CEWP 23-07

One Size Does Not Fit All: Co-Benefits of Congestion Pricing in the San Francisco Bay Area

Ekaterina Alekhanova
Carleton University

Kate Foreman
NERA Economic Consulting

Maya Papineau
Carleton University

Reid Stevens
Independent Consultant

November 12, 2023

CARLETON ECONOMICS WORKING PAPERS

Department of Economics

1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B

One Size Does Not Fit All: Co-Benefits of Congestion Pricing in the San Francisco Bay Area

Ekaterina Alekhanova* Kate Foreman† Maya Papineau‡ Reid Stevens§

November 12, 2023

Abstract

On July 1, 2010, congestion pricing during peak traffic times was implemented on the San Francisco-Oakland Bay Bridge. In response to the toll, automobile traffic on the bridge declined. Exploiting a quasi-experimental approach, the study finds that although public transit ridership increased after the new road toll policy went into effect, congestion pricing did not cause a change in traffic-related air pollution and respiratory illness incidence in the bridge vicinity, in contrast with the past work on the topic in other settings. This points to the importance of considering the heterogeneous place-based factors that drive the welfare effects of environmental policy.

Keywords: Air Pollution, Respiratory Health, Congestion Pricing, Public Transit

*Department of Economics, Carleton University
†NERA Economic Consulting
‡Department of Economics, Carleton University
§Independent Consultant

1 Introduction

Congestion pricing, which raises road toll fees during periods of high demand, was implemented on the San Francisco-Oakland Bay Bridge, also known as the Bay Bridge, on July 1, 2010. Though the primary motivation for the policy was to raise funds to pay for bridge repair and maintenance, it was also expected to reduce negative externalities associated with private automobile use. Foreman (2016) shows that the new toll policy decreased traffic volume and travel times during peak hours, and some of that traffic shifted towards off-peak shoulder hours and, to a lesser extent, a lower-priced nearby bridge. Our paper extends the analysis in Foreman (2016) by examining how the change in driving patterns caused by congestion pricing affected traffic-related air pollution in the vicinity of the bridge and the respiratory health of patients living close to the bridge. In addition, we explore if commuters shifted from private automobiles to higher public transit usage, which has a lower per-person carbon footprint (MIT CLimate Portal, 2023).

It is not clear, a priori, how overall levels of air pollution, respiratory health, and public transit ridership could be influenced by modified commuting behavior due to the non-uniform toll. If drivers simply reduce the number of trips they take ("trip reduction"), mainly during peak hours, this will likely have no effect on public transportation but will lower air pollution levels and may improve respiratory health in the areas close to the bridge. If drivers make the same number of trips but shift the time of those trips away from peak periods ("time shifting"), this usually has no effect on overall air pollution, respiratory health, and public transportation. There will be no pollution, health, and public transit effects if drivers continue to travel during peak periods but avoid the route with congestion pricing by using alternative roads ("route shifting"). Finally, drivers can continue to commute during peak periods but use public transportation instead of driving ("mode shifting"), which will reduce vehicle emissions, as well as improve air quality and respiratory health. The overall effect of these responses to congestion pricing on pollution, health, and public transit has important implications for the design of both climate and transportation policy. A precise estimate of the effect of decreased driving on air pollution and respiratory health will also refine welfare calculations for congestion pricing.

Existing literature estimates the changes in air pollution levels and public health due to either command-and-control type of driving restrictions, such as low-emission zones (Pestel and Wozny, 2019; Sun et al., 2014; Davis, 2008; Friedman et al., 2001), or market-based policies, including congestion pricing (Simeonova et al., 2021; Isaksen and Johansen, 2021; Gibson and Carnovale, 2015; Currie and Walker, 2011; Auffhammer and Kellogg, 2011). In addition, some studies specifically focus on the effect of public transportation on air pollution (Gendron-Carrier et al., 2022; Rivers et al., 2020; Beaudoin et al., 2015; Anderson, 2014; Chen and Whalley, 2012). Most of the papers find that driving restrictions are associated with lower air pollution and improved health.

Among the aforementioned research works, a growing number of papers use quasi-experimental approaches to estimate how changes in driving and commuting behavior affect air pollution and/or health (Simeonova et al., 2021; Isaksen and Johansen, 2021; Rivers et al., 2020; Pestel and Wozny, 2019; Gibson and Carnovale, 2015; Anderson, 2014; Chen and Whalley, 2012; Auffhammer and Kellogg, 2011; Currie and Walker, 2011; Davis, 2008).

In this study, we exploit the quasi-natural experiment created by the introduction of congestion pricing on the San Francisco Bay Bridge, located in the 13th largest metropolitan area in the United States with a population of 4.5 million people (The U.S. Census Bureau, 2022). Using a regression discontinuity approach, we find that while congestion pricing led to an increase in public transit ridership, it did not change local air pollution levels and respiratory illness hospital admissions.

The rest of the paper is organized as follows. We describe the policy background in Section 2 and review the data in Section 3. Section 4 discusses the research design. In Section 5, we demonstrate the results from the regression discontinuity models. Concluding remarks are presented in Section 6.

2 Background

In the San Francisco Bay Area in California (see Figure 1), the communities in the East Bay are connected with the San Francisco Peninsula via three bridges and a metro rail system.

Figure 1: San Francisco Bay Area

Source: OnTheWorldMap (2023).

The Bay Bridge is the main bridge that is heavily used by drivers to commute and/or travel between downtown San Francisco and cities in the East Bay. The traffic volume on the bridge is around 124,000 vehicles per weekday in each direction. There are also two smaller bridges located south of the Bay Bridge that link the San Francisco Peninsula with the East

Bay (see Figure 2): the San Mateo Bridge, 18 miles of the Bay Bridge, carries roughly 2.7 times less vehicles than the Bay Bridge, as well as the Dumbarton Bridge, 25 miles of the Bay Bridge, which is used by 4 times less vehicles (Foreman, 2016).

Figure 2: Bridges in the San Francisco Bay Area

Source: Wikipedia (2023).

The main public transportation option for people traveling between San Francisco and the East Bay is the light rail system (BART). In addition, there is the AC Transit bus service, which also has local transbay lines[1].

Foreman (2016) provides a detailed description of the road toll policy before and after July 1, 2010. For the Bay, San Mateo, and Dumbarton Bridges, a road toll only exists for any vehicle on the westbound trip. There are FasTrak lanes, cash lanes, and carpool lanes. A FasTrak lane is for Electronic Toll Collection tag holders (FasTrak tags), while a cash lane allows drivers to pay with either cash or their FasTrak tags. Using FasTrak lanes decreases the delay from toll collection since drivers do not have to stop to pay the toll, as compared to those paying in cash. Finally, there are carpool lanes available during peak hours only. Carpools are three or more people on the Bay Bridge and two or more people on the San Mateo and Dumbarton Bridges. Motorcycles, two-seat vehicles with two passengers,

[1]As noted in Foreman (2016), there are other options which include the ferry and two roundabout driving routes. Only few people take the ferry, and the roundabout routes are either crossing two bridges (the Richmond and Golden Gate Bridges) or driving south around the Bay. The first roundabout option would make drivers pay two road tolls, while the second one would require them to drive over 50 miles out of their way. Therefore, same as in Foreman (2016), we assume these other ways to cross the Bay are not reasonable substitutes for the Bay Bridge.

and vehicles with DMV-issued Clean Air decals are also allowed to use carpool lanes (DMV stands for Department of Motor Vehicles). In the off-peak time, carpool lanes are either closed, or they revert to bus, FasTrak, or cash lanes. Before July 1, 2010, the road toll was a $4 per vehicle fee for vehicles with 2 axles (vehicles with more axles face a higher rate), and the price was identical across all hours of the day for the three bridges. There was no toll for carpool lanes during peak hours.

On July 1, 2010, the tolls on the Bay Area bridges were raised with the purpose of getting additional funding for maintenance, transport projects, and seismic retrofitting. There was a uniform increase in the tolls on the San Mateo and Dumbarton Bridges: the tolls increased from $4 to $5 per vehicle, with the reduced $2.50 toll for carpools during peak hours (carpools were now required to have FasTrak tags). As for the Bay Bridge, it implemented a non-uniform change in the toll rates. During weekdays, the toll became $6 per vehicle in peak hours (from 5 am to 10 am and from 3 pm to 7 pm) and $4 per vehicle in the off-peak time; the weekend toll was $5 per vehicle. The toll for carpool lanes was the same as on the other bridges.

We use the introduction of the congestion pricing policy on the Bay Bridge as a quasi-experimental setting that provides an opportunity to empirically measure the effect of the toll on air pollution, respiratory health, and public transportation ridership.

3 Data

This paper uses several separate data sources. First, we use daily and hourly air pollution and weather data for our sample area. Second, we exploit daily emergency department visits and hospital admission data for patients residing in the sample area. Finally, we utilize a comprehensive dataset of transbay BART and AC Transit trips at the hourly level. The sample area should be located close enough to the Bay Bridge so that we can capture the direct effect of congestion pricing implemented on the bridge. Our definition of "close enough" is within 6 miles of the Bay Bridge.

All the observations are for the period between May 1st, 2010, and August 31st, 2010, i.e. 60 days before and 60 days after July 1, 2010, the day when the new toll policy went into effect. The sample does not include weekends, as well as federal and state holidays since the congestion pricing is in effect only on weekdays; in addition, we drop the period from July 2 until July 9 to account for additional vacation days around Independence Day that many Americans take (July 2 is Friday before Independence Day, and July 5–9 is the week following the holiday). We also drop Spare the Air days which are usually a few days in August and September declared by the Bay Area authorities as days on which the local residents are urged to drive less since the concentrations of ground-level ozone exceed federal air quality standards (Bay Area Air Quality Management Disctrict, 2023).

3.1 Air Pollution

Chronic exposure to ambient air pollution, even at low levels, is associated with respiratory illness and other negative health effects (Ransom and Pope, 1992; Schlenker and

Walker, 2016; Cohen et al., 2017; Deryugina et al., 2019; Jans et al., 2018; Anderson, 2020; Manisalidis et al., 2020; Bala et al., 2021; Simeonova et al., 2021). We investigate one of the criteria air pollutants, fine particulate matter ($PM_{2.5}$); it is linked to vehicle traffic and detrimental to respiratory health.

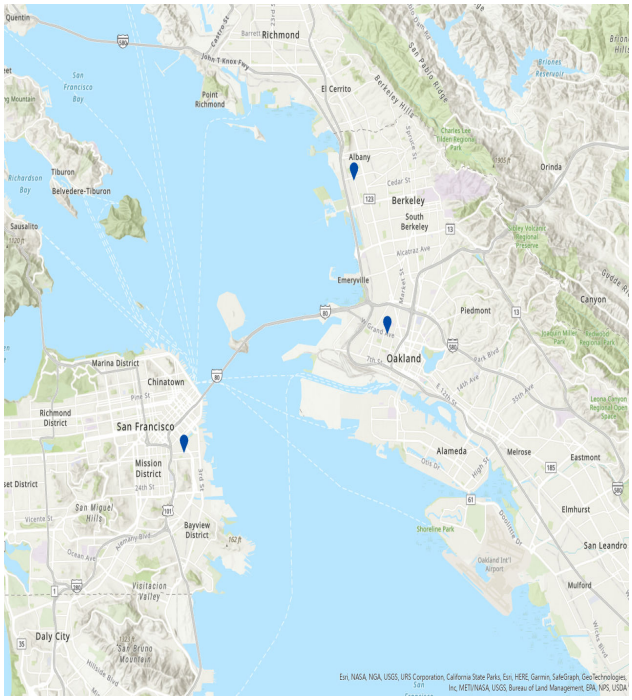
Ambient air pollution, specifically in the form of $PM_{2.5}$[2], has been identified as a leading cause of mortality and morbidity around the world (The Institute For Health Metrics And Evaluation, 2019). Particulate matter (PM) refers to the concentration of small airborne liquid and solid particles, such as dust, dirt, soot, smoke, and liquid droplets, that are classified by size. PM may originate from both natural sources, such as windblown soil or sea salt spray, and anthropogenic sources, including fossil fuel burning, various industrial processes, agricultural activity, and road dust (Isphording and Pestel, 2021; Health Canada, 2022a). The particles less than 10 microns (μm) in diameter can penetrate into the respiratory tract and are those of main concern for human health; the particles that are less than 2.5 μm in diameter are known as fine particulate matter, $PM_{2.5}$ (Health Canada, 2022a). The major mechanism for removing $PM_{2.5}$ is via precipitation. $PM_{2.5}$ can be transported for several days downwind, affecting populations up to approximately 600 miles away from the point of emission; on the other hand, primary $PM_{2.5}$ emitted in urban areas will have a large impact in the immediate vicinity (Gilmore et al., 2019).

Ambient air pollution data is obtained using from the Air Quality and Meteorological Information System of the California Air Resources Board (California Air Resources Board, 2023). Non-mobile monitoring stations measure hourly levels of hazardous air pollutants in the San Francisco Bay Area. The sample area includes three pollution monitoring sites within 6 miles of the Bay Bridge[3]: one monitor is in San Francisco (Arkansas Street) and two monitors are in the East Bay (West Oakland and Berkeley), shown in Figure 3.

[2]Ground-level ozone is the other indicator that is used to measure exposure to air pollution in the Global Burden of Disease study (Cohen et al., 2017), and it leads to similar inflammatory reactions as $PM_{2.5}$ (Isphording and Pestel, 2021). Moreover, along with $PM_{2.5}$, ozone is a major traffic-related air pollutant. However, ozone is not emitted directly by vehicles – it is created by chemical reactions of certain pollutants under sunlight (Isphording and Pestel, 2021; Health Canada, 2022b), and measuring the effect of vehicles on ozone concentrations can be challenging (Rivers et al., 2020). As a result, in the paper, we conduct a single-pollutant analysis, evaluating the association between ambient $PM_{2.5}$ and congestion pricing.

[3]We calculate the geographical distance between pollution stations in the Bay region and the east Bay Bridge toll entrance, the west Bay Bridge toll entrance, as well as the middle of Bay Bridge (Treasure Island). A given pollution station is located in the 6-mile vicinity of the Bay Bridge if it lies within 6 miles of the west toll entrance, or 6 miles of the east toll entrance side, or 6 miles of the center of the bridge.

Figure 3: Pollution Monitoring Sites in the Sample Area



Notes: The sample area includes 3 pollution monitoring sites within 6 miles of the Bay Bridge.

In the analysis, we use the daily mean measurements of $PM_{2.5}$. Within a day, $PM_{2.5}$ levels tend to peak during the morning commute and then gradually decline until a slight rise later in the evening. Daily mean $PM_{2.5}$ measurements display a seasonal pattern, peaking during the winter months and falling during the summer months.

Since weather, including wind, temperature, and precipitation, affects pollution, we augment the air pollution data with daily weather data for the San Francisco Bay Area. We use gridded weather data produced by PRISM Climate Group (2023). This provides daily precipitation, minimum temperature, and maximum temperature data for four-kilometer grids[4]. For each pollution monitoring station, we keep the nearest weather grid. For wind speed and direction, we exploit data from the National Oceanic and Atmospheric Administration's (NOAA) Center for Operational Oceanographic Products and Services; we obtain wind data from a single weather station closest to the Bay Bridge, the one in Alameda, which is on the East Bay side of the bridge (National Oceanic and Atmospheric Administration, 2023a)[5].

[4]We use only monitor-days for which observations are recorded for at least 21 (87.5%) hours per day (if we only kept the monitoring stations with all 24 non-missing hourly observations per each day of the sample period, we would have just few stations left). In the gridded weather data, we drop days with both air temperature and precipitation missing; in the wind data, we only keep days with both wind speed and wind direction non-missing.

[5]We also use hourly weather data, which is obtained from a different source. We exploit data on hourly precipitation, temperature, wind direction, and wind speed from NOAA's National Weather Service (National Oceanic and Atmospheric Administration, 2023b). We use two airport weather stations near the sample area: San Francisco International Airport and Oakland International Airport (the rest of the weather stations in the sample area have too many missing wind speed and wind direction observations). We assign

3.2 Respiratory Health

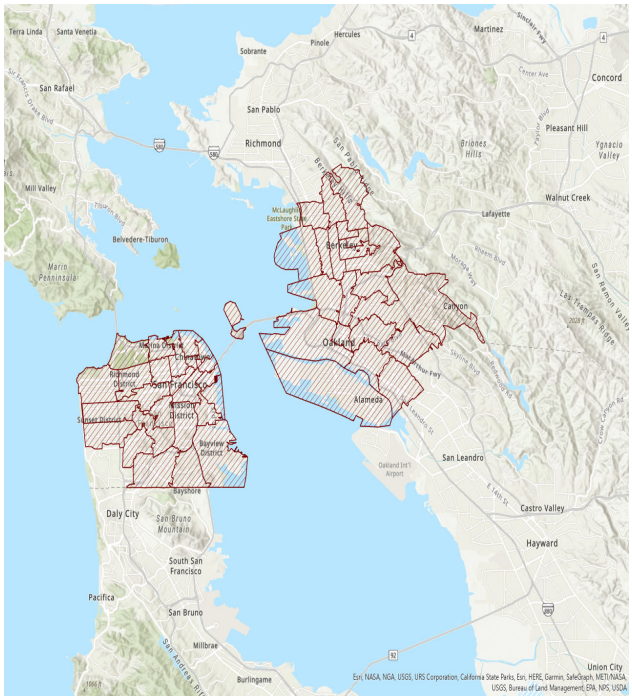
We use non-public data from California's Department of Health Care Access and Information (California Department of Health Care Access and Information, 2023). The dataset contains emergency department visits, overnight hospital stays, as well as ambulatory surgeries (i.e. surgeries that do not require overnight hospital admission). For each patient visit, the data lists the service date, the primary diagnosis code, the secondary diagnosis codes, as well as patient and hospital ZIP codes. We use the data on admission for respiratory illness (either as a primary or secondary diagnosis code) for patient ZIP codes in the sample area, i.e. we study the daily total numbers of emergency department visits, overnight hospital stays, and ambulatory surgeries related to respiratory health.

Each ZIP code of a patient residence is joined with its corresponding ZIP Code Tabulation Area (ZCTA)[6]. There are 47 ZCTAs within 6 miles of the Bay Bridge (Figure 4). We calculate the geographical distance between ZCTAs' centroids and three points on the Bay Bridge (the east Bay Bridge toll entrance, the west Bay Bridge toll entrance, and the middle of Bay Bridge) to allocate ZCTAs to the sample region. As for the weather data, we calculate the geographical distance between the weather grids' centroids and ZCTAs' centroids, and we keep the nearest weather grid for each ZCTA.

each pollution station to its closest airport: one pollution station in San Francisco is connected to San Francisco International Airport, and the two pollution monitoring sites in the East Bay are joined with the data from Oakland International Airport.

[6]ZCTAs are areal representations of ZIP code service areas. For the 2010 United States Census, each ZCTA aggregates the census blocks whose addresses are associated with a given ZIP code. Since a ZCTA represents the ZIP code used by most addresses in a census block, addresses can sometimes be assigned to a ZCTA code that differs from their ZIP code (The U.S. Census Bureau, 2023). In the sample, there were only few ZIP codes like that. Therefore, in the paper, we may use the terms 'ZCTA' and 'ZIP code' interchangeably.

Figure 4: ZCTAs in the Sample Area



Notes: The sample area includes 47 ZCTAs within 6 miles of the Bay Bridge.

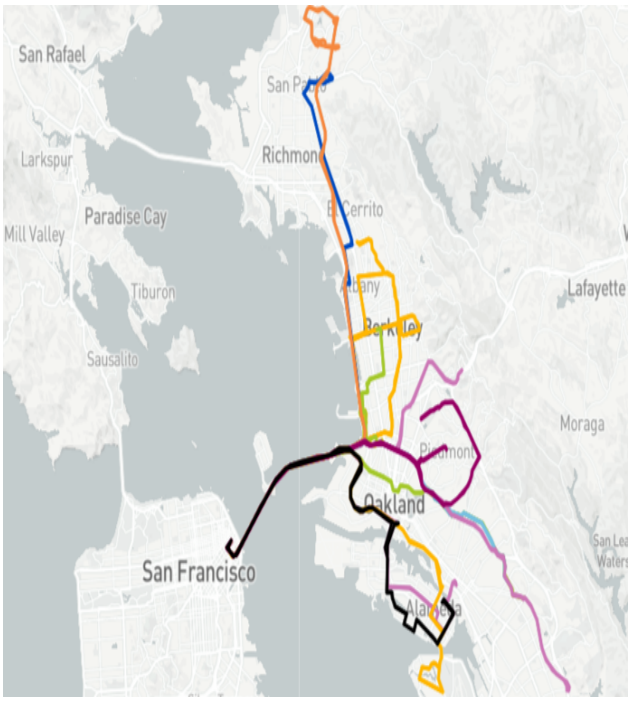
3.3 Public Transit

The paper uses non-public data on public transportation usage from BART (Bay Area Rapid Transit), the metro rail system, and AC Transit (Alameda-Contra Costa Transit District), the bus system, servicing the San Francisco Bay Area (Bay Area Rapid Transit, 2023; AC Transit, 2023). The BART map is shown in Figure 5, and AC Transit routes (transbay only) are displayed in Figure 6.

Figure 5: BART Map

Source: Bay Area Rapid Transit (2023).

Figure 6: AC Transit Map (Transbay Routes)



Source: AC Transit (2023).

We only examine ridership on transbay routes that follow the Bay Bridge[7]. The transbay routes are a viable substitute for people who switch from driving to commuting via public transportation. In addition, since there is no toll for any vehicle on the eastbound trip, we study westbound ridership only, i.e. trips from the East Bay to San Francisco. Within a day, transbay BART ridership is highest during morning commute times. Unlike BART, AC Transit ridership tends to peak around mid-day. It appears that AC transit carries fewer traditional commuters. However, BART riders constitute the majority of public transit passengers. Additionally, during the period of observation in this study, there were no BART or AC Transit fare changes.

The unit of observation in the BART dataset is the number of riders, each hour, entering a station and exiting from another station. The AC Transit dataset provides a sample of ridership levels by hour for each of the commuter bus routes. In the two datasets, we calculate the total number of riders per hour per day across all transbay station pairs (BART) or bus routes (AC Transit) and then merge the datasets to obtain the hourly number of public transit riders per each day in the sample. The public transit does not operate 24 hours a day; the official transit schedule does not match fully our data since the schedule depends on a route/station and there are some days when the transit officials temporarily change the schedule (sports games, days around holidays, etc.). For example, Red and Green BART Lines operate from 5 am (some start around 4:30 am) until around 10 pm; the closing time of the other lines (Orange, Blue, Yellow) can vary by station: some stations operate until 10 pm, but some are open until 1 am. At the same time, our data shows that BART has non-zero passengers in each hour even when the metro is supposed to be closed. Based on the data, there are only few passengers on average from 1 am until 4 am for Red and Green Lines and from 2 am until 4 am for the rest of the lines. Therefore, the final dataset does not include the 1 am – 4 am hours for Red and Green Lines, as well as 2 am – 4 am for the other lines. AC Transit has zero passengers from 2 am to 5 am and a small number of passengers from 1 to 2 am; we drop the 1 am – 5 am time period from the AC Transit data.

4 Model

We measure the effect of congestion pricing using a sharp regression discontinuity design. This approach uses the abrupt implementation of congestion pricing on July 1, 2010, to estimate the short-run effect of the policy change. We restrict the sample to a relatively narrow time interval around the date the new policy went into effect. In this short time window, any factors affecting air quality, respiratory illness, or transit ridership are likely to be similar, so that observations before the policy serve as a comparison group for observations after the road toll implementation[8] (Davis, 2008).

[7]For BART, we restrict our analysis to BART traffic between the San Francisco Central Business District, defined as the area serviced by the Embarcadero, Montgomery Street, Powell Street, and Civic Center BART stations, and the East Bay through the transbay tube.

[8]Obtaining causal inference about the effectiveness of congestion pricing using another quasi-experimental approach, such as difference in differences, may be more challenging because it would require us to construct a valid counterfactual for air pollution, hospital admission, and public transit ridership in the absence of the new toll policy.

More specifically, we regress air pollution ($PM_{2.5}$) concentrations, respiratory illness hospital admissions, or public transit ridership on a dummy variable for the new road toll policy in a regression discontinuity framework with time as the running variable.

We estimate the following specification for air pollution and hospital admissions:

$$\mathrm{Y}_{it} = \beta_0 + \beta_1 \mathrm{After}_t + \boldsymbol{\beta_2} \sum_{n=1}^{3} (t-c)^n + \boldsymbol{\beta_3} \mathrm{After}_t \times \sum_{n=1}^{3} (t-c)^n + \boldsymbol{\beta_4} W_t + \theta_i + \phi_d + \varepsilon_{it}, \quad (1)$$

where Y_{it} is the dependent variable, i.e. mean daily air pollution or daily respiratory illness hospital visits per 10,000 population on date t; index i indicates either a pollution station (air pollution data) or a patient residence ZIP code (health data). The sample area includes air pollution monitoring stations or patient residence ZIP codes within 6 miles of the Bay Bridge. $After_t$ is a dummy variable for the congestion pricing policy: it equals 0 before July 1, 2010, and 1 afterwards. The cut-off date of July 1, 2010, is represented by c. We use a 60-day bandwidth, meaning the sample includes daily observations 60 days before and 60 days after the road toll implementation: $(t-c)$ is negative before the policy went into effect, and it becomes positive after July 1, 2010. The sample excludes weekends and holidays, so the 60-day bandwidth should provide just enough observations in the selected neighborhood around the cut-off point; we test the sensitivity of our results to the choice of bandwidth in Appendix A. Following relevant studies by Yang et al. (2018) and Zhang et al. (2020), we include a third-order polynomial in the model, $\sum_{n=1}^{3}(t-c)^n$; higher-order polynomials give less weight to samples near the breakpoint (Gelman and Imbens, 2019). This functional form takes into account how air pollution and hospital admission evolve over time. Since weather conditions are important in explaining ambient air pollution concentrations and respiratory health incidence patterns[9], the model includes the vector of weather controls, W_t. The weather covariates are quadratic in precipitation, minimum temperature, maximum temperature, and wind speed, as well as wind direction. Finally, we add pollution site or ZIP code fixed effects, θ_i, as well as day-of-week fixed effects, ϕ_d, in order to adjust for geographic and temporal variation in the outcomes, respectively. Standard errors, ε_{it}, are clustered by date. The coefficient of interest, β_1, tells us the effect of congestion pricing on daily air pollution or hospital admissions, i.e. it shows the local average treatment effect of the policy at the break point.

Leveraging hourly ridership data, we use a different regression discontinuity specification for public transit. We stay in line with Foreman (2016) who uses a non-parametric regression discontinuity model to assess the policy effect on traffic volume. Having hourly ridership as the dependent variable enables us to estimate the treatment effect heterogeneity for peak and off-peak hours[10]. The model for public transit is as follows:

$$\mathrm{R}_{th} = \gamma_0 + \gamma_1 \mathrm{After}_t + \gamma_2 (t-c) + \gamma_3 \mathrm{After}_t \times (t-c) + \phi_d + \omega_h + \varepsilon_{th}, \quad (2)$$

[9]In addition, the weather covariates also serve as a proxy for air pollution that aggravates symptoms of respiratory illness.

[10]Ambient $PM_{2.5}$ pollution and respiratory illness are related, and since we only have daily health data, Specification (1) estimates the policy effect on air pollution at the daily level as well. In Appendix A, we provide the results for hourly air pollution.

R_{th} is the total number of public transit riders in hour h on date t. Since there is no toll for vehicles crossing eastward onto the East Bay, the dependent variable measures westbound ridership only, as in Foreman (2016). The model excludes the weather controls and adds ω_h, the hour-of-the-day fixed effect. The rest of the variables are the same as in Specification (1)[11]. The coefficient of interest, γ_1, measures the discontinuous change in public transit ridership at the threshold.

The identifying assumption of the regression discontinuity study design is that in the absence of congestion pricing, the change in our variables of interest (air pollution, hospital admission, and public transportation) would have been smooth. This assumption is reasonable as long as the running variable is not manipulated and all other factors influencing the variables of interest change continuously in the vicinity of the toll start date, thereby making any discontinuous changes in the outcomes at the time of the policy implementation attributable to the sudden change in the congestion pricing scheme. Below we provide supportive evidence for the validity of the research design, checking that the running variable is not subject to manipulation and there is not jump in each of the control variables at the breaking point.

First, as discussed in Hausman and Rapson (2018), the McCrary (2008) 'manipulation test' that checks whether the density function of the running variable is continuous at the cut-off becomes irrelevant when the distribution of the running variable is uniform, which is the case with time. Therefore, we can evaluate the presence of the sorting effects only indirectly. We are not familiar with any evidence on sorting into or out of treatment with respect to the new congestion pricing policy implementation around the July 1, 2010, threshold. Yet, if any of the sorting effects are present, our results should be interpreted as a combination of the causal treatment effect of interest and any unobserved anticipation, avoidance, and other types of effects (Hausman and Rapson, 2018).

Second, if the research design is valid, we should not observe any discontinuities in the covariates as the running variable crosses the threshold. To test for such discontinuities, we replace the dependent variable in Specification (1) with each of the weather controls and run the model[12]. Although Specification (1) includes the same set of weather controls for the air pollution and health data, the weather observations are assigned to either the pollution stations or the ZCTAs in the sample area, making the estimation results slightly differ for the two datasets. Tables 1 and 2 report the regression results of the covariates in Specification (1). The coefficients of the control variables are statistically insignificant, meaning that there are no discontinuities in the covariates near the cut-off.

[11]Clustering standard errors by pre- and post-policy dates separately or using two-way clustering in Specification (1), i.e. clustering by date and pollution monitoring station or ZIP code, do not change the estimation results presented in Section 5.

[12]We do not perform the test for the public transit data since we do not have weather covariates in Specification (2).

Table 1: Discontinuities in Covariates: Air Pollution Data

VARIABLES	(1) Precip	(2) T°_{min}	(3) T°_{max}	(4) $Wind_{dir}$	(5) $Wind_{speed}$
After	0.234	0.867	-1.866	-10.64	0.869
	(0.744)	(0.702)	(1.742)	(17.63)	(0.835)
Constant	-2.646	5.670*	38.91***	161.6*	-1.866
	(6.055)	(2.966)	(6.162)	(82.75)	(4.784)
Observations	235	235	235	235	235
R-squared	0.408	0.730	0.745	0.368	0.432
Days	60	60	60	60	60
Site FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES

Notes: Each column of the table reports the results of estimating Specification (1) for the pollution dataset with one of the weather covariates as the dependent variable. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Discontinuities in Covariates: Health Data

VARIABLES	(1) Precip	(2) T°_{min}	(3) T°_{max}	(4) $Wind_{dir}$	(5) $Wind_{speed}$
After	0.0572	0.987	-2.256	-10.79	0.751
	(0.709)	(0.686)	(1.730)	(16.87)	(0.766)
Constant	1.495	5.887**	28.32***	142.8**	1.558
	(4.408)	(2.715)	(5.764)	(60.93)	(3.491)
Observations	3,854	3,854	3,854	3,854	3,854
R-squared	0.384	0.768	0.763	0.357	0.402
Days	60	60	60	60	60
ZIP Code FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES

Notes: Each column of the table reports the results of estimating Specification (1) for the health dataset with one of the weather covariates as the dependent variable. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

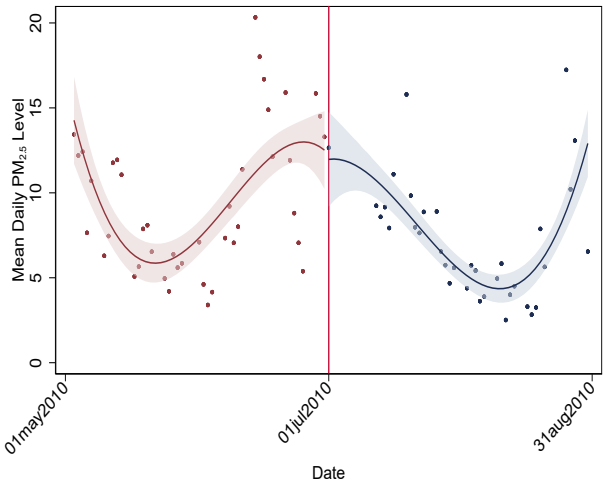
5 Results

In this section, we provide the results of estimating the regression discontinuity models for air pollution, respiratory health, and public transportation ridership[13].

5.1 Air Pollution

Figure 7 shows mean daily $PM_{2.5}$ concentrations during the sample period. There is no change in the air pollution levels after the congestion pricing policy.

Figure 7: Daily Air Pollution: Graphical Evidence



Notes: The figure shows mean daily $PM_{2.5}$ pollution concentrations; the curve is obtained using third-degree polynomial smoothing.

In Table 3, we present the results for the air pollution regression, Specification (1). Consistent with the graphical evidence presented in the figure above, average daily pollution did not change following the introduction of the new road toll.

[13]Before proceeding with the estimation, we perform a Fisher-type test for unit roots in our panel datasets of air pollution, public health, as well as BART transit (see Appendix Table A4). The test rejects the null hypothesis that all panels contain a unit root, confirming the stationary of the panel data. For the public transit data, we perform a Dickey–Fuller test which also confirms the stationary of the time series.

Table 3: Daily Air Pollution: Regression Results

VARIABLES	(1) $PM_{2.5}$	(2) $PM_{2.5}$	(3) $PM_{2.5}$
After	-0.411	0.222	1.197
	(2.193)	(2.409)	(2.089)
Constant	12.45***	12.25***	14.94
	(1.947)	(1.946)	(14.20)
Observations	220	220	220
R-squared	0.425	0.433	0.565
Days	60	60	60
Site FE	YES	YES	YES
Time FE	NO	YES	YES
Weather Controls	NO	NO	YES

Notes: The table contains the results of estimating Specification (1) with mean daily $PM_{2.5}$ air pollution as the dependent variable and a 60-day bandwidth. The three columns present different combinations of fixed effects and control variables: in Column (1), we only use pollution site fixed effects; Column (2) adds day-of-week fixed effects; Column (3) includes the weather controls as well. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

As a possible explanation for the results in Table 3, the drop in traffic volume[14] was not large enough to change the pollution levels by any statistically significant amount, which could be associated with the relative size of the toll and its salience. As mentioned in Foreman (2016), the toll is just a small part of the total driving cost across the bridge and collecting the toll electronically (many drivers pay the toll with their Fast Trak tags) makes the toll less salient. Therefore, although raising the peak-hour toll by 50%, from $4 to $6, may be considered a large increase, the lack of salience may have affected responsiveness to the toll. In particular, while the toll reduces traffic volume, the magnitude of this reduction may not be substantial enough to cause second-order effects associated with air pollution.

In addition, important determinants of local $PM_{2.5}$ are wind direction, which influences where the pollutant blows from and disperses to, and wind speed, which impacts the pollutant concentrations (Hart et al., 2020)[15]. Specifically, wind affects air pollution by changing the way pollution originating from local sources, such as traffic-related pollution, is distributed over the given area or by transporting pollution produced externally into the area (Deryugina et al., 2019). In summer months, the Bay Area typically sees winds blowing from the west, i.e. from the ocean inland, while Deryugina et al. (2019) show that the region is characterized by the lowest $PM_{2.5}$ levels when the wind is blowing from the west and the north. Moreover, the Bay Area is relatively windy, with the strongest winds in summer: during the sample period, the average wind speed was 11 mph (with wind gusts of around 30 mph), and wind speeds of 9 mph or more favor clearing the air from pollutants (The Economic Times, 2019). Therefore, any relatively small changes in air pollution can be quickly dispersed by wind, and

[14] Foreman (2016) finds that the road toll caused a maximum drop of 7% in the hourly number of vehicles crossing the bridge in peak shoulder hours.

[15] $PM_{2.5}$ is removed by precipitation, but precipitation is minimal in the Bay Area in summer.

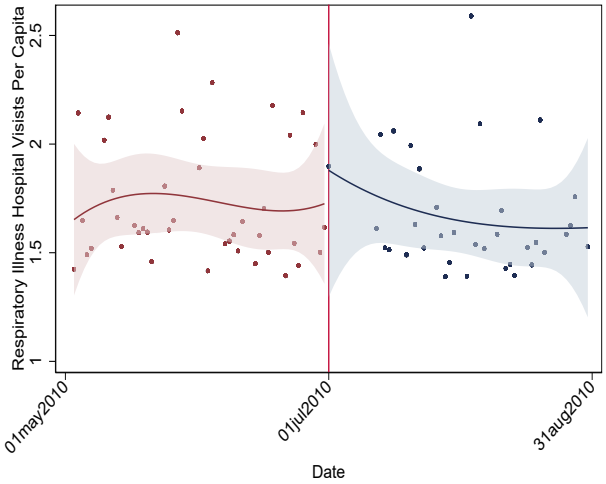
the effect on traffic and pollution should be large enough for the estimates from Specification (1) to be statistically significantly different from zero.

The results are robust to a number of changes in the dependent variable and the model. First, we use hourly pollution as the dependent variable in Specification (1). A possible reason for not observing any statistically significant results in Table 3 could be a relatively small sample size. The hourly pollution dataset contains almost 20 times more observations than our sample with daily pollution. The estimation results reported in Tables A1 and A2 show that there is no effect of the road toll on hourly $PM_{2.5}$ concentrations either, meaning that the absence of the policy effect with respect to daily air pollution is likely not a sample size issue. In addition, in Table A5, we test a sample area located far away from the Bay Bridge (placebo effect), as well as experiment with different polynomial degrees and bandwidths in Specification (1) (see Tables A8 and A9, respectively)[16].

5.2 Respiratory Health

In Figure 8, we plot patient counts for respiratory diseases per ZIP code per day during the sample period. We do not observe a change in respiratory health illnesses after the date when the toll was implemented.

Figure 8: Daily Respiratory Illness Hospital Visits: Graphical Evidence



Notes: The figure shows mean daily respiratory illness hospital visits per capita per ZIP code; the curve is obtained using third-degree polynomial smoothing.

The results of estimating Specification (1) are presented in Table 4: the policy has no effect on respiratory health.

[16]Moreover, there is no effect on other criteria air pollutants associated with road traffic, such as carbon monoxide and nitrogen oxides, and the results do not change when using the logarithm of the dependent variable or maximum daily $PM_{2.5}$ pollution instead of the mean daily pollution on the left-hand side of Specification (1). The results are also robust to using July 1, 2009, or July 1, 2011, as the false date of the policy change (placebo effect). The estimation results are available upon request.

Table 4: Daily Respiratory Illness Hospital Visits: Regression Results

VARIABLES	(1) Hosp. Visits	(2) Hosp. Visits	(3) Hosp. Visits
After	0.148	0.142	-0.0150
	(0.171)	(0.194)	(0.189)
Constant	1.730***	1.740***	3.505***
	(0.148)	(0.172)	(0.896)
Observations	3,619	3,619	3,619
R-squared	0.199	0.200	0.202
Days	60	60	60
ZIP Code FE	YES	YES	YES
Time FE	NO	YES	YES
Weather Controls	NO	NO	YES

Notes: The table contains the results of estimating Specification (1) with daily hospital visits per 10,000 population per ZIP code as the dependent variable and a 60-day bandwidth. The three columns present different combinations of fixed effects and control variables: in Column (1), we only use ZIP code fixed effects; Column (2) adds day-of-week fixed effects; Column (3) includes the weather controls as well. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Previously, we have shown that there is no evidence of $PM_{2.5}$ levels changing after the congestion pricing is introduced. Since exposure to $PM_{2.5}$ air pollution is directly linked to human health, we do not see any changes in hospital visits caused by the new policy. In addition, the health results serve as a double-check for our pollution estimates. As shown above, we do not observe a drop in $PM_{2.5}$ concentrations, but we must admit that the pollution levels can be imperfectly measured: the wind patterns described above and, possibly, the size of the pollution effect itself may require us to have more pollution stations to capture it. If there was in fact a pollution effect, for some reason not captured by our model, we would see a statistically significant change in hospital visits that are measured more precisely than the ambient $PM_{2.5}$ level. However, the results of estimating the pollution and health models are consistent, meaning our pollution estimates are likely not driven by any measurement error.

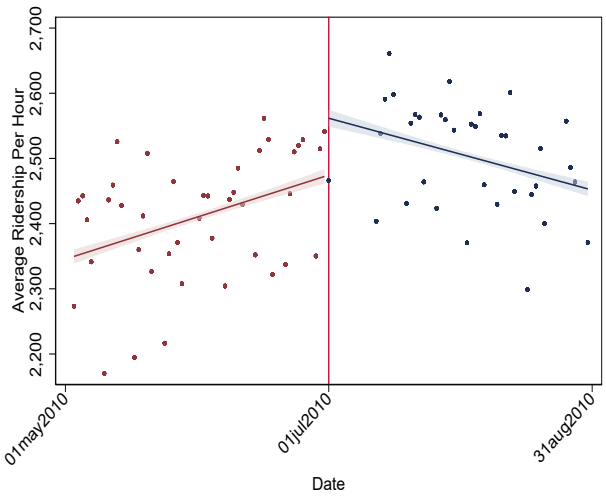
In Appendix A we present the following checks demonstrating the robustness of the health effect: we estimate the effect of congestion pricing on hospital admission for broken bones as a placebo condition in Table A6; different polynomial degrees and bandwidths are in Tables A8 and A9, respectively[17].

[17] Chen et al. (2018) study the effect of air quality alerts on health, including hospital admissions or emergency department visits, using a non-parametric regression discontinuity model. The effect of the road toll policy on hospital admission estimated with a non-parametric model is not statistically significantly different from zero, similar to the results shown in Table 4. In addition, we estimate Specification (1) for a larger sample area to see if the number of observations influences the results. We use a sample area consisting of ZCTAs located 10 miles around the Bay Bridge; the sample contains around 5,000 observations. The estimation results are the same as those reported in Table 4. However, we should be cautious about using

5.3 Public Transit

Figures 9-11 plot westbound public transit ridership for all hours, as well as peak and off-peak hours separately. The figures demonstrate a sharp increase in public transit ridership after the new road toll policy went into effect. Interestingly, in Figure 10, an immediate jump in peak transit ridership is followed by a drop back to the pre-policy ridership levels. This is consistent with Foreman (2016) who employs regression discontinuity design and finds that peak traffic volume drops immediately after July 1, 2010, but then it quickly recovers to the level that is below the pre-policy one but above the July 2010 number of vehicles.

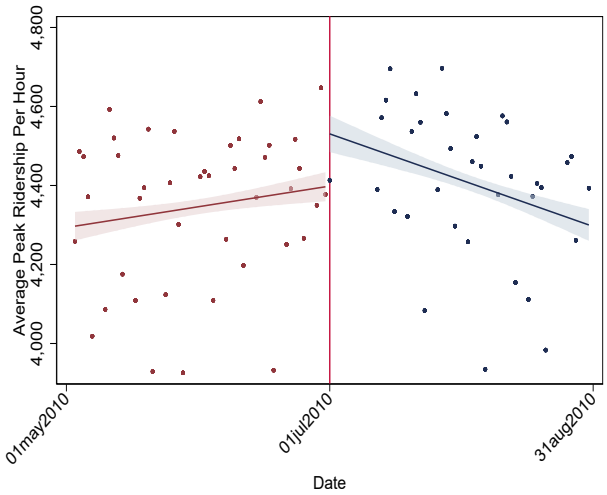
Figure 9: Graphical Evidence: Westbound Public Transit, All Hours



Notes: The figure shows the scatter plot of mean hourly westbound public transit ridership and date, along with the line corresponding to the prediction for the ridership from a linear regression of the ridership on date.

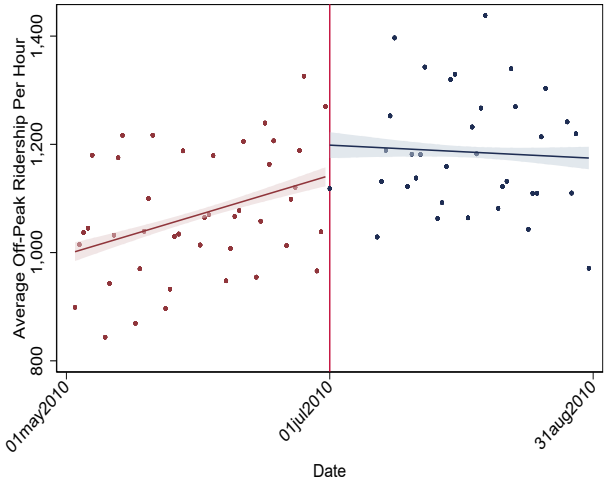
larger sample areas since there could be more factors influencing respiratory health hospital admissions than the congestion pricing alone. Finally, the results do not change when using the logarithm of the dependent variable or heart disease hospital admission per capita as the dependent variable in Specification (1). The results are also robust to using July 1, 2009, or July 1, 2011, as the false date of the policy change (placebo effect). The estimation results are available upon request.

Figure 10: Graphical Evidence: Westbound Public Transit, Peak Hours



Notes: The figure shows the scatter plot of mean hourly westbound public transit ridership during peak hours and date, along with the line corresponding to the prediction for the ridership from a linear regression of the ridership on date.

Figure 11: Graphical Evidence: Westbound Public Transit, Off-Peak Hours



Notes: The figure shows the scatter plot of mean hourly westbound public transit ridership during off-peak hours and date, along with the line corresponding to the prediction for the ridership from a linear regression of the ridership on date.

In Table 5, we present the results of estimating Specification (2). We see an increase in public ridership, with its magnitude larger for peak hours.

Table 5: Hourly Westbound Public Transit Regression Results

VARIABLES	(1) Riders: All Hours	(2) Riders: All Hours	(3) Riders: Peak Hours	(4) Riders: Off-Peak Hours
After	87.19**	83.11***	106.3**	67.03**
	(41.50)	(30.78)	(52.82)	(32.16)
Constant	2,474***	2,472***	4,390***	1,144***
	(24.77)	(13.32)	(29.80)	(16.96)
Observations	1,694	1,694	693	1,001
R-squared	0.000	0.983	0.979	0.966
Days	60	60	60	60
Time FE	NO	YES	YES	YES

Notes: The table contains the results of estimating Specification (2) for a 60-day bandwidth, where the dependent variable is hourly westbound public transit ridership, which is the number of passengers across all BART lines and AC Transit routes traveling westward onto the San Francisco Peninsula during a certain hour on a certain day. Column (1) does not include time fixed effects (day of the week and hour of the day); in Column (2), the time fixed effects are added in Specification (2). The model used to obtain results shown in Columns (3) and (4) is identical to that in Column (2), but the dependent variable is hourly westbound public transit ridership for peak hours in Column (3) or off-peak-hours in Column (4). Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Some of the reduction in the number of vehicles crossing the bridge could result from shifting to public transit. Substitution towards public transit is likely the case for carpoolers. Foreman (2016) finds that carpool traffic on the bridge decreased by 300-500 vehicles per hour during peak times. Carpoolers are more likely to be former public transit users rather than solo drivers and they also may well take transit as a backup for their return trip (Deakin et al., 2012; Public Transport Users Association, 2016). Therefore, it makes sense for some carpoolers to switch to public transit once carpooling is no longer an option, which is why we see an increase in transit ridership after the new policy was introduced. However, the jump in public ridership is quite small compared to a drop in carpool traffic: the latter leads to a minimum of 900 people off the road (carpools are three or more people on the Bay Bridge, so 300 is multiplied by 3), while the former is 100 people maximum (see Column (3) of Table 5). While it is "induced" carpools, i.e. those created to save commuting time, whose driver and passengers are more likely to switch to public transit, they constitute the minority of the carpool traffic since most of the carpools are "natural" carpools, such as a parent with children, that would have occurred anyway (Forbes, 2019).

As shown in Appendix A, the estimated policy effect on transit ridership is robust to the following: daily public transit (Table A3) or hourly BART transit per station pair (Table A4) as the dependent variable; using July 1, 2009, as the false date of the policy change (placebo) in Table A7; different bandwidths in Table A10[18].

[18]Moreover, one could assume that weather patterns may influence public transit ridership. Adding weather controls (temperature, precipitation, wind speed, wind direction - all four variables or just precipitation) to Specification (2) does not change the transit effect. Next, estimating the public transit model for AC Transit ridership only (as a time series or a panel, with the latter consisting of hourly AC Transit ridership per route) yields no effect of the congestion pricing policy on the number of bus passengers. The results do not change when using the logarithm of the dependent variable in Specification (2). The results are also robust to exploring July 1, 2011, as the false date of the policy change (placebo effect).

Following recommendations on evaluating a regression discontinuity in time discussed in Hausman and Rapson (2018), we conduct some additional robustness checks for the three datasets. Specifically, we test an augmented local linear model, a longer time frame, such as up to 2 years before and after the toll (with the month-of-year fixed effects included); we add a lag of the dependent variable to the right-hand side of the model, estimate the model with Newey-West standard errors, conduct a "donut" regression discontinuity dropping a week before July 4 as well, or do not drop any days around the new toll implementation. For the public transit analysis, depending on the robustness check, the coefficients on ridership (across all hours or for some sets of hours) may become statistically insignificant. However, this is still consistent with the main results since the jump in public ridership (Table 5) is relatively small in magnitude. The estimation results are available upon request.

6 Conclusion

We use regression discontinuity design to estimate the causal relationship between the congestion pricing policy introduced on the Bay Bridge on July 1, 2010, and air pollution, respiratory health, as well as public transportation ridership. Congestion pricing provides a quasi-experimental setting where the increase in the cost of transportation is plausibly exogenous to other factors that determine the three outcomes of interest.

The paper finds that the new road toll policy, which led to a drop in rush hour traffic volume on the Bay Bridge (Foreman, 2016), was associated with a moderate increase in public transit ridership (mainly, metro rail ridership), but, contrary to much of the existing literature on the topic, it did not affect local air pollution levels and respiratory health of people residing in the Bay Bridge vicinity. One reason for this result could be associated with the relatively small size of the toll and its lack of salience. While the toll induces reductions in traffic volume, the magnitude of this decline may not be large enough to cause second-order or third-order effects associated with air pollution and respiratory health. Another explanation for our findings could be that the effect of congestion pricing on pollution and health likely depends on the region under study. The Bay Area has certain distinct climate characteristics, such as wind patterns or proximity to the Pacific Ocean, which may mute the effects of the toll on air pollution.

From a public policy perspective, our study underscores that the social welfare effects of congestion pricing are driven by heterogeneous, context-specific factors. As a result, pricing mechanisms to attain environmental goals should be tailored to local conditions rather than applied as a one-size-fits-all approach. A further insight from our work is that road toll implementation requires robust evaluation and timely adjustment to new information to ensure positive impacts on air quality, public health, and, more generally, the attainment of public policy goals.

References

AC Transit, 2023. https://www.actransit.org.

Anderson, Michael L., "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion," *The American Economic Review*, 2014, *104* (9), 2763–2796.

_ , "As The Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," *Journal of the European Economic Association*, 2020, *18* (14), 1886–1927.

Auffhammer, Maximilian and Ryan Kellogg, "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality," *The American Economic Review*, 2011, pp. 2687–2722.

Bala, Gabriel-Petrica, Ruxandra-Mioara Rajnoveanu, Emanuela Tudorache, Radu Motisan, and Cristian Oance, "Air Pollution Exposure — the (In)Visible Risk Factor for Respiratory Diseases," *Environmental Science and Pollution Research*, 2021, *28*, 19615–19628.

Bay Area Air Quality Management Disctrict, "Spair the Air," 2023. https://www.baaqmd.gov/about-air-quality/spare-the-air.

Bay Area Rapid Transit, 2023. https://www.bart.gov.

Beaudoin, Justin, Y. Hossein Farzin, and C.-Y. Cynthia Lin Lawell, "Public Transit Investment And Sustainable Transportation: A Review Of Studies Of Transit's Impact On Traffic Congestion And Air Quality," *Research in Transportation Economics*, 2015, *52*, 15–22.

California Air Resources Board, "Air Quality and Meteorological Information System," 2023. https://www.arb.ca.gov/aqmis2/aqmis2.php.

California Department of Health Care Access and Information, 2023. https://hcai.ca.gov.

Chen, Hong, Qiongsi Li, Jay S. Kaufman, Jun Wang, Ray Copes, Yushan Su, and Tarik Benmarhnia, "Effect of Air Quality Alerts on Human Health: a Regression Discontinuity Analysis in Toronto, Canada," *The Lancet: Planetary Health*, 2018, *2* (1), E19–E26.

Chen, Yihsu and Alexander Whalley, "Green Infrastructure: The Effects of Urban Rail Transit on Air Quality," *American Economic Journal: Economic Policy*, 2012, *4* (1), 58–97.

Cohen, Aaron J, Michael Brauer, Richard Burnett, H Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan, Bert Brunekreef, Lalit Dandona, Rakhi Dandona, Valery Feigin, Greg Freedman, Bryan Hubbell, Amelia Jobling, Haidong Kan, Luke Knibbs, Yang Liu, Randall Martin, Lidia Morawska, C Arden Pope III, Hwashin Shin, Kurt Straif, Gavin Shaddick, Matthew Thomas, Rita van Dingenen, Aaron van Donkelaar, Theo Vos, Christopher J L Murray, and Mohammad H Forouzanfar, "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015," *Lancet*, 2017, *389* (10082), 1907–1918.

Currie, Janet and Reed Walker, "Traffic Congestion and Infant Health: Evidence from E-ZPass," *American Economic Journal: Applied Economics*, 2011, *3* (1), 65–90.

Davis, Lucas W., "The Effect of Driving Restrictions on Air Quality in Mexico City,"

Journal of Political Economy, 2008, *116* (1), 38–81.

Deakin, Elizabeth, Karen Trapenberg Frick, Robert Cervero, Alexander Skabardonis, Ian Barnes, Karla Kingsley, James Rubin, Jin Murakami, Javier Amaro, and Erik Jensen, "Bay Bridge Toll Evaluation: Final Report," May 2012. University of California Berkeley.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif, "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," *American Economic Review*, 2019, *109* (12), 4178–4219.

Forbes, "Carpool Cheats May Be Helping Traffic; How HOV Lanes Can Fail," 2019. https://www.forbes.com/sites/bradtempleton/2019/08/06/carpool-cheats-may-be-helping-traffic-how-hov-lanes-can-fail/?sh=39533d763d5d.

Foreman, Kate, "Crossing the Bridge: The effects of Time-Varying Tolls on Curbing Congestion," *Transportation Research Part A*, 2016, pp. 76–94.

Friedman, Michael S., Kenneth E. Powell, Lori Hutwagner, LeRoy M. Graham, and W. Gerald Teague, "Impact of Changes in Transportation and Commuting Behaviors During the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma," *The Journal of the American Medical Association*, 2001, *285* (7), 897–905.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business Economic Statistics*, 2019, *37* (3), 447–456.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A. Turner, "Subways and Urban Air Pollution," *American Economic Journal: Applied Economics*, 2022, *14* (1), 164–196.

Gibson, Matthew and Maria Carnovale, "The Effects of Road Pricing on Driver Behavior and Air Pollution," *Journal of Urban Economics*, 2015, *89*, 62–73.

Gilmore, E.A., N.Z. Heo J.and Muller, C.W. Tessum, J.D. Hill, J.D. Marshall, and P.J. Adams, "An inter-comparison of the social costs of air quality from reduced-complexity models," *Environmental Research Letters*, 2019, *14* (7), 074016.

Hart, Ronan, Lu Liang, and Pinliang Dong, "Monitoring, Mapping, and Modeling Spatial–Temporal Patterns of $PM_{2.5}$ for Improved Understanding of Air Pollution Dynamics Using Portable Sensing Technologies," *International Journal of Environmental Research and Public Health*, 2020, *17* (14).

Hausman, Catherine and David S. Rapson, "Regression Discontinuity in Time: Considerations for Empirical Applications," *Annual Review of Resource Economics*, 2018, *10*, 533–552.

Health Canada, "Canadian Health Science Assessment for Fine Particulate Matter (PM2.5)," 2022. https://publications.gc.ca/site/eng/9.907208/publication.html.

_ , "Health impacts of air pollution in Canada: estimates of morbidity and premature mortality outcomes – 2021 Report," 2022. https://www.canada.ca/en/health-canada/services/publications/healthy-living/2021-health-effects-indoor-air-pollution.htmla8.3.

Isaksen, Elisabeth and Bjørn Johansen, "Congestion Pricing, Air Pollution, and Individual-Level Behavioural Responses," *The Journal of Human Resources*, June 2021. Centre for Climate Change Economics and Policy Working Paper No. 390; Grantham Research Institute on Climate Change and the Environment Working Paper No. 362.

Isphording, Ingo E. and Nico Pestel, "Pandemic meets pollution: Poor air quality increases deaths by COVID-19," *Journal of Environmental Economics and Management*, 2021, *108.*

Jans, Jenny, Per Johansson, and J. Peter Nilsson, "Economic status, air quality, and child health: Evidence from inversion episodes," *Journal of Health Economics*, 2018, *61*, 220–232.

Manisalidis, Ioannis, Elisavet Stavropoulou, Agathangelos Stavropoulos, and Eugenia Bezirtzoglou, "Environmental and Health Impact of Air Pollution: A Review," *Frontiers in Public Health*, 2020, *8.*

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

MIT CLimate Portal, "Public Transportation," 2023. https://climate.mit.edu/explainers/public-transportation.

National Oceanic and Atmospheric Administration, "Air Quality and Meteorological Information System," 2023. https://tidesandcurrents.noaa.gov/stationhome.html?id=9414750info.

_ , "National Weather Service," 2023. https://www.weather.gov.

OnTheWorldMap, 2023. https://ontheworldmap.com.

Pestel, Nico and Florian Wozny, "Low Emission Zones for Better Health: Evidence from German Hospitals," *IZA Institute of Labor Economics: Discussion Paper Series*, August 2019. IZA DP No. 12545.

PRISM Climate Group, "PRISM Gridded Climate Data," 2023. http://prism.oregonstate.edu.

Public Transport Users Association, "Myth: It's Just as Effective to Promote Carpooling," 2016. https://www.ptua.org.au/myths/carpool.

Ransom, M. R. and C.A. Pope, "Elementary school absences and PM10 pollution in Utah Valley," *Environmental Research*, 1992, *58*, 204–219.

Rivers, Nicholas, Soodeh Saberian, and Brandon Schaufele, "Public Transit and Air Pollution: Evidence from Canadian Transit Strikes," *Canadian Journal of Economics*, 2020, *53* (2), 496–525.

Schlenker, Wolfram and W. Reed Walker, "Airports, Air Pollution, and Contemporaneous Health," *The Review of Economic Studies*, 2016, *83* (2), 768–809.

Simeonova, Emilia, Janet Currie, Peter Nilsson, and Reed Walker, "Congestion Pricing, Air Pollution, and Children's Health," *The Journal of Human Resources*, 2021, *56* (4), 971–996.

Sun, Cong, Siqi Zheng, and Rui Wang, "Restricting Driving for Better Traffic and Clearer Skies: Did It Work in Beijing?," *Transport Policy*, 2014, *32*, 34–41.

The Economic Times, "How Weather Determines Air Quality," 2019. https://economictimes.indiatimes.com/news/politics-and-nation/how-weather-determines-air-quality/wind-speed/slideshow/72159795.cms?from=mdr.

The Institute For Health Metrics And Evaluation, "Global Burden of Disease," 2019. https://www.healthdata.org/research-analysis/gbd.

The U.S. Census Bureau, "Metropolitan and Micropolitan Statistical Areas Population Totals: 2020-2022," 2022. https://www.census.gov/data/datasets/time-series/demo/popest/2020s-total-metro-and-micro-statistical-areas.html.

_ , "ZIP Code Tabulation Areas (ZCTAs)," 2023. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html.

Wikipedia, "San Francisco–Oakland Bay Bridge," 2023. https://en.wikipedia.org/wiki/San$_{F}rancisco$

Yang, Jun, Shuai Chen, Ping Qin, Fangwen Lu, and Antung A. Liu, "The Affect of Subway Expansions on Vehicle Congestion: Evidence from Beijing," *Journal of Environmental Economics and Management*, 2018, *88*, 114–133.

Zhang, Ming, Cheng Shan, Wenwen Wang, Jingru Pang, and Shu Guo, "Do Driving Restrictions Improve Air Quality: Take Beijing-Tianjin for Example?," *Science of the Total Environment*, 2020, *712*.

A Robustness Checks

Table A1: Hourly Air Pollution

VARIABLES	(1) $PM_{2.5}$	(2) $PM_{2.5}$	(3) $PM_{2.5}$
After	-0.420	0.201	0.433
	(2.135)	(2.333)	(2.321)
Constant	12.46***	12.27***	54.48**
	(1.895)	(1.881)	(20.95)
Observations	5,233	5,233	5,138
R-squared	0.273	0.300	0.318
Days	60	60	60
Site FE	YES	YES	YES
Time FE	NO	YES	YES
Weather Controls	NO	NO	YES

Notes: The table contains the results of estimating Specification (1) with hourly $PM_{2.5}$ air pollution as the dependent variable. The three columns present different combinations of fixed effects and control variables: in Column (1), we only use pollution site fixed effects; Column (2) adds day-of-week and hour-of-the-day fixed effects; Column (3) includes weather controls as well. The weather covariates are the same as those used when estimating Specification (1) for the daily pollution data, except that we do not include minimum and maximum hourly ambient temperature - we use mean hourly temperature instead. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Hourly Air Pollution: Different Sets of Hours

VARIABLES	(1) $PM_{2.5}$: Peak	(2) $PM_{2.5}$: Off-Peak	(3) $PM_{2.5}$: $Peak_{Non-Sh}$	(4) $PM_{2.5}$: $Peak_{Sh}$	(5) $PM_{2.5}$: $Off\text{-}Peak_{Sh}$
After	1.034	0.0648	0.989	1.171	1.179
	(2.237)	(2.393)	(2.193)	(2.490)	(2.622)
Constant	40.74	62.25***	35.56	46.83*	66.53***
	(27.24)	(19.47)	(32.17)	(26.19)	(19.38)
Observations	1,919	3,219	1,062	857	864
R-squared	0.323	0.322	0.356	0.301	0.349
Days	60	60	60	60	60
Site FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES

Notes: The table contains the results of estimating Specification (1) with hourly $PM_{2.5}$ air pollution as the dependent variable. Each column corresponds to hourly $PM_{2.5}$ concentrations measured during specific hours. Column (1) reports the estimation results for hourly $PM_{2.5}$ pollution in peak hours (from 5 am to 10 am and from 3 pm to 7 pm); Column (2) uses hourly $PM_{2.5}$ pollution in off-peak hours (the hours except for the peak ones); Column (3) shows the results for peak non-shoulder hours (from 6 am to 9 am and from 4 pm to 6 pm); Column (4) includes $PM_{2.5}$ for peak shoulder hours (from 5 am to 6 am, from 9 am to 10 am, from 3 pm to 4 pm, as well as from 6 pm to 7 pm); Column (5) shows $PM_{2.5}$ for off-peak shoulder hours (from 4 am to 5 am, from 10 am to 11 am, from 2 pm to 3 pm, as well as from 7 pm to 8 pm). The fixed effects and weather controls are the same as reported in the notes under Table A1. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Daily Westbound Public Transit

VARIABLES	(1) Riders Per Day	(2) Riders Per Day
After	1,918**	1,829**
	(930.8)	(704.7)
Constant	54,438***	54,385***
	(555.6)	(304.9)
Observations	77	77
R-squared	0.322	0.778
Days	60	60
Time Controls	NO	YES

Notes: The table contains the results of estimating Specification (2), where the dependent variable is daily westbound public transit ridership, which is the daily count of passengers across all BART lines and AC Transit routes traveling westward onto the San Francisco Peninsula. Column (1) does not include time fixed effects (day of the week); in Column (2), the time fixed effects are added in Specification (2). Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Hourly Station-to-Station Westbound BART Transit

VARIABLES	(1) Riders: All Hours	(2) Riders: All Hours	(3) Riders: Peak Hours	(4) Riders: Off-Peak Hours
After	0.962**	0.917***	1.182**	0.726**
	(0.464)	(0.342)	(0.553)	(0.364)
Constant	27.06***	27.03***	46.95***	12.77***
	(0.281)	(0.149)	(0.316)	(0.186)
Observations	152,768	152,768	63,756	89,012
R-squared	0.104	0.577	0.607	0.633
Days	60	60	60	60
Station Pair FE	YES	YES	YES	YES
Time FE	NO	YES	YES	YES

The table contains the results of estimating Specification (2), where the dependent variable is westbound BART ridership per hour per station pair, which is the number of passengers per station pair traveling westward onto the San Francisco Peninsula during a certain hour on a certain day. Column (1) includes station-pair fixed effects only; in Column (2), we add time fixed effects (day of the week and hour of the day). The model used to obtain results shown in Columns (3) and (4) is identical to that in Column (2), but the dependent variable is westbound BART ridership per hour per station pair for peak hours in Column (3) or off-peak-hours in Column (4). Standard errors are clustered by date (using two-way clustering, i.e. clustering by date and station pair, does not change the estimation results), shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Placebo Test: Daily Air Pollution in the South Bay

VARIABLES	(1) $PM_{2.5}$	(2) $PM_{2.5}$	(3) $PM_{2.5}$
After	-2.710	-1.983	-0.407
	(2.270)	(2.490)	(1.652)
Constant	10.92***	10.80***	-4.365
	(1.731)	(1.794)	(16.07)
Observations	76	76	76
R-squared	0.266	0.285	0.613
Days	60	60	60
Time FE	NO	YES	YES
Weather Controls	NO	NO	YES

Notes: The table contains the results of estimating Specification (1) with mean daily $PM_{2.5}$ air pollution as the dependent variable. We select a "false" sample area which is 6 miles around Downtown San Jose in the South Bay. In the new sample area, there is one pollution monitoring site located near the Guadalupe River Park. The three columns present different combinations of fixed effects and control variables: in Column (1), we do not use any fixed effects (we omit pollution-site fixed effects from Specification (1) since there is only one pollution monitor in the sample area); Column (2) adds day-of-week fixed effects; Column (3) includes the weather controls as well. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Placebo Test: Daily Hospital Admission for Broken Bones

VARIABLES	(1) Hosp. Visits	(2) Hosp. Visits	(3) Hosp. Visits
After	0.219	0.193	0.337
	(0.356)	(0.357)	(0.301)
Constant	1.802***	1.825***	-0.0551
	(0.301)	(0.285)	(1.019)
Observations	3,619	3,619	3,619
R-squared	0.164	0.165	0.167
Days	60	60	60
ZIP Code FE	YES	YES	YES
Time FE	NO	YES	YES
Weather Controls	NO	NO	YES

Notes: The table contains the results of estimating Specification (1) with daily hospital visits for broken bones per 10,000 population per ZIP code as the dependent variable. The three columns present different combinations of fixed effects and control variables: in Column (1), we only use ZIP code fixed effects; Column (2) adds day-of-week fixed effects; Column (3) includes the weather controls as well. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Placebo Test: Hourly Westbound Public Transit with the Cut-Off Date of July 1, 2009

VARIABLES	(1) Riders: All Hours	(2) Riders: All Hours	(3) Riders: Peak Hours	(4) Riders: Off-Peak Hours
After	-32.57	-31.83	-39.40	-26.59
	(38.44)	(30.09)	(62.17)	(28.59)
Constant	2,450***	2,444***	4,312***	1,151***
	(24.24)	(15.20)	(28.76)	(20.42)
Observations	1,672	1,672	684	988
R-squared	0.000	0.985	0.981	0.968
Days	60	60	60	60
Time FE	NO	YES	YES	YES

Notes: The table contains the results of estimating Specification (2), where the dependent variable is hourly westbound public transit ridership, which is the number of passengers across all BART lines and AC Transit routes traveling westward onto the San Francisco Peninsula during a certain hour on a certain day. We use a "false" cut-off date, which is July 1, 2009. Column (1) does not include time fixed effects (day of the week and hour of the day); in Column (2), the time fixed effects are added in Specification (2); the model used to obtain results shown in Columns (3) and (4) is identical to that in Column (2), but the dependent variable is hourly westbound public transit ridership for peak hours in Column (3) or off-peak-hours in Column (4). Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Different Polynomial Degrees (Daily Air Pollution and Daily Hospital Visits)

Dependent variable:	$PM_{2.5}$ (1)	Hosp. Visits (2)
Panel A: Degree 4		
After	4.137	-0.0264
	(-3.169)	(-0.252)
R-squared	0.587	0.202
Panel B: Degree 5		
After	0.891	-0.00514
	(-1.899)	(-0.245)
R-squared	0.649	0.202
Days	60	60
Time FE	YES	YES
Weather Controls	YES	YES
Number of observations	220	3,619

Notes: The table contains the results of estimating Specification (1), where the dependent variable is mean daily $PM_{2.5}$ air pollution or daily hospital visits per 10,000 population per ZIP code – see Column (1) or (2), respectively. In Panel A, instead of a third-order polynomial, we include a fourth-order polynomial in the model, and Panel B presents the results for a fifth-order polynomial. The model includes pollution site or ZIP code fixed effects, day-of-week fixed effects, as well as the weather controls. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Different Bandwidths (Daily Air Pollution and Daily Hospital Visits)

Dependent variable:	$PM_{2.5}$ (1)	Hosp. Visits (2)
Panel A: 90 Days		
After	0.602	0.107
	(-2.707)	(-0.191)
R-squared	0.392	0.187
Number of observations	320	5,358
Panel B: 120 Days		
After	1.228	0.177
	(-2.043)	(-0.159)
R-squared	0.397	0.174
Number of observations	436	7,285
Time FE	YES	YES
Weather Controls	YES	YES

Notes: The table contains the results of estimating Specification (1), where the dependent variable is mean daily $PM_{2.5}$ air pollution or daily hospital visits per 10,000 population per ZIP code – see Column (1) or (2), respectively. In Panel A, instead of a 60-day bandwidth, we include the 90-day one, and Panel B presents the results for a 120-day bandwidth. The model includes pollution site or ZIP code fixed effects, day-of-week fixed effects, as well as the weather controls. Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Different Bandwidths (Hourly Westbound Public Transit Ridership)

Dependent variable:	Riders: All Hours (1)	Riders: Peak Hours (2)	Riders: Off-Peak Hours (3)
Panel A: 90 Days			
After	58.78***	-63.65	143.5***
	(-22.05)	(-44.92)	(-26.82)
R-squared	0.983	0.979	0.963
Number of observations	2,508	1,026	1,482
Panel B: 120 Days			
After	45.09**	-39.31	103.5***
	(-22.46)	(-43.57)	(-22.49)
R-squared	0.982	0.978	0.955
Number of observations	3,432	1,404	2,028
Time FE	YES	YES	YES

Notes: The table contains the results of estimating Specification (2), where the dependent variable is hourly westbound public transit ridership, which is the number of passengers across all BART lines and AC Transit routes traveling westward onto the San Francisco Peninsula during a certain hour on a certain day. In Panel A, instead of a 60-day bandwidth, we use the 90-day one, and Panel B presents the results for a 120-day bandwidth. All the columns include day-of-week and hour-of-the-day fixed effects. The model used to obtain results shown in Columns (2) and (3) is identical to that in Column (1), but the dependent variable is hourly westbound public transit ridership for peak hours in Column (2) or off-peak-hours in Column (3). Standard errors are clustered by date, shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.