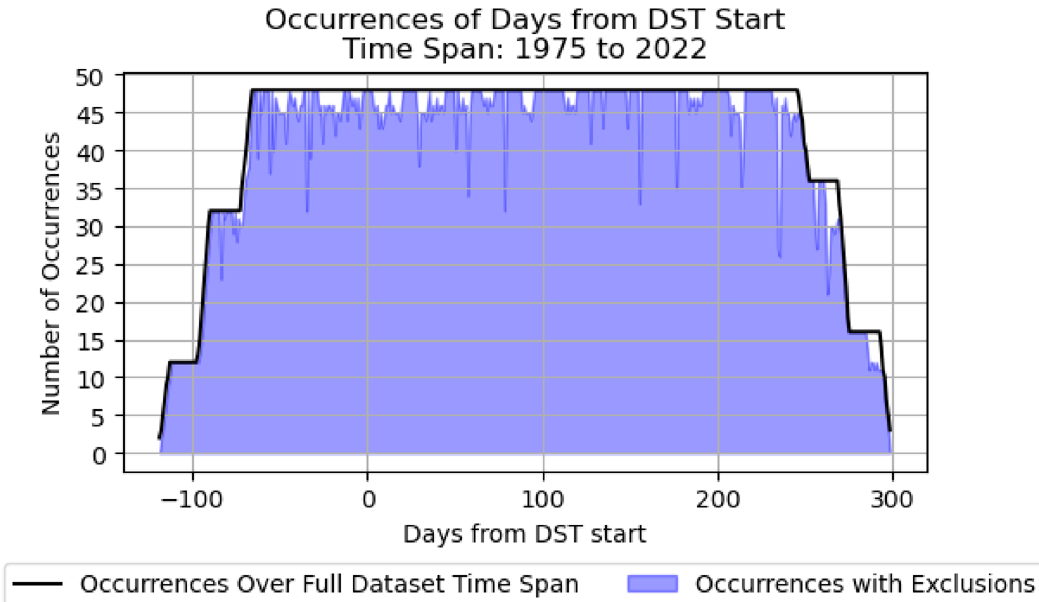

Figure 2. Visualization of the data missing after removing holidays from the data set.

Note in *Figure 2* that the excluded days are not uniformly distributed with respect to the spring transition (zero on the x-axis of the plot). The start and end of daylight savings time is pinned to a specific day of the week: Sunday, so the day of the year that this occurs changes year-to-year. Also, the start and/or end of daylight savings time has changed two times (in 1987 and 2006, which is the source of the two steps on the left and right extremes of the plot) over the time span of the data set due to amendments to the Uniform Time Act of 1966. In fact, there are multiple days before (negative) and after (positive) the start of daylight savings time that are always excluded from the analysis (around Christmas and New Years for example). Luckily there is a significant amount of non excluded data in the week(s) before and after the start of daylight savings time, which is of primary interest.

Consistency of Observance

Indiana did not consistently observe DST over the time period of the data set. Additionally, the state was split into two time zones. To avoid the need for county-level analysis, Indiana is excluded from the analysis.

Arizona and Hawaii did not observe daylight savings time over the time period of the data set, providing an opportunity to compare the number of fatal crashes in states that observe daylight savings time with those that do not.

COVID-19 Pandemic

The pandemic has significantly changed the way people live and work. As shown in *Figure 1*, an increase in the number of fatal crashes occurred after the start of shutdowns in the United States. Because the shutdowns occurred near the spring transition to daylight savings time in March 2020, and we only have data up to 2022 in the data set, only pre-pandemic data is used in the analysis.

Should a correction be made to account for the decrease of day length during spring transition to DST?

One of the main references, and a source of prior research and information into the subject of daylight savings time and car collisions was Smith 2016, “Spring Forward at Your Own Risk: Daylight Saving Time and Fatal Vehicle Crashes.” Smith investigated potential causality between daylight savings time and car collisions, but manipulated the data in a questionable way that became a place of difficulty during this project: should a correction be made to account for the decrease of day length during the spring transition to daylight savings time?

On the day of daylight savings time, an hour of the day is lost, resulting in a missing hour of crashes. The question then arises as to if we should correct for this missing hour in some way—should we add crashes to fill in for the missing hour? This is a potential issue for regression discontinuity because there is no 23-hour day just before daylight savings time.

Smith’s solution to this is to add data to the missing hour—he specifically mentions that “I adjust the crash count by counting the 3-4 am hour twice, using it as a proxy for the missing 2-3 am hour. For the 25-hour fall transition date, I divide the fatalities occurring from 1-2 am by two, because this hour occurred twice.” (Smith 2016).

Since the initial Sunday of daylight savings time is 23 hours long, whereas other days are 24 hours long, it can be argued that since there is one hour less time, the crash count should be adjusted when aggregating data on a daily or weekly basis. As quoted, Smith’s solution to this is to adjust the crash count by doubling the number of crashes occurring in the 3-4AM hour, in effect using the 3-4AM hour as a proxy for the missing 2-3AM hour. Similarily, for the fall transition out of daylight savings time, Smith divided the fatalities occurring from 1-2AM by two, because this hour occurred twice.

It is not clear that this correction is appropriate. One can make the argument that the crash count is not significantly impacted, or even potentially increased, by the Spring transition missing hour since the traffic volume may increase due to the missing hour. The number of commuters probably does not go down because it happens to be the day daylight savings time goes into effect (it is not a holiday). Perhaps the crash count should not be adjusted and in fact the crash count may already be slightly elevated due to an increase in traffic volume given the same number of commuters traveling over a shorter time span. This is particularly important because if the crash count the day of the spring transition is doubled for the 3-4AM hour, as done by Smith, it may introduce an artificial increase in the crash count for that day, leading to a misleading conclusion that the transition to daylight savings time is associated with an increase in the number of fatal crashes.

To investigate this question, it’s first necessary to look at the number of crashes that occur on each hour of the day of the spring transition to daylight savings time (Sunday), as well as on the Sunday the week before:

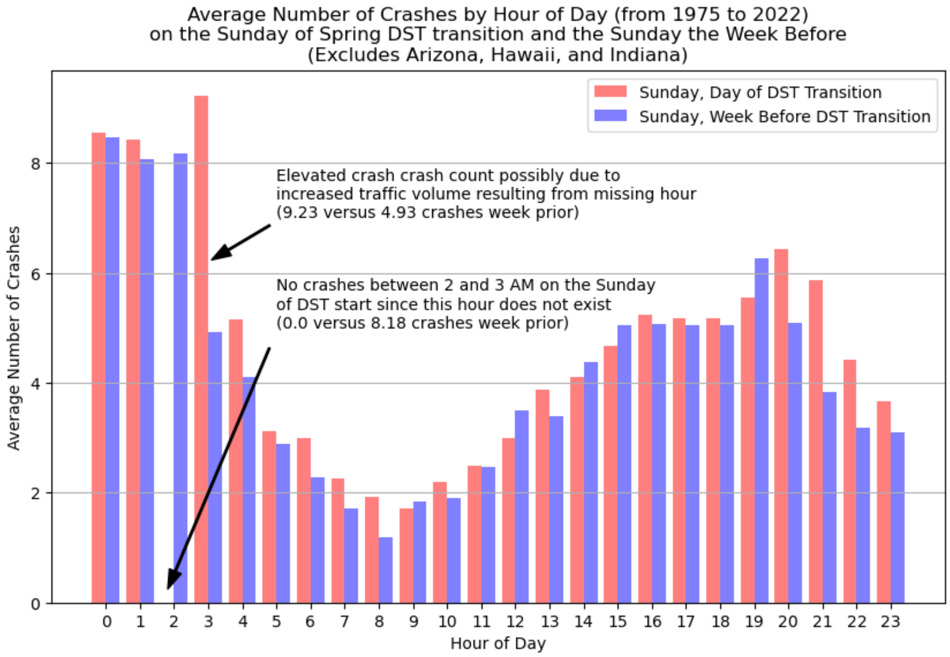

Figure 3. Histogram of the mean crashes per hour in the United States from 1975 to 2022.

Note that in the 2-3AM hour, there are no crashes on the day of the spring transition to daylight savings time. This is because this hour does not exist as mentioned prior (time jumps from 2 to 3AM due to the start of daylight savings time). If the assumption made by Smith is correct and the 3-4AM hour should be used as a proxy for a missing 2-3AM hour, then the number of crashes in the 3-4AM hour on the day of spring transition should be similar to the number of crashes in the 3-4AM hour the week before. If the 3-4AM hour the day of the daylight savings time change has significantly more crashes than the 3-4AM hour the week before, then it is indicative that the crash count is already inflated due to more people on the road making up for the lost hour.

As shown in *Figure 3*, there is clearly a large increase in the crash count in the 2-3AM hour on the Sunday of daylight savings time relative to the prior Sunday. In the 3-4AM hour there is nearly a doubling of the number of crashes on the day of the Spring transition to daylight savings time compared to the week before. This suggests that the assumption made by Smith is not correct, or should at least be subject to further scrutiny. It's a fair conclusion to draw that the doubling of the crash count in the 3-4AM hour the day of the daylight savings time transition likely introduces a significant error with an incorrect bias towards the conclusion that daylight savings time has a negative impact on fatal crashes.

For this analysis, the crash count data is not altered to account for the 23 hour day on the Sunday of the spring transition to daylight savings time. The data clearly shows that the crash count is already elevated, perhaps in part due to an increase in traffic volume peaking in the 3-4AM hour (due to the missing 2-3AM hour), but perhaps persisting to some degree over the day.

Variation in Traffic Volume by Day of Week and Day of Year

The number of commuters on the road varies significantly with the day of the week and year. For example, the number of commuters on the road may be significantly higher on a Friday than a Monday. Similarly, the number of commuters on the road may be significantly higher during the spring and summer relative to the winter.

These confounders are further addressed below.

Day of Week and Day of Year Trends

It's important to recognize that in order to perform a regression discontinuity analysis, it is necessary for the conditions before and after the event of interest be the same. This is not necessarily the case, as it's possible the vehicle collisions occur with varying frequencies depending on the day of year and day of week.

In order to investigate this, two plots were made. One that looks at the daily mean crashes over the course of the entire year, and one that looks at the mean crashes over the day of the week.

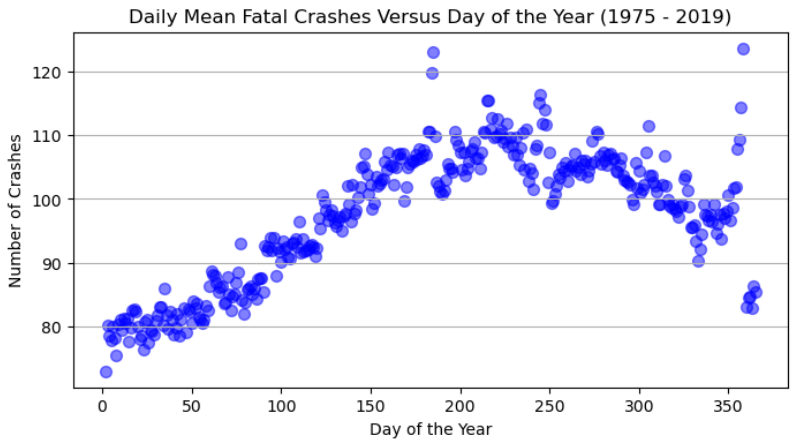

Figure 4. Fatal car crashes per day during the year.

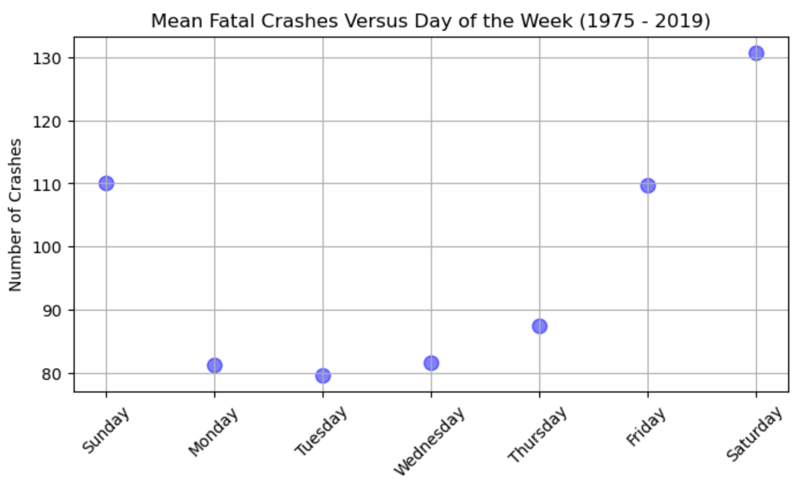

Figure 5. Fatal car crashes per day during the week.

There is significant variation in the data as a function of day-of-week and day-of-year. This presents a difficulty with a regression discontinuity analysis. Ideally the conditions just before and after spring daylight savings time transition should be the same, however this is not the case since the Saturday just before spring transition cannot be compared to the Sunday just after given there is clearly more accidents on Saturday relative to Sunday regardless of the daylight savings time transition. Also, the crash count appears to have a positive slope in the March/April time frame, which is in the time frame of the spring transition to daylight savings time.

In an attempt to deal with this issue, the data will be normalized to a model that accounts for the day-of-week and day-of-year structures in the data. Specifically, Locally Weighted Scatter-Plot Smoothing (LOWESS) will be used as implemented by the statsmodels package in Python. The bandwidth, or fraction of data used, is 0.5 (half of 1 year), which is much larger than the bandwidth that will be used for later regression discontinuity analysis. Hence, the impact of this smoothing on a discontinuity just before and after the spring transition to daylight savings time should be minimal (the impact was not quantified due to time constraints, but could be in a more detailed analysis using, for example, a boostrap approach).

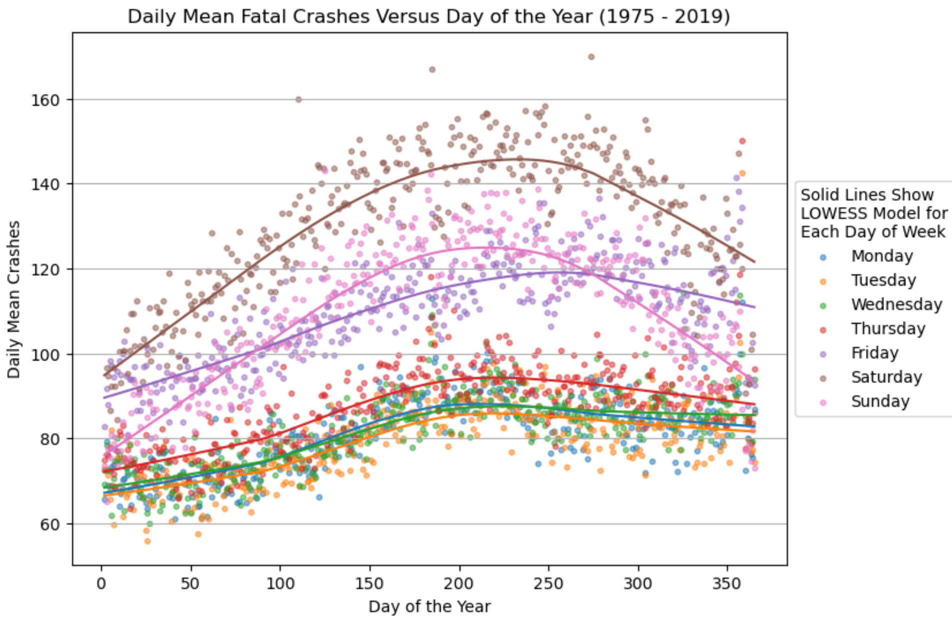

Figure 6. Daily mean crashes.

The data is normalized to the model, that is divided by the model prediction, then scaled so the overall mean prior to normalization is recovered. In effect, the variation due to day-of-week and day-of-year is removed, but any variation over later regression bandwidth timespans are retained because the bandwidth used for the LOWESS model is much larger.

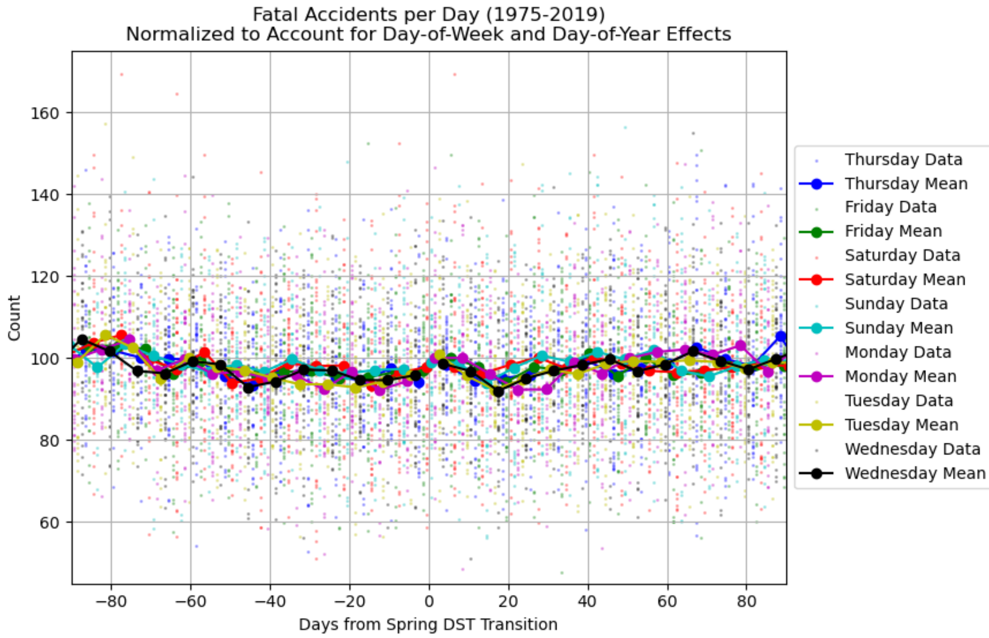

Figure 7. Normalized, then scaled, crashes per day.

Note that *Figure 7* extends over more than 365 days since it is plotted relative to Spring daylight savings time transition, and over the span of the data set the start of daylight savings time has changed two times. Also, note that the normalization to the model appears to have effectively removed day of year and day of week effects.

Finally, the data around the Spring transition to daylight savings time is plotted.

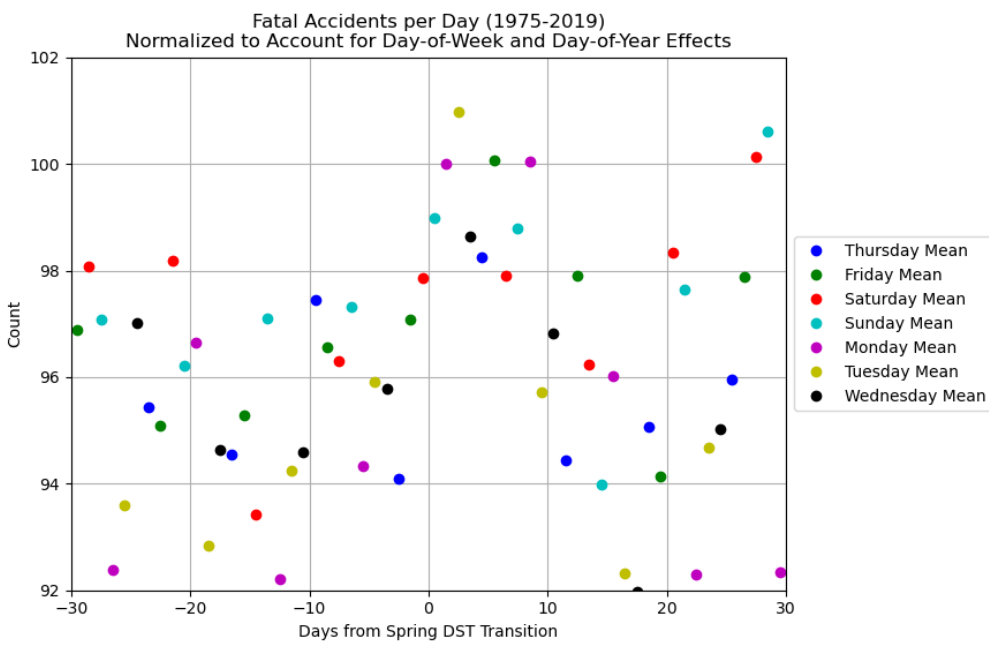

Figure 8. A zoomed in view of the normalized and scaled crashes per day to show the time around the Spring transition to daylight savings time.

Regression Discontinuity Analysis

Now the data is prepared for regression discontinuity analysis. To do this, the normalized and scaled mean fatal crash data was looked at close to the Spring transition to daylight savings time. Lines were fit to either sides of the transition and confidence intervals for each fit were also plotted. Following this, the regression discontinuity between the two fitted lines was measured. The results are shown below.

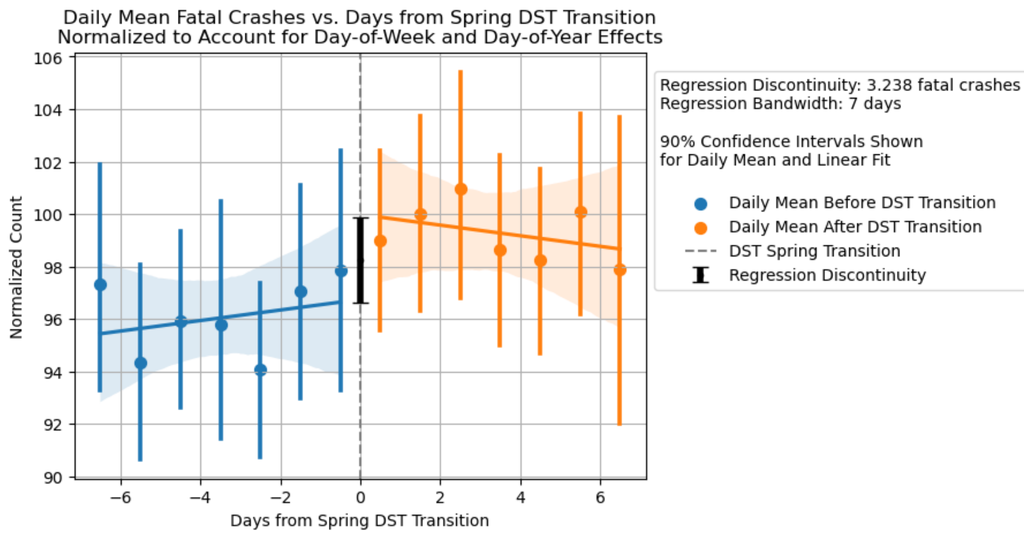

Figure 9. Regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time.

Following this, the bootstrap method was used to create confidence intervals for the results from the regression discontinuity analysis. Those results are shown below. This was done by sampling with replacement the data over the bandwidth, fitting to determine the regression discontinuity, and repeating for ten-thousand bootstrap samples. This is a relatively simple implementation of the boostrap method and there may be more appropriate implementations, but due to time constraints it was not possible to learn about this technique in as much detail as may be necessary.

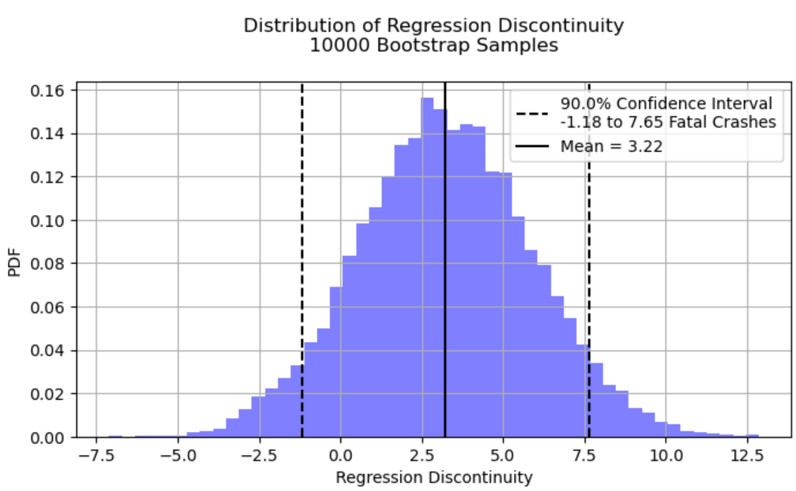

Figure 10. Bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time.

This process was repeated for different regression bandwidths ranging from 7 days to 28 days.

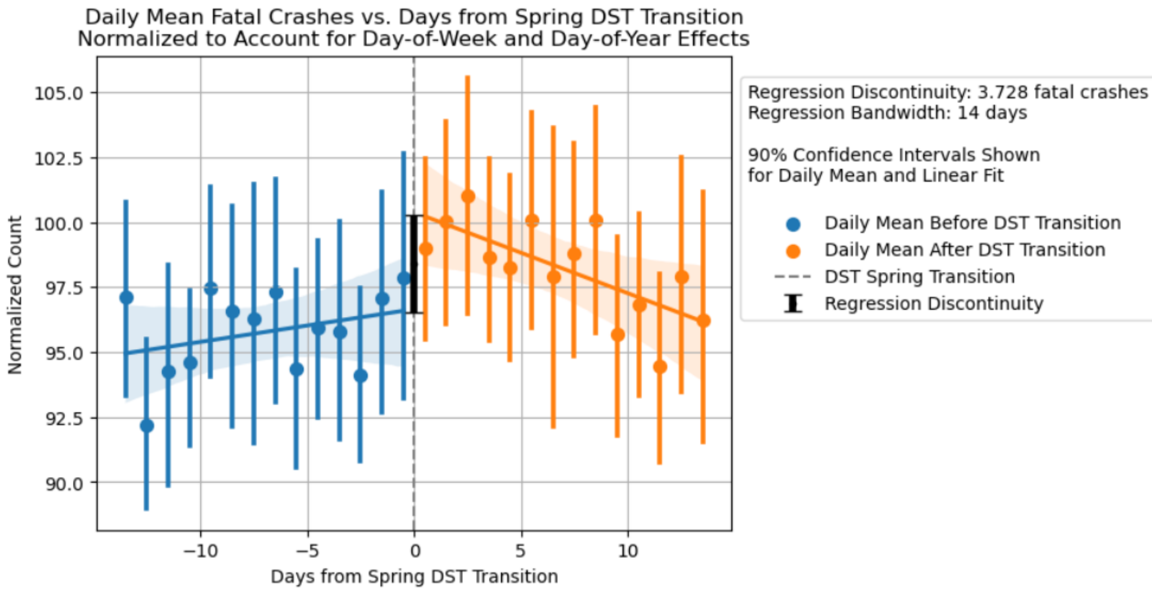

Figure 11. Regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 14 day regression bandwidth.

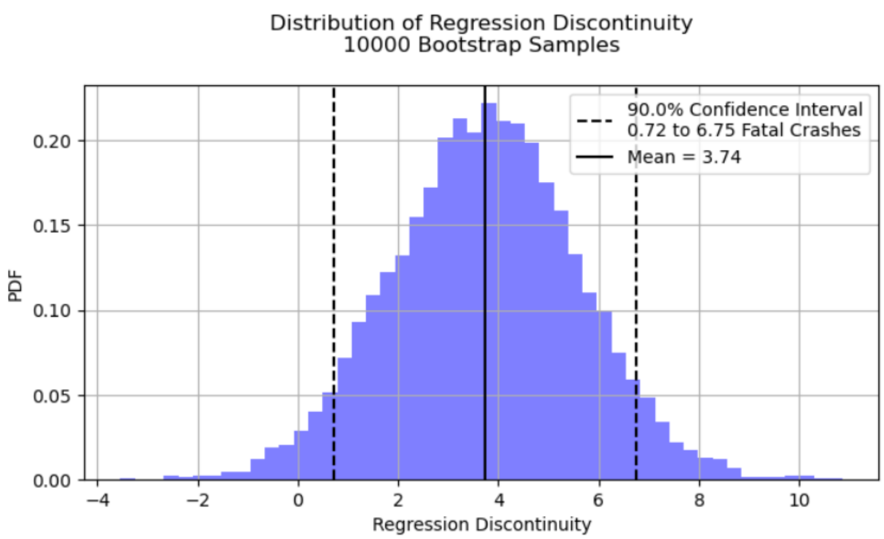

Figure 12. Bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 14 day regression bandwidth.

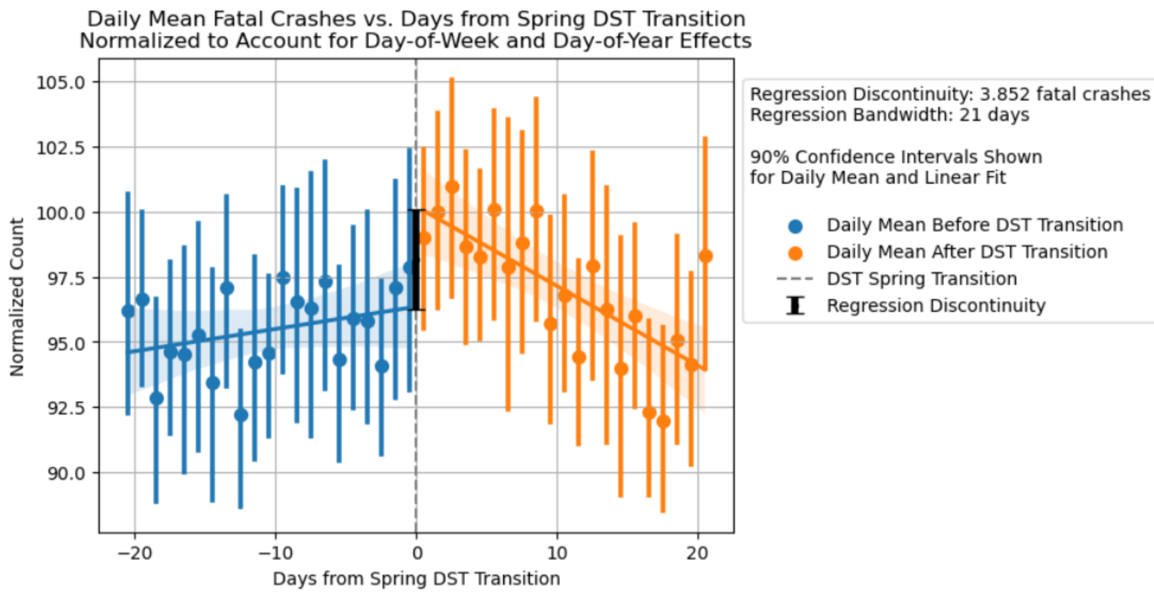

Figure 13. Regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 21 day regression bandwidth.

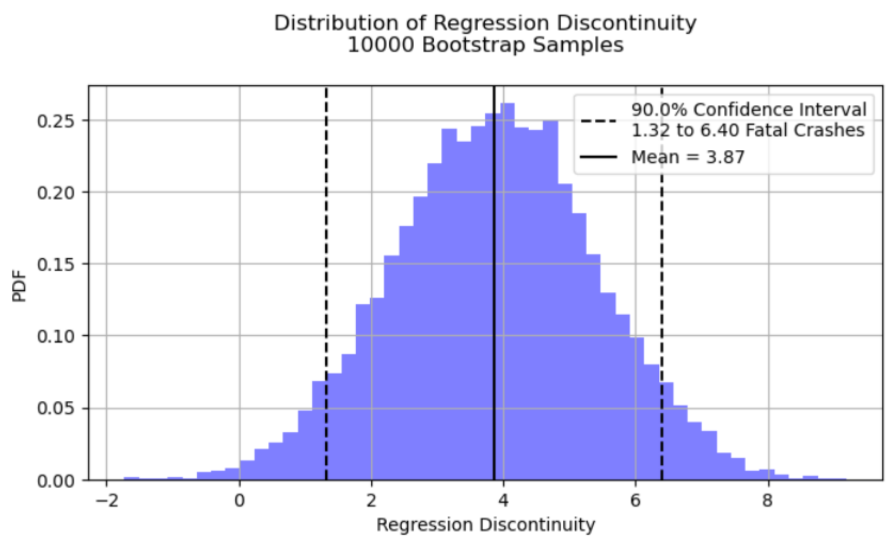

Figure 14. Bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 21 day regression bandwidth.

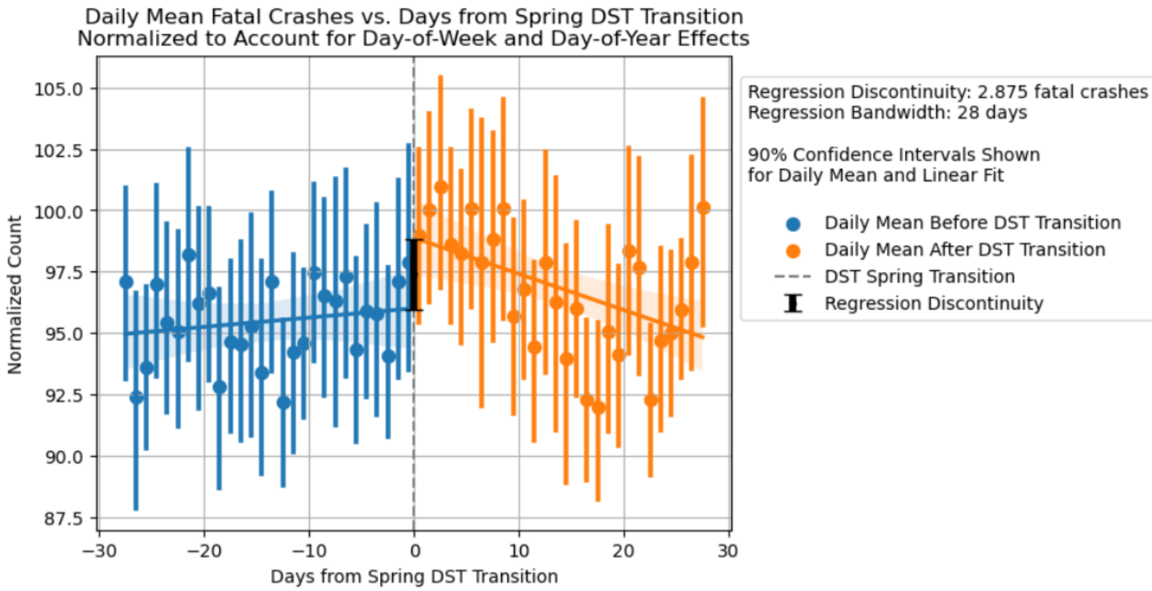

Figure 15. Regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 28 day regression bandwidth.

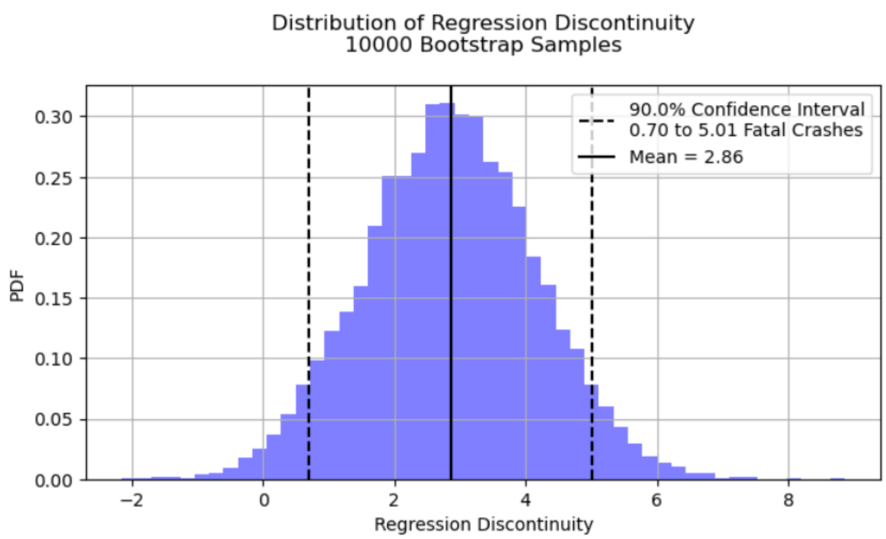

Figure 16. Bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 28 day regression bandwidth.

Analysis of Arizona

Arizona does not observe daylight savings time. This provides an opporunity to compare the number of fatal crashes in states that observe daylight savings time with those that do not. First we look at a regression discontinuity analysis using the same approach as above. If the Spring transition to daylight savings time is associated with an increase in the number of fatal crashes, then we would expect to see no significant discontinuity in the number of fatal crashes just before and after the Spring transition to daylight savings time.

First, *Figure 6* for daily mean fatal crashes is repeated but just for the Arizona crash data:

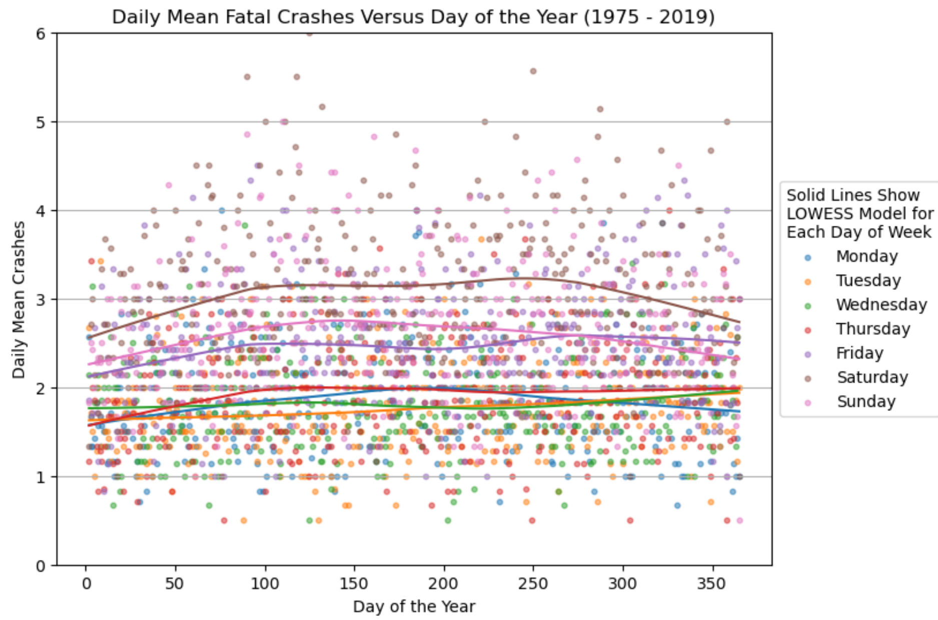

Figure 17. Daily mean fatal car crashes in Arizona.

Note that in *Figure 17* there are significantly less crashes per day in Arizona relative to the prior analysis because the population of Arizona is significantly less than the population of the whole of the states that observe daylight savings time. This is important to keep in mind when comparing the number of fatal crashes in Arizona to the prior analysis. To enable a comparison between the Arizona results with the prior analysis, the number of fatal crashes in Arizona are rescaled by using a ratio of the daily mean crashes for all daylight savings time observing states to the daily mean crashes for Arizona, which is equal to 44.14. As a cross-check, the ratio of the population of all daylight savings time observing states to the population of Arizona is 43.37, which is surprisingly close to the ratio above.

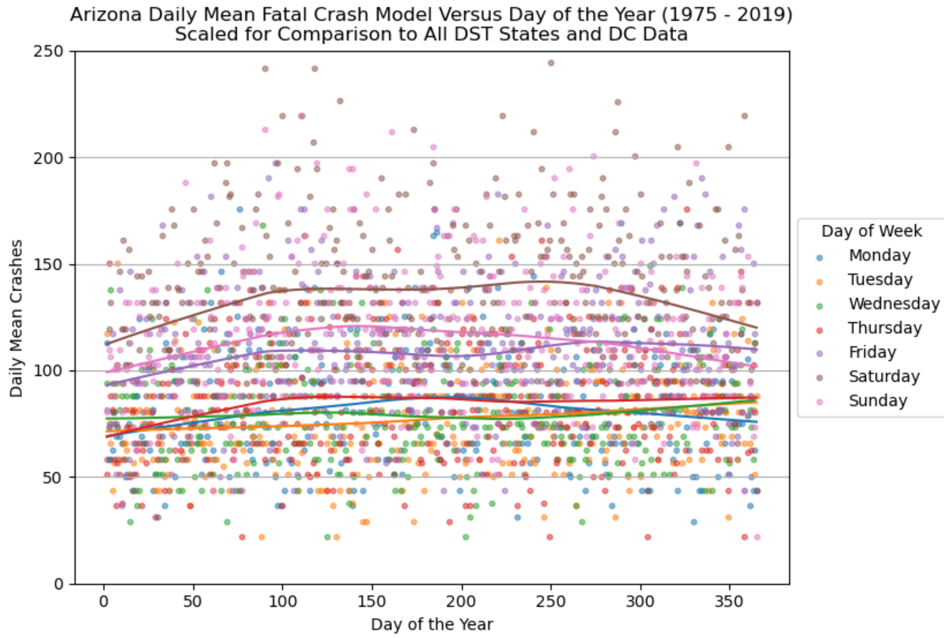

Figure 18. Scaled version of the daily mean fatal car crashes in Arizona.

Like before, the crash data is then normalized and scaled to account for day-of-week and day-of-year effects on the number of crashes that occur.

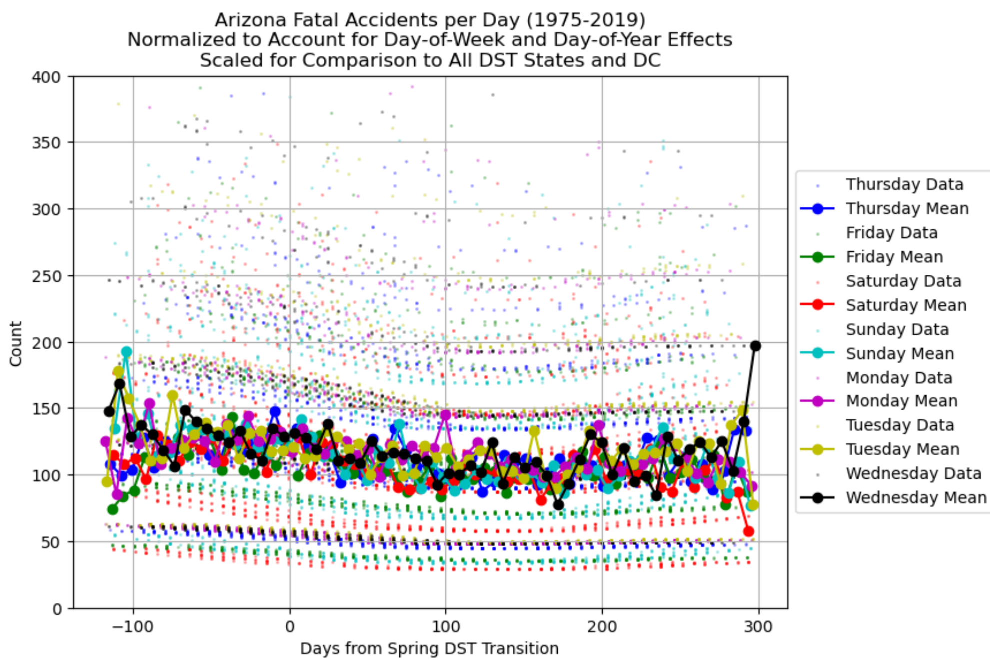

Figure 19. Normalized scaled daily mean fatal car crashes in Arizona.

Again, like before, below is a zoomed in plot of the normalized Arizona crash data showing what the data looks like around the Spring transition to daylight savings time.

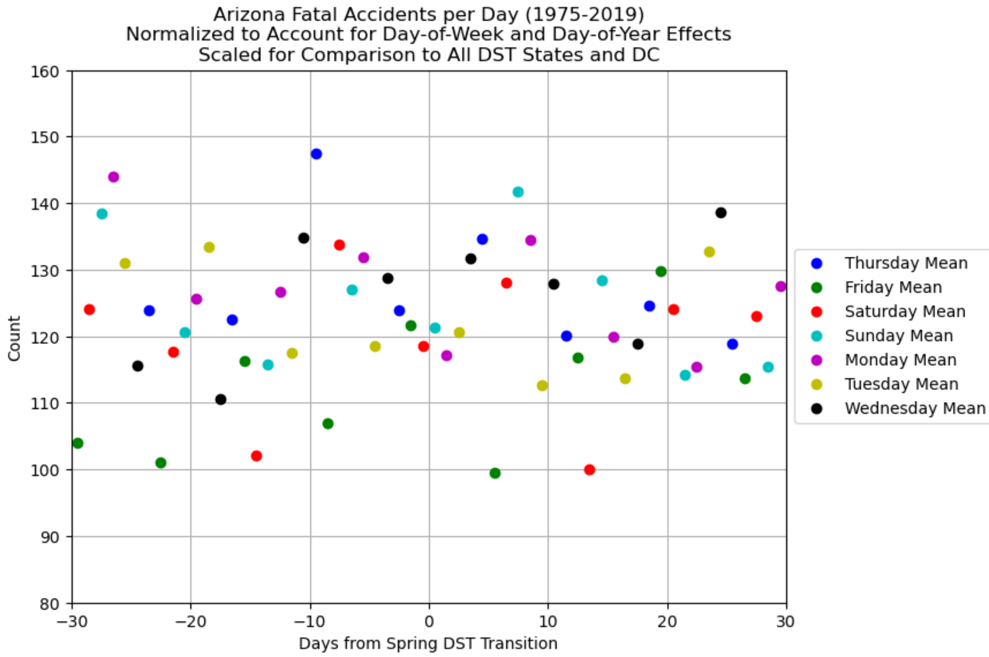

Figure 20. Zoomed in plot of the nomalized Arizona fatal crash data around the time of the Spring transition to daylight savings time.

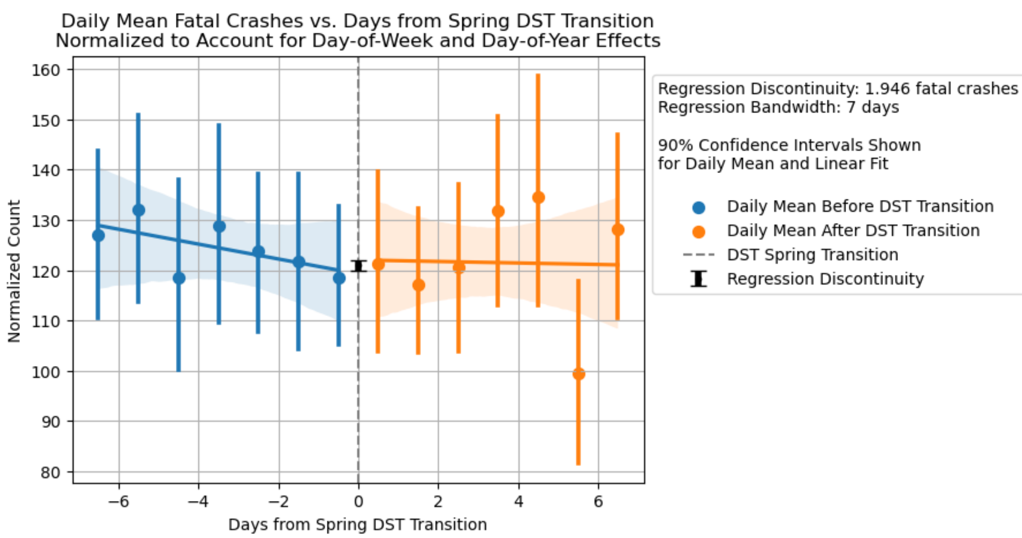

Figure 21. Regression discontinuity analysis with a 7-day bandwith on normalized Arizona fatal crash data.

Finally, a bootstrap analysis was performed again to gain insight into the confidence interval for these results.

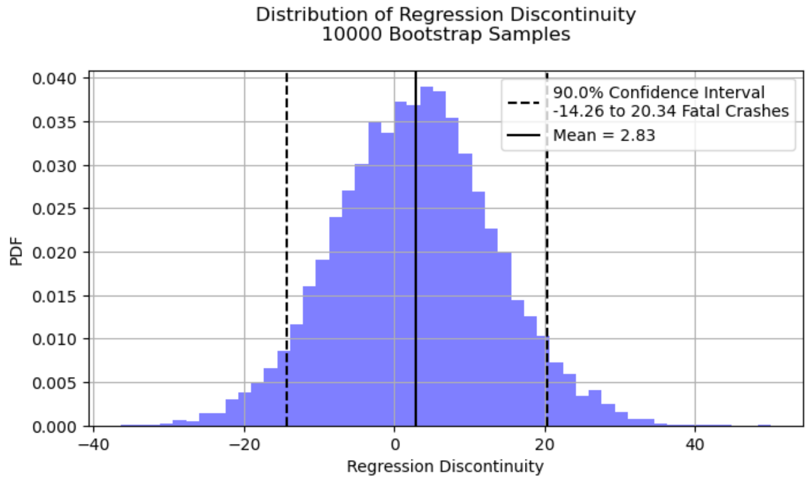

Figure 22. Bootstrap analysis for Arizona data only.

Note that these results have a very large confidence interval, meaning that drawing conclusions for Arizona is difficult. This is primarily due to the lack of data for a single state.

This process was repeated for different regression bandwidths ranging from 7 days to 28 days.

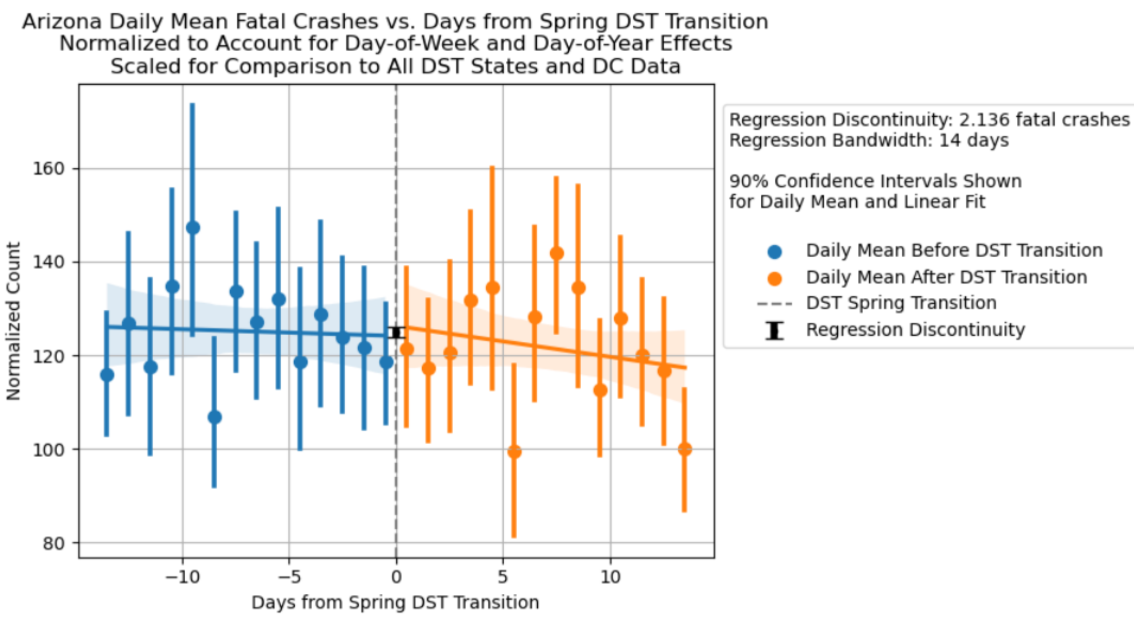

Figure 23. Arizona regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 14 day regression bandwidth.

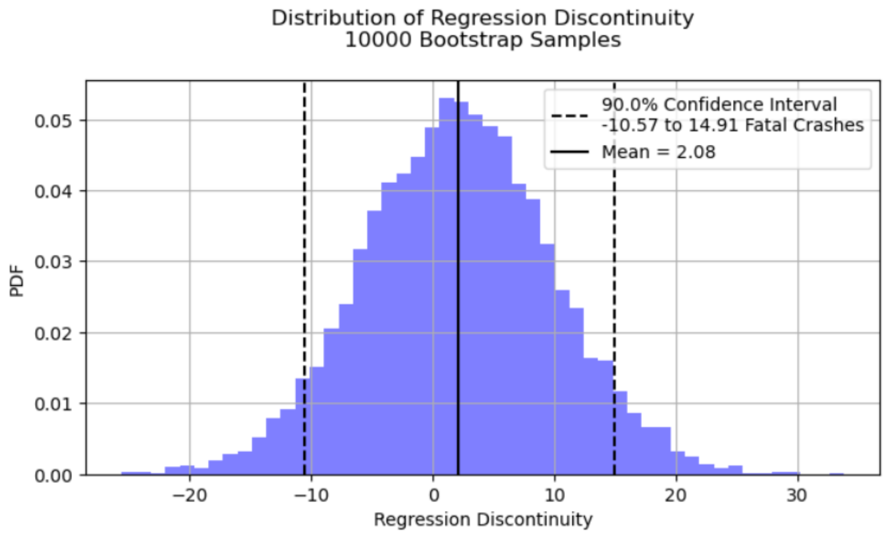

Figure 24. Arizona bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 14 day regression bandwidth.

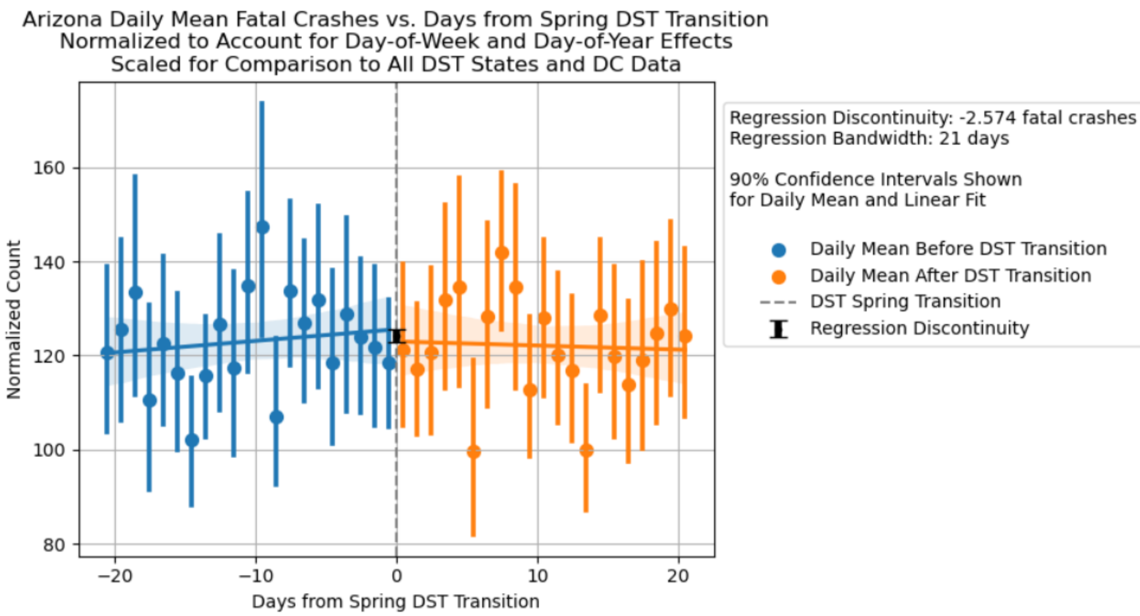

Figure 25. Arizona regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 21 day regression bandwidth.

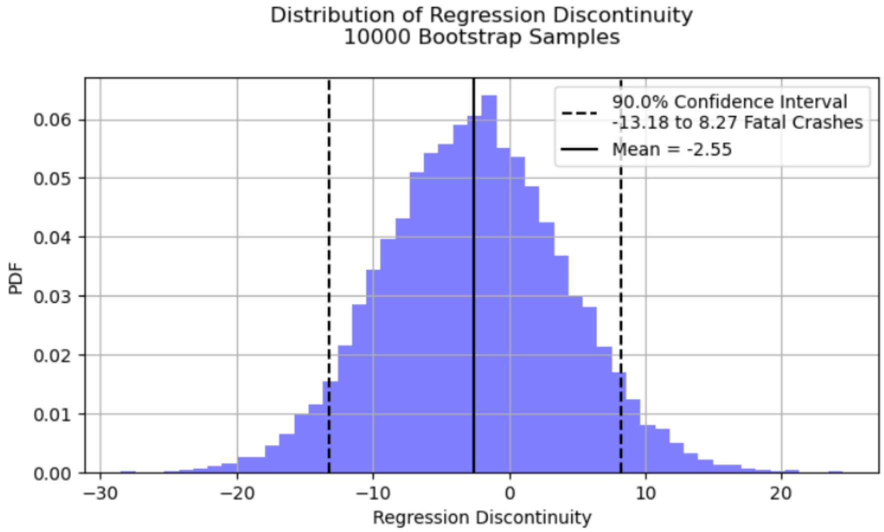

Figure 26. Arizona bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 21 day regression bandwidth.

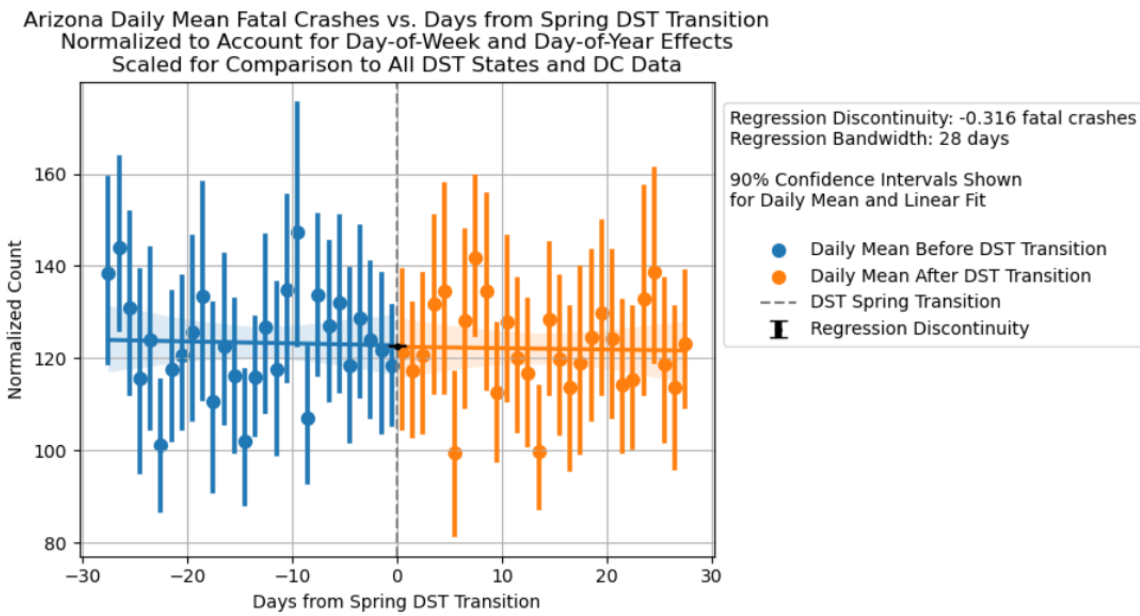

Figure 27. Arizona regression discontinuity analysis around the Spring transition of daylight savings time of states that observe daylight savings time – 28 day regression bandwidth.

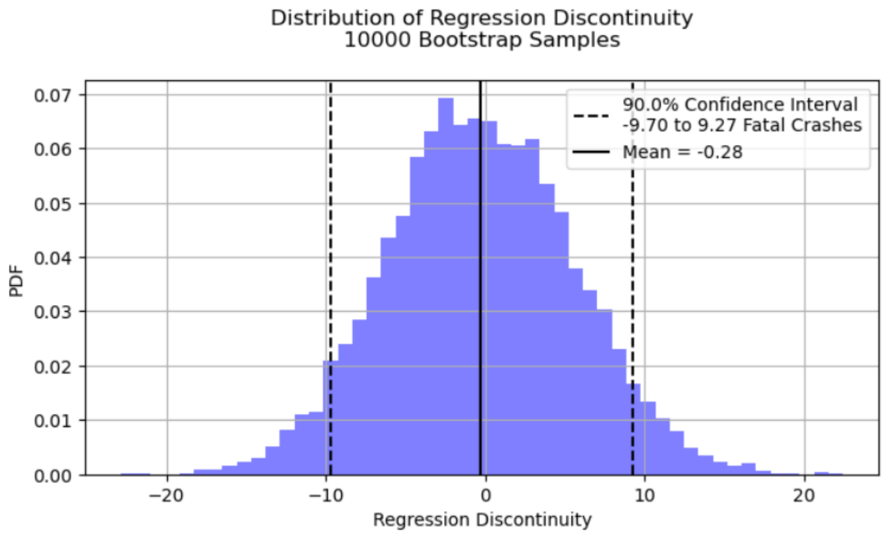

Figure 28. Arizona bootstrap confidence interval analysis of regression discontinuity analysis of fatal car crashes in states that observe daylight savings time – 28 day regression bandwidth.

Summary

As a final summary, to collate the results of regression discontinuity, analyses were done by using different regression bandwidths. Removing a day before and after the holidays was also explored, but the results did not change significantly.

Holidays Removed*	Bandwidth** 7 days			Bandwidth 14 days			Bandwidth 21 days			Bandwidth 28 days		
	Mean	**LCB*** **	**UCB*** **	**Mean**	**LCB**	**UCB**	**Mean**	**LCB**	**UCB**	**Mean**	**LCB**	**UCB**
0	3.59	-0.78	8	3.83	0.8	6.78	3.83	1.38	6.28	2.65	0.53	4.76
1	3.22	-1.17	7.7	3.74	0.72	6.75	3.87	1.32	6.4	2.86	0.7	5.01
3	3.23	-1.18	7.67	3.76	0.72	6.85	**3.97**	**1.37**	**6.6**	3.5	1.34	5.7

* 0 corresponds to no holidays removed, 1 corresponds to the day of the holiday being removed, 3 corresponds to the holiday and +/-1 day removed
** Bandwidth of linear regression used for regression discontinuity analysis
*** Lower and Upper Confidence Bounds (LCB, UCB) are based on a 90% confidence interval

Table 1. Mean and 90% confidence bound on excess deaths due to daylight savings time transition for various regression discontinuity bandwidths and days surrounding holidays that are removed.

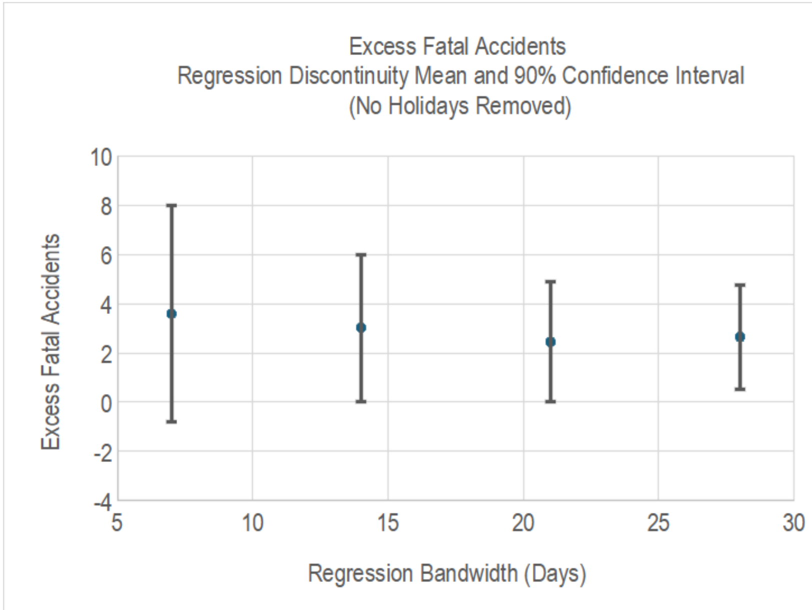

Figure 29. Excess deaths as measured by regression discontinuity for differing regression bandwidths. No holidays removed.

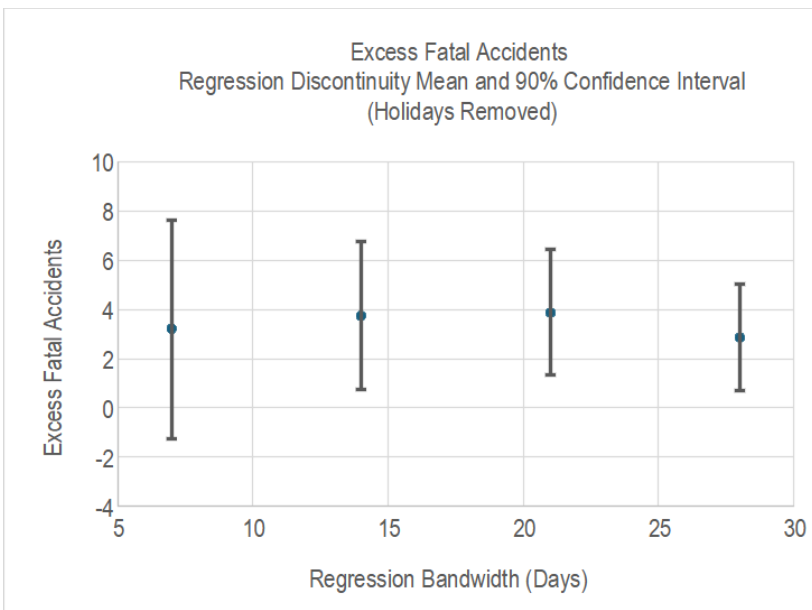

Figure 30. Excess deaths as measured by regression discontinuity for differing regression bandwidths. Holidays removed.

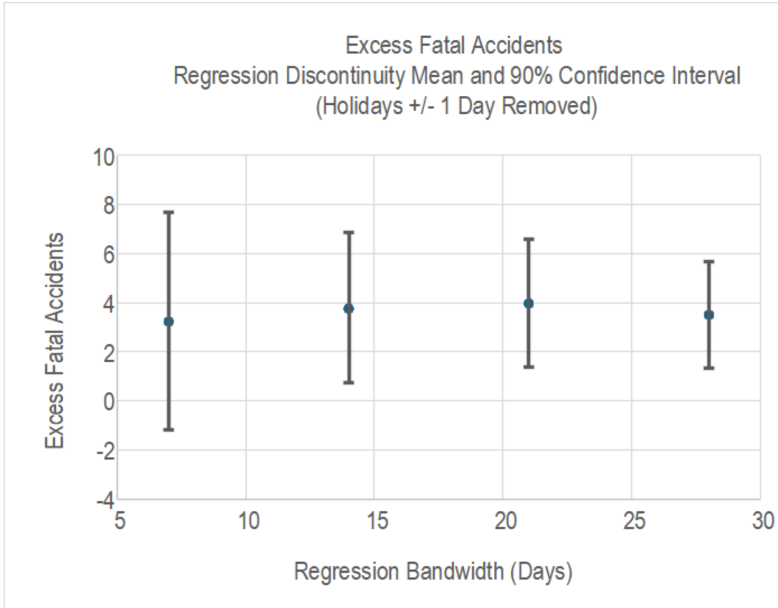

Figure 31. Excess deaths as measured by regression discontinuity for differing regression bandwidths. Holidays and one surrounding day on either side removed.

Conclusions

An analysis using regression discontinuity with various bandwidths, and a bootstrap approach to calculate discontinuity confidence intervals suggests that there is influence of daylight savings time on the number of fatal vehicle crashes in the United States. This is evidenced by an increase in daily crashes for states that observe daylight savings time between 1.37 and 6.60 (90% confidence interval). See *Table 1*.

Contrary to the regression discontinuity analysis for states observing daylight savings time, the Arizona analysis was inconclusive. This is a result of how little data there is present for Arizona, despite there being data from the years 1975 to 2022. The Spring transition to daylight savings time only occurs once per year, and only approximately two fatal car crashes per day occur in Arizona, so that adds up to much less data around Spring daylight savings time transition than one might first imagine when considering how many fatal car collisions there are over the course of 20+ years. For this reason, the confidence intervals for the regression discontinuity in Arizona are significantly larger than what is found in the data containing the rest of the United States, and it is harder to draw a firm conclusion on whether or not there is any discontinuity present.

References

1. Coate, Douglas, and Sara Markowitz. "The effects of daylight and daylight saving time on US pedestrian fatalities and motor vehicle occupant fatalities." Accident; analysis and prevention vol. 36,3 (2004): 351-7. doi:10.1016/S0001-4575(03)00015-0

2. Ferguson, S A et al. "Daylight saving time and motor vehicle crashes: the reduction in pedestrian and vehicle occupant fatalities." American journal of public health vol. 85,1 (1995): 92-5. doi:10.2105/ajph.85.1.92

3. Hicks, G J et al. "Fatal alcohol-related traffic crashes increase subsequent to changes to and from daylight savings time." Perceptual and motor skills vol. 86,3 Pt 1 (1998): 879-82. doi:10.2466/pms.1998.86.3.879

4. Huang, Arthur, and David Levinson. "The effects of daylight saving time on vehicle crashes in Minnesota." Journal of safety research vol. 41,6 (2010): 513-20. doi:10.1016/j.jsr.2010.10.006

5. Laliotis, Ioannis et al. "Summertime and the drivin' is easy? Daylight saving time and vehicle accidents." Health economics vol. 32,10 (2023): 2192-2215. doi:10.1002/hec.4715

6. Lambe, M, and P Cummings. "The shift to and from daylight savings time and motor vehicle crashes." Accident; analysis and prevention vol. 32,4 (2000): 609-11. doi:10.1016/s0001-4575(99)00088-3

7. Liu, Tony et al. "Quantifying causality in data science with quasi-experiments." Nature computational science vol. 1,1 (2021): 24-32. doi:10.1038/s43588-020-00005-8

8. Orsini, Federico et al. "Driving simulator performance worsens after the Spring transition to Daylight Saving Time." iScience vol. 25,7 104666. 24 Jun. 2022, doi:10.1016/j.isci.2022.104666

9. Orsini, Federico et al. "Driving fatigue increases after the Spring transition to Daylight Saving Time in young male drivers: A pilot study." Transportation research. Part F, Traffic psychology and behaviour vol. 99 (2023): 83-97. doi:10.1016/j.trf.2023.10.014

10. Poteser, Michael, and Hanns Moshammer. "Daylight Saving Time Transitions: Impact on Total Mortality." International journal of environmental research and public health vol. 17,5 1611. 2 Mar. 2020, doi:10.3390/ijerph17051611

11. Smith, Austin, C. "Spring Forward at Your Own Risk: Daylight Saving Time and Fatal Vehicle Crashes." American Economic Journal: Applied Economics, 8,2 (2016): 65-91. doi: 10.1257/app.20140100

12. Sullivan, John M, and Michael J Flannagan. "The role of ambient light level in fatal crashes: inferences from daylight saving time transitions." Accident; analysis and prevention vol. 34,4 (2002): 487-98. doi:10.1016/s0001-4575(01)00046-x

13. Sullivan, John M, and Michael J Flannagan. "Determining the potential safety benefit of improved lighting in three pedestrian crash scenarios." Accident; analysis and prevention vol. 39,3 (2007): 638-47. doi:10.1016/j.aap.2006.10.010

14. Uttley, Jim, and Steve Fotios. "The effect of ambient light condition on road traffic collisions involving pedestrians on pedestrian crossings." Accident; analysis and prevention vol. 108 (2017): 189-200. doi:10.1016/j.aap.2017.09.005

15. Zhao, Shi et al. "All cause and cause specific mortality associated with transition to daylight saving time in US: nationwide, time series, observational study." BMJ medicine vol. 3,1 e000771. 6 Mar. 2024, doi:10.1136/bmjmed-2023-000771