

# The Route 508 Streetcar and Congestion on Toronto's Shoreline, 2011 - 2020\*

Examining Traffic Patterns Before and After Public Transit Improvements

Maria Mangru

April 1, 2024

This study investigates the impact of enhanced public transit infrastructure on car congestion. It specifically focuses on the reintroduction of the Route 508 Lake Shore streetcar between 2014 and 2020. The data, sourced from the City of Toronto's Transportation Services Division, includes detailed traffic volumes for various modes at intersections across the city. The method of analysis includes comparing the negative binomial regression models of daily car traffic before and after transit upgrades. The results highlighted a significant reduction in the relationship between bus traffic and car congestion. Additionally, it was observed that car traffic decreased over time. These findings highlight the usefulness of public transit improvements in reducing urban congestion.

## 1 Introduction

Traffic congestion represents a significant challenge for urban cities. It is a non-productive activity which negatively affects economic efficiency, environmental sustainability and overall quality of life. To address this challenge, many cities like Toronto have undergone improvements to public transit infrastructure with the aim of reducing car dependency and mitigate traffic congestion. However, it is still highly debatable whether or not such improvements lead to a decrease in car congestion.

This study specifically investigates the 508 Lake Shore Streetcar route in Toronto. It utilizes two datasets provided by the City of Toronto's Transportation Services Division. These datasets capture traffic volumes for various modes at city intersections. To analyze this data, R Core Team (2023) was used along with several R packages. Through the use of a negative

---

\*Code and data are available at: <https://github.com/MariaMangru/Traffic-Congestion-on-Toronto-s-Shoreline>.

binomial regression, we can assess daily car traffic relative to bus traffic, pedestrian, bicycle flows and a time trend. Models were created to investigate traffic patterns before and after public transit improvements.

The estimand of this study is quantitatively evaluated through the change in the relationship between daily bus traffic and daily car traffic, as well as the evolution of car traffic trends over time. The findings from this study indicates a significant reduction in the relationship between bus traffic and car congestion, following the transit improvements. Additionally, a reversal from an increasing to a decreasing trend in car traffic over time was noted.

This research begins with an in-depth exploration of the dataset and the analytical methods employed, highlighting the use of R Core Team (2023) for data analysis. It then goes on to explain the results of the regression model, comparing and contrasting traffic patterns prior to and following the installation of transportation upgrades. The discussion interprets these findings within the broader framework of managing urban congestion and the importance of public transit. Through this approach, I hope to illustrate the important role of public transit improvements in combating urban congestion, as well as its significance for promoting sustainable urban mobility and improving the standard of living in urban areas.

## 2 Data

### 2.1 Data Source

This research relies on traffic volume data sets from the City of Toronto’s Transportation Services Division. It accessible through the Toronto OpenData portal and is titled “Traffic Volumes at Intersections for All Modes” which is free and accessible for public use. The information collected in the dataset are of two main types:

1. Automatic Traffic Recorder Counts (ATR): These are segment-level volumes which capture the total number of vehicles, cyclists, or pedestrians moving in a specific direction on a street.
2. Turning Movement Counts (TMCs): These detail the volume observed at each leg of an intersection, including the turning movement by mode (car, truck, bus, pedestrian, cyclist, other).

The data set covers various intersections across Toronto, providing a detailed view of the city’s traffic dynamics. It includes data spanning 2010-2019 and 2020-2024 which was combined into one data set. The data set is segmented by direction of approach, turning movement, and mode, in 15-minute intervals. This level of detail allows for a thorough investigation of traffic patterns, which is especially important for analyzing how changes to public transport affect traffic congestion on the 508 Lake Shore route.

The data set was cleaned and prepared, with yearly data files combined, relevant date periods filtered, and traffic volumes aggregated to get daily totals for every mode of transportation. This preprocessing step made sure the data was suitable for analysis, with a particular emphasis on the time frame prior to and following the 508 Lake Shore route’s streetcar upgrades.

## 2.2 Variables of Interest

In addition to a temporal trend variables, specific variables were chosen for this analysis, with an emphasis on the dynamics of car traffic in relation to bus, pedestrian, and bicycle traffic volumes:

- Daily Car Traffic (**daily\_cars**): Total volume of car traffic recorded at selected intersections along the 508 Route.
- Daily Bus Traffic (**daily\_bus**), Pedestrian Traffic (**daily\_peds**) and Bicycle Traffic (**daily\_bike**): Bus, pedestrian and bike traffic volumes used to indicate the presence and frequency of non personal car usage.
- Time Trend (**time\_trend**): A constructed variable to analyze traffic pattern changes over time.

Using R Core Team (2023), the dataset was processed by using tools like **ggplot2** for visualization, **dplyr** for data manipulation, and **lubridate** for date management. These packages made it easier to clean, aggregate, and analyze traffic volumes, which helped to provide a more complex picture of traffic patterns and flow.

Table 1: Summary statistics table for daily traffic volume

Transport Mode	Mean	Median
Daily Car Traffic	14169.85135	13291.5
Daily Bus Traffic	223.91441	190.5
Daily Pedestrian Traffic	686.05856	344.0
Daily Bicycle Traffic	50.48649	22.0

The summary table above provides a quantitative comparison across transportation modes by condensing the traffic data for each mode into average and median statistics. The data reveals that car traffic surpasses other modes in volume, with the mean and median values for daily car traffic being the highest.

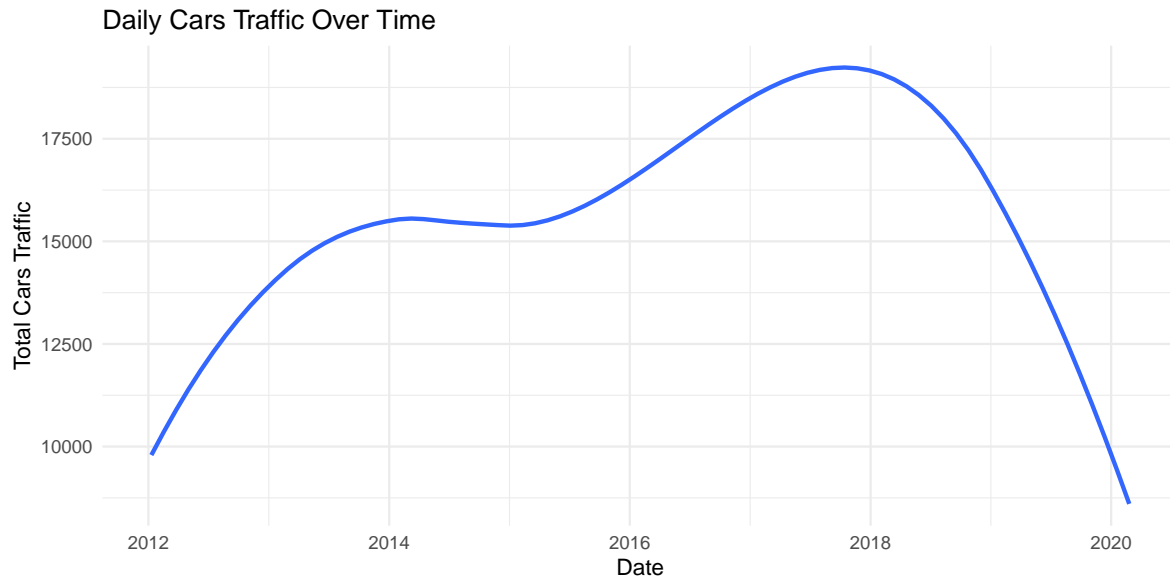


Figure 1: Daily Car Traffic Over Time. This graph displays the trend in number of cars within the area of interest between December 2011 to December 2014 and November 2018 to February 2020.

The daily variation in car traffic from 2011 to 2020 is shown in Figure 1. The traffic volume is depicted on a smooth line graph that first increases and peaks about 2016, after which it sharply declines. The fall that follows may point to the benefits of the Route 508 Lake Shore streetcar improvements on easing traffic congestion. The peak denotes a period of increased traffic.

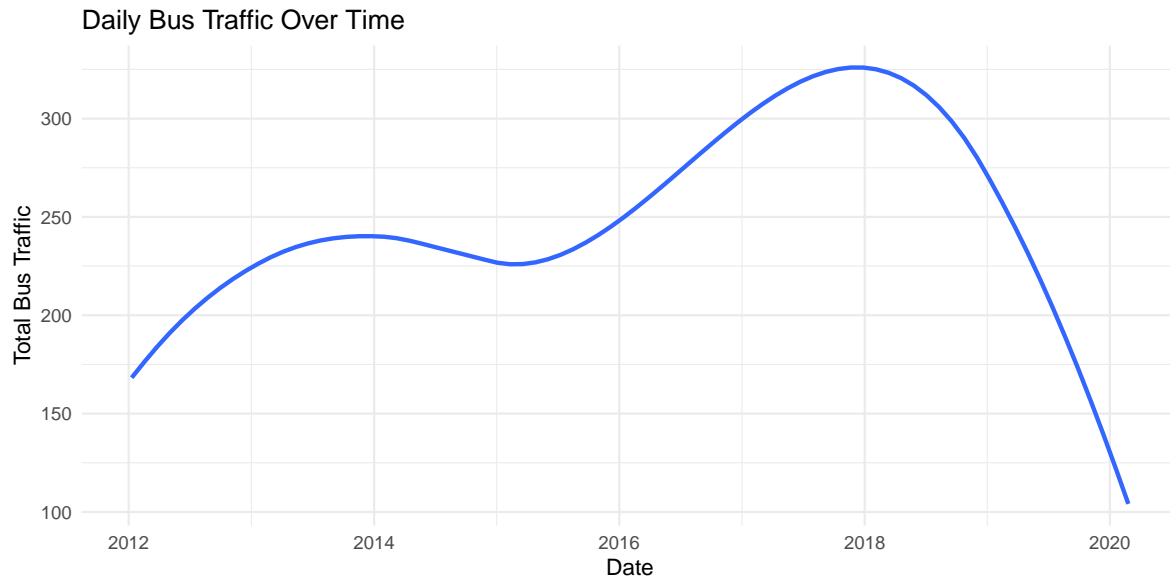


Figure 2: Daily Bus Traffic Over Time. This graph displays the trend in number of buses within the area of interest between December 2011 to December 2014 and November 2018 to February 2020.

Figure 2 depicts the daily patterns in bus traffic, which follow a similar pattern to that of car traffic. It is a smooth line graph that rises to a high around in 2016, then declines. This pattern might be an indication of how bus utilization has changed, which could be related to the streetcar route's service modifications and other improvements to public transportation.

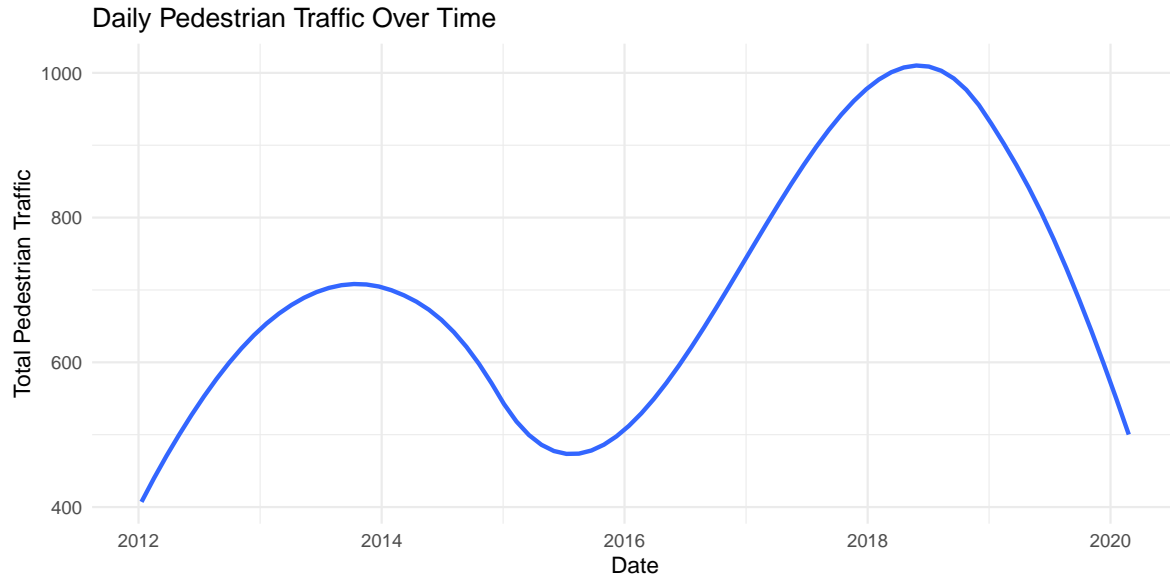


Figure 3: Daily Pedestrian Traffic Over Time. This graph displays the trend in number of pedestrian within the area of interest between December 2011 to December 2014 and November 2018 to February 2020.

Variations in pedestrian traffic over is shown in Figure 3's line graph. The middle of the time period has a noteworthy decrease, which is followed by a rise and another decline. This could be attributed to a number of things, such as changes in walkability, urban development, and perhaps the availability of better public transportation alternatives that encourage walking as a mode of transportation.

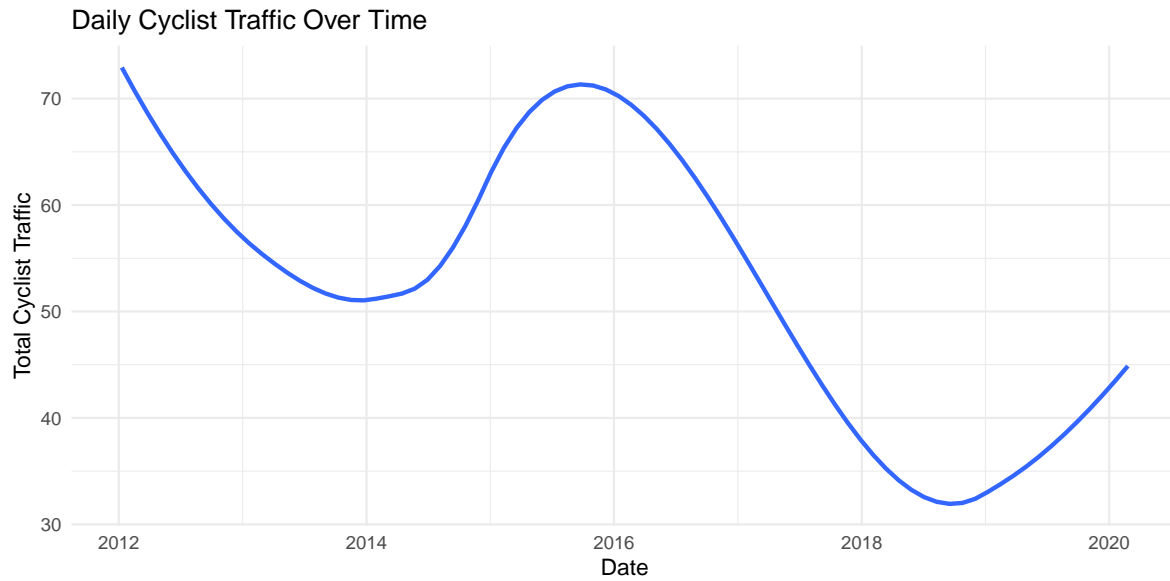


Figure 4: Daily Bike Traffic Over Time. This graph displays the trend in number of bikes within the area of interest between December 2011 to December 2014 and November 2018 to February 2020.

The pattern of bicycle traffic over time is depicted in Figure 4, where a notable rise and fall are shown on the line graph, indicating huge variations. The variations may be attributed to changing public opinions on riding, the implementation of bike infrastructure, or seasonal variations in weather patterns, which may be exacerbated by changes made to the transit system.

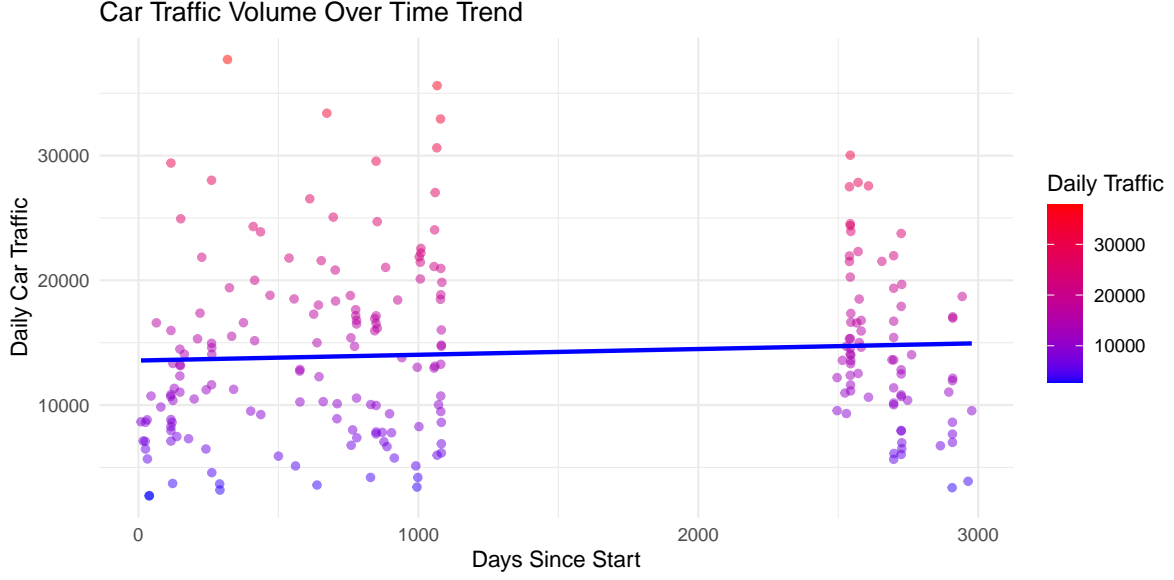


Figure 5: Daily Car Traffic Over Time. This graph displays the general trend in number of cars from the beginning of data collection period.

The scatter plot illustrates the fluctuating daily car counts over time, with a trend line suggesting an initial increase in traffic followed by a decline.

### 3 Model

This study uses negative binomial regression to address the problem of overdispersion, which occurs when observed variance is greater than the mean and is frequently seen in count data. The dependent variable in the model is the daily count of cars, which is assumed to be impacted by temporal variations and the counts of alternative transportation modes (such as buses, bicycles, and pedestrians).

As such, we estimate the following model:

$$\log(\text{Count\_Cars}) = \beta_0 + \beta_1 \times \text{Bus} + \beta_2 \times \text{Pedestrians} + \beta_3 \times \text{Bike} + \text{time\_trend}$$

where

$\log(\text{Count\_Cars})$  is the natural logarithm of the count of cars,

$\beta_0$  is the intercept of the model,

$\beta_1, \beta_2, \beta_3$  are the coefficients for the counts of buses, pedestrians, and bicycles,

$\text{time\_trend}$  is the variable capturing the longitudinal changes.



### 3.1 Model justification

There are two reasons for using negative binomial regression. First, since events tend to cluster together in both place and time, traffic count data usually show overdispersion. For such data, the negative binomial model makes sense since it allows the variance to surpass the mean. Second, given that traffic counts frequently deviate from a normal distribution, this model can account for their skewed nature.

The analysis is divided into two temporal models in order to evaluate the effects of the Route 508 Lake Shore streetcar service upgrades: `model_before`, which is the model before the improvements, and `model_after`, which is the model after the improvements. This method clarifies the relationship that existed before and after the policy intervention between the independent and dependent variables.

Together with a temporal trend, the `model_before` captures the baseline relationship between automobile traffic and other modes of transportation. As an illustration of possible changes in traffic patterns brought about by the improved streetcar service, the `model_after`, on the other hand, depicts the altered dynamics following the enhancement. By use of comparative analysis, these models demonstrate the effectiveness of the Route 508 enhancements in mitigating car congestion.

## 4 Results

## 5 Discussion

### 5.1 First discussion point

### 5.2 Second discussion point

### 5.3 Third discussion point

### 5.4 Weaknesses and next steps

## **Appendix**

### **A Additional data details**

### **B Model details**

#### **B.1 Posterior predictive check**

#### **B.2 Diagnostics**

## References

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.