Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
An analysis of the New York City traffic volume, vehicle collisions, and safety under COVID-19

Paolo Cappellari, Bryan S. Weber *
City University of New York, College of Staten Island, 2800 Victory Blvd, Staten Island 10316, NY, USA

1. Introduction

The analysis of traffic congestion and associated collisions has been the interest of research for a number of decades (Cui, Henrickson, Ke, & Wang, 2018; De Fabritiis, Ragona, & Valenti, 2008; Jain, Sharma, & Subramanian, 2012), and continues to be an area of great interest (Faghii-Imani, Anowar, Miller, & Eluru, 2017; Mammen, Shim, & Weber, 2019). Studies focus on a number of aspects such as traffic patterns predictions (Jain et al., 2012), traffic flow speed estimation (Cui et al., 2020; De Fabritiis et al., 2008), relationship between traffic volume and collisions (Noland & Quddus, 2004; Shefer & Rietveld, 1997; Wang, Quddus, & Ison, 2009), and more (Ragheabani, Taryani, Allahviranloo, & Gao, 2020; Liu, Zheng, Chawla, Yuan, & Xing, 2011).

Many of these studies rely on potentially endogenous factors (Cullinane, 2004), such as the introduction of car sharing services (Dills & Mulholland, 2018), taxi ridership (Faghii-Imani et al., 2017), public transit usage (Iyer, Boxer, & Subramanian, 2018), policy making (Abouk & Adams, 2013; Mammen et al., 2019), or economic impact (Parry & Bento, 2002), whose effects are usually visible over a relatively long period of time, making it hard to understand if, and which, other factors may contribute to the changes.

In this paper, we use the data associated with the COVID-19 pandemic as an exogenous factor to analyze the traffic volume and vehicle collisions in the boroughs of New York City. This is not because we are interested in pandemic traffic patterns per se, but because the pandemic provides a substantive adjustment to traffic volume without other associated traffic rules changing, allowing for a consistent estimation of the association between traffic volume and vehicle collisions (Cullinane, 2004). The New York area was hit with the pandemic starting March 2020. The COVID-19 pandemic and the associated health and public policy implications created an unprecedented scenario that gives us the ability to observe the largest shift in vehicle traffic that is due to an exogenous factor. This is novel in a literature where changes to traffic volume are mostly linked with new transportation or economic policies (Green, Heywood, & Navarro, 2016), or the inter-temporal shifts in volume that may be accompanied by other traffic changes (Shefer & Rietveld, 1997; Zhou & Sisiopiku, 1997). In particular, we analyze the relationship between the number of COVID-19 cases, its impact on traffic, and the associated collision between vehicles. The source data for our analysis are data sets from both New York city and state. Specifically, we consider the following distinct data sets: traffic volumes, vehicle collisions, traf-
fic speeds, traffic summons issued, and COVID-19 testing reports. These data sets are combined together to generate a comprehensive and unique view of traffic changes around the time period the exogenous factor is observable. All data sets are publicly available to enable full verifiability and reproducibility of our study.

For this study, we used an instrumental variable (IV) approach, where we infer the traffic volume by the number of COVID-19 cases on the week prior (Cullinane, 2004). We control for standard variables, including time based-fixed effects, the number of traffic summons, as well as for several possible systemic breaks at the time of the first COVID-19 case and the stay-at-home order. We also consider additional robustness checks such as changing the data window, additional leads of potentially obscuring events, and variations on the functional form of the instrument.

Among the findings, we observed that the reduction in traffic volume is associated with a significant reduction in collisions, at a rate of approximately 1.7% fewer collisions for each 1% reduction in traffic volume. Conversely, we measure an increase in fatalities and injuries as traffic volume decreases, suggesting higher volumes of traffic has a dampening effect on collisions. We also find the suggestion that the casualty/collision rate has increased in the lower volume traffic. Corroborating on this finding, we point out that speed is generally higher in lower-volume traffic, and significantly increased speeds in each borough are associated with declines in traffic volume. The simple elasticity of speed with respect to volume is approximately equal to –0.13, suggesting that the policies still may be double-edged. Lastly, we estimate that the estimated monetary value in collisions, injuries, and casualties from the COVID-19 traffic declines across all four boroughs (excluding Staten Island).

Despite the reduction in collisions, we find an increase in social costs during the period. We calculate a reduction of $453,000 in property costs (excluding casualties) and the counter-veiling value of increased fatalities and injuries during this period are approximately $2.6 million. This highlights the complex relationship between traffic volume, speed, and safety, but indicates that the overriding effect of substantive volume changes does not guarantee improved safety.

The paper is organized as follows: in Section 2 we discuss the state of the art in this research area; in Section 3 we present the data sets used in this study; in Section 4 we illustrate the methodology to analyze the data; in Section 5 we discuss the findings and observations; finally, in Section 6 we draw our conclusions.

2. Literature review

Many works focus on the relationship between traffic, safety, and economics. In Shefer and Rietveld (1997), authors highlight the complexity of the relationship between congestion and collisions. The authors highlight three stages of congestion, named Stage 1, Stage 2, and Stage 3. Stage 1 describes low density traffic at a presumably high speed with substantial variance in speed. In this stage, there are so few cars that fatalities are rare, but additional cars greatly contribute to the fatality rate. Stage 2 describes moderately dense traffic, where additional cars lead to additional fatalities, but mitigate the speed (and speed variance). Stage 3 describes gridlocked traffic, such that fatalities rarely occur since the speed (and its variance) is essentially zero. Shefer and Rietveld appeal to several national hour-of-day traffic decompositions to make their case; and highlight the fundamental work of Vickrey (1969) and Pigou (1920) on appropriate approaches to obtain an optimal level of congestion. The general association between traffic density and collisions was identified independently around that time period by Zhou and Sisiopiku (1997), who analyzed traffic patterns on the interstate I-94 in Detroit. Still, this leaves a great deal of empirical work to do in identifying the marginal effect of additional traffic volume.

More recent work, Green et al. (2016) uses a difference in differences approach against a public congestion tax in order to highlight the broad safety improvements of the congestion-reducing tax. Several other researchers investigate the role of additional public transit services in safety improvements (Anderson, 2014; Bauernschuster, Hener, & Rainer, 2017; Jackson & Owens, 2011; Lichtman-Sadot, 2019), and have managed to find that public transit services simultaneously reduce traffic volume and improve safety. Edlin and Karaca-Mandic (2006) further builds upon a spatial analysis of London and found suggestive evidence of casualties (fatalities or injuries) being associated with congestion, with the latter inferred through regional population and employment. Finally, Wang et al. (2009) found no evidence of an association of traffic congestion on fatalities in the M25 motorway in England.

Compared to the state of the art, our work differs from the existing research for the following main factors: first, our between-day variation in vehicle traffic is substantially larger than the previous papers we mention (we see a decline in traffic that falls to nearly 16% of the original peak), giving us a broad base for estimation; second, our traffic variation comes not from public policy targeting traffic or the potentially endogenous traffic density, but instead from the entirely uncorrelated daily variation in the severity of the COVID-19 crisis the day prior.

3. Data sets

In this section, we present the data sources used in this study. We considered a total of five different data sets, specifically: motor vehicle collision, number of vehicles on the road, vehicle speeds, traffic related summons, and COVID-19 test reporting. In total, we collect these data for the period between 2019/01/01 and 2021/02/21, approximately 1 year before and after the first case of COVID-19. In the reminder of this section, we briefly discuss each data set and its nature. The Motor Vehicle Collision data set contains the records of collisions occurring between vehicles within the city of New York (NYPD, 2020). The data set is maintained and provided by New York Police Department under the Open-Data (2020) initiative. The NYC Open Data initiative is meant to provide free and transparent access to data from the city and the administration to residents and beyond. By New York City law it is mandatory to report collisions where someone is injured or killed, or where there is at least $1000 worth of damage, which makes this data set a fairly complete records of all collisions. Each record in the data set reports a single collision, specifically: the date and time of when it occurred; its location; the number of people injured or killed, broken down into motorists, cyclists, and pedestrians; the factors that contributed to the accident; and, the type of vehicles involved. Note that the data set does not include sensitive information that would allow one to trace back the people or cars involved in the collision or the report.

The second data set, the number of vehicles on the road, is also provided under the NYC Open Data initiative and it is maintained by the Metropolitan Transportation Authority (2020) of New York. This data set tracks the vehicles passing through the city bridges and tunnels on an hourly basis. Specifically, the data set reports on the number of vehicles that go through individual toll plazas, every hour, and broken down by: the number of vehicles using an electronic toll collection system; and, the number of vehicles
not using the electronic toll collection system or for which such system malfunctioned. These vehicle counts are also associated with the direction of the traffic, which allows us to aggregate the counts to the individual boroughs, which is the geographical aggregation level of reference in our study. Note that this data set does not carry the exact number of vehicles on every single road: we use the counts in this data set to evaluate the relative change in the traffic volume in the boroughs.

The third data set is the traffic speed, provided by the City of New York Department of Transportation (2020a). The data set is a collection of vehicle speed records observed from a multitude of sensors disseminated throughout the boroughs of the city, mostly on major arterials and highways. Each record in the data set contains the following information: the date of the observation, the identifier of the sensor detecting the speed, the speed of the vehicle (measured as the ratio of the segment of road observed over the travel time to cover it), and the location of the sensor. Speed sensors only retain in memory the speed percentiles of vehicles, not the count of vehicles, therefore speed and volume must be measured at different places.

The fourth data set is the traffic summons issued, provided by the City of New York Department of Transportation (2020b). This data set is a record of all traffic summons issued by the city, organized by borough on a daily basis. This is intended to be a proxy for measures of enforcement, since some might be concerned that enforcement has declined during this period.

Finally, the COVID-19 testing data set contains the historical records of the results from the COVID-19 testing campaign. The data set is maintained and provided by New York State Department of Health (2020). The data set begins on March 1st, and it is currently ongoing. Each record in the data set carries data for a specific date and New York State county, and contains the following information: the number of new positive cases discovered during the past 24 hours, with the time cut off set at 12am of the day for which the report is provided; the number of tests performed in past 24 hours, including positive, negative, and inconclusive; and, the cumulative number of new positive cases and of tests performed since the beginning of the testing campaign.

3.1. Data summary

A summary of the data is presented in Table 1. Data are separated into two coarse groups based on the number of new COVID-19 cases in the week prior; we will later instrument based on the number of new cases. We can see a clear difference in days preceded by no new COVID-19 cases, Table 1a, with a much larger number of vehicles on the road and a much larger number of collisions, injuries, and fatalities. There does not appear to be a large difference in the number of traffic summons on these days (13.7 vs 13.5), which suggests traffic enforcement is relatively unchanged. On days following no new COVID-19 cases, we reason that drivers believe there is low risk associated with leaving the home, since there are fewer infections and the pandemic spread appears to have slowed in the short run. Conversely, Table 1b, on days following many new infections, drivers tend to avoid leaving their homes, since the spread appears to be relatively rapid. At the same time, collisions drop precipitously on days after new COVID-19 cases, though no direct traffic safety regulations have been passed (only an indirect change in the daily traffic volume of automobiles). Finally, every day has at least one collision, regardless of circumstance, suggesting safety is a major problem.

Fig. 1 shows that the number of vehicles on the road (as measured by tolls) dramatically drop around the time period of the first COVID-19 case, and again during the second wave (though the drop is less dramatic). We also indicate the stay-at-home order issued by the New York State Governor on March 22, 2020, by the dotted vertical line in the figure. We observe that the stay-at-home order was preceded, to a large extent, by individuals choosing to stay home ahead of the legal order as COVID-19 initially began its onset in the region. This reduction in mobility preceding state and local stay-at-home policies matches the observations of others using cell phone records (Badr et al., 2020). We later take the possibility of a systemic break in time trends around the stay-at-home order into careful consideration (Section 5.2). One might also anticipate that the level of collisions afterwards varies because of policies altering the type of individuals on the road. To mitigate this we include fixed effects for each day, which capture changes in the level of collisions before and after the policy. A quadratic spline fits the collisions in each borough before and after the first NYC COVID-19 case. Speed-measuring observations that have been flagged with an “error” by the NYC Department of Transportation have been omitted, but there are still errors clearly visible in the simple mean speed, such as the spike in speed for Brooklyn around 2019-07, so we consider these data somewhat tenuous.

Fig. 1 suggests that the COVID-19 cases caused a sudden and exogenous shock to the number of vehicles passing through tolls in the NYC area. A similar but smaller decline follows the so-called “second wave” during the winter of 2021. We observe that this precipitous drop follows the first NYC COVID-19 case (vertical black line) but precedes the stay-at-home order: individuals seemed to have already stopped driving prior to the announcement to a large extent. We assume the decline in vehicles passing through tolls serves as a proportional measure of the daily traffic volume.

We also observe an increase in speed in Fig. 1 coinciding with COVID-19 cases, which conforms with the thesis of Shefer and Rietveld (1997). Lower densities of traffic will have higher speeds (and higher variances between the speed of vehicles), and therefore one might anticipate more collisions or more dangerous collisions from the countervailing consequences of speed. We highlight that there has been no major change in the posted speed in NYC in the recent time frame, though there have been major adjustments 6 years ago, in 2014 (Mammen et al., 2019).

In the following sections we use these data to measure the relationship between daily traffic volume, measured by the volume of cars passing through tolls, and the number of collisions. For our experimental variation, we use the number of COVID-19 cases on the week prior in each borough in order to instrument for the current day’s traffic.1 Note that in analyzing and merging the data sets, we realized that there are no data on traffic volume going to Staten Island, which we then removed from our analysis.

4. Methodology

In order to identify the relationship between daily traffic volume and collisions, we use the standard instrumental variables (IV) technique (Cullinean, 2004) with the primary specification:

\[
\ln(\text{Collisions})_{i,t} = \beta_0 + \beta_1 \ln(\text{VehicleCount})_{i,t} + \beta_2 \text{boroughFE}_i + \\
\quad \beta_3 \text{timeFE} + \beta_4 \text{boroughFE} \times \text{trend}_i + \\
\quad \beta_5 \text{boroughFE} \times \text{trend} + \beta_6 \text{PostCOVID} + \\
\quad \beta_7 \ln(\text{SummonsCount})_{i,t} + \epsilon_{i,t}
\]

1 We consider in the robustness section several other specifications rather than simply the 7-day response and find it does not meaningfully alter the estimations.
Comparison between days one week after new COVID-19 cases and after no new cases.

To resolve this issue, we exploit plausibly random variation in the number of cases that occurred exactly one week prior (a lag) as a measure of the first COVID-19 case (or the stay-at-home order), in case there was a systemic break in commuting patterns during those periods. We also include the log of traffic summons for each day, since one might believe police are altering their enforcement or presence along this data window.\(^2\)

The instrument \(z_{i,t-7}\) is a simple count of the number of COVID-19 cases that occurred exactly one-week prior (a lag) as a measure of the severity of the pandemic, which exploits the fact that individuals appear to have stayed at home of their own volition in response to the pandemic (Badr et al., 2020). We note that the lag is appropriate because new cases on day \(t\) do not alter driving patterns on day \(t\) (they have already left the home), but cases from the week prior (\(t-7\)) appear to be strongly associated with changes in automobile traffic.\(^3\) We anticipate that while additional COVID-19 cases alters the number of cars on the road, the virus itself does not directly alter how people drive (e.g., individuals do not drive slower or more cautiously because of the existence of COVID-19). Any remaining change in driving habits during this period is likely, therefore, a result of the altered volume of traffic on the road. Measuring the association between traffic volume and collisions is critical for forecasting the consequences of such policies like congestion taxes in urban areas, particularly if they are severe (Green et al., 2016).

5. Results

In this section we present the results of our work. First, we discuss the changes on the collision patterns, then we elaborate on the robustness checks built in the methodology of the data analysis.

### 5.1. Primary results: change in collisions

In Table 2 we show the association between the count of collisions and traffic as we sequentially increase the controls:

We have data measured daily (\(t\)) for each (and all) borough (\(i\)) with complete data. Our interest is in the consistent estimation of the parameter \(\beta_i\), which represents the elasticity between vehicle counts and collisions, the percentage change in vehicle collisions for each 1% increase in vehicle counts (Cullinane, 2004). We use the IV technique because we are concerned there could be omitted factors that affect both the count of vehicles and collisions. It could even be the case that collisions lead to changes in the volume of traffic, resulting in inconsistent estimation of the \(\beta_i\) parameter.

To resolve this issue, we exploit plausibly random variation in the count of vehicles, where fewer cars are on the road the week following high-infection days. This change in vehicle counts is uncorrelated with other potentially problematic factors - the presence of the disease does not alter driving behavior, except as mediated through the volume of traffic. We operate with the understanding that drivers do drive differently during the pandemic, but this is derived from the fact that the density of the traffic is lower and there is additional space (Tucker & Marsh, 2021), not because drivers have fundamentally changed during the period. Therefore, COVID-19 serves as an important source of experimental variation for studying the traffic collisions in NYC, and allows for the derivation of consistent parameter estimates for \(\beta_i\) through this technique.

The vector of controls \(brough\) contains simple fixed effects for each borough, controlling for the fact that each borough has different average populations and commuting patterns. Similarly, the vector \(timeFE\) contains fixed effects for each day, since certain days may have unusual weather or be prone to more collisions (new years). These individual and time fixed effects will be partialled out by the fixed effect estimation. In our robustness checks, we also consider that the borough FE undergo some systemic shift during the stay at home period in order to account for possible changes in the constitution of drivers during that period. In our preferred specifications, we include adding borough-specific time trends (\(trend_i\)), and further break these trends along important dates like the first COVID-19 case (or the stay-at-home order), in case there was a systemic break in commuting patterns during those periods. We also include the log of traffic summons for each day, since one might believe police are altering their enforcement or presence along this data window.\(^2\)

The instrument \(z_{i,t-7}\) is a simple count of the number of COVID-19 cases that occurred exactly one-week prior (a lag) as a measure of the severity of the pandemic, which exploits the fact that individuals appear to have stayed at home of their own volition in response to the pandemic (Badr et al., 2020). We note that the lag is appropriate because new cases on day \(t\) do not alter driving patterns on day \(t\) (they have already left the home), but cases from the week prior (\(t-7\)) appear to be strongly associated with changes in automobile traffic.\(^3\) We anticipate that while additional COVID-19 cases alters the number of cars on the road, the virus itself does not directly alter how people drive (e.g., individuals do not drive slower or more cautiously because of the existence of COVID-19). Any remaining change in driving habits during this period is likely, therefore, a result of the altered volume of traffic on the road. Measuring the association between traffic volume and collisions is critical for forecasting the consequences of such policies like congestion taxes in urban areas, particularly if they are severe (Green et al., 2016).

### Table 1

Comparison between days one week after a COVID-19 case and after no new cases.

#### a. Characteristics of days 1 week after a COVID-19 case

| Variables                      | Mean       | Std. Dev. | Min      | Max       |
|--------------------------------|------------|-----------|----------|-----------|
| Count Vehicles\(^*\)           | 168194.6   | 82775.35  | 33,194   | 348,303   |
| Count Collisions               | 88.7522    | 30.24864  | 19       | 183       |
| Count Injured                  | 23.95947   | 11.43227  | 2        | 77        |
| Count Fatalities
| 0.0901826 | 0.30767 | 0 | 3 |
| Total Fatal\(\text{Injured})\ | 24.04966   | 11.48352  | 2        | 77        |
| Simple Mean Speed\(^*\)        | 36.33974   | 6.221759  | 18.91236 | 49.33846 |
| New Cases                      | 0.2111872  | 3.549588  | 0.072    | 123       |
| Count Traffic Summons          | 13.35499   | 23.65592  | 0        | 171       |

**Note:** N:1752, \(i = 4\) boroughs, \(t = 438\) days, “Missing 16 observations,” “Missing 184 observations.

#### b. Characteristics of days 1 week after no COVID-19 cases

| Variables                      | Mean       | Std. Dev. | Min      | Max       |
|--------------------------------|------------|-----------|----------|-----------|
| Count Vehicles                 | 121933.3   | 69626.33  | 10,891   | 303,348   |
| Count Collisions               | 41.11735   | 19.03718  | 3        | 125       |
| Count Injured                  | 16.0809    | 9.792077  | 0        | 54        |
| Count Fatalities               | 0.0867347  | 0.3225972 | 0        | 41        |
| Total Fatal\(\text{Injured})\ | 16.16764   | 9.855463  | 0        | 55        |
| Simple Mean Speed\(^*\)        | 40.91199   | 6.389576  | 24.74412 | 55.46777  |
| New Cases                      | 472.8761   | 480.0707  | 2        | 2722      |
| Count Traffic Summons          | 13.73174   | 23.6801   | 0        | 190       |

**Note:** N:260, \(i = 4\) boroughs, \(t = 65\) days, “Missing 12 observations.”

With a first stage of:

\[
\ln(\text{VehicleCount})_{i,t} = \beta_0 + \beta_1 z_{i,t-7} + \beta_2 \text{boroughFE}_i + \beta_3 \text{timeFE}_i + \\
\beta_4 \ln(\text{SummonsCount})_{i,t-7} + \beta_5 z_{i,t-7} + \beta_6 \text{boroughFE}_i \times \text{trend}_i + \\
\beta_7 \text{boroughFE}_i \times \text{trend}_i \times \text{PostCovid}_i + \\
\beta_8 \ln(\text{SummonsCount})_{i,t-7} + \epsilon_{i,t}
\]  

\(^{(2)}\)
individual and time fixed effects, individual borough time trends, and systemic breaks in time trends at the day of the first case. The first stage is extremely strong, with an F-statistic of over 75 in all specifications, suggesting that our instrument, the number of cases, is an effective predictor of the number of cars passing through tolls the following day (Stock & Yogo, 2005). It suggests, as one might expect, that people have a strong reaction to protect their health and safety from COVID-19 and, as a result, fewer peo-

Fig. 1. Traffic volume and speed following the first COVID-19 cases (solid line) and the stay-at-home order (dotted line) in New York City. The trend lines are second-degree polynomial fits of the variable with respect to time. The periods before and after the first COVID-19 case, and for each borough are fit independently.
ple drive on days after many people have tested positive. Since our specification is log–log, the coefficients represent estimations of the elasticity of the number of collisions relative to the volume of traffic (i.e., they are the percentage change in collisions over the percentage change in traffic volume).

In Table 2, column 1 reports on the values of the elasticity between collisions and vehicle traffic when only fixed effects for borough and time periods are included, for which we observe a value of 1.5, roughly. These time fixed effects take into account day of week traffic patterns, and seasonal commuting patterns. The estimated elasticity remains similar at 1.4 when adding individual time trends, (see column 2). These individual time trends take into account the fact that each borough may be trending upward or downward in traffic safety independently, due to factors such as gentrification. In column 3, we add a systemic break for the first COVID-19 case, since the borough-specific traffic trends may have altered at that time, perhaps as individuals work from home at different rates in different regions. The estimated elasticity between vehicle count and collisions increases to 1.7, though the standard error increases. This is robust to the addition of ln(SummonsCount) in column 4. Critically, this inclusion does not substantively alter the coefficient of interest. Our final estimation measures that an increase in the number of vehicles by 1% is associated with a significant 1.7% increase in the number of collisions, and the final R² is around 30%. Using a short calculation, the nearly 28% reduction in vehicle traffic in the Covid-19 period we see in Fig. 1, is associated with nearly a 37% reduction in collisions.4

5.2. Robustness checks

In this section we run a number of tests to highlight that the estimation in Table 2, column 4, is robust to a variety of specification changes. Results of the robustness checks are reported in Table 3. The changes that we take into account are: we modify the data window’s size; we add a second systemic break in trends on the day of the governments stay-at-home order. and, we explored several variations of lags for the COVID-19 case instrument. The general pattern of the result remains consistently positive and significant, and the estimated elasticity between traffic volume and collisions falls typically between 1.2 and 1.7, and none are more than 3 standard deviations away from our preferred specification of 1.661.

In Table 3 column 1, we reduce the size of the data window by about 6 months, such that it begins on June 1st 2019 instead of January 1st 2019. We note that the estimated elasticity only modestly increases in size and it is still significant at about 1.4. In column 2 we consider that the breaks in time trends were insufficient, and includes another systemic break in trends at the time of the governments stay-at-home order. We further allow each borough to have an additional break in the value of their associated fixed effect within the stay-at-home period, in case the rule triggered.

Table 2
Estimated elasticity between collisions and vehicle counts.

| Variables          | (1)     | (2)     | (3)     | (4)     |
|--------------------|---------|---------|---------|---------|
| ln(Vehicle Count)  | 1.480***| 1.415***| 1.695** | 1.661***|
| Daily Fixed Effects| YES     | YES     | YES     | YES     |
| Borough Fixed Effects| YES     | YES     | YES     | YES     |
| Borough Trends     | YES     | YES     | YES     | YES     |
| Break After First Case| YES     | YES     | YES     | YES     |
| Enforcement Control| YES     | YES     | YES     | YES     |
| Observations      | 3,108   | 3,108   | 3,108   | 3,108   |
| R-squared (Centered)| 0.226   | 0.337   | 0.323   | 0.332   |
| Log Likelihood    | 1272    | 1513    | 1479    | 1501    |
| Number of Boroughs| 4       | 4       | 4       | 4       |

Standard errors are clustered by borough.

* p < 0.1.
** p < 0.05.
*** p < 0.01.

Table 3
Robustness checks.

| Variables          | (1)                  | (2)                  | (3)                  | (4)                  |
|--------------------|----------------------|----------------------|----------------------|----------------------|
| ln(Vehicle Count)  | 1.379***              | 1.228***              | 1.339**              | 1.619***              |
| Daily Fixed Effects| YES                  | YES                  | YES                  | YES                  |
| Borough Fixed Effects| YES                 | YES                  | YES                  | YES                  |
| Borough Trends     | YES                  | YES                  | YES                  | YES                  |
| Break After First Case| YES               | YES                  | YES                  | YES                  |
| Enforcement Control| YES                  | YES                  | YES                  | YES                  |
| Observations      | 2,500                | 3,108                | 3,108                | 2,940                |
| R-squared (Centered)| 0.362               | 0.310                | 0.390                | 0.347                |
| Number of Boroughs| 4                    | 4                    | 4                    | 4                    |

Standard errors are clustered by borough.

* p < 0.1.
** p < 0.05.
*** p < 0.01.

4 1 = (1–0.01660956)²8 ≈ 0.37.
garded particular changes in the constitution of travelers from each borough. While the stay-at-home order applies to have been preceded to a large extent by New Yorkers already staying home (see Fig. 1), it is possible that there were still transitions to the driving (beyond simply daily traffic volume) during the post-announcement period. Jointly, these breaks are significant, reinforcing that this is another plausible specification. Still, including this second systemic break only decreases the estimated coefficient slightly to 1.2, while the significance remains. In column 3, we consider alternatives to individuals waiting 7 days to respond to new COVID-19 cases. Instead, we consider a fourth-degree polynomial of the 7-day instrument. Our point with this specification, and the next, is that individuals have a generalized pattern of reducing driving in response to past cases of COVID-19 and that our results do not appear to be dependent on a particular functional form in the first stage. The instruments remain strongly significant, with an F-statistic of 52. The estimated elasticity increases to a significant 1.3, still in the same general direction and within one SD of the primary estimates in Table 2. In column 4 we test another set of instruments, exploring the idea that individuals respond to 7, 14, and 21 day lags of COVID-19 cases, that is they have a memory of just under a month before they return to normal habits. The instruments remain strongly significant, with an F-statistic of 25. The estimated elasticity remains at a significant 1.62, within 0.05 of the original estimates in Table 2, column 4.

5.3. Discussion and further elaboration

So far, we have focused heavily on the percentage change in the number of collisions as a consequence of daily traffic volume. However, not all collisions are equally dangerous. In this section, we consider how the daily traffic volume is associated with the total count of fatalities and injuries. Previous research highlights that the direction of the effect appears to be complicated by potentially reducing the distance between cars, therefore softening the collisions (Shefer & Rietveld, 1997; Wang et al., 2009). We examine these concerns, particularly the notion that denser traffic leads to “cushioning” and safer collisions. We find significant support for the idea that higher traffic volumes are associated with safer collisions.

In Table 4, we continue to use the same IV specification as in column 3 of Table 2, though in each of the 4 columns we change the independent variable in each of the estimations to the count of: injuries, fatalities, the sum of both (casualties), and casualties per collision. In each of the 3 rows, we consider a different subset of victims: total, pedestrians, or cyclists, which are contained as subsets of: injuries, fatalities, the sum of both (casualties), and casualties per collision. In each of the 3 rows, we consider a different subset of victims: total, pedestrians, or cyclists, which are contained as subsets of casualties (row 2), but the coefficients are insignificant with large standard errors. For cyclists the standard errors are even larger, however, the rate of casualties/collision is found to be negative for all categories. This suggests that collisions may be generally safer in high-volume roadways. We conclude, therefore, that higher volumes of vehicle traffic are more likely to have collisions, but those collisions tend to have fewer injuries and potentially fewer deaths than their low-volume collision counterparts.

We find that collisions have declined during this COVID-19 period, but the safety of those collisions also declined, matching the mixed effects of Shefer and Rietveld (1997). As a result, we would like to investigate if the other elements of the “cushioning” hypotheses hold true in the relevant range, in particular that vehicles will slow down under higher densities of traffic, which leads to safer accidents.

We note it is visually apparent that the simple mean speed of traffic has drastically increased in all boroughs post-COVID-19, see Fig. 1, which according to Shefer and Rietveld (1997) is a consequence of the reduced density. To corroborate on this observation, we run a simple regression of log vehicle volume against log speed (controlling only for time and borough fixed effects). We find a significant estimated elasticity between volume and speed of about −13%. This component of the evidence seems to reinforce the plausibility of traffic volume as a double-edged sword: while the higher traffic volumes lead to more collisions, sudden decreases in traffic volumes are associated with higher speeds. This suggests a rich, complex relationship, despite this factor being outweighed in the relevant range by the safety benefits we measure in Table 4.

In total, we calculate that there were nearly 356 collisions each day prior to COVID-19, combined across all four boroughs. Prior to COVID-19, the value of these collisions, injuries, and deaths each day are approximately $4.2 million per day.7 In the post-COVID-19 period, we calculate that the 28% reduced traffic volume leads to about a 37% reduction in collisions, resulting in approximately $453,000 in social property costs avoided daily, excluding any injuries or fatalities prevented.8

---

6 $0.4367 change in injuries for each 1% change in traffic volume / 0.245 average injuries over data window + 100% = −2.1 injuries per day for each 1% change in traffic volume.
7 88.75 collisions/borough/4*1$3,447 + 23.95 injuries/borough/4*2 $28,299 + 0.0901 fatalities/borough/4*3 $3,186,408 [approximately equal symbol goes here] $4.2 million/day.
8 We use the value of $3,447 for property damage-only collisions from (Parry, 2004), so 37% reduction in collisions 88.75 collisions per day 4 boroughs $3, 447 $452, 763.
However, the increased injury rate of those remaining collisions more than offsets the benefits of the reduced collisions. We further calculate that the increase in fatalities and injuries are worth approximately $2.6 million in social cost a day, though this will vary substantially depending on the statistical value of a human life (Parry, 2004), and any variation in the fatalities estimate. This suggests that policymakers may want to be careful in considering policies like congestion taxes that may adjust the traffic volume, since there may be unintended consequences for safety in terms of speed. These consequences may not appear when adjustments to traffic volume are modest, but are visible when there are sudden and substantive traffic volume changes, such as in the case of COVID-19.

### 6. Conclusion

In this paper, we utilize a unique instrumental variable: the number of COVID-19 cases, to instrument for the traffic volume on the following day. The use of the instrumental variable helps address the concerns of Wang et al. (2009), Shefer and Rietveld (1997), and Noland and Quddus (2004), about the challenges of identifying a relationship between congestion and traffic safety (Cullinean, 2004). We find that the exogenous shocks to traffic volume substantially reduce the number of collisions at a rate of roughly 1.7% fewer collisions for every 1% reduction in traffic volume. This finding is robust to various specifications, including different time windows, borough-specific time trends, alternative specifications of the instrument itself, and a structural break in time trends at the time of the stay-at-home order. On the other hand, our estimates suggest lower volumes of traffic are associated with net increase in injuries and fatalities over the relevant range, specifically an increase of 0.003 fatalities for each 1% decrease in traffic volume, and an increase of 0.44 for injuries for each 1% decrease in traffic volume.

Both of these findings suggest that the remaining collisions are of a more dangerous type when traffic volume is reduced, a concern highlighted by Shefer and Rietveld (1997) and Zhou and Sissipiku (1997). We find other corroborating evidence in our data: we find an increase in casualties per collision as traffic volume decreases. Though this association is not significant, the pattern remains persistently negative for pedestrians. Accordingly, we estimate a simple elasticity of speed/volume approximately equal to −0.13.

We also provide an estimates of the value of this transition. Using a back-of-the-envelope calculation, we estimate that the approximate 28% decline in traffic volume during the COVID-19 period is associated with about $453,000 per day in savings from the property damage of collisions (Parry, 2004). However, the increased bodily harm from collisions is estimated to be worth nearly $2.6 million a day across the boroughs, though this estimate may vary dramatically depending on the difficult valuation of human life and the effects on external participants (Parry, 2004). Despite variations in valuation, the estimated value of the injuries and fatalities are an order of magnitude larger than the measured reduction in property damage. This highlights a complex relationship between traffic volumes, speeds, and safety that is worth further study.

### References

Abouk, R., & Adams, S. (2013). texting bans and fatal accidents on roadways: Do they work? Or do drivers just react to announcements of bans? American Economic Journal: Applied Economics, 5, 179–195.

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on congestion. American Economic Review, 104, 2763–2796.

Badr, H., Du, H., Marshall, M., Dong, E., Squire, M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. Online. The Lancet Infectious Diseases. https://doi.org/10.1016/S1473-3099(20)30553-3.

Baghestani, A., Tayarani, M., Allahviranloo, M., & Gao, H. O. (2020). Evaluating the traffic and emissions impacts of congestion pricing in New York City. Sustainability, 12, 3655.

Bauernschuster, S., Hener, T., & Rainer, H. (2017). When labor disputes bring cities to a standstill: The impact of public transit strikes on traffic, accidents, air pollution, and health. American Economic Journal: Economic Policy, 9, 1–37.

City of New York Department of Transportation (2020a). Traffic speeds. https://data.cityofnewyork.us/Transportation/DOT-Traffic-Speeds-NBE/i4qj-fg9j (accessed May, 2021).

City of New York Department of Transportation (2020b). Traffic summons. https://data.cityofnewyork.us/Public-Safety/NYPD-B-Summons-Historic-/bnme-57ty4 (accessed August, 2021).

Cui, Z., Heinrichsson, K., Ke, R., & Wang, Y. (2020). Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. IEEE Transactions on Intelligent Transportation Systems, 21(11), 4883–4894. https://doi.org/10.1109/TITS.2019.2950416.

Cullinean, K. (2004). Statistical and econometric methods for transportation data analysis. Springers.

De Fabris, C., Ragona, R., & Valenti, G. (2008). Traffic estimation and prediction based on real time floating car data. In 2008 11th international IEEE conference on intelligent transportation systems (pp. 197–203). IEEE.

Dills, A. K., & Mulholland, S. E. (2018). Ride-sharing, fatal crashes, and crime. Southern Economic Journal, 84, 956–961.

Edlin, A. S., & Karaca-Mandic, P. (2006). The accident externality from driving. Journal of Political Economy, 114, 931–955.

Faghhi-Imani, A., Anowar, S., Miller, E. J., & Eliu, N. (2017). Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City. Transportation Research Part A: Policy and Practice, 101, 11–21.

Green, C. P., Heywood, J. S., & Navarro, M. (2016). Traffic accidents and the London congestion charge. Journal of Public Economics, 133, 11–22.

Iyer, S., Boxer, K., & Subramanian, L. (2018). Urban traffic congestion mapping using bus mobility data. In CEUR workshop proceedings (Vol. 2227, pp. 7–13). CEUR-WS.

Jackson, C. K., & Owens, E. G. (2011). One for the road: Public transportation, alcohol consumption, and intoxicated driving. Journal of Public Economics, 95, 106–121.

Jain, V., Sharma, A., & Subramanian, L. (2012). Road traffic congestion in the developing world. Proceedings of the 2nd ACM symposium on computing for...
Lichtman-Sadot, S. (2019). Can public transportation reduce accidents? Evidence from the introduction of late-night buses in Israeli cities. *Regional Science and Urban Economics*, 74, 99–117.

Liu, W., Zheng, Y., Chawla, S., Yuan, J., & Xing, X. (2011). Discovering spatio-temporal causal interactions in traffic data streams. In *Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining* KDD 2011 (pp. 1010–1018). New York, NY, USA: Association for Computing Machinery.

Mammen, K., Shim, H. S., & Weber, B. S. (2019). Vision Zero: Speed limit reduction and traffic injury prevention in New York City. *Eastern Economic Journal*, 1–19.

Metropolitan Transportation Authority (2020). Hourly traffic on Metropolitan Transportation Authority (MTA) bridges and tunnels. https://data.ny.gov/Transportation/Hourly-Traffic-on-Metropolitan-Transportation-Authority/qzve-kjga (accessed May, 2021).

New York State Department of Health (2020). New York state statewide COVID-19 testing. https://health.data.ny.gov/Health/New-York-State-Statewide-COVID-19-Testing/xdss-u53e (accessed May, 2021).

Noland, R. B., & Quddus, M. A. (2004). A spatially disaggregate analysis of road casualties in England. *Accident Analysis & Prevention*, 36, 973–984.

NYPD (2020). Motor vehicle collision. https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95 (accessed May, 2021).

OpenData, N. (2020). New York City OpenData initiative. https://opendata.cityofnewyork.us/ (accessed May, 2021).

Parry, I. W. (2004). Comparing alternative policies to reduce traffic accidents. *Journal of Urban Economics*, 56, 346–368.

Parry, I. W., & Bento, A. (2002). Estimating the welfare effect of congestion taxes: The critical importance of other distortions within the transport system. *Journal of Urban Economics*, 51, 339–365.

Pigou, A. C. (1920). The economics of welfare. Palgrave Macmillan, Reprinted in 2013. Shefer, D., & Rietveld, P. (1997). Congestion and safety on highways: Towards an analytical model. *Urban Studies*, 34, 679–692.

Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and Inference for Econometric models: Essays in Honor of Thomas Rothenberg*, 109–120.

Tucker, A., & March, K. (2021). Speeding through the pandemic: Perceptual and psychological factors associated with speeding during the covid-19 stay-at-home period. *Accident Analysis & Prevention*, 159. https://doi.org/10.1016/j.aap.2021.106225 https://www.sciencedirect.com/science/article/pii/S0001457521002566.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, 59, 251–260.

Wang, C., Quddus, M. A., & Ison, S. G. (2009). Impact of traffic congestion on road accidents: A spatial analysis of the m25 motorway in England. *Accident Analysis & Prevention*, 41, 798–808.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Zhou, M., & Sisiopiku, V. P. (1997). Relationship between volume-to-capacity ratios and accident rates. *Transportation Research Record*, 1581, 47–52.

Paolo Cappellari is an Associate Professor at the College of Staten Island, CUNY. His research focuses the effective management, interpretation, and utilization of data, with specific attention to the study and transformation of schema and data between models, keyword search over semantic datasets and data management in sensor networks. He has published his research results in the major journals and conferences, including ACM-SIGMOD, VLDB, EDBT, ER, ICDE, FoIKS.

Bryan S. Weber is an Assistant Professor of Economics at the College of Staten Island, CUNY. His research focuses on transportation, crime, and statistical methods. Currently, his work explores the introduction of new and novel transportation systems into urban areas, helping enumerate the costs and benefits to residents. His published research results include Transportation Research Part A: Policy and Practice, Regional Science and Urban Economics, and Economic Inquiry.