

# Analysis of the Potential Causal Effect of Biannual Time Change on the Amount of Traffic Collisions in the City of Los Angeles

Juan Acosta

January 12, 2022

## Abstract

The biannual time change takes place in several countries around the world, where the clocks are turned forward or back, in tune with the shifting daylight hours of the fall and spring. This seasonal change in time can lead to restlessness and fatigue, and may affect activities such as driving. In order to analyze if the time change causes an increase in traffic collision incidents in the City of Los Angeles, traffic collision data from the City of Los Angeles, as well as time change data generated by myself, was used. In order to possibly infer causality from observational data, the difference-in-differences method, in conjunction with a multiple linear regression model, was used. The parallel trend assumption failed, and so causality could not be inferred. Furthermore, the most surprising result was that traffic collision incidents actually decreased during the time periods involving a time change. This led to the conclusion that the biannual time change may not directly cause a change in the number of traffic collisions. However, this conclusion may or may not be limited to Los Angeles, and so more analyses with more cities that also experience the biannual time change would be needed to come to a more informed conclusion.

**Keywords:** Causal Inference, Difference-In-Differences, Parallel Trend Assumption, Biannual Time Change, Los Angeles, Traffic Collisions.

## Introduction

Daylight Saving Time was originally introduced in the United States during the First and Second World Wars, and then uniformly enacted throughout most of the United States, including California, by the Uniform Time Act of 1966 [1]. This was meant to align daylight hours with the times that most people would be awake and going about their daily activities [1]. Thus, the clocks move forward an hour in the spring and move back an hour in the fall [1]. However, much debate has raged over this biannual time change and its impacts on our health and well-being [2, 3]. Transitioning between Daylight Saving Time and Standard Time can affect our circadian rhythms and cause fatigue and sleeplessness that can lead to an increase in traffic accidents on the day of the time change [2]. This is a significant reason why lawmakers in several jurisdictions, including California, have pushed for the elimination of the biannual time change [3].

In this report, the focus will be on the effect of the time change on the number of traffic collisions in the City of Los Angeles, from the years 2010 to 2019, from the day before the time change to the day after the time change. More formally, the goal will be to answer the following question: **Does the biannual time change cause the amount of traffic collisions reported in the City of Los Angeles to significantly increase?**

In order to answer this question, two data sets will be used. The first is traffic collision data from the City of Los Angeles for the years 2010 to 2021 [4]. The second is a data frame (generated by myself) comprised

of several dates in relation to the start and end of DST for the years 2010 to 2019. Both data sets will be used and referenced throughout the rest of this report.

## Terminology

For the remainder of this report, the following terminology and abbreviations will be used:

- **DST:** Short-hand for “Daylight Saving Time”; the time of year in most areas of the United States (including California) when the clocks are set an hour ahead of Standard Time.
- **Standard Time:** The time used in most areas of the United States (including California) during the winter months; when transitioning from DST to Standard Time, the clocks are set an hour back.
- **Difference-in-Differences:** A statistical method that attempts to infer causality in observational data by comparing the differences over time between a control group and a treatment group [5].
- **Parallel Trend Assumption:** One of the most critical assumptions necessary to ensure that the difference-in-differences method is the most suitable for the given observational data. This assumption requires that, given no treatment, the difference between the treatment and control groups is constant, or parallel, over time [5].
- **Treatment Group:** The group that is being exposed to the treatment. In this case, the treatment group is the dates immediately before and after a time change.
- **Control Group:** The group that is not being exposed to the treatment, and is being used as a control to measure the response of the treatment group to the treatment. In this case, the control group is the dates that are four and two days before a time change.
- **Los Angeles:** When “Los Angeles” is used in this report, it is referring to the City of Los Angeles, and not the Greater Los Angeles Area.
- **LAPD:** Short-hand for the “Los Angeles Police Department”.

## Hypothesis

The hypothesis is that there will be a significant increase in the number of traffic collisions in Los Angeles as a direct result of the time change. This can be attributed to fatigue and restlessness caused by the time change, and as a result drivers may be less alert on the road, leading to more accidents compared to a time period without a change in the clocks [2].

## Overview of Following Sections

A brief overview of the following sections of this report is as follows:

- **Data:** This section covers the data collection and cleaning processes, a summary of the important variables, numerical summaries, and important plots and figures for the data analysis.
- **Methods:** This section covers the statistical methodology used in this report, which is the difference-in-differences method to determine if causality can be inferred from the observational data used in this report. The use of this method will be justified by showing that the Parallel Trend Assumption holds for the data used in the analysis. Parameters of interest will be explained and justified by practical rationale.

- **Results:** This section covers the results obtained by running the difference-in-differences method on the traffic collision data, including an interpretation of the results and their significance to the goal of this report. Commentary on whether or not the results seem reasonable is also provided.
- **Conclusions:** This section gives a summary of the initial hypothesis, the statistical causal inference method of difference-in-differences used for the analysis, and the results obtained by running the method. Conclusions drawn from the results will be stated here with respect to the goal of this report. Commentary on limitations and weaknesses, as well as next steps, will also be provided here.
- **Bibliography:** This section lists all the references used throughout this report.
- **Appendix:** This section includes an ethics statement, as well as a supplementary subsection with a glimpse of the data used in this report.

## Data

### Data Collection Process

#### City of Los Angeles Traffic Collision Data from 2010 to 2021

The data were collected on paper by the Los Angeles Police Department, digitized, and published on the online Los Angeles Open Data Portal [4]. The data collected were traffic collision details, including the dates and times of traffic collisions, the location where they occurred, and some demographic data on the victim. Note that only accidents that were reported to, and documented by, the Los Angeles Police Department appear in this data frame.

There are some limitations to this set of data. Since the data was transcribed from paper reports to the digital database, there may be some inaccuracies which is typical with this kind of process. Also, accidents not reported to the LAPD were not recorded in the database, and so the data may not be as complete as it could have been.

#### Time Change Data

The data were manually inputted by myself, using the time change dates listed on the *Time and Date* website [6], as well as other dates in relation to the time change and the dates that will form the control group for the analysis.

### Data Summary

#### City of Los Angeles Traffic Collision Data from 2010 to 2021

The data is presented as a table consisting of 568,118 rows and 18 columns. Each row represents a specific traffic collision incident, and each column represents a specific variable. The values of those variables assigned to each row include the following [4]:

- A Division of Records Number, which is the official file number.
- The date the traffic collision incident was reported.
- The date the traffic collision incident occurred.
- The time (in military time) the traffic collision incident occurred.
- The area ID assigned to a geographic area by the LAPD.
- The area name assigned to a geographic area by the LAPD.
- The reporting district, which is a geographic sub-area within an area, assigned by the LAPD.

- The crime code for the crime committed. In this data frame, the crime code is always 997.
- The crime code description. In this data frame, the crime code description is always “Traffic Collision”.
- The Modus Operandi code, where Modus Operandi is the activities associated with the suspect in relation to the crime, which in the case of this data frame is a traffic collision incident.
- The age of the victim.
- The sex of the victim.
- The ethnic descent of the victim.
- The premise code, which denotes the type of structure or location where the traffic collision incident took place.
- A description of the premise where the traffic collision incident took place.
- The street address, rounded to the nearest hundred block for anonymity, where the traffic collision incident took place.
- The cross street of the rounded street address.
- The location, whose coordinates are rounded to the nearest hundred block for anonymity, where the traffic collision incident took place.

Some of the variables above will be omitted in order to work with a more focused set of data in pursuit of the answer to the initial research question. Also, data from the years 2020 and 2021 will be excluded from the analysis via the data cleaning process, because these were the years of the COVID-19 pandemic, and traffic patterns may have been interrupted during this time, which may lead to a misinterpretation of the final results.

## Time Change Data

The data is presented as a table consisting of 10 rows and 11 columns. Each row represents both time changes, as well as dates that will be used as part of the control group, and time change related information, for a specific year. Each column represents a specific variable. The values of those variables assigned to each row include the following:

- The year.
- The date that is four days before the start of DST.
- The date that is two days before the start of DST.
- The date before the start of DST.
- The date when DST starts.
- The date after the start of DST.
- The date that is four days before the end of DST.
- The date that is two days before the end of DST.
- The date before the end of DST.
- The date when DST ends.
- The date after the end of DST.

## Data Cleaning Process

The data cleaning process for this report makes use of the Tidyverse set of functions in the R programming language [7, 8], and is as follows:

1. The Tidyverse function `select()` was used to select only the most pertinent variable in the collision data for the analysis, which in this case is `Date.Occurred` (the date that the traffic collision incident occurred), to create the template for the clean collision data, which will be used for the rest of the data cleaning process.
2. The Tidyverse function `rename()` was used to rename the variable `Date.Occurred` in the clean collision data to `date_occurred` for a more streamlined cleaning process.

3. First, the Tidyverse function `mutate()` was used to create a variable in the clean collision data called `num_collisions`, whose values will initially be all integer 1s. Second, the Tidyverse function `group_by()` was used in tandem with the Tidyverse function `summarise(across(everything(), sum))` in order to group the clean collision data by `date_occurred` and obtain the value of the variable `num_collisions`, which is now updated to show the count of the number of collisions that occurred on the given date.
4. The Tidyverse function `mutate()` was used to create a new variable in the clean collision data called `year`. Then, the Tidyverse function `filter()` was used to remove the rows where the value of the variable `year` is either “2020” or “2021”, since those years will not be used for the analysis.
5. First, the Tidyverse function `mutate()` was used to create a new variable in the clean collision data called `include_date`, which will initially have a string value of “Blank” for all entries. Second, a for-loop over the range from 1 to 3652 was used, along with if-else if-else statements inside the for-loop, to update the values of the variable `include_date` in the clean collision data so that if the corresponding date is also found in any of the time change data variables (except the variables `year`, `dst_start`, and `dst_end`) from rows 1 to 10, the value of `include_date` updates to “Yes”, otherwise it updates to “No”. Third, the Tidyverse function `filter()` was used to keep only the rows in the clean collision data where the value of the variable `include_date` is “Yes”.
6. First, the Tidyverse function `mutate()` was used to create two variables called `time` and `treatment_group` in the clean collision data, each with initial values of the integer 0 for all entries. Second, a for-loop over the range from 1 to 80 was used, along with if-else if-else statements inside the for-loop, to assign dates in the clean collision data to the treatment group (`treatment_group = 1`) if the date is also found as a value of the variables `day_before_dst_start`, `day_after_dst_start`, `day_before_dst_end` or `day_after_dst_end` in the time change data from rows 1 to 10, and assign dates to the control group (`treatment_group = 0`) otherwise. Third, the Tidyverse function `arrange()` was used to sort the clean collision data by the variable `treatment_group`. Fourth, a for-loop over the range from 1 to 80 was used, along with if-else if-else statements inside the for-loop, to assign dates in the clean collision data to the pre-treatment time of 0 (`time = 0`) if the date is also found as a value of the variables `four_days_before_dst_start`, `day_before_dst_start`, `four_days_before_dst_end` or `day_before_dst_end` in the time change data from rows 1 to 10, and assign dates to the post-treatment time of 1 (`time = 1`) otherwise.
7. The Tidyverse function `arrange()` was used to sort the clean collision data, first by the variable `time` and then by the variable `treatment_group`.
8. The Tidyverse function `mutate()` was used to create a variable in the clean collision data called `id_number` to assign an ID number to match dates based on their time. In this case, the function is used as `mutate(id_number = rep(c(1:40), times = 2))`, where an ID from 1 to 40 is assigned to the dates with variable `time = 0`, and an ID from 1 to 40 is assigned to the dates with variable `time = 1`, and these ID numbers correspond with each other.
9. Finally, the Tidyverse function `select()` was used to select only the variables `id_number`, `time`, `treatment_group`, and `num_collisions` to remain in the clean collision data.

The time change data did not need cleaning, and will no longer be used beyond this point in the report. Only the clean collision data will continue to be used and referenced for the remainder of the report.

## Important Variables

After the cleaning is done, only four important variables remain in the clean collision data:

- **id\_number**: An integer, independent variable that is used to match corresponding dates together for the difference-in-differences analysis.

- **time:** A binary integer, independent variable that assigns a pre-treatment time (**time** = 0) and a post-treatment time (**time** = 1) to an ID number.
- **treatment\_group:** A binary integer, independent variable that assigns an ID number to either the control group (**treatment\_group** = 0) or the treatment group (**treatment\_group** = 1).
- **num\_collisions:** An integer, dependent variable that measures the number of collisions that occurred on a given date.

## Numerical Summaries

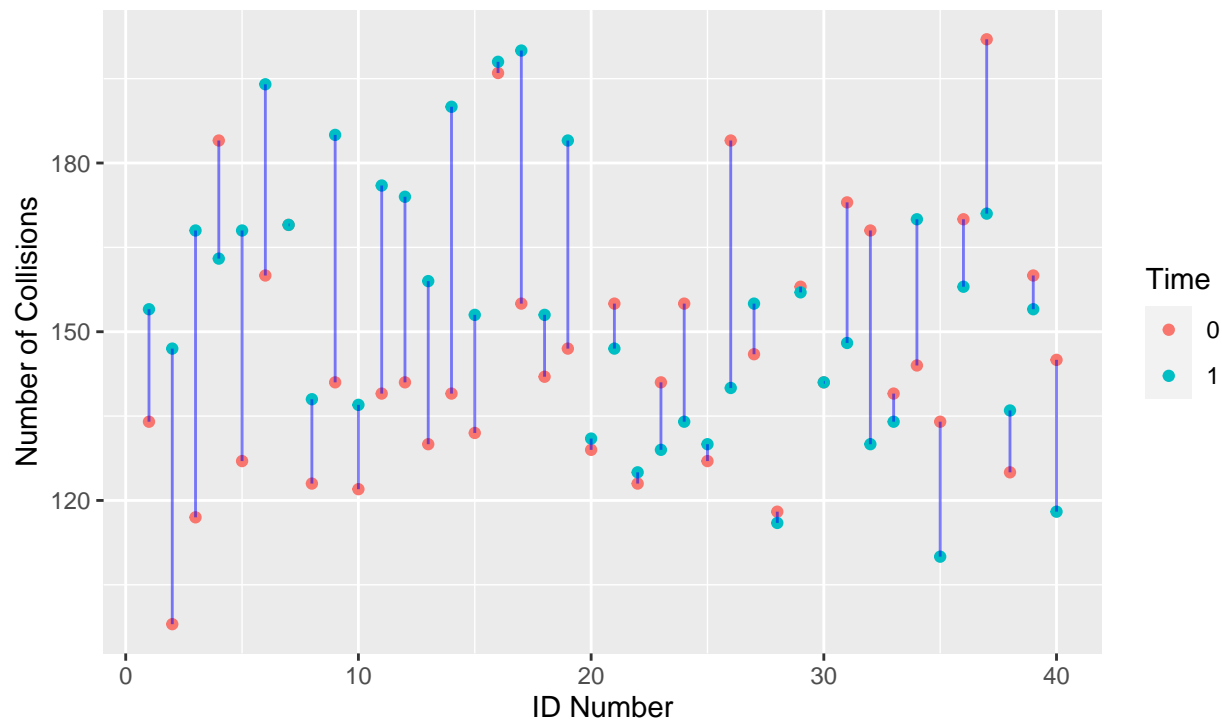
The following numerical summaries were taken on the variable of interest in this report, **num\_collisions** from the clean collision data:

- The mean of **num\_collisions** is 150. In other words, the average number of collisions on a given date, from the dates of interest that were kept in the clean collision data, is 150.
- The variance of **num\_collisions** is 529.
- The standard deviation of **num\_collisions** is 23.

## Data Plots

Plot 1

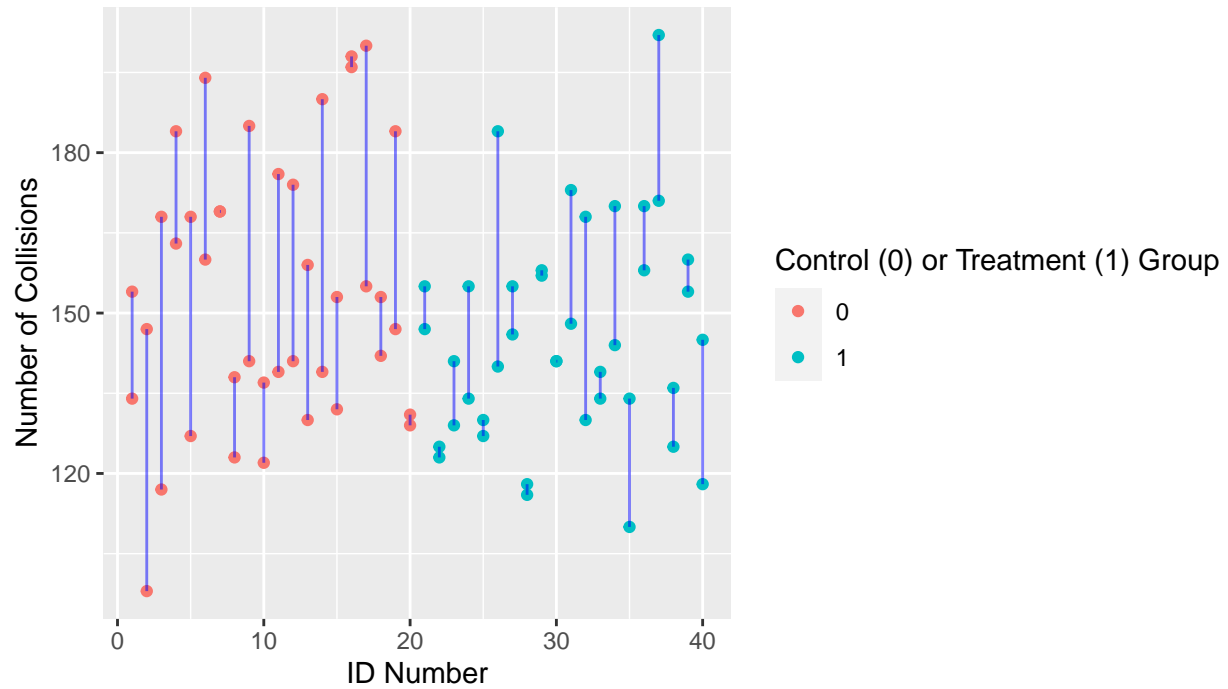
Measuring The Difference in The Number of Collisions Between Time 0 and Time 1 for Each ID Number



Plot 1 demonstrates that the number of collisions at time 0 and time 1 differ for each ID number (pair of pre-treatment and post-treatment dates in either the control group or treatment group). The gaps vary from very minuscule to extensive.

## Plot 2

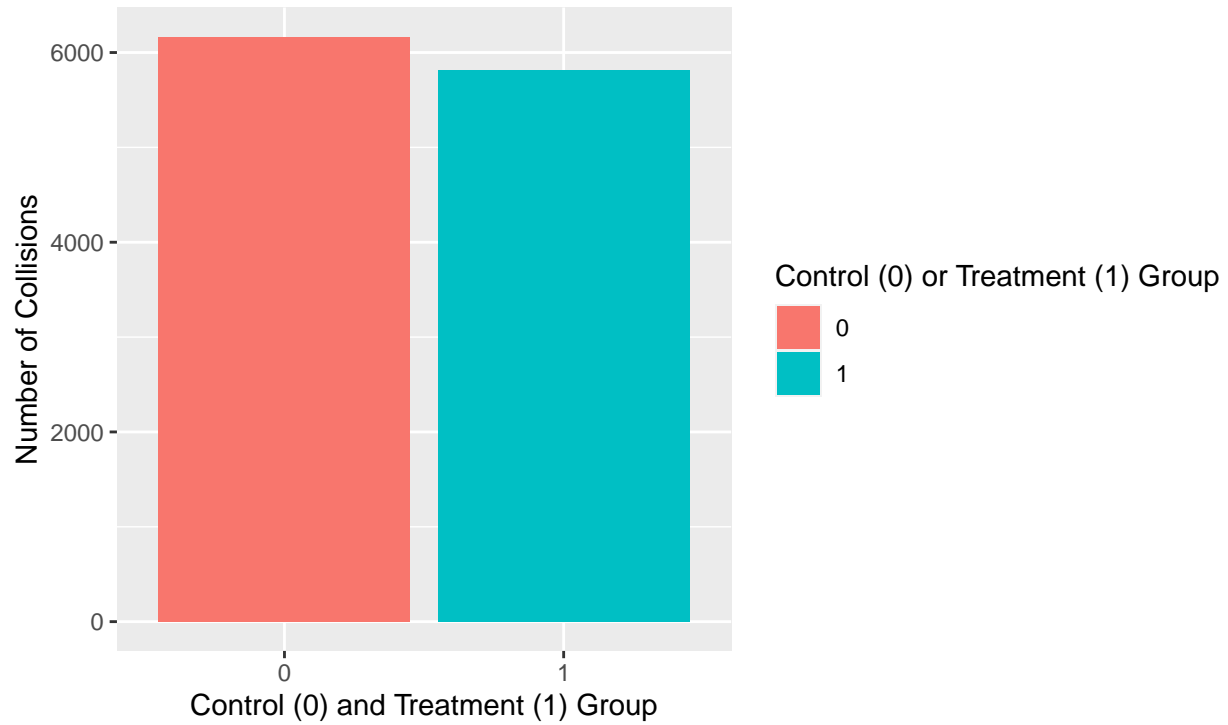
Measuring The Difference in The Number of Collisions Between Time 0 and Time 1 for Each ID Number, Grouped by Control (0) or Treatment (1) Group



Plot 2 demonstrates the same information as Plot 1, except this time the colours of the points highlight which ones belong to the control group, and which ones belong to the treatment group.

Plot 3

Measuring The Total Number of Collisions Between The Control (0) Group and The Treatment (1) Group

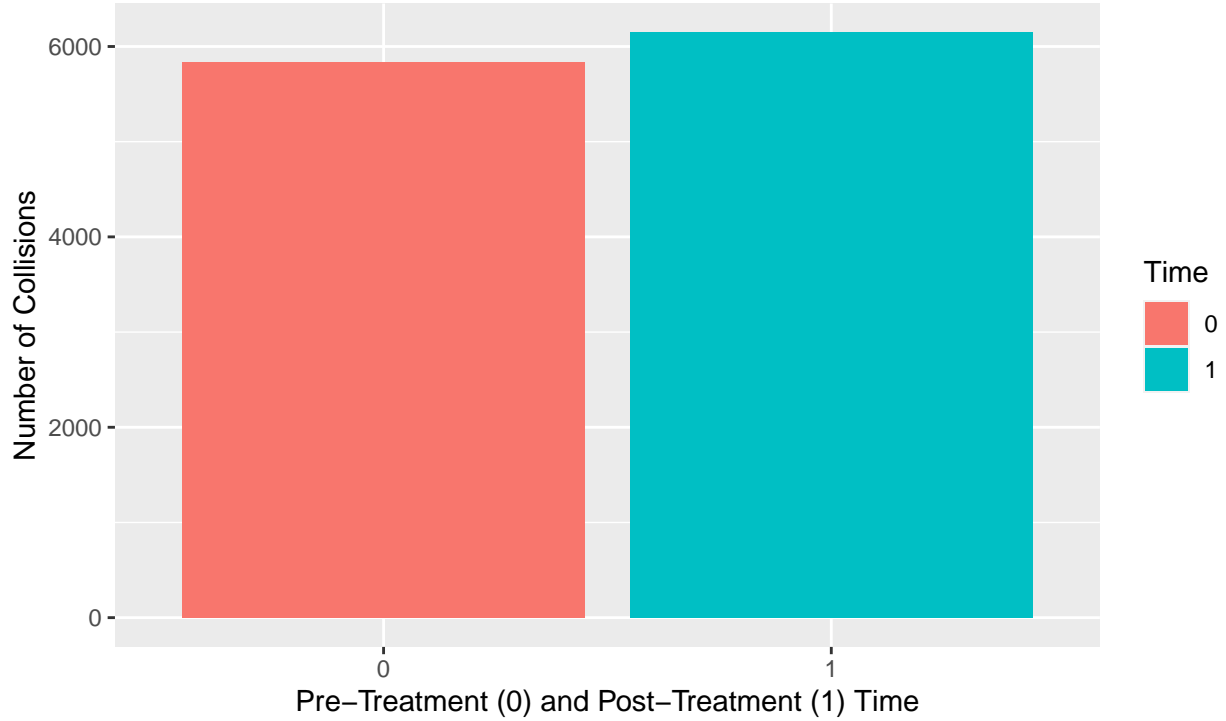


Plot 3 demonstrates that the total number of collisions in the control group is slightly higher than the total number of collisions in the treatment group.



Plot 4

Measuring The Total Number of Collisions Between The Pre-Treatment (0) and Post-Treatment (1) Time



Plot 4 demonstrates that the total number of collisions in the pre-treatment time is slightly lower than the total number of collisions in the post-treatment time.

All analysis for this report was programmed using R version 4.1.1.

## Methods

The methodology that is used in this report is the difference-in-differences method for inferring causality in observational data [5]. Difference-in-differences is implemented using a multiple linear regression model, which predicts the effect that several independent variables will have on the dependent variable.

The primary assumption that will need to be satisfied for the difference-in-differences method to produce a meaningful result is the parallel trend assumption, which states that if there were no treatment effect, the difference between the control and treatment groups is constant over a period of time [5]. This assumption can be confirmed visually using a plot, which will be shown in the **Results** section.

The multiple linear regression model used in this report is  $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \epsilon$ .

- $\beta_0$  represents the intercept of the control group [5].
- $\beta_1$  represents the time trend in the control group [5].
- $x_1$  represents the factorized value of the variable `time`.
- $\beta_2$  represents the difference between the control and treatment groups in the pre-treatment time [5].
- $x_2$  represents the factorized value of the variable `treatment_group`.
- $\beta_3$  represents the difference-in-differences over time [5].
- $x_1x_2$  represents the interaction between the factorized values of the variables `time` and `treatment_group`, respectively.

- $\epsilon$  represents an error term.

For the purposes of the difference-in-differences method, the value of interest will be the predicted value for  $\beta_3$ , which is  $\hat{\beta}_3$ . This will not only tell us if there is a significant difference-in-differences, but if the value of  $\hat{\beta}_3$  is negative, then the parallel trend assumption automatically fails.

The independent variables included in the multiple linear regression model have been selected using the practical rationale technique. Since **time** and **treatment\_group** are essential to the difference-in-differences method, they are included.

## Results

The results of the multiple linear regression model shown in the **Methods** section are summarized in Table 1 below:

Table 1: Multiple Linear Regression Model

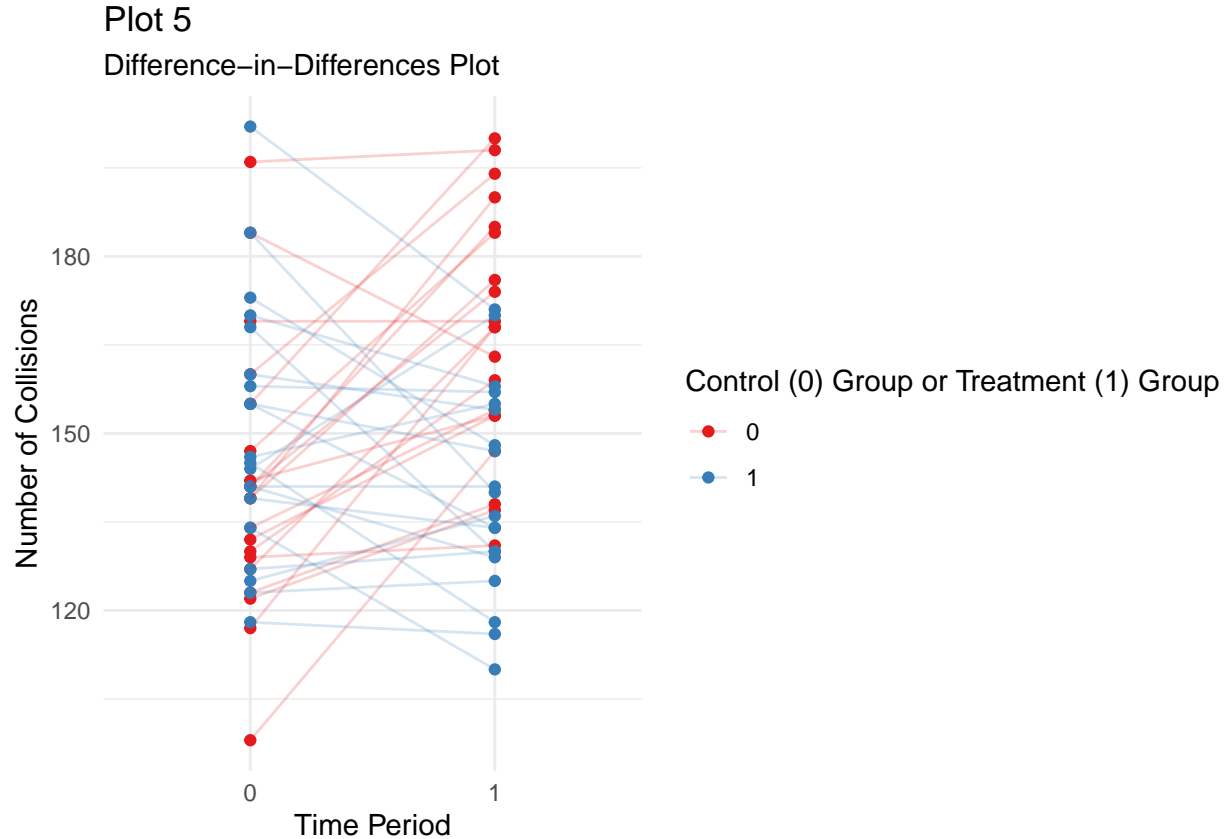
Parameter	Estimate	Standard Error	Statistic	P-value
(Intercept)	141.25	4.627154	30.526320	0.0000000
time1	25.80	6.543785	3.942673	0.0001778
treatment_group1	9.15	6.543785	1.398273	0.1660982
time1:treatment_group1	-36.05	9.254309	-3.895483	0.0002093

The estimated values of the parameters are the following:

- The value of  $\hat{\beta}_0$  is 141.25.
- The value of  $\hat{\beta}_1$  is 25.8.
- The value of  $\hat{\beta}_2$  is 9.15.
- The value of  $\hat{\beta}_3$  is -36.05.

Using a significance value of 0.05, the p-values show that the variable **time** and the interaction between the variables **time** and **treatment\_group** are the most significant in the multiple linear regression model, since their p-values are essentially 0 and so are lower than the significance value.

The value that is of most interest to the difference-in-differences method is the value of  $\hat{\beta}_3$ , which is -36.05. Since this value is negative, it can be safely determined that the parallel trend assumption fails. Not only that, but the value shows that there is a significant difference in the differences. The following plot, Plot 5, will visualize these findings.



Plot 5 demonstrates the previous findings. It can be seen visually that the parallel trend assumption fails, as it cannot be assumed that the control and treatment groups would be parallel to each other without the treatment effect.

Of interest to the analysis of this report, it can be visually seen that in the treatment group, the number of collisions at time 0 tend to be higher than the number of collisions at time 1. This goes against the hypothesis at the beginning of this report, which stated that the number of collisions in the treatment group would increase significantly between times 0 and 1. This, along with the failure of the parallel trend assumption, means that the hypothesis can be safely rejected.

## Conclusions

The initial hypothesis of this report was that there would be a significant increase in the number of traffic collisions in Los Angeles as a direct result of the time change. In order to determine if this was the case, the difference-in-differences method, along with a multiple linear regression model, would be employed. Two groups were created, a control group of dates that did not involve a time change, and a treatment group that involved a time change. Each of these groups had pre-treatment dates and post-treatment dates that corresponded with each other. The multiple linear regression model was run, and the parallel trend assumption of the difference-in-differences method was checked. The results concluded that the parallel trend assumption failed, and so the difference-in-differences method could not be used to infer causality. However, by looking at the difference-in-differences plot, it could be seen that traffic collision incidents actually tended to decrease, and in some cases significantly, during the time period involving a time change, contrary to the initial hypothesis. Thus, the hypothesis was safely rejected, and it cannot be said for certain that the biannual time change has a direct influence on the number of traffic collision incidents in Los Angeles.

## Weaknesses

One of the weaknesses in the collision data was that it only included the years from 2010 to 2021. The years 2020 and 2021 were not included in the analysis of this report because those years involved the COVID-19 pandemic, and so traffic patterns varied with shifting lockdowns and commuting habits. However, the analysis of this report could have been further enriched with more years of data to work with.

Another weakness in the collision data is that it was transcribed from paper reports to the Los Angeles Open Data Portal [4]. This means that there could have been an error, such as an incorrectly inputted date, that may interfere with the results of this report. Thus, the results of this report are only as accurate as the data inputted into the portal.

## Next Steps

A next step would be to gather collision data from other cities around the world where the biannual time change is used, run an analysis like the one in this report, compare results, and determine if there is a consensus or not, and what that signifies in terms of inferring causality.

A recommendation for future analyses would be to tread carefully if including collisions during the COVID-19 pandemic, since the change in commuting patterns may have an impact on how the results are interpreted.

## Discussion

In conclusion, it could not be determined if the biannual time change directly causes a change in the number of traffic collision incidents in Los Angeles. More surprisingly, it was shown that collisions tended to decrease after the time change occurred. It could be that the time change does not have as adverse an effect on the sleep patterns of Angelinos in comparison to cities in more northern latitudes. It could also be that drivers try to be more alert on the road, even if they may be restless, by drinking coffee or listening to music. More analyses would need to be conducted to reach a definitive conclusion. In the meantime, the biannual time change will continue to be a heated source of debate, and a twice-a-year nuisance for some [9].

## Bibliography

1. CBC News (2011, November 2). *Springing forward, falling back: the history of time change*. CBC. Retrieved December 13, 2021, from <https://www.cbc.ca/news/canada/springing-forward-falling-back-the-history-of-time-change-1.755925>.
2. Pacheco, D., & Rehman, A. (2021, November 5). *Daylight Saving Time*. Sleep Foundation. Retrieved December 1, 2021, from <https://www.sleepfoundation.org/circadian-rhythm/daylight-saving-time>.
3. Kwon, D. (2020, October 29). *Governments Worldwide Consider Ditching Daylight Saving Time*. Scientific American. Retrieved December 1, 2021, from <https://www.scientificamerican.com/article/governments-worldwide-consider-ditching-daylight-saving-time/>.
4. City of Los Angeles (2021, December 14). *Traffic Collision Data from 2010 to Present*. Retrieved December 14, 2021, from <https://data.lacity.org/Public-Safety/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w>.
5. Columbia University (2019). *Difference-in-Difference Estimation*. Population Health Methods. Retrieved December 14, 2021, from <https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation>.
6. Time and Date (2021). *Time Change 2021 in the United States*. Time and Date. Retrieved December 15, 2021, from <https://www.timeanddate.com/time/change/usa>.

7. *Tidyverse*. Retrieved December 19, 2021, from <https://www.tidyverse.org/>.
8. *The R Project for Statistical Computing*. Retrieved December 19, 2021, from <https://www.r-project.org/>.
9. Puleo, M. (2021, November 9). *The hidden health risks of daylight saving time*. AccuWeather. Retrieved December 1, 2021, from <https://www.accuweather.com/en/health-wellness/hidden-health-risks-of-daylight-saving-time/1043322>.

# Appendix

## A1: Ethics Statement

One ethical consideration that went into this report was the use of open data. Open data is great from an ethical standpoint because it allows people to use data freely (with the proper attribution, of course) and thus aids in the work of making analyses reproducible. The Los Angeles Open Data Portal offers open data for free and transparent use, and it is one of the main reasons that it was used in this report [4].

Another ethical consideration that went into this report revolved around privacy. The traffic collision data lists some demographic data from the victims of the collisions. In order to keep the anonymity of the victims intact, no names or birth dates appear in the data, and the location of the collision is rounded to the nearest 100 block for added anonymity. Since the demographic and location data was not pertinent to the analysis conducted in this report, it was removed in the clean collision data.

A final ethical consideration that went into this report was the proper citation in the bibliography. In the name of verification, all external sources and information used in this report were cited and included in the bibliography so that the reader may verify the information for themselves, and also if they would like to pursue a certain topic further. It also ensures that credit is given where credit is due.

## A2: Materials

### Glimpse of Uncleaned Collision Data

DR.Number	Date.Reported	Date.Occurred	Time.Occurred	Area.ID
190319651	08/24/2019	08/24/2019	450	3
190319680	08/30/2019	08/30/2019	2320	3
190413769	08/25/2019	08/25/2019	545	4
190127578	11/20/2019	11/20/2019	350	1
190319695	08/30/2019	08/30/2019	2100	3
190411883	07/06/2019	07/06/2019	950	4

Area.Name	Reporting.District	Crime.Code	Crime.Code.Description	MO.Codes
Southwest	356	997	TRAFFIC COLLISION	3036 3004 3026 3101 4003
Southwest	355	997	TRAFFIC COLLISION	3037 3006 3028 3030 3039 3101 4003
Hollenbeck	422	997	TRAFFIC COLLISION	3101 3401 3701 3006 3030
Central	128	997	TRAFFIC COLLISION	0605 3101 3401 3701 3011 3034
Southwest	374	997	TRAFFIC COLLISION	0605 4025 3037 3004 3025 3101
Hollenbeck	423	997	TRAFFIC COLLISION	3101 3401 3701 3003 3025 3029

Victim.Age	Victim.Sex	Victim.Descent	Premise.Code	Premise.Description
22	M	H	101	STREET
30	F	H	101	STREET
NA	M	X	101	STREET
21	M	H	101	STREET
49	M	B	101	STREET
60	M	H	101	STREET

Address	Cross.Street	Location
JEFFERSON BL	NORMANDIE AV	(34.0255, -118.3002)
JEFFERSON BL	W WESTERN	(34.0256, -118.3089)
N BROADWAY	W EASTLAKE AV	(34.0738, -118.2078)
1ST	CENTRAL	(34.0492, -118.2391)
MARTIN LUTHER KING JR	ARLINGTON AV	(34.0108, -118.3182)
MAIN	JOHNSTON	(34.066, -118.2102)

### Glimpse of Cleaned Collision Data

id_number	time	treatment_group	num_collisions
1	0	0	134
2	0	0	98
3	0	0	117
4	0	0	184
5	0	0	127
6	0	0	160

### Glimpse of Time Change Data

year	four_days_before_dst_start	two_days_before_dst_start	day_before_dst_start
2010	03/10/2010	03/12/2010	03/13/2010
2011	03/09/2011	03/11/2011	03/12/2011
2012	03/07/2012	03/09/2012	03/10/2012
2013	03/06/2013	03/08/2013	03/09/2013
2014	03/05/2014	03/07/2014	03/08/2014
2015	03/04/2015	03/06/2015	03/07/2015

dst_start	day_after_dst_start	four_days_before_dst_end	two_days_before_dst_end
03/14/2010	03/15/2010	11/03/2010	11/05/2010
03/13/2011	03/14/2011	11/02/2011	11/04/2011
03/11/2012	03/12/2012	10/31/2012	11/02/2012
03/10/2013	03/11/2013	10/30/2013	11/01/2013
03/09/2014	03/10/2014	10/29/2014	10/31/2014
03/08/2015	03/09/2015	10/28/2015	10/30/2015

day_before_dst_end	dst_end	day_after_dst_end
11/06/2010	11/07/2010	11/08/2010
11/05/2011	11/06/2011	11/07/2011
11/03/2012	11/04/2012	11/05/2012
11/02/2013	11/03/2013	11/04/2013
11/01/2014	11/02/2014	11/03/2014
10/31/2015	11/01/2015	11/02/2015