# Causal Inference Final Project: Investigating Causal Effect of Road Lighting on Severity of Traffic Accidents with Casualties

Uri Menkes 300691367
Code Repository [5]

March 2023

## 1 Introduction

The goal of this causal inference project is to shed light on the relationship between road lighting and *traffic accidents with casualties*. Adequate lighting on roads is crucial for the safety of drivers and pedestrians, yet the extent to which it influences the severity of accidents involving casualties is still uncertain. By utilizing causal inference methods, such as *inverse propensity score weighting* (IPW), *S-learner* etc., this project aims to establish a clear causal relationship between road lighting and traffic accident casualties. The findings of this study may provide valuable insights for policymakers, transportation planners, and other stakeholders to improve road safety and mitigate the severity of accidents involving casualties..

It is noteworthy that the question of causality addressed by this project is conditional in nature, rather than a standard one. Specifically, the question seeks to determine the impact of road lighting on the severity of an accident that has already occurred. Applying the target trial rule of thumb, it may prove challenging to create an experiment that can effectively test this causal effect. Nonetheless, in the subsequent sections, I will provide justifications for why this is indeed an appropriate causal question.

The data for this project was collected from police reports and other official sources. The data includes information on the location, time, and severity of traffic accidents, as well as the characteristics of the road (e.g., lighting conditions, road design, etc.).

The hope is to uncover a positive correlation between road lighting and accident severity, such that improved road lighting leads to a reduction in the severity of accidents with casualties.

## 2 The Data

**Traffic Accident with Casualties.** A traffic accident with casualties is defined ([1]) as a crash involving at least one vehicle and resulting in at least one person being injured or killed. This definition encompasses a variety of scenarios such as vehicle-to-vehicle collisions, vehicle-to-train collisions, vehicle-pedestrian accidents, and vehicle-animal or object accidents. However, accidents resulting solely in property damage, acts of terrorism, and incidents involving suicide or attempted suicide without causing harm to others are not included in this definition. Additionally, multiple vehicle collisions are considered a single traffic accident with casualties.

**The dataset.** The data used in this project is the *Traffic Accident with Casualties - 2019* data set [2], and is based on a monthly administrative file provided by the Israeli Police, as well as on monthly reports provided by Israeli hospitals. It consists of over 12,500 recorded traffic accidents with casualties, and includes over 40 features regarding the the accident and the environment it happened in. For example, the location of the accident (district, municipality, street/road number, etc.), the type of the road (municipal, highway, etc.), the time of accident (day of the week, month, hour of the day), weather and road condition, the type of the accident (head-on collision, rear-end collision, vehicle hit pedestrian, etc.), and more. Full description of the data can be found in [3].

**The outcome.** The feature which will correspond to the outcome in this project is the severity of the accident. It is measured by The Maximum Abbreviated Injury Scale (MAIS), "a globally accepted and widely used trauma scale used by medical professionals" [4]. In this data set it is categorized in 3 levels: minor, major and fatal.

**The treatment.** The feature which will be tested for having a causal effect on the outcome is the lighting on the location of the accident. The Lighting ("TEURA") column in the data set contains 9 different lighting categories, including daylight, twilight, night with proper road lighting, night without road lighting, night with limited visibility and more. The presence of multiple treatment categories or levels may pose challenges when attempting to establish causality during the identification stage, and this issue is addressed in section 2.2 of the discussion.

## 2.1 Exploratory Data Analysis

All features except for two were categorical, and Figure 1 displays the distribution of accident severity across several of these features.

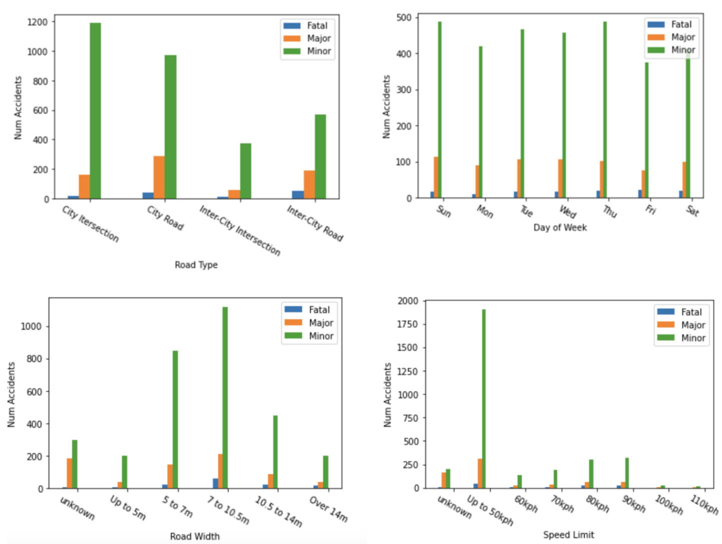


Figure 1: Distribution of accidents' severity with respect to some features.

To handle missing values, I imputed them using the mode value of the corresponding 'unknown' category for each feature, which was already present in the dataset.

The features that were not categorical were the location coordinates 'X' and 'Y'. As there were only a few missing values in these columns, I dropped the entire rows containing them.

I eliminated some features that were redundant or useless in the analysis, such as 'Year' since all accidents were recorded in the same year. This resulted in a final set of 38 features, excluding the treatment and outcome.

## 2.2 Data Pre-Processing for Causal Inference Justification

To find a causal effect of road lighting on the severity of accidents from this data, I filtered out accidents that did not occur during nighttime. Then, I was left with approximately 1000 rows where the lighting condition was labeled as 'unknown'. I attempted to deduce this information from rows of different accidents that occurred in the same location where the lighting condition was known. Among these 1000 rows, I found 60 candidates, but I was only able to make this deduction with high confidence for 8 rows, which all had a recorded lighting condition of "no lighting". Other candidates had different lighting conditions such as "improper lighting", and I could not assume with high certainty that this was the lighting condition for another accident that occurred in the same location. The lighting condition may have been fixed or broken between the two accidents. As a result, I removed the remaining rows with 'unknown' lighting conditions.

The last ambiguous lighting condition I encountered was "proper lighting with limited visibility". This condition could refer to an object that limits the driver's field of vision, such as a tree or heavy fog, which can be considered a hidden confounder. All the rows with this lighting condition have the 'rainy' weather condition recorded as well, but I could not determine if it affected visibility. Therefore, I decided to remove these rows as well.

After the filtering process, I was left with three lighting conditions: "proper lighting", "no lighting", and "lighting doesn't work". Since the last two conditions can be considered as "no lighting", I was able to reduce the treatment to the familiar two-category treatment of "with lighting" and "without lighting". This solution helped overcome a **SUTVA** violation that may have occurred if the treatment had different forms.

Table 1 shows the number of rows left in the data after this filltering.

| **Proper Lighting** | **No Lighting** |
|---|---|
| 2680 | 307 |

Table 1: Treatment distribution

To complete the SUTVA assumption, it's pretty clear that lighting conditions on one accident's location have no effect whatsoever on the severity of different accidents (chain accidents classified as such in the data, so there is no accident that caused another one).

Although the causal question addressed in this study is conditional, meaning that the outcome is based on the occurrence of an accident, a justification for it can be made by examining the data (after the aforementioned filtering) and naively analyzing the distribution through the treatment perspective. Figure 2 shows the proportion of severe accidents (major and fatal) increases when happened in no lighting conditions, which corresponds with the initial assumption.

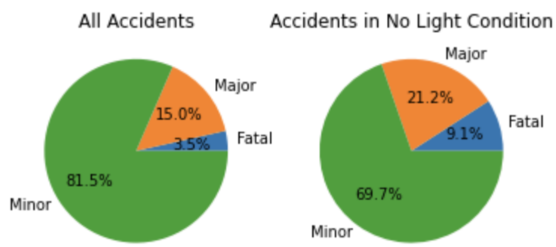


Figure 2: Accidents Severity Distribution

The **Consistency** assumption is also highly likely to be satisfied. The data was collected from police reports, and it is reasonable to assume that the reports were filled out by police officers who arrived at the accident location shortly after it occurred.

As for the **Ignorability** assumption, which concerns the presence of unmeasured hidden confounders, the only potential confounder that comes to mind is the socioeconomic index of the local authority (municipality or local council) in its territory the accident occurred. This is because it could potentially have an impact on both the treatment and the outcome: A wealthier authority may invest more resources in facilities such as road lighting, as well as in education related to safe driving. After conducting further investigation, I have come to the conclusion that the socioeconomic index of the local authority is unlikely to have a significant impact on the outcome. There are several reasons for this. Firstly, there is no guarantee that the drivers involved in the accident are residents of the local authority in which the accident occurred. Secondly, 33% of the recorded accidents occurred on intercity roads, which are not necessarily influenced by the local authority.

Finally, there are other features in the data, such as the region, county, and type of settlement, that can account for differences in lifestyle that may be implied by the aforementioned index.

As we have learned in class, there is no definitive method for verifying there are no unmeasured hidden confounders, however, based on the reasons mentioned above, and further supported by the treatment pre-processing that was previously discussed, I hypothesize that the Ignorability assumption is satisfied.

To address the **Overlap/Positivity** assumption, I trained a model to estimate propensity score. I used XGB classifier and tuned hyper parameters using grid search cross-validation. I added weights to overcome the imbalanced data issue, and the models were evaluated using auc-roc score. All the code used in this project can be found in [5].

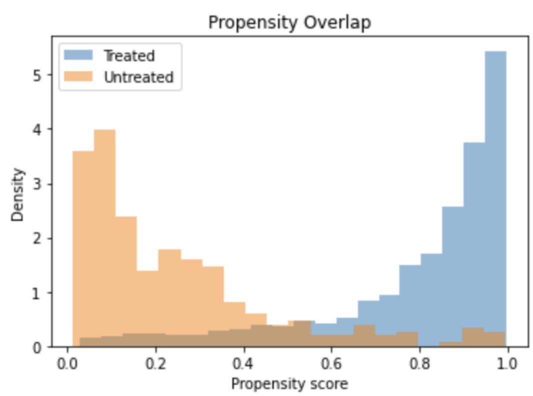


Figure 3: Propensity overlap. Treated means with proper lighting.

We can see in Figure 3 that the overlap is almost ideal, i.e., there are treated and untreated units in almost, every propensity category, except for 1 bin which lacks untreated units. To be precise, there are no untreated (no lighting) units in the propensity range [0.8, 0.86].

However, it's not perfect, and to satisfy the Overlap assumption I excluded units from the treated group that don't have common support from the untreated group. In addition to the above, to address the causal question, I removed units with extreme propensity scores. This was done because such units are likely to be accidents that occurred in locations where it is highly unlikely for the other type of treatment to be present. For instance, the highest propensity scores probably correspond to accidents that occurred at major intersections or on main roads, which would always have adequate lighting. Figure 4 displays the overlap post-trimming, and the results appear to be considerably satisfactory.

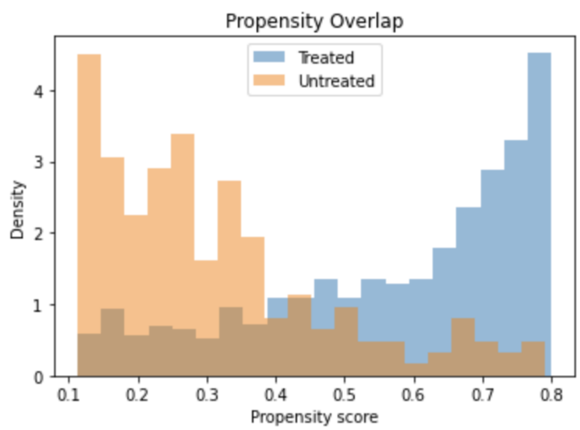


Figure 4: Propensity overlap after trimming.

# 3 Estimating ATE

I used 3 methods for estimating the Average Treatment Effect (ATE) of road lighting on the severity of traffic accidents with casualties.

Given observational data points $(x_1, t_1, y_1), ..., (x_n, t_n, y_n)$ s.t $x_i \in \mathbb{R}^d$ is the covariate, $t_i \in \{0, 1\}$ is the treatment, where 0 is no lighting and 1 is proper lighting, and $y_i \in \{0, 1, 2\}$ is the outcome, i.e., the severity of the accident, where 0 is fatal, 1 is major and 2 is minor. The methods I used are the following:

1. **S-Learner**. Fit 1 model, $f_s$, to estimate the outcome where both $x, t$ are the features, i.e., $\hat{y}_i = f_s(x_i, t_i)$.

$$\widehat{ATE}_s = \frac{1}{n} \sum_{i=1}^{n} f_s(x_i, 1) - f_s(x_i, 0) \tag{1}$$

2. **T-Learner**. Fit 2 models, $f_0, f_1$, to estimate the outcome given only $X$ as features, 1 for each treatment arm, i.e., $\hat{y}_i^0 = f_0(x_i), \hat{y}_i^1 = f_1(x_i)$,

where to fit $f_0$ we use only $x_i$ such that $t_i = 0$ ,and to fit $f_1$ we use only $x_i$ such that $t_i = 1$.

$$\widehat{ATE}_t = \frac{1}{n}\sum_{i=1}^{n} f_1(x_i) - f_0(x_i) \tag{2}$$

3. **Inverse Propensity Weighting (IPW)**. Fit 1 model, $f_{ps}$, to estimate propensity score, i.e., the probability of covarite $x$ to receive a treatment $t = 1$. Formally, $\hat{e}(x) = P[t_i = 1|x_i] = f_{ps}(x_i)$.

$$\widehat{ATE}_{ipw} = \frac{1}{n}\sum_{i=1}^{n} \frac{t_i y_i}{\hat{e}(x_i)} - \frac{1}{n}\sum_{i=1}^{n} \frac{(1-t_i)y_i}{1-\hat{e}(x_i)} \tag{3}$$

The Average Treatment Effects (ATEs) were estimated using a bootstrapping approach. Specifically, for each iteration, a sample was randomly drawn (with replacement) from the data after the trimming process. Then, an XGBClassifier (or two for the second method) was trained using the grid-search cross-validation method and 'F1' score for evaluation, on the sample, and the estimated ATE was calculated using the presented formulas (1), (2), (3).

## 3.1 Results

Table 2 displays the outcomes of the bootstrap procedure, which encompasses the average ATE and its corresponding confidence interval.

| **Method** | **ATE** | **95% CI** |
|---|---|---|
| S-Learner | 0.0243 | (0.0, 0.072) |
| T-Learner | 0.0736 | (-0.0495, 0.2266) |
| IPW | -1.1656 | (-15.3755, 1.5767) |

Table 2: Bootstrapped ATE in different methods

We can see that the average ATEs are inconsistent, and further more that the confidence interval for the ATE includes 0. It suggests that the treatment effect is not statistically significant at the chosen level of confidence. In other words, the results indicate that the treatment may not have a meaningful impact on the outcome, or that the observed effect could be due to chance or other confounding factors.

# 4 Weaknesses

There are several potential areas for improvement when estimating Average Treatment Effects (ATEs). Firstly, there is always a possibility of unmeasured confounding factors that were not accounted for in the analysis. Secondly, although the models used to estimate the ATEs may have provided satisfactory results, there may be opportunities to further optimize the models, such as through improved feature selection or tuning of hyperparameters. Additionally, the dataset used to train these models was very imbalanced (in terms of the treatment), and it may be necessary to adjust for any imbalances in the data to ensure accurate estimates of treatment effects. Overall, it is important to be aware of these potential limitations and to take steps to address them when estimating ATEs.

In relation to the results obtained using the IPW method, which appear to differ significantly from the other two methods: it should be noted that this method relies on the use of probabilities and can be more sensitive to extreme results, such as values that are close to 0 or 1. As the extreme propensity scores were not trimmed during each iteration of the bootstrap process, this is likely to be the main factor contributing to the differences observed in the results obtained with the IPW method.

# 5 Discussion

Despite my initial expectation to discover a causal effect between road lighting and the severity of traffic accidents, the findings indicate otherwise. However, this outcome is not entirely surprising. It is more reasonable to assume that the factors that truly influence accident severity are physical details related to the event itself, such as the type of vehicles involved, their safety features, and their speed at the time of the accident. Therefore, it is important to take into account these physical variables when studying the relationship between road lighting and accident severity.

Having said that, it is still possible that a causal effect between road lighting and the severity of traffic accidents could be discovered through the use of alternative analysis methods, such as Do-Calculus. Additionally, collecting more data or limiting the analysis to specific types of accidents, such as those occurring at intersections, may also reveal evidence of a causal relationship.

Finally, I think it would be worthwhile to explore the causal effect of road lighting on the likelihood of traffic accidents occurring in the first place. This avenue of investigation holds promise for yielding meaningful insights, although it would require collecting new data and developing a separate project.

# References


[1] The Israeli Central Bureau of Statistics. URL https://www.cbs.gov.il/he/pages/default.aspx.

[2] The Israeli Government Database. URL https://data.gov.il/dataset/teunot2019.

[3] Used Data Description. URL https://data.gov.il/dataset/teunot2019/resource/595955cf-914f-42e1-bd97-cf6211c6e4e5.

[4] Association for the Advancement of Automotive Medicine. Abbreviated Injury Scale (AIS). URL https://www.aaam.org/abbreviated-injury-scale-ais-position-statement/.

[5] Uri Menkes. Code Repository for This Project, 2023. URL https://github.com/uriMen/Causal_inference_project.