ECON 573 Research Proposal: Vehicle emissions outcomes of the New York City congestion pricing program

Hunter Tyszka

April 21, 2025

Abstract

On January 5, 2025, New York City became the first city in North America to introduce congestion pricing. With the intent on reducing vehicle traffic, increasing driving speeds, while reducing emissions and improving quality of life, all cars entering Manhattan south of 60th street will be subject to a toll. Since the policy's implementation, the effect on emissions is unclear. This study will use a difference in difference and event study approach to quantify the impact on vehicle emissions both within the New York Metropolitan area, and between major US cities. It is expected that daily emissions within the cordon zone, and its immediate surrounding areas, will decrease relative to other boroughs. Similarly, New York City emissions on aggregate should fall relative to other US cities. Thus, through this study, we should be able to answer the question as to whether this congestion pricing program achieved its pollution reduction goal.

1 Introduction

On January 5, 2025, after years of debate, and years of chronic traffic jams, New York City finally introduced a congestion charge. This policy imposed a toll on vehicles entering the Manhattan central business district. It is the first of its kind in North America, and will serve as a bellwether for other highly congested cities. After years of extreme congestion that saw vehicle travel times slow to walking pace (De Avila, Rust, 2024), coupled with high levels of noise and pollution, the state government intervened. The purpose of the policy was to increase vehicle travel speeds and reduce travel time, while simultaneously reducing vehicle emissions and hence improving quality of life. Now, after three months of the policy, it is unclear whether these benefits have been realized.

In response to this lack of understanding, this paper will study the emissions implications of the congestion pricing program in New York City. I will investigate whether particulate matter emissions, specifically PM2.5 which has been shown to be highly associated with vehicle congestion (Bedoya-Maya et al, 2022), has decreased in the central business district, and the immediate surrounding areas since the policy was implemented. I will consider several variations of geographic boundary, including the central business district, and the CBD plus immediately adjacent districts, to acount for possible spillovers and changes in traffic behaviour. More generally, I will then study whether New York City in aggregate, has a sustained difference in overall PM2.5 pollution compared to other large cities in the United States, since January 5.

The data used to measure particulate matter air pollution in New York City will be sourced from Purple Air. Purple Air is an air quality monitoring community that uses a series of private, consumer owned air quality monitoring sensors to measure real-time ambient air quality over space. This provides a series of hundreds of air quality monitors that can measure PM2.5 pollution across the New York City Metropolitan area, both within and outside of the congestion pricing region. Data is also available for other US cities. Moreover, data relating to weather variables which may affect pollution will be sourced from the National Center for Environmental Information covering temperature, precipitation, and wind over time. Finally, traffic data from TomTom will be used to understand traffic flow patterns in New York City, to better construct a treated and untreated zone when measuring emissions outcomes, and to understand the impacts on congestion.

I will propose two research designs to analyze the impact of the policy on PM2.5 emissions in New York City. First, both a difference in difference and event study design will beused to study emissions within the New York Metropolitan area. The study period will be between October 1, 2024, and April 1, 2025. Using traffic data from TomTom to understand which regions experienced a rapid and sustained change in traffic volume and flow immediately after the introduction of the policy, monitors in these areas would be considered quasi-treated. Other areas in the region would be considered untreated. Then, using the policy date of January 5, 2025, a difference in difference regression would be run to understand the overall impact on emissions. Subsequently, the event study would measure the effect on these emissions over time.

Given that there may be spillovers of emissions and traffic volume as a result of the program, I also propose analyzing emissions in New York City versus those in other large metropolitan areas in the United States. Once again, I propose using a difference in difference approach to understand the overall impact over the study period. Subsequently, an event study to understand long term trends. Examples of cities of interest include Philadelphia, Chicago, Boston, Washington, Baltimore, etc. In these cases, controls for weather variables will be necessary, to account for factors that influence emissions concentration over time.

The remainder of this paper will be laid out as follows: Section 2 will discuss the background and other work in the literature. Section 3 will delve into the data that will be used to construct the empirical specification. Section 4 will discuss empirical approaches and identifying assumptions. Section 5 will present some preliminary findings and hypothetical results. Finally, Section 6 will conclude.

2 Background

Congestion pricing programs are not new policies for urban areas. Major cities around the world have introduced programs to reduce traffic volumes within urban areas, lessen the emissions and noise externalities of private automobiles, and increase travel speeds. Cities such as Singapore (1998), London (2003), Oslo (2007), Milan (2012) have used these policies for years and have been successful in improving traffic flow in central business districts. There has been a substantial amount of research that has been conducted in the Transport Policy, Traffic Engineering and Urban and Regional economics fields that study the theory and outcomes of the policies in these cities. Much of this research focuses on changes in traffic patterns, shifts in transport mode of choice and shifts in traffic related emissions. There is also some research that focuses on the health outcomes of traffic related emissions within cities.

2.1 Brief history on New York's program

New York City has long suffered from severe traffic congestion in the heart of the city. This has led Lower Manhattan to experience high levels of noise, pollution and frustration. In 2023, it was estimated that this congestion results in around $9.1 Billion in lost time (Kaske, 2024). The average speed when driving through Manhattan was around 4.5mph, approaching walking speeds (De Avila, Rust, 2024). At the same time, there is an increase in the volume of vehicles trying to reach these roads, since COVID-19.

In response to this crisis, the City of New York introduced the congestion pricing program, which charges vehicles, by type, a daily fee for entering Manhattan south of 60th street (NYC 311, n.d.). Introduced on January 5, 2025, this program seeks to increase vehicle speeds, reduce traffic and emissions, improve air quality and quality of life within the region. In the short span of time since its introduction, the average speed over the Manhattan bridge has nearly doubled from 13mph to 23mph and speeds within the district have increased considerably (Roeder, Learner, 2025). The effect on overall emissions, especially over a longer time frame, has yet to be investigated.

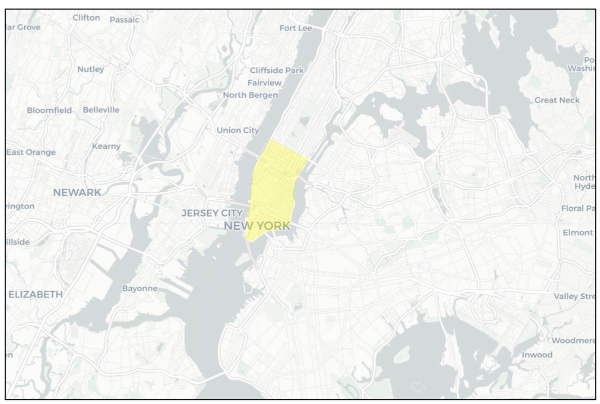


Figure 1: Manhattan congestion pricing zone

2.2 Vehicle congestion and pollution

To understand the impact of the program, we must understand the impact of traffic on air pollution. This is especially important given a major focus of the New York plan is to improve air quality. (Bedoya-Maya et al, 2022) studied the impact of road congestion in Latin America on air quality, and found that high levels of road congestion lead to significant, non-linear increases in PM2.5 pollution. They find that in Latin American cities, some of the most congested in the world, there is a dynamic temporal relationship between vehicle traffic and air quality. The longer a road is congested, the longer emissions compound and thus exacerbate pollution damages. They additionally find that weather, including wind speed, temperature, humidity and precipitation tend to be associated with lower values of pollution. They refer to previous literature which shows high association with PM2.5 pollution and vehicle congestion, as well as increases in traffic fatalities and accidents. As such, it is reasonable to assume that New York's policy to reduce vehicle congestion may have a strong link to air quality, and PM2.5 pollutants.

Additional research by (Lagravinese et al, 2014) investigated the impact of air pollution on hospital admissions from Italy. This is relevant to connect air pollution to overall quality of life and well-being. They refer to a wide array of previous research that has documented severe health effects of vehicle based air pollutants such as PM10 and NO2. They connect PM10 pollutants to increased hospital admissions for children and the elderly and found an increase of 6.15% in admissions for young children.

2.3 Related literature on pricing programs

There are several published papers on other congestion pricing programs around the world and their effect on emissions. (Green et al, 2020) discuss whether the London congestion charge reduced pollution. They hypothesized that because there is high potential for spillover effects of the program into adjacent areas that were not 'treated', that they could not compare within London. Thus, they exclude monitoring stations outside of greater London and consider all monitors within London as 'treated'. They use a difference in difference model to compare pollutants within the congestion cordon, to other cities in the UK. As a result, the authors concluded that while the program reduced PM10, CO and NO, there was an increase in NO2 as commuters shifted from private vehicles to diesel-fueled public transport.

In Bergen, Norway, (Isaken, Johansen, 2021) study the connection between congestion pricing, air pollution and transport mode of choice. They exploited a quirk in the policy that did not impose rush hour charges on weekends, thus designating weekdays as treated and weekends as control. They used a difference in difference approach, with toll gate data, weather data and hourly air quality monitoring data to estimate the effects of the policy on air quality. They find that the congestion charge leads to significant reductions in NO2, but did not find a statistically significant change in PM2.5 and PM10 pollution.

Finally, in Milan (Percoco, 2014) takes advantage of a natural experiment in Milan, where the congestion pricing was paused for 50 days, as the result of a lawsuit by parking owners within the region. This paper used a regression discontinuity design, by taking the threshold as the suspension data of the charge, and the outcome as traffic flow, they study the change before and after the program was paused. The authors find that because the program charged vehicles differently by type, there was not a marked change in vehicle congestion and pollution, but instead, a change in vehicle type. More people tended to drive motorcyles into the city as opposed to private cars.

My research paper, will be the first of its kind to investigate the emissions impact on a congestion pricing regulation within North America. Like the studies in Norway and London, it will use a difference in difference approach. However, I will compare emissions both within the city and between cities. By designating treated and untreated monitoring vicinities based on changes in traffic flow after the policy implementation, I will better be able to account for spillovers, and determine a local effect of the program on pollution. This is different from how the London paper strictly excluded parts of London and instead directly compared to other cities. This approach will also help to better link changes in traffic to changes in pollution at a more local level. Unlike these studies, I will also employ an event study design, which will trace the policy's outcome over time. This will help to determine if there is some longer run equilibrium that may occur as transport mode or business activity responds to the program. My study will also focus on particulate matter pollution, as it is associated with vehicle emissions, and has been linked to health concerns as outlined in the literature.

This paper will not only fill a hole in our understanding of the New York City congestion pricing experience, but it may also help to inform other cities in North America, about policies and strategies to improve air quality and traffic flow too.

3 Data

In order to analyze the outcomes of the pricing policy, I will use data on air quality in New York and other US cities, traffic flow data at the real time level, and weather data that measures key variables of interest in the related cities. Ideally, these historical data will be as close to real time as possible, given traffic and weather varies considerably hour by hour and air pollution closely follows these variables.

3.1 Purple Air - Air Quality Data

To measure ambient air quality over time in New York, and other US cities, Purple Air data will be used for the period of October 1, 2024 to April 1, 2025. Purple Air is a community-based air quality monitoring network, where consumer-owned monitors collect air quality data (Purple Air, 2025). This data is then published on the Purple Air platform. Historical data is available from when a sensor is activated, and is published in various increments, from 15 minutes to a year. This data is easily collected using the Purple Air download tool and API. For the purpose of this study, I would prefer to use one hour averages, as travel patterns vary substantially hour by hour, and emissions would follow suit. The variables of interest include the time of collection, and pm2.5_alt, which measures current PM2.5 pollution.

Within New York, and similarly for other US cities, there are over one hundred monitors on the network. A limitation to this however, is that they are not dispersed equally over space, and there could be confounders associated with their location. Also, not all sensors are available for the whole study period, thus the data is not balanced. For example, wealthier individuals may be more likely to purchase these sensors, and these people may live in regions that are less likely to be polluted. Furthermore, only a certain number of free 'tokens' are available using the API. Realistically, without any external source of funding, I would have to use six hour increments for pollution monitoring, to collect data from all the necessary locations. With more tokens, this study could be expanded to include PM10 data, and a larger array of sensors.

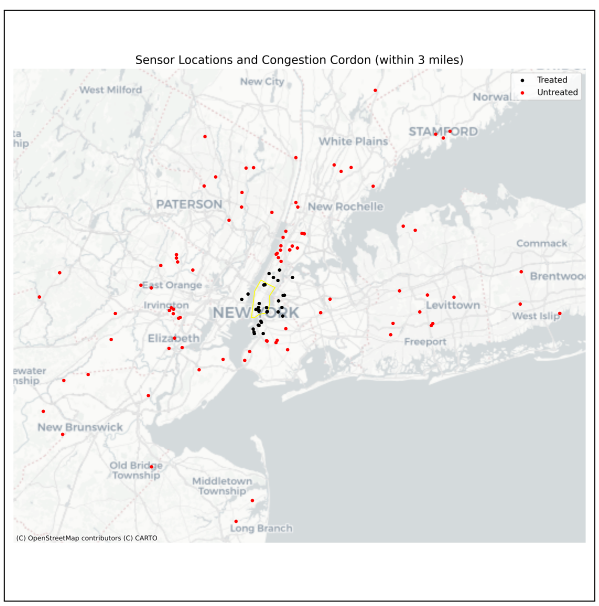


Figure 2: Air pollutions monitors in metropolitan NYC (Purple Air, 2025)

In the above figure, I have plotted most of the outdoor sensors in the New York area. Black dots refer to sensors within the congestion cordon, and 3 miles of its border. This is an assumption that these areas likely experience less through traffic due to the zone, which will be verified with TomTom data below. Red dots refer to sensors outside of this zone. I have not plotted other cities due to limited tokens on the API.

3.2 National Center for Environmental Information - Weather Data

Given that PM2.5 emissions and pollution in general is correlated with weather conditions, which can vary over time and within a small geographic area, I will use weather data to control for these factors. Daily weather data is available at a disaggregated level from the National Centers for Environmental Information (NOAA, 2025). To source this daily summary data for Oct 1, 2024 to Apr 1, 2025, I selected the counties in New York, and for other US cities, that coincide with the Purple Air, air quality, monitors. The variables of interest are average wind speed (AWND), precipitation (PRCP) and average temperature (TAVG). Unfortunately, not all weather stations contain all of this information. Thus, I will drop weather stations that do not contain the needed variables, and subsequently match air quality monitors to the nearest weather station with full data, using longitude and latitude. The weather data is only available as an average over an entire day, and thus will not vary in the same six hour increments as the purple air data. This limitation may require the Purple Air data to be taken as a daily average, in order to match the weather information.

3.3 TomTom - Traffic and Road Congestion Data

Given that traffic entering the congestion pricing zone must come from outside of it, TomTom Traffic Stats data will be used to determine which routes into the congestion zone are most popular (TomTom 2025). This dataset measures historical traffic volume, and allows for the measurement of vehicle travel time and speed. Thus, by understanding which traffic routes into and through the zone are most congested, I can determine which air quality monitoring stations in regions adjacent to the treated area should also be considered treated. This is because traffic through these adjacent zones and into the congestion pricing cordon is likely to change. The TomTom traffic stats data is available on a free trial basis, with more data available after paying a fee. It can be downloaded from the TomTom web portal or through the API with a key.

4 Empirical Approach

I will use two designs to compare the emissions outcomes both within the New York City Metropolitan area, and also between New York City and other large US cities.

4.1 Within New York City - Empirical Specification

4.1.1 Difference in Differences

First, I will attempt to estimate the impacts on PM2.5 emissions within New York City using a difference in difference estimator. The outcome of interest is PM2.5 emissions. The treatment variable will be a dummy for whether or not the monitoring station is within the congestion pricing zone, or an immediately adjacent connected region. These immediately adjacent regions will be selected if they have high levels of traffic entering the congestion zone. I will use a post dummy for air quality readings before and after January 5, 2025, the beginning of the policy.

The immediately adjacent 'treated areas' would include for example, the following places:

- Jersey City, NJ, Adjacent to the Holland Tunnel
- Union City, NJ, Adjacent to the Lincoln Tunnel
- Parts of Queens, NY, adjacent to the Ed Koch Queensboro bridge and Queens-Midtown tunnel.
- Parts of Brooklyn, NY, adjacent to the Williamsburg, Manhattan and Brooklyn bridges.

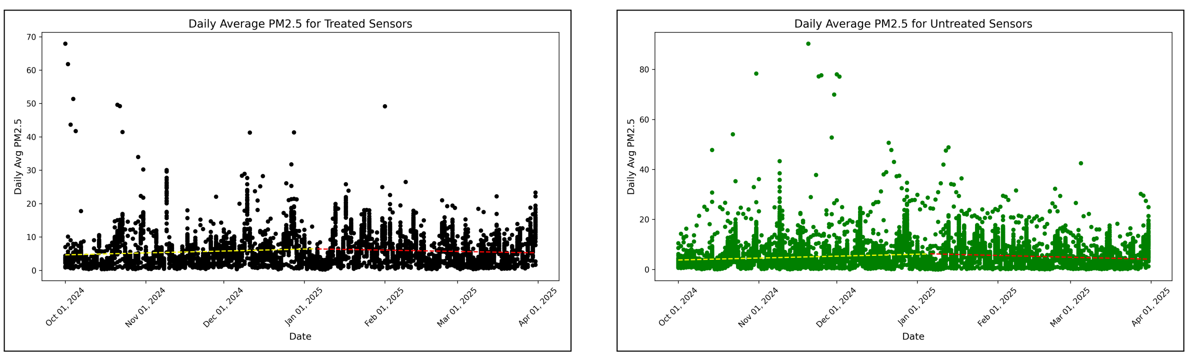
The identifying assumption for this difference in difference estimator is that, in the absence of the congestion pricing program, the treated regions would have their emissions follow a similar evolution as the untreated regions. Given there were no other major confounding policies occurring at the same time, it is likely that the congestion pricing region would have continued on the same path. Thus, it is plausible that the parallel trend assumption holds. If, however, for some reason outside of the congestion pricing policy, the emissions in New York City would have followed a different trend in the absence of treatment, then this strategy would be violated.

Similarly, if there are spillovers to regions within the New York metro area, as a result of this policy, this estimator may be violated. This is plausible in this case. It is possible that the congestion pricing program causes traffic to be diverted into other parts of New York, or causes a rearrangement in transport and business activity, which leads to 'untreated' regions experiencing a change in their emissions, even without a toll. However, as we move further away from the congestion zone, this becomes less likely, and so treating some of the adjacent regions as 'treated' will help to account for these possible spillovers.

The estimating equation for this difference in difference model is as follows:

$$PM2.5_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treated_i + \beta_3 Post_t * Treated_i + \gamma X_{it} + \epsilon_{it}$$

β_3 is the coefficient of interest, which measures the effect on PM2.5 emissions after January 5, 2025 in the congestion pricing cordon and the immediately adjacent regions. The index i refers to individual air quality monitors. The index t refers to days, starting from October 1, 2024 to April 1, 2025. X_{it} is a vector of weather controls, including temperature, precipitation and wind speed, which is matched to the nearest air quality monitoring station. These controls help to pin down the effect of congestion on emissions that is not determined by local weather patterns. Figure 3 plots the data for the treated and untreated regions. At a summary level, the line of best fit shows a gradual decrease in emissions for both regions, and in absolute terms emissions are generally higher outside of the congestion zone.



(a) Treated sensors

(b) Untreated sensors

Figure 3: Data visualization: air pollution monitors in metropolitan NYC

4.1.2 Event Study

Given that the introduction of congestion pricing may lead to changes in driver behaviour, and possibly business and economic location over time, an event study will be useful in estimating long-run trends. This setup will help to identify gradual changes in overall emissions patterns over time. The use of this event study will also help to inform of the plausibility of the parallel trends assumption, if pre-trends are also parallel for treated and untreated monitors.

The estimating equation for this event study model is as follows:

$PM2.5_{it} = \alpha_0 + \alpha_1 Treated_i + \alpha_2 Post_t + \sum_{j=-q}^{-1} \beta_j T_t^j + \sum_{j=0}^{m} \delta j T_t^j + \gamma X_{it} + \epsilon_{it}$

The series of coefficients β_j and δ_j represent the lead and lags of the event study, which show the difference in pollution between the treated and untreated regions over time. I would expect there to be a consistent or zero difference beforehand, and then a rise in the magnitude of the difference after the introduction of the policy. Once again, this framework may fail if there are spillovers between the treated and untreated regions, or if the parallel trend assumption does not hold.

4.2 Between US Cities - Empirical Specification

Given there is a likelihood of spillover and endogeneity in the former empirical design, I suggest a second model to serve as an overall comparison between several large metropolitan areas. In this model, I suggest once again using a difference in difference estimator and event study design to compare New York City, to other cities, all of which do not have congestion charges. In this case, it is almost certain that there are no spillovers between the congestion charge in New York, and driving behaviour and related emissions in other cities.

Once again, Purple Air, air quality data and weather data from NOAA will be used as outcome variables and control variables respectively. Cities to compare to New York include: Baltimore, Boston, Chicago, Detroit, Philadelphia, and Washington.

4.2.1 Difference in Differences

Similar to the previous difference in difference specification, PM2.5 will be the outcome variable. Now, 'Treated' refers to air quality monitoring stations within the New York City Metropolitan area. The 'Post' variable will continue to refer to time periods after January 5, 2025. The identifying assumption, parallel trends, requires that in the absence of treatment, New York City's PM2.5 emissions would have followed a similar trend to the other untreated cities. If there are confounding factors that would change New York emissions simultaneously, this estimation would not be valid.

The estimating equation is as follows:

$PM2.5_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treated_i + \beta_3 Post_t * Treated_i + \gamma X_{it} + \epsilon_{it}$

The variables hold the same meaning as in the New York specification, except for Treated which refers strictly to all monitors within metro New York.

4.2.2 Event Study

As in the New York specification, it is possible that PM2.5 emissions may reach a long-run equilibrium over time. Thus, an event study will be a useful tool in plotting this trend. The specification will take the same form, and will require the same assumptions, however, Treated will refer to all monitors in New York City.

The specification is as follows:

$$PM2.5_{it} = \alpha_0 + \alpha_1 Treated_i + \alpha_2 Post_t + \sum_{j=-q}^{-1} \beta_j T_t^j + \sum_{j=0}^{m} \delta j T_t^j + \gamma X_{it} + \epsilon_{it}$$

5 Results

As a result of the empirical models, I expect to find a decrease in emissions in and around the Manhattan cordon zone as a result of the policy. Since there will be less traffic in these areas, PM2.5 emissions should decrease. Relative to untreated areas, however, the overall effect could be ambiguous, and therefore the regression result may not be statistically significant. This is because a decrease in traffic entering the zone may either decrease emissions in regions outside of it, or may increase as economic activity or trip patterns change in these 'untreated regions'. I have run a preliminary regression, comparing air quality monitors inside and within 3 miles of the congestion zone, to those beyond that threshold. The results are as follows:

Table 1: Regression Results for Daily Average PM2.5

	Intercept	post	treated_sensor	post_x_treated	avg_wind	avg_rain	avg_temp
Estimate	7.959***	0.679***	0.666***	0.068	-0.301***	0.840***	0.002
SE	(0.184)	(0.083)	(0.102)	(0.143)	(0.008)	(0.128)	(0.003)
R-squared				0.080			
Adj. R-squared				0.079			
No. observations				20,655			

Note: Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The coefficient of interest, post_x_treated, shows a positive, but statistically insignificant value. This does not pin down the causal effect of the policy on emissions in this form. The regression used data aggregated at the daily average level, which may not properly estimate changes throughout the day, or at peak travel times. Thus, more granular data may be necessary, or using a daily max emissions instead.

In terms of the second specification, the comparison of New York to other US cities, I would expect New York to have less emissions around the central business district compared to other cities. The regression coefficient of interest, post_x_treated would represent the change in emissions for New York City after Jan 5, 2025, compared to other cities. For this result, I would expect a negative and statistically significant change.

6 Discussion

Given that traffic congestion is a prevalent problem in urban areas globally, understanding the tools and policies that governments can use to combat it is vital. A congestion pricing program has the potential to reduce emissions, make urban areas more livable, and raise quality of life for residents. This study will seek to pin down the emissions outcomes of these decisions.

That being said, it is challenging to overcome the spillover and endogeneity effects that may be present in this type of policy. Neighbourhoods close to the urban core have different characteristics from those far from it. Relative location to industrial areas, and changing travel behaviour as a result of the policy creates roadblocks to estimation. Thus, the within city approach may estimate only a causal effect for certain times of day in very specific geographical locations. The adoption of difference and difference and event study model between cities, however, may help to analyze some of the broader level emissions trends that result from congestion pricing, without spillovers.

There is also potential for a broad set of future research arising from this analysis. This includes quantifying the health impact of the reduction in vehicle emissions, analyzing long-term trends on the change in retail and residential sorting patterns, and measuring changes to transport mode choice. Hence, this paper paves the way for future opportunities in understanding congestion-related policy.

7 Works Cited

Bedoya-Maya, F., Calatayud, A., Gonzalez Mejia, V. (2022). Estimating the effect of road congestion on air quality in Latin America. Transportation Research Part D. https://doi.org/10.1016/j.trd.2022.103510

De Avila, J., Rust, M. (Jun 12, 2024). New York City Traffic is Horrendous. There's No Relief in Sight. Delayed congestion-pricing plan sparks frustration. 'We are falling behind here in New York'. Wall Street Journal (online). Retrieved from: https://www.proquest.com/blogs-podcasts-websites/new-york-city-traffic-is-horrendous-theres-no/docview/3066903162/se-2?accountid=14656

Green, G.P., Heywood, J.S., Paniagua, M.N. (2020). Did the London congestion charge reduce pollution? Regional Science and Urban Economics. 84. https://doi.org/10.1016/j.regsciurbeco.2020.103573

Isaken, E.T., Johansen, B.G. (2021) Congestion pricing, air pollution, and individual-level behavioural responses. University of Oslo, Department of Economics, Oslo. https://www.econstor.eu/handle/10419/250303

Kaske, M. (June 24, 2024) NYC Has the World's Worst Traffic Congestion, Costing $9 Billion. Bloomberg CityLab. Retrieved from: https://www.bloomberg.com/news/articles/2024-06-25/nyc-has-the-world-s-worst-traffic-congestion-costing-9-billion

Lagravinese, R., Moscone, F., Tosetti, E., Lee, H. (2014). The impact of air pollution on hospital admissions: Evidence from Italy. Regional Science and Urban Economics. 49 278-285. https://doi.org/10.1016/j.regsciurbeco.2014.06.003

NOAA. (2025). Climate Data Online [Data set]. NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/cdo-web/datatools/selectlocation

NYC 311. (n.d.) Congestion Pricing Program. NYC 311 The Official Website of the City of New York. Retrieved from: https://portal.311.nyc.gov/article/?kanumber=KA-03612

Percoco, M. (2014). The effect of road pricing on traffic composition: Evidence from a natural experiment in Milan, Italy. Transport Policy 31 55-60. https://doi.org/10.1016/j.tranpol.2013.12.001

Purple Air (2025). Purple Air Data Download Tool [Data set]. Purple Air. https://community.purpleair.com/t/purpleair-data-download-tool/3787

Roeder, O., Learner, S. (Jan 18, 2025) New York congestion scheme speeds traffic: Transport. Toll impact City implements pay-to-drive zones and joins a select club of metropolises with restrictions. Financial Times. Retrieved from: https://www.proquest.com/newspapers/new-york-congestion-scheme-speeds-traffic docview/3167423780/se-2?accountid=14656

TomTom. (2025). Traffic Stats [Data set]. TomTom International BV. https://www.tomtom.com/products/traffic-stats/