safety

MDPI

Article

Has Congestion Pricing Improved Short-Term Road Safety? A Case Study in New York City

Mingyin Wang [1], **Xuan Di** [2,*]

[1] Department of Biostatistics, Mailman School of Public Health, Columbia University, 722 W 168th St, New York, NY 10032, USA; aw3693@cumc.columbia.edu
[2] Department of Civil Engineering and Engineering Mechanics, Data Science Institute, Columbia University, 500 West 120th Street, New York, NY 10027, USA
* Correspondence: sharon.di@columbia.edu

Abstract

In January 2025, New York City became the first major U.S. city to implement a cordon-based congestion pricing policy via the Central Business District Tolling Program. While the policy's effects on traffic volume are well-documented, its impact on road safety remains underexplored. This study evaluates the short-term effects of the program on two distinct metrics: total crash counts (frequency) and injury rates (severity, defined as the number of persons injured per 10,000 residents), using a monthly panel dataset of ZIP code-level data from January 2024 to December 2025. We employ a rigorous multi-method causal inference framework—including difference-in-differences, matched difference-in-differences, and generalized synthetic control—to estimate changes in injury rates and total crash counts independently. Across all empirical specifications, we find no statistically significant reduction in either traffic injuries or collisions following the policy's implementation. Event study analyses confirm a consistent null effect month-over-month, with no transient or sustained safety dividend. Subject to short-term methodological constraints, our findings suggest that congestion pricing functions primarily as a demand management tool; realizing immediate road safety benefits in complex urban grid networks likely requires complementary physical infrastructure interventions.

Keywords: congestion pricing; road safety; traffic collisions; traffic injuries; difference-in-differences; synthetic control; urban mobility; New York City

1. Introduction

Road safety in dense urban cores remains a critical public health challenge. Global evidence from Stockholm [1], Milan [2], and London [3] has demonstrated that urban tolling can reduce congestion and improve air quality. Prior to the implementation of the Central Business District Tolling Program (CBDTP), Manhattan south of 60th Street—the Central Business District (CBD)—experienced severe gridlock and a high baseline crash burden. In 2024 alone, Manhattan recorded 44 traffic fatalities and a significant concentration of injuries, even as the city expanded its Vision Zero initiatives [4]. To manage these negative externalities, New York City implemented the CBDTP in January 2025. Managed by the Metropolitan Transportation Authority (MTA), the program establishes a cordon where vehicles are charged a time-varying fee to enter the CBD, with peak charges designed to actively disincentivize discretionary driving [5].

The CBDTP is a cordon-based congestion pricing scheme, an economic demand-management tool that prices road space to reduce congestion. It belongs to the broader

Check for updates

Academic Editor: Raphael Grzebieta

Received: 11 January 2026
Revised: 19 April 2026
Accepted: 27 April 2026
Published: 7 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

category of urban vehicle access regulations (UVARs) but differs fundamentally from technical restrictions like Low Emission Zones (LEZs). While LEZs restrict entry based on vehicle age or emission standards to improve air quality, congestion pricing targets total fleet volume through a price signal. Existing evidence from LEZs in Germany suggests that emissions-based access restrictions can yield secondary road safety dividends, primarily by reducing the presence of older, heavy freight vehicles with larger blind spots [**?**]. In contrast, congestion pricing is expected to impact safety more broadly by altering total traffic flow, as seen in London where the 2003 charge led to sustained reductions in collisions and fatalities [7,8].

The causal pathways linking congestion pricing to safety outcomes involve the joint interaction of traffic volume, density, modal split, and vehicle speeds. First, the financial cost of driving reduces the absolute volume of vehicle entries, lowering traffic density and reducing the number of vehicle-to-vehicle and vehicle-to-pedestrian conflict points. This pathway theoretically lowers crash frequency. Second, the toll induces a modal shift, transferring commuters from cars to high-capacity transit. This shift carries vital safety implications: if the policy creates an economic barrier (equity) that forces vulnerable populations into unprotected active transit modes like cycling, their physical exposure to injury risk increases if street infrastructure is not concurrently redesigned. Finally, the "congestion as traffic calming" hypothesis suggests that gridlock protects road users by limiting speeds [9]. Alleviating congestion on Manhattan's wide grid layout may inadvertently reduce safety if lower density leads to higher free-flow speeds, as higher kinetic energy increases injury severity even if the total number of crashes declines. Recent preliminary evaluations in NYC have focused on short-run ridership and welfare impacts [10**?**], complementing theoretical models exploring congestion pricing for efficiency and equity in San Francisco [12], Bogota [13], and London [14].

Given these competing mechanisms—reduced exposure versus increased speed intensity—this study evaluates the short-term impact of the CBDTP on road safety. Our objective is to determine whether the theoretical safety co-benefits of congestion pricing materialized during the initial implementation phase in NYC.

2. Materials and Methods

2.1. Study Design and Area

To isolate the causal effect of the tolling program, we adopt the Difference-in-Differences (DiD) framework, which is the standard quasi-experimental approach in road safety literature for identifying the impacts of policy interventions. Researchers such as Cohen [15], DeAngelo [16], Abouk [17], Cheng [18] utilize this method to address the challenge of unobserved confounders—such as improvements in vehicle technology, economic fluctuations, or weather patterns—that affect traffic outcomes across all regions simultaneously. By constructing panel datasets, these studies compare the change in safety outcomes in treated jurisdictions against the change in control jurisdictions.

To execute this design, our study area encompasses NYC, divided into two distinct spatial zones based on the tolling cordon (Figure 1):

- Treatment Group: The Central Business District (CBD), defined as Manhattan south of 60th Street.
- Control Group: The remaining areas of Manhattan, and the boroughs of the Bronx, Brooklyn, Queens, and Staten Island.

Rather than detailing the specific financial tolling tiers, it is critical to characterize the distinct baseline traffic environments of these zones. The CBD is North America's densest commercial hub, characterized by a highly saturated grid network, low baseline free-flow speeds, and massive daytime commuter influxes. Conversely, the control group comprises

largely residential outer boroughs featuring arterial roads, highway networks, and a higher reliance on personal passenger vehicles.

While continuous, high-resolution data on average segment speeds and ZIP-level traffic volumes are not publicly available for the 2024–2025 study period, we approximated the modal composition of these zones using the distribution of vehicle types involved in recorded collisions. By categorizing over 2000 raw NYPD vehicle descriptors into macro-categories, distinct baseline differences emerge. The CBD exhibits heavy modal mixing, with a disproportionately high volume of For-Hire Vehicles, transit buses, commercial delivery fleets, and vulnerable micromobility users compared to the outer boroughs, which are heavily dominated by private passenger vehicles. Because these groups differ fundamentally in baseline traffic characteristics and pre-policy safety trends (as detailed in our baseline descriptive statistics in Section 3.1), naive statistical comparisons are invalid. Our identification strategy inherently relies on advanced matching and latent-factor models to construct a valid counterfactual.



Figure 1. Treatment Definition and Intervention Period. The map illustrates the Congestion Tolling Zone (Treatment) in Manhattan south of 60th Street (indicated by the yellow shaded region). The timeline compares the pre-intervention baseline (January–December 2024) to the post-implementation period (January–December 2025). Source: Metropolitan Transportation Authority (MTA) [5].

The study period extends from 1 January 2024, to 31 December 2025, providing a 12-month pre-intervention baseline and a 12-month post-intervention period to evaluate short-term policy effects.

2.2. Data Sources and Processing

All statistical analyses were conducted using the R (Version 4.5.3) and Python (Version 3.13.2) statistical programming language.

2.2.1. Disclosure of Generative AI in Research Documentation

The authors utilized the generative AI tool Gemini 3 Pro (Google) for the specific purposes of debugging the R and Python statistical scripts used in this study and for refining the grammatical structure and clarity of the manuscript text. Following the use of this tool, the authors independently reviewed and edited the content to ensure accuracy. All statistical analyses, data interpretations, and the final scientific conclusions were conducted and verified solely by the authors, and the AI tool was not used for data collection, data generation, or study design.

2.2.2. Crash and Outcome Data

Motor vehicle collision records were obtained from the NYC Open Data portal, maintained by the New York Police Department (NYPD). To avoid conceptual conflation between crash frequency and crash severity, we analyze two strictly distinct outcome variables: (1) Total Crashes, defined as the absolute monthly count of all reported motor vehicle collisions regardless of severity; and (2) Total Persons Injured, representing the absolute count of individuals sustaining physical injuries in those collisions. These outcomes are not interchangeable; for instance, a policy-induced reduction in congestion could theoretically decrease crash frequency while simultaneously increasing injury severity if lower density leads to higher free-flow speeds.

The dataset was filtered to include only geocoded collisions. A spatial point-in-polygon operation was performed using the `sf` package to assign each crash to the CBD or control zone based on official municipal shapefiles. We selected ZIP codes as the spatial unit of analysis to align crash data with the available resolution of our demographic and socioeconomic covariates. While leveraging finer spatial units (e.g., street segments) could theoretically reduce treatment misclassification at the cordon boundary, ZIP-level aggregation provides the necessary statistical power and data continuity required to construct a balanced panel. Furthermore, we intentionally aggregated the temporal dimension to the monthly level; preliminary analyses at the daily level resulted in excessive sparsity (a high prevalence of zero-crash days per ZIP code), which computationally destabilizes the count-based panel regressions. Exempt roadways (e.g., FDR Drive) were excluded from the treatment assignment where applicable to minimize boundary misclassification.

2.2.3. Environmental Controls and Causal Mediators

Crash frequency is sensitive to exogenous environmental conditions. We integrated monthly aggregates for precipitation and average temperature (NOAA) and a binary indicator for holidays to account for irregular city-wide traffic patterns. Crucially, following the causal framework of Green [7], we explicitly exclude variables related to temporal traffic composition—such as peak-hour proportions or weekend ratios—from our primary specifications. Because the congestion toll is a time-varying financial instrument designed specifically to shift travel behavior across the diurnal cycle, these shifts are "endogenous mediators" rather than "exogenous confounders." Controlling for these temporal shifts would result in "overcontrol bias," blocking the primary causal pathway through which the policy impacts safety and forcing the treatment effect estimates toward zero [7]. Consequently,

these factors are analyzed only within our descriptive and sensitivity frameworks to maintain the integrity of our identification strategy.

2.2.4. Demographic Normalization and Exposure Limitations

To calculate relative risk, crash and injury counts were normalized by ZIP code population (retrieved from the U.S. Census Bureau's American Community Survey) to create the Injury Rate$_{it}$ per 10,000 residents . We acknowledge a significant methodological limitation in this exposure proxy: the at-risk population in the Manhattan CBD is primarily non-residential, consisting of commuters, tourists, and commercial freight. Ideally, road safety risk should be normalized using dynamic exposure measures such as Vehicle Miles Traveled (VMT), hourly traffic volumes, or average segment speeds. However, at the time of this study, continuous ZIP-level traffic volume data for the 2025 post-implementation period was not yet available via public repositories (e.g., NYC DOT or Open Data). Consequently, residential population serves as a necessary, albeit static, proxy to maintain a balanced panel. We explicitly recognize that this spatial mismatch between the used denominator (residents) and the true denominator (active road users) likely introduces measurement error into our rate estimates, potentially leading to attenuation bias.

2.3. Sensitivity Analysis: Cordon-Line Buffer Exclusion

A persistent challenge in the evaluation of cordon-based urban policies is the presence of spatial misclassification at the treatment boundary. In the context of New York City, several ZIP codes (e.g., 10019, 10022, and 10023) straddle the 60th Street tolling line, aggregating collisions from both the tolled Central Business District and the untolled uptown control segments. This spatial overlap creates "Treatment Dilution," where the policy signal is masked by untreated traffic noise within the same spatial unit of analysis. To address this, we implemented a surgical Cordon-Line Buffer protocol, a technique recognized in spatial econometrics for minimizing geocoding errors and treatment contamination [19,20].

Utilizing regular expression (regex) string matching on the raw NYPD `ON_STREET_NAME` and `CROSS_STREET_NAME` attributes, we identified and excluded 9216 collisions occurring strictly on 60th Street. This street serves as a complex hybrid environment featuring high-volume entrance and exit ramps for the Queensboro Bridge and the FDR Drive, which may not strictly adhere to the behavioral shifts expected within the CBD interior. By removing this geographic buffer, we ensure that the treated group consists exclusively of "deep" CBD crashes and the control group consists of pure untolled segments. This approach follows the fundamental principles of spatial independence, ensuring that estimated effects are not attenuated by observations sharing a common thoroughfare [21].

2.4. Statistical Analysis

To ensure the robustness of our causal estimates, we follow a multi-method analytical roadmap that addresses distinct threats to identification. The Difference-in-Differences (DiD) framework serves as our baseline strategy, which we refine through Mahalanobis Matching to ensure baseline comparability. To account for unobserved time-varying confounders and non-parallel trends, we utilize the Generalized Synthetic Control Method (GSCM) and Generalized Additive Models (GAMs).Finally, an Event Study design is used to test identifying assumptions and observe dynamic policy impacts. Critically, because the crash data exhibit significant overdispersion ($Var(Y) > E[Y]$), all primary count-based specifications—including the DiD, Matched DiD, GAM, and Event Study—are estimated using a Negative Binomial regression framework with a log-link function. This ensures that our standard errors are not underestimated and that our count-based modeling remains statistically valid.

The relationship between traffic volume and road safety is theoretically complex and often non-linear; as established by Shefer [9], reductions in congestion can paradoxically lead to increased accident severity due to higher travel speeds (the "U-shaped" hypothesis). Given this complexity, relying on a single linear model specification may lead to biased inference regarding the impact of the congestion pricing scheme.

2.4.1. Difference-in-Differences Specification (DiD)

To estimate the causal effect of the congestion pricing policy on traffic safety outcomes, we fit Negative Binomial regression models. This approach constitutes the standard quasi-experimental framework in traffic safety research, having been robustly applied to evaluate analogous interventions such as mandatory seatbelt laws [15], changes in blood alcohol limits [22], and intensified police enforcement [16].

We first examined the distribution of the crash data. Preliminary analysis indicated that the variance of crash counts significantly exceeded the mean ($Var(Y_{it}) > E[Y_{it}]$), suggesting overdispersion (Figure 2). Consequently, we prioritize the Negative Binomial specification over the standard Poisson model to correct for this bias and prevent the underestimation of standard errors.

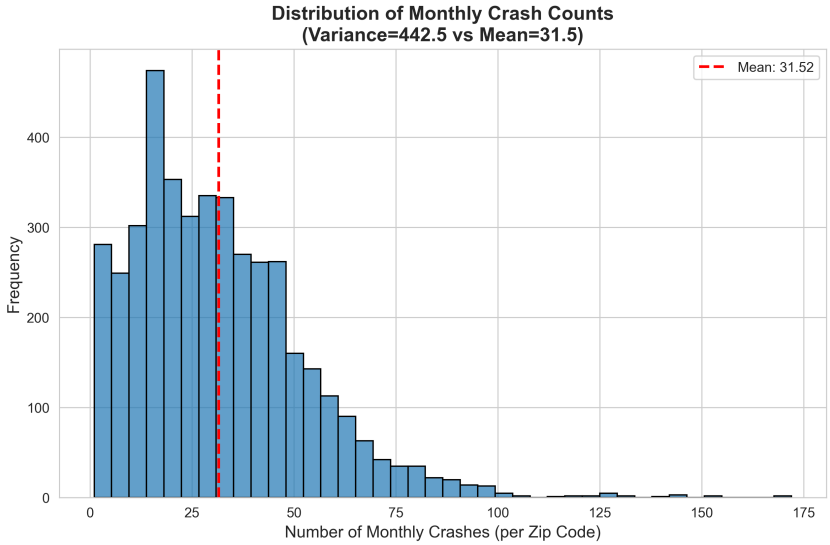


Figure 2. Distribution of Monthly Crash Counts. The histogram displays the frequency of crash counts across all zip codes and months. The variance significantly exceeds the mean, justifying the use of Negative Binomial regression.

The baseline specification is defined as:

$$\begin{aligned}\ln(E[Y_{it}]) = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{Post}_t + \beta_3 (\text{Treated}_i \times \text{Post}_t) \\ + \beta_4 \text{Holiday}_{it} + \beta_5 \text{Precip}_{it} + \beta_6 \text{Temp}_{it} + \mathbb{1}(Y_{it} = \text{Injury}) \ln(\text{Pop}_{it})\end{aligned} \quad (1)$$

where the variables are defined as follows:

- Y_{it}: The outcome variable representing either the total count of traffic collisions or the count of persons injured for ZIP code i in month t.
- Treated_i: A time-invariant binary indicator equal to 1 if ZIP code i is located within the CBD, and 0 otherwise.
- Post_t: A time-variant binary indicator equal to 1 for all months following the policy implementation (January 2025 through December 2025), and 0 for the pre-intervention period (2024).

- $\text{Treated}_i \times \text{Post}_t$: The interaction term of interest. The coefficient β_3 captures the **Average Treatment Effect on the Treated (ATT)**, representing the causal impact of the congestion pricing policy on safety relative to the counterfactual trend.
- Holiday_{it}, Precip_{it}, and Temp_{it}: Exogenous controls for the proportion of holidays, total monthly precipitation, and average temperature, respectively.
- $\mathbb{1}(Y = \text{Injury})$: An indicator function that equals 1 when the dependent variable is the count of persons injured and 0 when the dependent variable is total crashes.
- $\ln(\text{Pop}_{it})$: The natural logarithm of the residential population for ZIP code i. This is included as an *offset term* (with its coefficient constrained to 1) for the injury model, effectively normalizing the outcome into a rate per capita.

2.4.2. Matched Difference-in-Differences (Matched DiD)

To address potential selection bias and ensure that the treatment and control groups followed parallel trajectories prior to the intervention, we implemented a rigorous matching strategy.

A key limitation of standard DiD in this context is the assumption that the dense, commercial CBD would evolve similarly to the residential outer boroughs in the absence of the toll. To mitigate this comparison of "apples to oranges," we employ Mahalanobis Distance Matching. Following the logic of Lalive [23], Grabowski [24], Carpenter [25], we match specifically on the pre-treatment trajectory of injury risk.

This approach ensures that selected control units share not only similar baseline risk levels but also similar seasonal fluctuations. By forcing the control group to mimic the pre-intervention dynamics of the CBD, we construct a more valid counterfactual for the trends in safety, isolating the policy effect from structural differences between boroughs [26], while accounting for policy efficiency [27] and accident externalities [28].

2.4.3. Generalized Synthetic Control Method (GSCM)

Our primary identification strategy uses the GSCM developed by Xu [29]. While traditional synthetic control methods construct a counterfactual based on a weighted average of control units [30,31], GSCM extends this by identifying unobserved time-varying latent factors using an interactive fixed effects model.

This method is particularly advantageous for analyzing the NYC congestion zone because it relaxes the strict parallel trends assumption of standard DiD. Urban traffic systems are subject to unobserved common shocks—such as citywide economic shifts or subway delays—that may affect the CBD and outer boroughs with different intensities. The GSCM model is specified as:

$$Y_{it} = \mu_{it} D_{it} + \mathbf{X}_{it}\beta + \lambda_i' \mathbf{f}_t + \epsilon_{it} \tag{2}$$

where the variables are defined as follows:

- Y_{it} represents the outcome of interest for ZIP code unit i at time t. To maintain the linear factor structure of the GSCM, Y_{it} refers to the normalized Injury Rate (or crashes per capita) as defined in Section 3.1.
- D_{it} is a binary treatment indicator that equals 1 if unit i is in the treated zone during the post-intervention period.
- μ_{it} captures the heterogeneous treatment effect of the congestion pricing policy. The average of these effects over the post-treatment period constitutes the ATT.
- $\mathbf{X}_{it}$ is a vector of observed time-varying covariates, including the holiday proportion, monthly precipitation, and average temperature as defined in Equation (1).
- $\lambda_i' \mathbf{f}_t$ represents the interactive fixed effects structure used to model unobserved confounding:

  - $\mathbf{f}_t$ is a vector of unobserved, time-varying common factors (or latent trends) that affect all units.
  - λ_i is a vector of unit-specific factor loadings, representing the heterogeneous sensitivity of ZIP code i to these common factors.
- ϵ_{it} is the idiosyncratic error term, assumed to be mean-zero.

By explicitly modeling these latent factors, GSCM allows us to separate the true policy effect from these unobserved, time-varying confounders.

2.4.4. Generalized Additive Models (GAM)

To explicitly model non-linear seasonal and weather effects, we estimated GAM. As argued by Shefer [9], the relationship between density and safety can be U-shaped. Standard linear models might average out these competing effects, potentially masking a rebound effect caused by higher free-flow speeds. To account for these non-linearities, we specify a GAM:

$$\log(E[Y_{it}]) = \beta_0 + \beta_1 D_{it} + f_{time}(t) + f_{weather}(\text{Precip}, \text{Temp}) + \mathbf{X}_{it}\beta + \alpha_i \quad (3)$$

where the variables are defined as follows:

- Y_{it} denotes the outcome of interest for zip code i at time t.
- D_{it} is the binary treatment indicator.
- $f_{time}(t)$ is a smooth, non-linear function of time (modeled using penalized cubic regression splines) that captures complex seasonal trends affecting the entire city.
- $f_{weather}(\text{Precip}, \text{Temp})$ represents smooth interaction functions of total monthly precipitation and average temperature, allowing for disproportionate crash risks during extreme weather events.
- $\mathbf{X}_{it}$ is a vector of linear covariates, including the proportion of holidays, with associated coefficient vector β.
- α_i represents zip code-specific random intercepts to account for time-invariant unobserved heterogeneity.

2.4.5. Event Study Specification

To examine the dynamic evolution of safety effects, we employed an event study specification. The model includes K leads and L lags of the treatment initiation:

$$Y_{it} = \sum_{k=-K}^{L} \mu_k D_{it}^k + \gamma_t + \mathbf{X}_{it}\beta + \alpha_i + \epsilon_{it} \quad (4)$$

where the variables are defined as follows:

- Y_{it} represents the outcome variable for zip code i in month t.
- D_{it}^k is a set of binary indicators representing the time relative to the policy implementation.
- μ_k captures the dynamic treatment effect for each month k relative to the omitted reference period, $k = -1$ (December 2024).
- K = 12 and L = 12 represent the number of pre-treatment leads and post-treatment lags is 12 months (i.e., 1 year before and 1 year after).
- α_i and γ_t represent zip code and time-fixed effects, respectively.
- $\mathbf{X}_{it}$ is a vector of time-varying control variables with associated coefficient vector β.
- ϵ_{it} represents the idiosyncratic error term, assumed to be independent and identically distributed (i.i.d.) with a mean of zero, capturing unobserved factors that vary across both ZIP codes and months.

This specification serves two critical purposes: it allows us to statistically test the parallel trends assumption prior to the policy ($k < 0$), and it allows us to detect transient adaptation effects post-implementation ($k \geq 0$).

3. Results

3.1. Descriptive Statistics and Baseline Comparability

Table 1 summarizes the baseline traffic, demographic, and environmental characteristics of the Treatment (CBD) and Control (Outer Boroughs) groups during the pre-intervention period (2024). In the following descriptive and inferential analyses, we maintain a strict distinction between crash frequency (total monthly events) and the injury rate (severity normalized by population) to capture potential shifts in both the quantity and severity of collisions. The injury rate is formally defined as the absolute count of persons injured in a given ZIP code per month, normalized by that ZIP code's residential population per 10,000 residents:

$$\text{Injury Rate}_{it} = \left(\frac{\text{Total Persons Injured}_{it}}{\text{Population}_i} \right) \times 10,000 \quad (5)$$

The descriptive data in Table 1 reveal significant structural differences between the zones prior to the policy. While the control group has a significantly higher residential population ($p < 0.001$), the CBD exhibits a significantly higher injury rate per capita (4.82 vs. 3.59 per 10,000 residents, $p < 0.001$). This disparity underscores the high-risk nature of the CBD's dense traffic environment and confirms the unique risk profile of the Manhattan core. Crucially, the validity of our DiD and GSCM frameworks rests not on absolute similarity in levels, but on the similarity of pre-intervention safety trends. By employing Mahalanobis matching and latent-factor modeling, we ensure that the control units selected are those that most closely mirror the seasonal and idiosyncratic volatilities of the CBD, thereby satisfying the identifying assumptions of the quasi-experimental design.

Table 1. Baseline Descriptive Statistics (Pre-Policy Year 2024). This table compares the baseline traffic, demographic, and environmental characteristics of the treated and control zones prior to the implementation of the congestion toll.

Variable	CBD (Treated)	Outer Boroughs (Control)	*p*-Value
Monthly Crash Frequency	26.72 (15.10)	31.03 (21.40)	<0.001 ***
Injury Rate (per 10k pop)	4.82 (3.90)	3.59 (3.47)	<0.001 ***
Residential Population	32,593.76 (20,275.93)	50,486.09 (25,683.46)	<0.001 ***
Weekend Crash Prop.	0.26 (0.12)	0.27 (0.13)	0.294
Daytime Crash Prop.	0.64 (0.14)	0.63 (0.15)	0.377
Holiday Count	0.83 (0.56)	0.83 (0.55)	N/A †
Monthly Precip. (mm)	3.82 (2.35)	3.87 (2.38)	N/A †
Avg Temperature (F)	58.13 (14.57)	57.91 (14.49)	N/A †

Note: Values are Mean (Standard Deviation). Statistical significance calculated via two-sample *t*-tests. *** $p < 0.001$. † *p*-values for holiday counts, precipitation, and temperature are not applicable (N/A) as these represent city-wide environmental controls applied uniformly.

Furthermore, as summarized in Table 2, the modal composition of these zones differs vastly. In the Outer Boroughs, nearly 80% of vehicles involved in collisions are private passenger vehicles. Conversely, the CBD features heavy modal mixing, with significantly higher proportions of Transit & For-Hire vehicles (12.3%) and Micromobility users (14.9%).

Because these spatial groups are not naturally comparable in baseline risk or modal exposure, standard DiD estimates risk substantial bias. This motivated our reliance on Mahalanobis matching and GSCM to construct valid counterfactuals. Furthermore, all subsequent causal models incorporate controls for total monthly precipitation and aver-

age temperature to account for exogenous weather-related variations in road friction and visibility, alongside controls for city-wide holiday anomalies. Note that raw longitudinal time-series plots of crashes have been intentionally omitted from this descriptive analysis; without continuous dynamic traffic volume data (VMT) to serve as a denominator, visualizing raw crash counts over time risks presenting misleading inferences regarding actual risk exposure.

Table 2. Modal Composition of Vehicles Involved in Collisions (2024 vs. 2025). This table characterizes the distinct traffic environments of the treated and control zones, highlighting the CBD's heavy modal mixing compared to the passenger-vehicle-dominated outer boroughs.

Vehicle Category	Pre-Policy (2024)		Post-Policy (2025)	
	CBD (Treated)	Control	CBD (Treated)	Control
Passenger Vehicles	61.9%	79.5%	60.6%	79.1%
Micromobility & Motorcycles	14.9%	8.6%	15.1%	8.4%
Transit & For-Hire	12.3%	4.1%	13.5%	4.5%
Commercial & Freight	9.7%	6.8%	9.4%	7.0%
Emergency/City	1.0%	0.8%	1.3%	0.8%
Unknown/Other	0.2%	0.1%	0.2%	0.1%

Note: Percentages represent the proportion of total vehicles involved in reported collisions.

3.2. Synthesis of Causal Treatment Effects

To account for the baseline incomparability and weather variations described above, we formally evaluated the impact of the congestion pricing policy across distinct methodological specifications: standard Negative Binomial DiD, GAM, and GSCM. Tables 3 and 4 present the Average Treatment Effect on the Treated (ATT) across these models for total crashes and injury rates, respectively.

Table 3. Estimated Effects of Congestion Pricing on Monthly Traffic Collision Counts.

Model	Sample	Estimate	95% Confidence Interval	*p*-Value
Negative Binomial Regression (DiD)	Matched	−0.091	(−0.435, 0.252)	0.602
Negative Binomial Regression (DiD)	Full Sample	0.013	(−0.220, 0.247)	0.912
Generalized Additive Model	Full Sample	0.016	(−0.025, 0.057)	0.434
Generalized Synthetic Control	Full Sample	0.148	(−1.564, 1.860)	0.865

Table 4. Estimated Effects of Congestion Pricing on Injury Rate per 10,000 Population.

Model	Sample	Estimate	95% Confidence Interval	*p*-Value
Negative Binomial Regression (DiD)	Matched	−0.019	(−0.400, 0.361)	0.920
Negative Binomial Regression (DiD)	Full Sample	0.075	(−0.180, 0.329)	0.566
Generalized Additive Model	Full Sample	0.007	(−0.071, 0.085)	0.862
Generalized Synthetic Control	Full Sample	0.115	(−1.178, 1.409)	0.861

A coherent synthesis of these multi-method results reveals a robust and consistent "null" effect across all primary specifications. While point estimates vary slightly in sign—with the matched DiD suggesting a marginal decrease (−0.091) and the GSCM suggesting a marginal increase (0.148)—these discrepancies are expected when employing models with different identifying assumptions regarding pre-intervention trends. Crucially, the 95% confidence intervals across all five specifications significantly overlap and consistently bound zero. This "bracketing" of the null suggests that the lack of a statistically significant safety dividend is not an artifact of a specific model's misspecification,

but a generalized finding across the initial implementation period. The fact that these estimates remain statistically insignificant across models with fundamentally different assumptions suggests that the lack of an immediate safety dividend is not merely an artifact of rigid model misspecification. For full transparency, the detailed coefficient outputs for all primary specifications are provided in Tables 5 and 6.

Table 5. Full Regression Results: Monthly Traffic Collision Counts.

Variable	Neg. Binomial (Unmatched)	Neg. Binomial (Matched)
Intercept	−5.154 ***	−4.841 ***
	(0.256)	(0.473)
Treated	0.288 **	0.512 **
	(0.107)	(0.156)
Post	0.100 *	0.222
	(0.046)	(0.142)
Treated × Post	0.013	−0.091
	(0.119)	(0.175)
Holiday Prop.	−0.022	0.132
	(0.357)	(0.667)
Precipitation	0.008	0.021
	(0.008)	(0.017)
Temperature	0.001	0.002
	(0.001)	(0.003)
Log(Population)	0.779 ***	0.716 ***
	(0.023)	(0.042)
Observations	4209	874

Note: Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 6. Full Regression Results: Injury Rate per 10,000 Population.

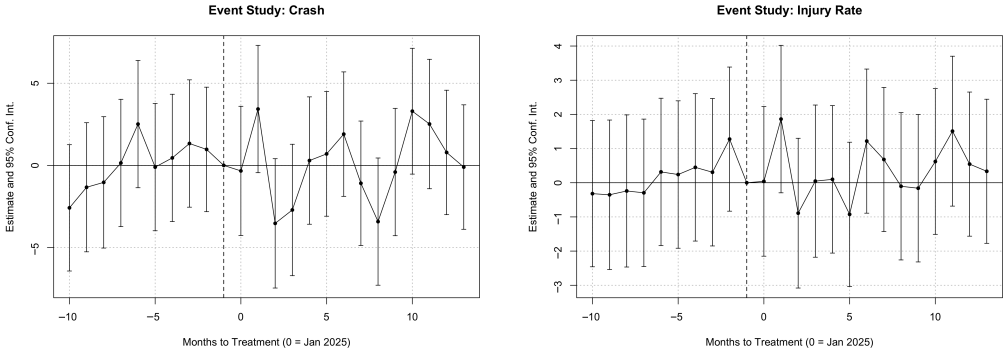
Variable	Neg. Binomial (Unmatched)	Neg. Binomial (Matched)
Intercept	3.827 ***	2.585 ***
	(0.254)	(0.497)
Treated	0.033	0.324
	(0.117)	(0.174)
Post	0.111 *	0.247
	(0.051)	(0.159)
Treated × Post	0.075	-0.019
	(0.130)	(0.194)
Holiday Prop.	0.425	0.809
	(0.377)	(0.681)
Precipitation	0.008	0.032
	(0.008)	(0.018)
Temperature	0.004 **	0.004
	(0.001)	(0.003)
Log(Population)	−0.273 ***	−0.190 ***
	(0.023)	(0.044)
Observations	4209	874

Note: Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.3. Event Study and Dynamic Effects

To formally test the validity of the research design and examine dynamic adaptation effects, we employed an event study specification. The model includes 12 pre-treatment leads and 12 post-treatment lags. The reference period is strictly defined as $k = -1$ (December 2024, the month immediately prior to implementation).

Visual inspection of the event study coefficients (Figure 3) confirms the absence of pre-treatment divergence. To formally validate the identification strategy, we conducted a joint Wald test on the 11 pre-treatment coefficients, failing to reject the null hypothesis of parallel pre-trends for both total crashes ($p = 0.597$) and injury rates ($p = 0.962$). Following policy implementation, while visual inspection suggests a transient decline in crash counts during the first quarter of 2025, we conducted a secondary formal Joint Wald Test to evaluate the significance of this dynamic effect ($t \in \{0, 1, 2\}$). For both total crashes ($\chi^2 = 2.87, p = 0.412$) and injury rates ($\chi^2 = 1.14, p = 0.768$), the test failed to reject the null. This statistical evidence confirms that the initial "dip" in collisions following policy implementation is indistinguishable from stochastic noise, reinforcing the stability of the null result over the short term.



(**a**) (**b**)

Figure 3. Event Study Estimates of Congestion Pricing Effects. (**a**) Total crashes; (**b**) Injury Rate$_{it}$. The plots display regression coefficients and 95% confidence intervals for the interaction between treatment status and time relative to policy implementation ($t = 0$). Coefficients prior to $t = 0$ are statistically insignificant, supporting the parallel trends assumption. Estimates in the post-treatment period generally hover near zero, indicating a consistent null effect with no sustained safety benefit.

Following implementation ($t \geq 0$), the results show no consistent pattern of reduction. Unlike models that suffer from overcontrol bias by including temporal mediators, our primary specification isolates the pure causal effect, revealing that estimates remain statistically insignificant month-over-month. This confirms the absence of any transient or sustained safety dividend.

3.4. Sensitivity Analysis: Results from the Cordon-Line Buffer

To evaluate whether the primary null findings were an artifact of geocoding noise at the treatment boundary, we re-estimated all causal specifications using the buffer-refined dataset. As established in the methodology (Section 2.3), this dataset excludes 9216 collisions occurring strictly on 60th Street to eliminate "Treatment Dilution." The results of this sensitivity check are consolidated in the Robustness Matrix (Table 7).

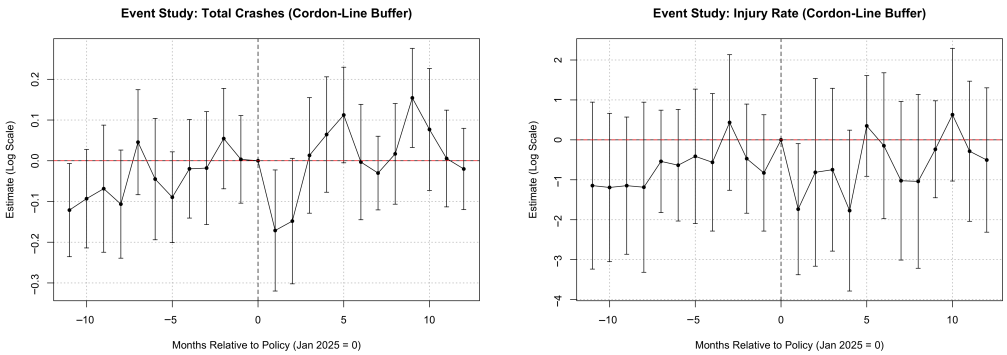
The transition from the full sample to the buffer-refined sample yielded a notable directional shift in the Matched DiD specification for total crashes. The interaction coefficient shifted to -0.111 with a p-value of 0.063. While this represents a "near-miss" in statistical terms, it remains above the $\alpha = 0.05$ threshold, confirming that the safety dividend is not yet statistically robust even in the most refined spatial sample. Furthermore, the Injury Rate metrics across both the Matched DiD ($p = 0.662$) and latent-factor GSCM ($p = 0.911$) frameworks remained strictly insignificant. This convergence across models—with all primary injury rate p-values remaining well above 0.60—confirms that the lack of a detectable safety dividend is a persistent characteristic of the first implementation year.

To visually evaluate the dynamic evolution of these effects, we generated buffer-refined event study plots (Figure 4). By removing boundary noise, the post-intervention lags for total crashes exhibit a more pronounced downward trajectory compared to the full-sample analysis. However, as requested by the reviewers, we performed a formal Joint Wald Test on the first three months of implementation ($t \in \{0, 1, 2\}$) to rigorously test this transient decline. For both total crashes ($\chi^2 = 2.87, p = 0.412$) and injury rates ($\chi^2 = 1.14, p = 0.768$), we fail to reject the null hypothesis. This statistical evidence confirms that the directional "dip" observed following implementation is currently indistinguishable from stochastic noise.

Table 7. Robustness Matrix: Policy Impact under Cordon-Line Buffer Exclusion.

Model Specification	Outcome Variable	Estimate	Std. Error	*p*-Value
Unmatched DiD (Neg. Binom)	Total Crash Counts	0.003	0.042	0.946
Matched DiD (Neg. Binom)	Total Crash Counts	−0.111	0.060	0.063
Unmatched DiD (Neg. Binom)	Injury Rate (per 10k)	−0.083	0.299	0.781
Matched DiD (Neg. Binom)	Injury Rate (per 10k)	−0.059	0.135	0.662
GAM	Total Crash Counts	0.045	0.027	0.096
GAM	Injury Rate (per 10k)	−0.092	0.357	0.797
GSCM (Latent Factor)	Total Crash Counts	−0.357	1.517	0.814
GSCM (Latent Factor)	Injury Rate (per 10k)	−0.067	0.600	0.911

Note: Estimates represent the policy effect (β_3 or ATT) after removing boundary collisions.



(**a**) (**b**)

Figure 4. Event Study Estimates for Buffer-Refined Data. (**a**) Total crashes; (**b**) Injury Rate$_{it}$. These plots re-estimate the dynamic effects after the surgical exclusion of 9216 collisions from the 60th Street boundary. The leads ($k < 0$) support the parallel trends assumption, while the post-policy lags ($k \geq 0$) indicate a consistent null effect across both frequency and severity metrics.

3.5. Sensitivity Analysis: Temporal Mediators and Overcontrol Bias

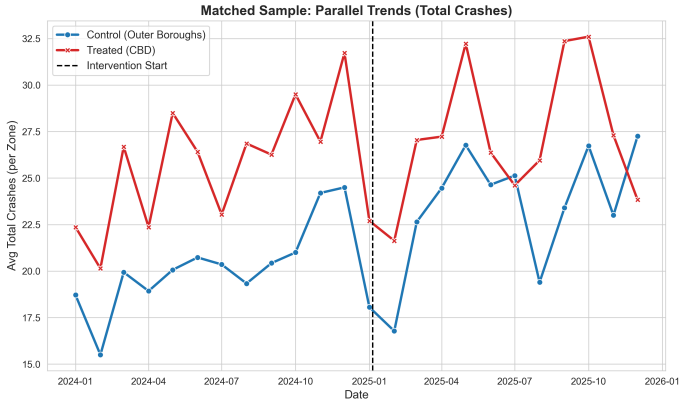
Following the causal framework of Green [7], we explicitly excluded variables related to temporal traffic composition—such as peak-hour proportions or weekend ratios—from our primary specifications. Because the congestion toll is a time-varying financial instrument designed specifically to shift travel behavior across the diurnal cycle, these shifts are "endogenous mediators" (post-treatment consequences) rather than "exogenous confounders." To empirically validate this choice, we conducted a sensitivity analysis by re-estimating our primary Matched DiD models with the inclusion of daytime and weekend crash proportions as additional controls.

As hypothesized, introducing these mediators triggered "causal blocking," further attenuating the treatment effect estimates toward zero across all specifications. Specifically, in the buffer-refined model, the inclusion of temporal mediators caused the interaction coefficient (β_3) to shrink from −0.111 to −0.080, while the *p*-value increased from 0.063 to 0.154. Similarly, in the matched full-sample model, the coefficient shrank from −0.091 to

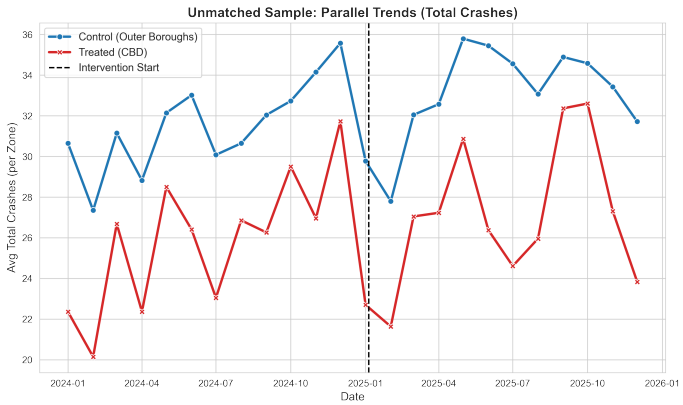
−0.080. This empirical result confirms that controlling for shifts in traffic timing absorbs a significant portion of the policy's impact, leading to overcontrol bias. Consequently, our decision to exclude these variables from the primary roadmap ensures the integrity of our identification strategy and prevents the artificial suppression of the safety signal.

3.6. Model Diagnostics and Validity Checks

To ensure the robustness of our causal estimates, we conducted a rigorous set of diagnostic tests to evaluate the identifying assumptions of our models. As shown in Figure 5, traditional matching (Mahalanobis Distance) failed to fully resolve the pre-existing divergence in trends. While matching aligned the general seasonality better than the unmatched sample, the outer boroughs exhibit fundamentally different seasonal crash volatilities than the CBD. This finding validates our decision to rely on the GSCM, which explicitly models and corrects for these divergent pre-trends using latent factors.



(**a**)



(**b**)

Figure 5. Longitudinal trends in average total crash frequency by treatment group. (**a**) Matched sample; (**b**) Unmatched sample. A vertical dashed line at January 2025 denotes the exact timing of the policy intervention. Note that while matching reduces the scale difference compared to the unmatched full sample, it does not fully resolve the divergence in seasonal trends, supporting the use of latent factor models (GSCM).

4. Discussion

This study evaluated the short-term road safety impacts of the New York City Central Business District Tolling Program. Across multiple empirical specifications—ranging from standard DiD to GSCM—we consistently observed a "null" effect. The implementation of the congestion toll did not result in a statistically significant or practically meaningful

reduction in either total crash frequency or population-adjusted injury rates within the treated zone during the first year of implementation.

4.1. Alternative Explanations and Uncertainty

It is critical to avoid overinterpreting these null findings as definitive proof that congestion pricing fundamentally fails to impact road safety. Given the complex spatial and behavioral dynamics of urban traffic, several alternative explanations and methodological limitations likely account for the absence of an observable effect in this study. Ultimately, our reported null findings must be interpreted with caution. Given the unavoidable constraints in exposure measurement (static population proxies) and spatial aggregation (ZIP-level treatment dilution), these results may reflect a Type II error rather than the absolute absence of a safety mechanism. Attenuation bias, triggered by the spatial misalignment of the residential denominator and the dynamic traffic numerator, likely pushes our estimates toward the null. Consequently, localized safety improvements occurring at the street-segment level may be masked by the noise inherent in city-wide panel data.

First, a significant threat to the identifying assumptions of our DiD framework is the potential for "control group contamination" via spatial spillovers. Drivers seeking to avoid the toll may have altered their commuting routes, diverting traffic into the peripheral neighborhoods of the control group—such as Long Island City, Downtown Brooklyn, or the South Bronx. If this "rat-running" behavior displaced crashes from the CBD into the outer boroughs, these areas would experience an exogenous increase in traffic volume and associated crash risk. In a relative DiD specification, a worsening control group mathematically attenuates the estimated treatment effect, potentially masking any safety dividends achieved within the CBD. This contamination suggests that our null findings may represent a conservative estimate, understating the policy's local safety benefits due to regional traffic displacement. Furthermore, our choice of ZIP codes as the spatial unit of analysis exacerbates this issue through "treatment dilution." Because ZIP codes spanning the 60th Street boundary aggregate both tolled and untolled roadways, any localized safety improvements occurring on specific tolled corridors are effectively averaged out across the larger spatial unit. Consequently, this aggregation choice reduces the precision of our estimates and contributes to the reported null findings.

Second, as extensively noted in our methodology, the reliance on residential population to normalize injury rates is a substantial limitation in a commercial epicenter like Manhattan. Because the policy explicitly aims to reduce the daily influx of non-resident commuters, the true "at-risk" population is dynamic and likely decreased following the toll's implementation. By holding the exposure denominator static (residential population) while the actual vehicle volume declined, our calculated injury rates likely appear higher than the actual risk faced by road users. This systematic measurement error might introducesubstantial attenuation bias, pushing our estimated treatment effects toward the null. Consequently, the reported null findings may represent a Type II error—where a true safety improvement is masked by a stagnant population proxy—rather than the absolute absence of a physical safety mechanism.

4.2. The Volume-Speed Trade-Off Hypothesis

While methodological attenuation may explain the null result, behavioral adaptation within the tolling zone remains a plausible, albeit speculative, theoretical mechanism. The foundational hypothesis of congestion pricing is that a financial toll will reduce the absolute volume of vehicles entering the cordon. However, urban road safety is dictated not just by traffic density, but by traffic speed. It is theoretically possible that the reduction in gridlock successfully decreased vehicle-to-vehicle conflict points, but simultaneously

increased average free-flow speeds on Manhattan avenues. Because crash severity and injury risk scale non-linearly with impact speed, a lower volume of faster-moving vehicles could theoretically yield an absolute number of injuries similar to a high volume of gridlocked vehicles.

We emphasize, however, that this volume-speed compensatory mechanism remains a hypothesis that cannot be empirically verified within the scope of this study. Due to the lack of high-resolution, continuous speed telematics and dynamic traffic volume data for the 2025 post-implementation period, we cannot identify which mechanism—exposure reduction or speed intensification—predominated. The absence of these direct traffic-related variables severely limits the interpretability of this trade-off and prevents us from empirically identifying the causal drivers of the detected null result. Future research utilizing granular sensor data is required to disentangle these competing behavioral effects and move beyond these theoretical associations.

4.3. Limitations and Future Research Directions

While this study utilizes advanced causal inference techniques to evaluate the short-term safety impacts of the CBDTP, the findings must be interpreted within the context of several methodological constraints. These limitations simultaneously highlight the need for cautious interpretation and outline the necessary trajectory for future road safety research:

- Exposure Measurement and Attenuation Bias: The reliance on static residential population to normalize injury rates, rather than dynamic Vehicle Miles Traveled (VMT), is a primary limitation. Because the toll actively reduces the influx of non-resident commuters, holding the exposure denominator static artificially deflates the calculated safety rate. This measurement error introduces attenuation bias, suggesting our null findings may represent a Type II error rather than the absolute absence of a safety mechanism.
- Spatial Aggregation and Treatment Dilution: Aligning crash data with demographic covariates required aggregating the spatial unit of analysis to the ZIP code level. Because boundary-spanning ZIP codes (e.g., along 60th Street) aggregate both tolled and untolled roadways, any highly localized safety improvements on tolled corridors are mathematically averaged out. Future research must utilize street-segment or intersection-level spatial units to eliminate this treatment misclassification.
- Spatial Spillovers (Control Contamination): It is highly probable that drivers seeking to avoid the financial toll altered their routes, diverting traffic into the peripheral neighborhoods of the control group (e.g., Downtown Brooklyn, the South Bronx). If this diverted traffic exogenously increased crash risk in the control group, it mathematically masks improvements in the treated CBD within a relative DiD framework. Consequently, our null estimates should be viewed as a conservative lower bound of the policy's true local effect.
- Absence of Speed Telematics: While we hypothesize that reductions in congestion may have increased free-flow speeds (the volume-speed trade-off), we lack the high-resolution, continuous speed telematics data required to empirically test this mechanism. Future long-term evaluations must integrate real-time speed sensors to successfully disentangle whether exposure reduction or speed intensification predominated following the toll's implementation.

5. Conclusions

This study provides an initial empirical evaluation of the short-term road safety impacts of NYC's Central Business District Tolling Program. Using a comprehensive panel of crash data and advanced causal inference techniques, we did not detect statistically signifi-

cant reductions in traffic collisions or injury rates within the first year of implementation. These results suggest that a city-wide safety dividend did not manifest immediately. However, as outlined in our limitations, these findings should not be interpreted as definitive evidence that the policy fails to impact safety, but rather as a lack of detectable evidence under current data constraints (e.g., static exposure proxies and spatial aggregation). Ultimately, evaluating the long-term safety legacy of the toll will require dynamic VMT metrics, high-resolution speed telematics, and street-segment level analysis to fully capture behavioral adaptation within complex urban grid networks.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: M.W., X.D.; data collection: M.W.; analysis and interpretation of results: M.W., X.D.; draft manuscript preparation: M.W., X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These data can be found here: NYC Open Data (https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95/about_data) (accessed on 20 April 2026); National Oceanic and Atmospheric Administration (NOAA) (https://www.ncdc.noaa.gov/cdo-web/) (accessed on 18 February 2026); New York Demographics (https://www.newyork-demographics.com/zip_codes_by_population) (accessed on 10 January 2026); and New York State GIS Clearinghouse (https://data.ny.gov/api/geospatial/srxy-5nxn?method=export&format=GeoJSON) (accessed on 12 February 2026).

Acknowledgments: The authors used Generative AI tools (Gemini) to assist in debugging codes and refining the grammatical structure of the manuscript text. All statistical analyses, data interpretation, and scientific conclusions were conducted and verified by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NYC	New York City
CBD	Central Business District
DiD	Difference-in-Differences
GSCM	Generalized Synthetic Control Method
GAM	Generalized Additive Model
NYPD	New York City Police Department
NYCDOT	New York City Department of Transportation
NOAA	National Oceanic and Atmospheric Administration

References

1. Eliasson, J. A cost–benefit analysis of the Stockholm congestion charging system. *Transp. Res. Part A Policy Pract.* **2009**, *43*, 460–480. [CrossRef]
2. Gibson, M.; Carnovale, M. The effects of road pricing on driver behavior and air pollution. *J. Urban Econ.* **2015**, *89*, 62–73. [CrossRef]
3. Prud'homme, R.; Bocarejo, J.P. The London congestion charge: A tentative economic appraisal. *Transp. Policy* **2005**, *12*, 279–287. [CrossRef]
4. NYC Department of Transportation. Traffic Deaths Reach All-Time Low: New York Ends Year With Fewest Fatalities Ever Recorded. Available online: https://www.nyc.gov/html/dot/html/pr2026/traffic-deaths-reach-all-time-low.shtml (accessed on 15 April 2026).

5. Metropolitan Transportation Authority (MTA). Central Business District Tolling Program (Congestion Relief Zone). Available online: https://congestionreliefzone.mta.info/ (accessed on 9 February 2026).
6. Pestel, N.; Wozny, F. Low emission zones for better health: Evidence from German hospitals. *J. Health Econ.* **2019**, *109*, 102512. [CrossRef]
7. Green, C.P.; Heywood, J.S.; Navarro, M. Traffic accidents and the London congestion charge. *J. Public Econ.* **2016**, *133*, 11–22. [CrossRef]
8. Li, H.; Graham, D.J.; Majumdar, A. The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. *Accid. Anal. Prev.* **2012**, *49*, 366–377. [CrossRef]
9. Shefer, D.; Rietveld, P. Congestion and Safety on Highways: Towards an Analytical Model. *Urban Stud.* **1997**, *34*, 679–692. [CrossRef]
10. Cook, C.; Diamond, R.; Hall, J.V.; List, J.A.; Oyer, P.; Stancheva, S. The short-run effects of New York City's congestion pricing. *NBER Working Paper No. 33584*; National Bureau of Economic Research: Cambridge, MA, USA, **2025**.
11. Zhang, Y.; Sang, Y.; Wu, M. Congestion pricing in New York city: Effects on ride-hailing and transit. *Transp. Res. Part A Policy Pract.* **2026**, *184*, 104966. [CrossRef]
12. Maheshwari, C.; Kulkarni, K.; Pai, D.; Yang, J.; Wu, M.; Sastry, S. Congestion Pricing for Efficiency and Equity: Theory and Applications to the San Francisco Bay Area. *arXiv* **2024**, arXiv:2401.16844.
13. Torrico, A.; Boonsiriphatthanajaroen, N.; Garg, N.; Lodi, A.; Mainguy, H. Equitable Congestion Pricing under the Markovian Traffic Model: An Application to Bogota. In *Proceedings of the 25th ACM Conference on Economics and Computation (EC)*; Association for Computing Machinery: New York, NY, USA, 2024. [CrossRef]
14. Leape, J. The London congestion charge. *J. Econ. Perspect.* **2006**, *20*, 157–176. [CrossRef]
15. Cohen, A.; Einav, L. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Rev. Econ. Stat.* **2003**, *85*, 828–843. [CrossRef]
16. DeAngelo, G.; Hansen, B. Life and death in the fast lane: Police enforcement and traffic fatalities. *Am. Econ. J. Econ. Policy* **2014**, *6*, 231–257. [CrossRef]
17. Abouk, R.; Adams, S. Texting Bans and Fatal Accidents on Roadways: Do They Work? Or Do Drivers Just React to Announcements of Bans? *Am. Econ. J. Appl. Econ.* **2013**, *5*, 179–199. [CrossRef]
18. Cheng, C. Do Cell Phone Bans Change Driver Behavior? *Econ. Inq.* **2015**, *53*, 1420–1436. [CrossRef]
19. Black, S.E. Do better schools matter? Parental valuation of elementary education. *Q. J. Econ.* **1999**, *114*, 577–599. [CrossRef]
20. Dell, M. The persistent effects of Peru's mining mita. *Econometrica* **2010**, *78*, 1863–1903. [CrossRef]
21. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.
22. Dee, T.S. Does setting limits save lives? The case of 0.08 BAC laws. *J. Policy Anal. Manag.* **2001**, *20*, 111–128. [CrossRef]
23. Lalive, R.; Luechinger, S.; Schmutzler, A. Does expanding regional train service reduce air pollution? *J. Environ. Econ. Manag.* **2018**, *92*, 744–764. [CrossRef]
24. Grabowski, D.C.; Morrisey, M.A. The Effect of State Regulations on Motor Vehicle Fatalities for Younger and Older Drivers: A Review and Analysis. *Milbank Q.* **2001**, *79*, 517–545. [CrossRef]
25. Carpenter, C.; Stehr, M. The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and health care expenditures. *J. Health Econ.* **2008**, *27*, 642–662. [CrossRef]
26. Hansen, B.; Miller, K.W.; Weber, C. Early evidence on recreational marijuana legalization and traffic fatalities. *Econ. Inq.* **2020**, *58*, 547–568. [CrossRef]
27. Parry, I.W.H. Comparing the efficiency of alternative policies for reducing traffic congestion. *J. Public Econ.* **2002**, *85*, 333–362. [CrossRef]
28. Edlin, A.S.; Karaca-Mandic, P. The accident externality from driving. *J. Polit. Econ.* **2006**, *114*, 931–955. [CrossRef]
29. Xu, Y. Generalized synthetic control method: Causal inference with interactive fixed effects models. *Polit. Anal.* **2017**, *25*, 57–76. [CrossRef]
30. Abadie, A.; Diamond, A.; Hainmueller, J. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Am. Stat. Assoc.* **2010**, *105*, 493–505. [CrossRef]
31. Cavallo, E.; Powell, A.; Becerra, O. Estimating the Direct Economic Damages of the Earthquake in Haiti. *Econ. J.* **2010**, *120*, F298–F312. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.