The air quality effects of Uber

Luis Sarmiento[*] Yeong Jae Kim [†]

October 30, 2021

This study looks at the air quality effects of Uber in the United States. For this, we estimate Uber's impact on the air quality index with state-of-the-art difference-in-difference estimators accounting for Uber's staggered implementation and dynamic treatment effects. Results show that Uber improves air quality. The value of the air quality index and the number of unhealthy air quality episodes decrease after its introduction. We provide evidence that the bulk of the improvement comes from declining ozone levels during the summer. Notably, results hold for a plethora of different specifications, samples, and robustness exercises. To the best of our knowledge, this article is the first to estimate the air quality effects of ride-hailing technologies empirically. However, further research is needed to identify the mechanisms driving the relationship between ride-hailing technologies, the transportation sector, and air quality.

Keywords— Ride-hailing, Uber, Air pollution, United States, Difference-in-differences

JEL— R40, H42, O33, Q53

[*]RFF-CMCC European Institute on Economics and the Environment (EIEE), Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC). E-Mail: luis.sarmiento@eiee.org. Ph: +39 342 613 3782

[†]Corresponding author: RFF-CMCC European Institute on Economics and the Environment (EIEE) and Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC). E-Mail: yeongjae.kim@eiee.org. Ph: +39 342 613 3782

1. Introduction

Uber was founded and started operations in San Francisco's Bay Area in 2010; at this time, the company had a more luxurious business model (now called Uber Black) than regular cabs. However, in 2012 it allowed private car owners to provide ride services similar to traditional taxis with Uber-X. After San Francisco, the company started operations in Chicago, New York City, Boston, Washington DC, and Seattle in 2011 and Phoenix, Dallas, Philadelphia, Denver, Atlanta, Minneapolis, and Los Angeles in 2012. By 2017, it was present in more than nine hundred counties across the United States (US).

Uber disrupted the urban transportation system, creating a world of convenience and changing the way people move in urban environments. Instead of waiting on the streets or phone calling a taxi, people could now use an online platform to book an Uber and watch the car's progress towards the pick-up location. Among the main advantages of Uber vs. traditional taxi services is that it allows the user to track the trip, pay electronically, rate the journey, and know the ride fare beforehand. Currently, the company has operations in more than eighty-five countries across seven hundred and fifty metropolitan areas (Uber, 2021).

Although previous studies on the effects of digital matching technologies for ride-hailing services like Uber had examined its impact on the transportation sector (e.g. Clewlow and Mishra, 2017; Keating, 2019; Schaller, 2017) the existing literature has yet to understand its effects on overall air quality. In this study, we fill this gap by looking at the impact of Uber on US cities' air quality.[1] Furthermore, we also contribute to the broader literature on the relationship between the transportation sector and air pollution.[2]

[1] We define urban agglomerations as all counties part of metropolitan statistical areas (MSA). MSAs are urban areas with at least one urbanized region of more than 50,000 inhabitants (U.S. Census Bureau, 2020).

[2] Notably, most studies looking at the relationship between the transportation sector and air pollution primarily focus on the effects of public transit infrastructure (e.g. Lalive et al., 2018) or policy mechanisms for regulating road traffic (e.g. Sarmiento et al., 2021). And even when they look at the impact of disruptive

Studying the air quality effects of Uber is highly relevant as local politicians constantly debate its consequences on air quality, congestion, and social welfare. For instance, in 2015, New York City's mayor Bill de Blasio proposed a new bill to stop the company's growth. Among de Blasio's main arguments was an increase in traffic and its associated air quality effect. In a 2019 opinion article, the mayor wrote, "Uber added to our pollution, worsened our air quality, and crowded out bus riders, pedestrians, and cyclists. Traffic speeds in midtown fell to just above 4 miles per hour — barely faster than walking" (Bill de Blassio, 2019). Notably, this discussion is not unique to New York City, with several other politicians in cities like London, Milan, and Los Angeles also considering limiting Uber because of air pollution concerns.[3]

Identifying the effect of Uber on air quality is challenging because of the complexities of the transportation network and the interaction of air pollutants in the lower atmosphere. For instance, there are three possible mechanisms through which Uber's effects on the transportation system can affect air quality; scale, substitution, and complementarity. The scale effect refers to an increase in the number of cars in the street. For instance, Clewlow and Mishra (2017) and Ward et al. (2021b) provide evidence of an increase in total vehicle mileage after the introduction of ride-hailing services across seven major US cities.[4] The substitution effect refers to replacing alternative transport modes like private vehicles, taxis, and subways with Uber rides. Its impact on air quality depends on the transportation mode that ride-hailing replaces. If Uber substitutes public transport, the introduction of Uber could potentially worsen air quality. For example, Clewlow and Mishra (2017) find evidence that the introduction of Uber to major US cities results in a 6% drop in the use of public transport. However, if instead of substituting public transportation, Uber replaces old taxis

technological innovations, they often only concentrate on the environmental effects of electrification (e.g. Holland et al., 2016).

[3]For example, in 2018, London's mayor Sadiq Khan tried to pass legislation limiting the number of Uber drivers because of social, pollution, and congestion concerns.

[4]It is relevant to notice that the transportation literature is still not conclusive on the effects of Uber in vehicle ownership and transit. For instance, several studies find opposite results to Ward et al. (2021b) (see Yan et al., 2019; Feigon and Murphy, 2016; Hampshire et al., 2017).

or private vehicles, it can also improve air quality (Keating, 2019). Finally, the complementarity effect suggests that Uber can decrease air pollution by facilitating access to public transportation hubs, overcoming the last mile problem of public transit (Rayle et al., 2014).

Our study does not attempt to identify these mechanisms as they depend on each city's transportation infrastructure, commuting patterns, and income elasticity. For instance, overcoming the last mile problem in urban areas with good public transportation like New York City is easier than in less robust areas like Los Angeles, Houston, or Dallas. Still, we run some robustness exercises and show countrywide evidence of an increase in the number of available cars per household (scale effect) and an increment of persons commuting to work with public transport (complementarity effect).[5] Furthermore, figure 1 of the appendix provides descriptive evidence of a substitution effect between taxis and Ubers in New York City and Chicago.

An additional concern is that the relationship between air pollutants is not always straightforward. At times, increases in one particle do not necessarily translate to worse air quality. For instance, there is often an inverse relationship between ground-level ozone (O_3) and traffic-related contaminants like nitrogen dioxides (NO_2); in dense traffic areas with significant emissions of NO_2, O_3 values are lower than in rural regions with less traffic. This inverse relationship exists because, at high concentrations, NO_2 degrades O_3 back into oxygen (O_2). Thus, even if the scale effect dominates and there is an increase in traffic-related NO_2, the impact of higher NO_2 may decrease O_3, making the overall effect on air quality challenging to assess by only looking at its effects on one particle.

To provide a general assessment of Uber's effect on air quality and avoid capturing pollutant-

[5] In table A.1 of the appendix, we run our empirical design on yearly estimates of the number of available household vehicles per county from the American Community Survey. Unsurprisingly, results show that the introduction of Uber increases the number of available cars per household in the US. Also, in table A.1 we estimate the effect on the share of persons commuting to work with data from the American Community Survey. The results provide evidence of complementarities through increases in the number of persons reporting work commutes with public transit.

specific consequences, we concentrate on its impact on the Environmental Protection Agency (EPA)'s air quality index (AQI). The AQI proxies air quality by transforming the concentration of criteria pollutants into a single scale running between 0 and 500 units.[6] At each point in time, the AQI is the maximum across all measured particles in that county. An AQI value of 100 units corresponds to both the air quality standard for that particle and the threshold between moderate and dangerous levels of exposure for sensitive groups (EPA, 2021a). Thus, even if Uber affects the concentration of each criteria contaminant differently, the AQI allow us to recover a more holistic assessment of its air quality effects.

We infer the causal effect of Uber on the AQI by leveraging the spatio-temporal variation in its roll-out with difference-in-differences (DD) designs. In layman's terms, we compare the value of the AQI on treated and non-treated demarcations before and after Uber started operations. The identification assumption is that conditional on observables, the introduction date of Uber is orthogonal to unobserved determinants of air quality. Furthermore, current developments in the DD literature provide evidence that in the presence of staggered and dynamic treatment effects, two-way fixed effects difference-in-differences (TWFE-DD) can generate bias estimates of the true average treatment effect on the treated (ATT) (De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021).[7] We consider this potential source of bias with Callaway and Sant'Anna's difference-in-differences (CS-DD) methodology, allowing us to estimate and flexibly aggregate group-time average treatment effects across multiple groups and time periods (Callaway and Sant'Anna, 2020).

Results show that the introduction of Uber improves air quality. In the preferred specification, Uber decreases the maximum yearly value of the AQI by 10.69 units. This reduction translates to a

[6] The clean air act requires the EPA to set North American Air Quality Standards (NAAQS) for six common air pollutants often referred to as "criteria air pollutants." These are, ground-level ozone (O_3), coarse particulate matter (PM_{10}), fine particulate matter (PM_{25}), carbon monoxide (CO), sulfur dioxides (SO_2), and nitrogen dioxides (NO_2).

[7] In figure 2 of the appendix, we show that this bias can potentially affect our estimates with the Goodman-bacon decomposition.

7.3% drop concerning pre-treatment values. Concerning the number of days of bad air quality, i.e., days with AQI values higher than 100 units, we find that Uber decreases their average number by 2.53. Reassuringly, results are robust to three different definitions of control counties, i.e., never and still not treated, never treated, and still not treated. Looking at heterogeneous seasonal effects, we provide evidence that the air quality improvement is more significant during the summer, suggesting that the air quality improvement may come from reductions in the concentration of O_3. We confirm this by running contaminant-specific regressions. As expected, O_3 is the only contaminant that reports significant reductions in its AQI after the introduction of Uber.[8]

We perform several robustness tests to check the stability of our results to different samples and econometric designs. First, we examine the effect of Uber across each US Census Region. Results confirm that air quality improvements occur across the country. Next, to avoid the bias effect of unobservables, we exclude from the sample all counties that report changes in their power plant's fleet, forest fires, or violations of NAAQS. Results are robust to excluding all of these counties from the sample. Finally, we run a more typical TWFE-DD model on the effect of Uber on air quality. Reassuringly, results are not statistically different from the CS-DD design.

Although this is the first article examining the relationship between Uber and air quality, its results align with previous modeling studies in the environmental literature. For instance, a recent article simulates the effect of ride-hailing companies on air pollution and finds that ride-hailing can reduce PM_{25}, NO_2, and VOC by decreasing cold-start emissions and replacing old vehicles with relatively new fleets (Ward et al., 2021a). Our study empirically complements Ward et al. (2021a) by providing suggestive evidence of the PM_{25} decrease and further showing that the lump of the air quality improvement comes from ozone reductions. Furthermore, our research also contributes to the current policy debate on the effects of ride-hailing technologies and implies that policymakers

[8]In line with Ward et al. (2021a), we also uncover suggestive evidence of reductions in PM_{25}. However, point estimates are generally insignificant at conventional significance levels.

should be careful when using traffic estimates as proxies for the impact of these technologies on air quality.

2. Data

We obtained the introduction date of Uber from peer-reviewed studies like Ward et al. (2021b) and online sources like Uber blogs, local media outlets, and google quests, where we used keywords like "When did Uber start operations in the Bronx county NYC?" However, even after a thorough search, there was still a modest share (less than 0.5%) of counties where we could not find the specific introduction date. Most of these counties were small urban areas without enough media coverage. If a county has no introduction date, we exclude it from the data set.

Panel (A) of figure 1 shows the 2017 spatial distribution of counties with and without Uber (henceforth treated and control counties). As expected, there are more treated counties in highly populated areas like the East Coast, California, and the Midwest. Moreover, first-treated counties belong to large urban agglomerations like New York, Los Angeles, Chicago, San Francisco, and Dallas. Panel (B) plots the number of counties where Uber started operations per year.

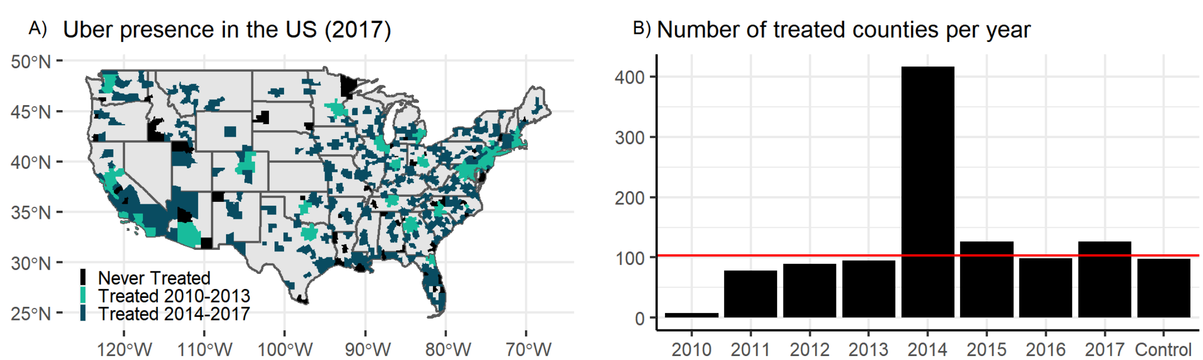

Figure 1: Descriptive statistics of Uber in the United States

Notes: A) Colored counties are all demarcations part of US metropolitan statistical areas (MSA). Control units are counties without Uber as of 2017. B) The vertical axis in panel (B) contains the number of newly treated counties. The red horizontal line indicates the average number of treated counties across the sample period.

Notably, Uber's market-driven roll-out makes first-treated counties different from later-treated and control units. Figure 2 portrays these cross-sectional differences in income, population density, and the transportation index with density plots.[9] As expected, first treated counties have higher incomes, are more densely populated, and use more public transportation. For instance, the average income per capita for counties treated in 2010 is 40,000 dollars higher than counties treated in 2017 and 43,000 higher than counties in the control group.

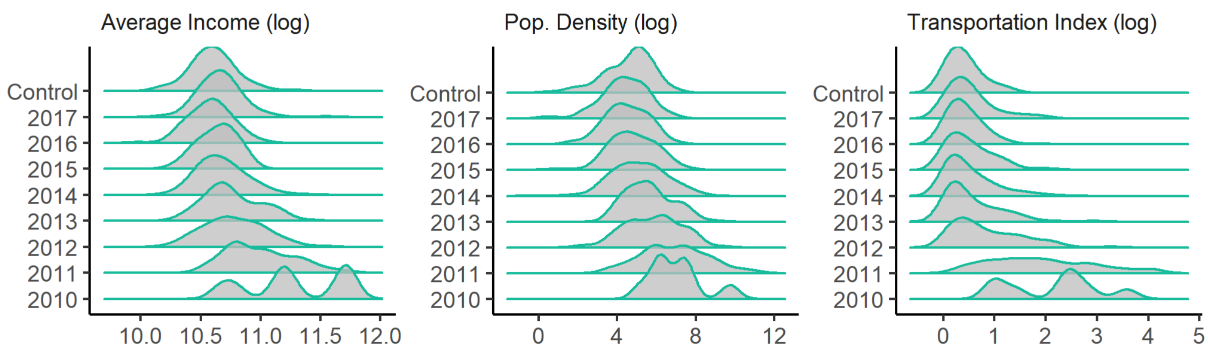

Figure 2: Differences between treatment and control groups

Notes: This figure portrays the density distribution of average income, population density, and the transportation index. The transportation index is the share of persons commuting to work with public transport. The vertical axis contains indicates the treatment group, i.e., the group of counties treated in a specific year.

We use the air quality index (AQI) as a proxy for the air quality of US urban agglomerations. The AQI normalizes the concentration of the six main criteria pollutants into a standardized measure between zero and five hundred units. The EPA divides the AQI into six categories based on its health risks; good for days between 0 and 50, moderate between 50 and 100, unhealthy for sensitive groups between 100 and 150, unhealthy between 150 and 200, very harmful between 200 and 300, and hazardous for days reporting values higher than 300.[10] The AQI for each county is the highest AQI across all measured contaminants and stations in that county.

[9]The transportation index is the share of persons commuting to work with public transport according to the American Community Survey. Appendix table A.3 shows the average of the air quality index, population density, income, and the transportation index across all periods for both treatment and control counties.

[10]See figure 3 of the appendix for additional information on each category.

The data-set comes from the EPA's yearly pre-generated data-files. It contains the annual maximum, 90th percentile, and median value of the AQI between 2000 and 2017 for all reporting counties in the US. For the rest of the study, we focus on the maximum as this is the measurement used by the EPA to assess the health effects of bad air quality. Furthermore, and contrary to the median and ninety percentile, the maximum AQI is informative regarding episodes of exacerbated air pollution.[11] The data-set also includes the number of days in which each county reports the index and the total number of days within each of the six risk categories.[12] Finally, we also obtain particle-specific AQIs to examine the effects of Uber across air contaminants.

Panel (A) of figure 3 compares the intertemporal AQI value between counties with and without Uber. This figure is slightly different from standard common trends plots because of the staggered introduction of Uber. For instance, in cases with a unique treatment date, researchers often center the event-time graph around treatment and look at differences in the average value of the dependent variable before and after treatment. Unfortunately, we cannot follow this approach because we have no unique treatment period. To fix this, we average the value of all possible event-time combinations across treatment groups. For instance, the value at -1 corresponds to the average value one year before the introduction date of Uber across all treatment groups, i.e., all groups of counties where Uber was introduced in the same year.[13] It is comforting to see that there is suggestive visual evidence of common trends between treated and control units before the introduction date of Uber. Panel (B) formally test for this pre-treatment difference between treated and control units with

[11]Throughout the study, we also provide results for the median and ninety percentile values of the AQI in the appendix.

[12]The number of days each county publishes the AQI can change because of malfunctions, maintenance, or administrative decisions.

[13]Specifically, the value of the AQI τ periods to the treatment date is: $\hat{AQI}_{\tau}^{Treated} \mid \hat{AQI}_{\tau}^{Control} = \frac{1}{N^{\tau}} \sum_{\tau=-15}^{6} \frac{1}{N^{c}} \sum_{c} AQI_{\tau}^{c} \quad \forall\ \tau = (Y - G)$ Where Y indicates the year of the observation and G the treatment group, i.e., the year that Uber started operations. N^{τ} is the number of times τ takes a specific value, e.g., for $\tau = -1$, there are eight different combinations of Y and G; (2009-2010, 2010-2011, ... 2016-2017). Finally, N^{c} referrers to the number of counties.

the methodology outlined in Callaway and Sant'Anna (2020).[14] The figure portrays coefficients and 95% confidence intervals for the effect of Uber on the pre-treatment difference between treated and control units. Reassuringly, we cannot reject the common trends assumption, as there is no significant coefficient before treatment.

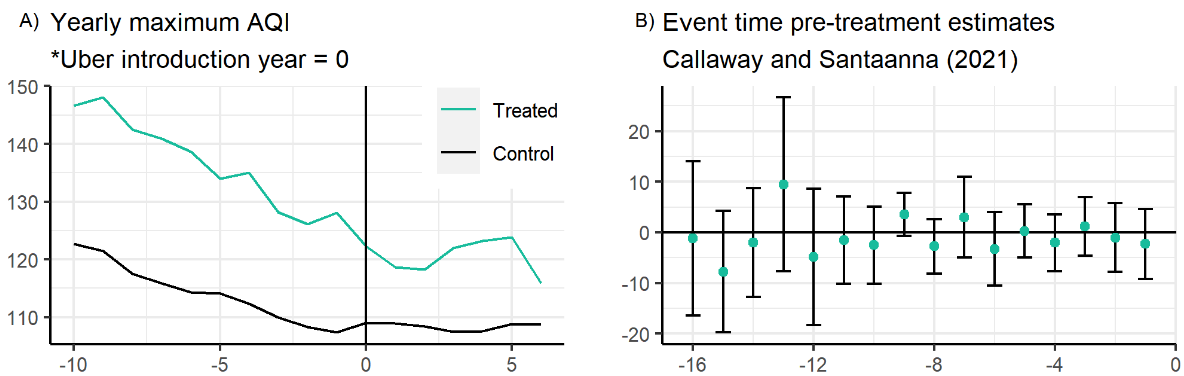

Figure 3: The common trends assumption

Notes: A) Inter-temporal comparison of the air quality index (AQI) between treated and control units. The vertical axis contains the average maximum of the AQI and the horizontal axis the time to treatment (τ). Each data-point for the treated and control group comes from: $\hat{AQI}_{\tau}^{Treated} \mid \hat{AQI}_{\tau}^{Control} = \frac{1}{N^{\tau}} \sum_{\tau=-16}^{6} \frac{1}{N^{c}} \sum_{c} AQI_{\tau}^{c} \quad \forall\ \tau = (Y - G)$, where Y indicates the year of the observation and G the treatment group. N^{τ} is the number of times τ takes an specific value, e.g., for $\tau = -1$, there are eight different combinations of Y and G; (2009-2010, 2010-2011, ... 2016-2017). Finally, N^{c} referrers to the number of counties. B) Contains Callaway and Sant'Anna (2020)'s difference-in-differences design point estimates and 95% confidence intervals on the effect of Uber on pre-treatment periods difference between treated and control stations' AQI.

Figure 4 plots the time series of each criteria pollutant AQI (panel A) and the number of bad air quality days by treatment group (panel B). Notably, trends between treated and control units also look similar across air pollutants. In general, there is a downward trend in the concentration of CO, NO_2, SO_2, and O_3 alongside smaller changes for the particulate matters. Panel B suggests that post-treatment air quality is significantly better across all sample groups regarding bad air quality days. However, because several other air pollution control policies were potentially concomitant with Uber, this should not be interpreted as a causal effect.[15]

[14] The methodology section contains a complete discussion of Callaway and Sant'Anna (2020)'s difference-in-differences design and their procedure to estimate event-time point estimates.

[15] Table A.2 of the appendix contains the exact values of the AQI across years and treatment groups; the treatment groups with the highest AQI are counties treated in 2011, 2012, and 2013.

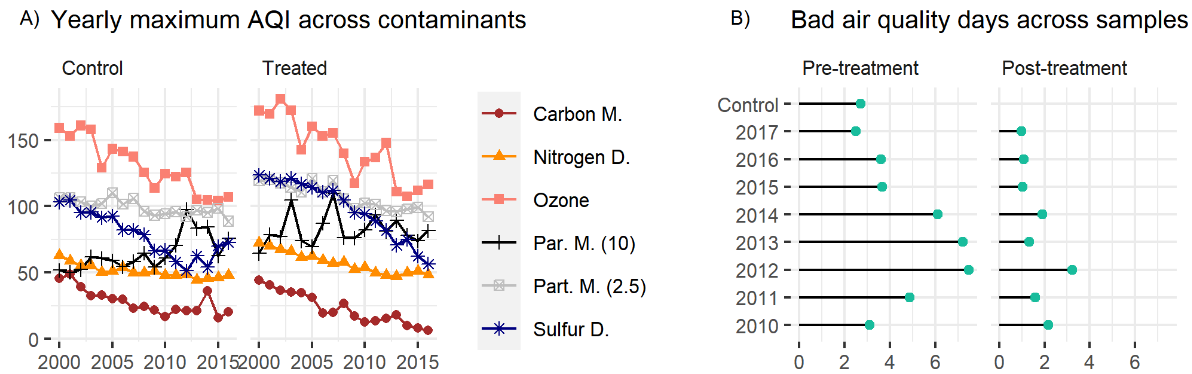

Figure 4: The air quality index across treated and control counties

Notes: A) Portrays the inter-temporal average maximum value of the air quality index (AQI) for all criteria pollutants. B) Compares the number of days that the AQI surpasses one hundred units for the control and treatment groups.

3. Empirical design

Identifying the effect of Uber on air quality is not trivial because the company's roll-out is far from random. Wealthier and more densely populated urban centers are more likely to have Uber services than poorer and scarcely populated agglomerations. If we do not account for these systematic differences between treated and control counties, they can lead to biased estimates of Uber's actual effect. We overcome this potential source of bias by exploiting the spatio-temporal variation in the company's roll-out with DD designs that estimate the difference in the difference between counties with and without Uber before and after treatment.[16] The primary assumption behind the DD strategy is that the difference between treated and control demarcations would have remained constant in the absence of Uber and that conditional on covariates, the introduction date of Uber is orthogonal to unobserved determinants of air quality.

Furthermore, when we estimate a DD design with more than two periods and variation in treat-

[16] Throughout the paper, we make no distinction between Uber black and Uber-X. Treatment only occurs when any Uber service enters the county. However, given that both the demand and supply for Uber black are lower than for Uber-X, table A.5 of the appendix shows the main regression results when considering the introduction date of Uber-X as the treatment trigger. Overall, all point estimates are qualitatively the same between the preferred and the Uber-X specification.

ment timing, the weights used to compute ATEs with standard TWFE-DD can lead to biased estimates (De Chaisemartin and d'Haultfoeuille, 2020).[17] We avoid this source of bias with Callaway and Sant'Anna (2020)'s staggered difference-in-differences methodology CS-DD.[18] In our preferred specification, treated and control counties are demarcations with and without Uber at time t; what Callaway and Sant'Anna refer to as the "never and still not treated" control group. Other options are to restrict the control group to never-treated counties or only consider still not treated units where eventually Uber starts operations. We favor the "never and still not treated" approach, as we do not have a sufficiently large pool of never-treated demarcations. However, point estimates are robust to the other two designs. Our primary CS-DD specification takes the form:

$$AQI_{cyg} = \beta_{eg} Uber_{cy} + \lambda_c + \omega_y + \varepsilon_{ct} \tag{1}$$

where AQI_{cyg} is the AQI in county c at year y for all counties treated at time g. β_{eg} is the point estimate of interest. It marks the ATT for counties in group g at time since treatment e, where e is the difference between the current period and the treatment date, i.e., $e = y - g$. $Uber_{cy}$ is an indicator variable equal to one if Uber is present at year y in county c. Notice that for never treated counties, this variable is always zero. λ_c are county fixed effects controlling for cross-sectional differences between counties, and ω_y are year fixed effects.

To estimate the average treatment effect for each period e, we aggregate β_{eg} according to equation 2. In it, $P[G = g | G + e = \leq T]$ is the probability of being first treated at period g and β_e is the average treatment effect on the treated e periods after treatment. The idea of this estimate is in the vein of a TWFE-DD event study design, with the advantage of avoiding the weighting issues associated

[17] In appendix A.2, we show the existence of this bias with the Goodman-Bacon decomposition (Goodman-Bacon, 2021).

[18] A further advantage of CS-DD is that it allows us to test for the common trends assumption while considering the potential pitfalls of particular treatment timing (Sun and Abraham, 2020).

with these models.

$$\beta_e = \sum_{g \in G} \omega_{gt}^e \beta_{eg} \quad \forall \quad \omega_{gt}^e = 1[g+e \leq T]P[G=g|G+e=\leq T] \tag{2}$$

Next, we determine the ATT across all groups and periods with equation 3. In it, β is the weighted sum of β_{ge} with strictly positive weights and larger weights for larger group sizes. As with equation 3, $\kappa = \sum_{g \in G} \sum_{t=2}^{T} 1[t \geq g]P[G=g|G \leq T]$ ensures that the weights in the second sum are positive and add up to one.

$$\beta = \frac{1}{\kappa} \sum_{g \in G} \sum_{t=2}^{T} \beta_{eg} \tag{3}$$

4. Results

4.1. Average Treatment Effects

Table 1 shows the effect of Uber on the maximum value of the AQI and the number of days of bad air quality across three different specifications of control units; never and still not treated, never treated, and still not treated.[19] The ATT of Uber on the maximum value of the AQI in the preferred "never and still not treated" specification is -10.69 or -7.3% of average pre-treatment values. Concerning the number of days of unhealthy air quality, The ATT suggests a decrease of 2.53 days. Reassuringly, point estimates and significance values hold for the "never treated" and "still not treated" samples. These results provide the first empirical evidence on the causal impact of Uber on urban air quality, and imply that the introduction of Uber improves the average air quality (proxied by the AQI) of treated counties.

[19] Table A.4 in the appendix presents additional results for the median and 90th percentile. Reassuringly, results remain in line with the effect of Uber on the maximum. The same table also presents estimates for the share of bad air quality days. The share of bad air quality days can be different if certain stations fail to report the AQI due to maintenance, malfunctioning, or strategic behavior. As with the median and 90th percentile, results align with the count estimates.

Table 1: Effects of Uber on the air quality index (AQI)

	Never and still not treated		Never treated		Still not treated	
	Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days
	-10.69^{***}	-2.53^{***}	-11.81^{***}	-2.94^{***}	-7.23^{***}	-2.24^{***}
	(2.47)	(0.41)	(3.30)	(0.52)	(2.03)	(0.44)
N.Counties	700	700	700	700	564	564
N.Groups	8	8	8	8	7	7
N.Periods	18	18	18	18	17	17
Parallel trends Wald Test (P-value)	1	1	1	1	1	1

Notes: This table contains the results of the Callaway and Sant'Anna's difference-in-differences (CS-DD) estimates of the impact of Uber on the maximum value of the air quality index (AQI) and the number of days of unhealthy air quality, i.e., days with AQI values greater than one-hundred units. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t. The "never treated" group only includes counties without Uber as of 2017. And the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 2 decomposes the count of bad air quality days into EPA's risk categories, i.e, unhealthy for sensitive groups between 100 and 150, unhealthy between 150 and 200, and hazardous for days higher than 201.

The number of unhealthy air quality days for sensitive groups (between 100 and 150 units) decreases by 2.1 in the preferred specification. Multiplying this estimate by the total number of treated counties (521) leads to 1,094 fewer bad air quality episodes per year. Regarding unhealthy days for the general population (between 150 and 200 units), Uber decreased their average number by 0.5 days or 270 fewer episodes of bad air quality. Finally, we see no significant effects on days with AQI values larger than 201.

4.2. Dynamic and group specific treatment effects

Figure 5 portrays the ATT for each period before and after the introduction date of Uber. Estimates suggest that, although not statistically different from each other, the effect of Uber takes time to

Table 2: Effect of Uber on the number of unhealthy air quality episodes by risk level

	Never and still not treated			Never treated			Still not treated		
	(100-150]	(151-200]	(201-]	(100-150]	(151-200]	(201-]	(100-150]	(151-200]	(201-]
	-2.1^{***}	-0.5^{***}	-0.1	-2.5^{***}	-0.5^{***}	-0.1	-1.9^{***}	-0.4^{**}	-0.1
	(0.4)	(0.1)	(0.1)	(0.4)	(0.1)	(0.1)	(0.4)	(0.1)	(0.1)
N.Counties	702	702	702	702	702	702	566	566	566
N.Groups	8	8	8	8	8	8	7	7	7
N.Periods	18	18	18	18	18	18	17	17	17
Parallel trends Wald Test (P-value)	1	1	1	1	1	1	1	1	1

Notes: This table contains the results of the Callaway and Sant'Anna's difference-in-differences (CS-DD) on the impact of Uber on the number of days with air quality index (AQI) values within three exposure intervals; (100-150], (151-200], and 201+. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t. The "never treated" group only includes counties without Uber as of 2017. And the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects and cluster standard errors at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

materialize.[20] Regarding pre-treatment ATTs, it is reassuring to see overall statistically insignificant coefficients.

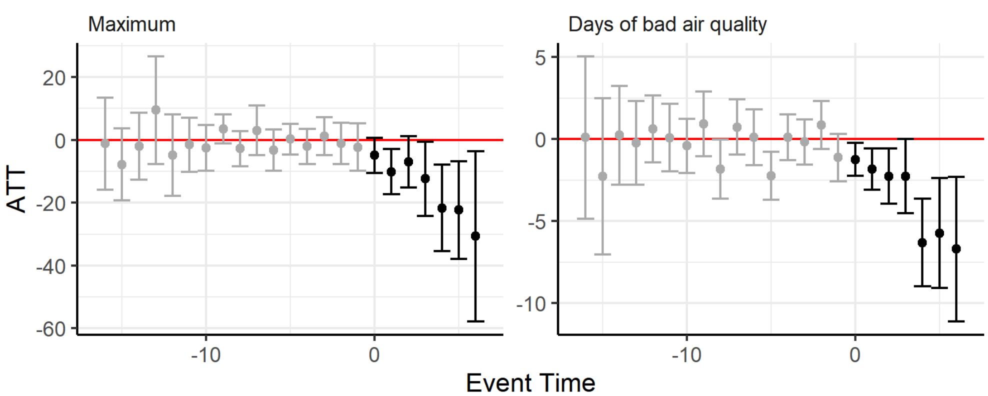

Figure 5: Dynamic estimates for the effect of Uber on the air quality index

Notes: This figure portrays seasonal point estimates and 95% confidence intervals on the impact of Uber on the maximum and ninety percentile values of the air quality index (AQI), as well as on the number of unhealthy air quality episodes, i.e., days with AQI values higher than one hundred units. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The treated group contains all counties where Uber started operations between 2010 and 2017. The control group refers to all counties without Uber at time t. The Callaway and Sant'Anna's difference-in-differences (CS-DD) model controls for county and year fixed effects. Standard errors are clustered at the county level.

[20] Figure 4 of the appendix contains equivalent estimates for the median, 90th percentile, and share of days with bad air quality.

Next, figure 6 contains the ATT for each treatment group, i.e., each group of counties treated in the same year. Although only statistically significant for counties treated between 2010 and 2013, we see negative coefficients across all samples regarding the maximum value of the AQI. Finding higher estimates for early treated counties makes sense as these counties were on average more densely populated, wealthier, and polluted than small demarcations treated from 2014 onwards (See tables A.2 and A.3 of the data section in the appendix). For the number of days of bad air quality, all point estimates are negative, and significant between 2011 and 2013.

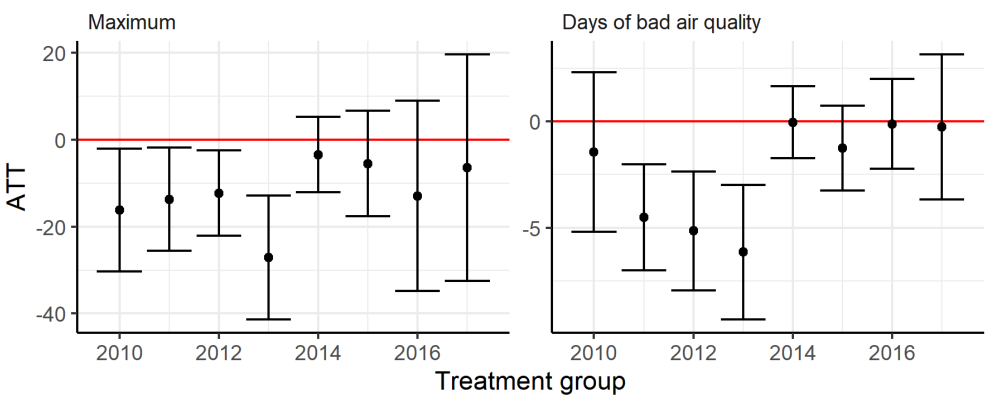

Figure 6: Group Specific Average Treatment Effect on the Treated for the effect of Uber on the air quality index

Notes: This figure portrays group-specific point estimates on the impact of Uber on the maximum-AQI and the number of days of bad air quality. Each group corresponds to all counties where Uber started operations in the same year. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The vertical lines are 95% confidence intervals. The treated group contains all counties where Uber started operations between 2010 and 2017. The control group refers to all counties without Uber at time t. The Callaway and Sant'Anna's difference-in-differences (CS-DD) model controls for county and year fixed effects. Standard errors are clustered at the county level.

4.3. Seasonal effects

Seasonal estimates can provide information on the mechanisms through which Uber improves air quality. For instance, ozone's seasonal behavior deviates from other criteria pollutants because of its dependence on solar radiation. If the air quality improvement relates to O_3, we should uncover higher effects during the summer months.

Equation 4 shows the empirical strategy behind the seasonal estimates. In it, AQI^s_{cyg} is the maximum value of the AQI for county c part of group g at time y and season s. The estimate of interest, β^s_{eg}, conveys the seasonal effect of Uber on the AQI.

$$AQI^s_{cyg} = \beta^s_{eg} Uber_{cy} + \lambda_c + \omega_y + \varepsilon_{ct} \tag{4}$$

Figure 7 shows the seasonal effects of Uber on the maximum value of the AQI and the number of bad air quality days.

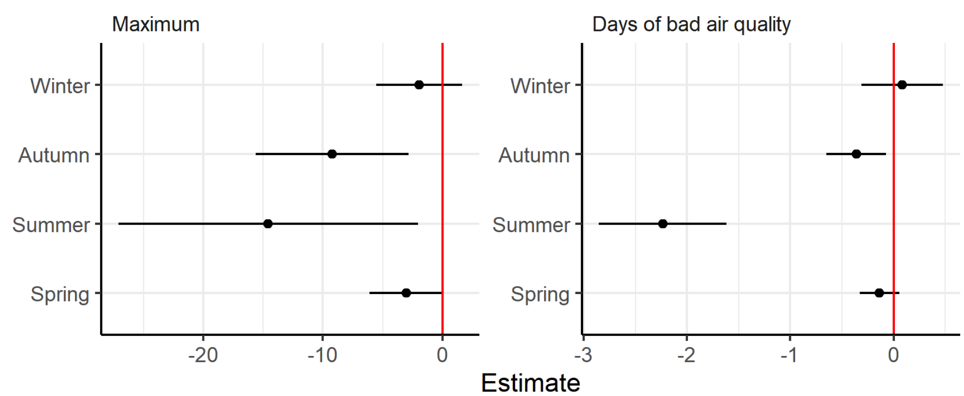

Figure 7: Seasonal effects of Uber on the air quality index (AQI)

Notes: This figure portrays the seasonal results of Callaway and Sant'Anna's difference-in-differences (CS-DD) design on the impact of Uber on the maximum AQI and the number of unhealthy air quality episodes, i.e., days with AQI values higher than 100 units. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. Treated and control counties refer to counties with and without Uber at time t. The CS-DD controls for county and year fixed effects. Standard errors are clustered at the county level.

Uber decreases the AQI during the spring, summer, and fall months by 3.02, 14.56, and 9.24 units. Notably, although not statistically different from the other seasons, the summer estimate is relatively larger. Concerning the number of days of bad air quality, estimates suggest a statistically significant reduction for autumn and summer. During the summer, bad air quality episodes decrease by 2.2 days. This summer reduction leads to 2,297 fewer summer days of bad air quality across all treated counties. Interestingly, the summer estimate is statistically larger than the estimates for the other seasons, suggesting that O_3 does play a role in the air quality improvement.[21]

4.4. Heterogeneous effect by air pollutant

This section estimates the effect of Uber on the concentration of the two most harmful criteria contaminants; O_3 and PM_{25}. Examining the impact of Uber on O_3 will allow us to disentangle if air quality effect comes from changes in this important secondary air contaminant. Furthermore, because changes in PM_{25} are closely associated with variations in traffic-related emissions, looking at their fluctuation allows us to disentangle the transit mechanism. Finally, because of the critical relationship between NO_2 and O_3, we also provide estimates for the latter. Table 3 contains the results of each pollutant-specific regression.

Table 3: Effect of Uber on the maximum AQI for selected contaminants

	Never and still not treated			Never treated			Still not treated		
	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}
	-1.23 (2.61)	-6.24^{***} (1.70)	-3.17 (2.78)	-0.93 (3.04)	-5.90^{**} (1.87)	-4.93 (3.44)	-0.42 (2.47)	-7.32^{***} (1.80)	2.68 (2.31)
N.Counties	249	565	533	249	565	533	164	485	423
N.Groups	8	8	8	8	8	8	7	7	7
N.Periods	18	18	18	18	18	18	17	17	17
Parallel trends Wald Test (P-value)	1	1	1	1	1	1	1	1	1

Notes: This table contains the results of the Callaway and Sant'Anna's difference-in-differences (CS-DD) estimates of the impact of Uber on the maximum value of the air quality index (AQI) across air pollutants. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t. The "never treated" group only includes counties without Uber as of 2017. And the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

We only find significant point estimates for O_3. In the preferred specification, Uber decreases the AQI of ozone by 6.24 units.[22] Conversely, we find no statistically significant effects for NO_2 or

[21] Figure 5 in the appendix portrays point estimates and 95% confidence intervals for the EPA risk categories.
[22] In additional results, table A.6 of the appendix shows the effect on the number of days of bad air quality;

PM_{25}.

Next, figure 8 shows the dynamic ATT for the maximum AQI value of O_3, PM_{25}, and NO_2. Post-treatment estimates for O_3 are all negative with significant coefficients for $e \in 1,2,4$. For PM_{25}, there is also suggestive evidence of a reduction, possibly due to reductions in cold-start emissions (Ward et al., 2021a). Finally, for NO_2 we do not find any significant effect, which may be due to the small number of stations measuring NO_2 across the US.

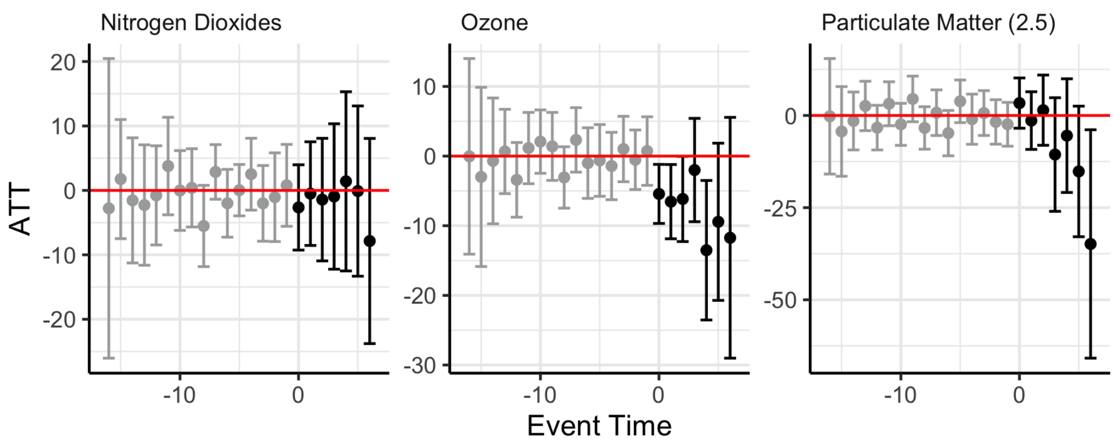

Figure 8: Dynamic average treatment effects of Uber

Notes: This figure portrays point estimates and 95% confidence intervals of Callaway and Sant'Anna's difference-in-differences (CS-DD) design on the impact of Uber on the maximum air quality index (AQI) value for nitrogen dioxides (NO_2), ground-level ozone (O_3), and fine particulate matter (PM_{25}). The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The treated and control groups contain all counties with and without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level.

4.5. Robustness checks

4.5.1. Regional heterogeneity

Even though we see an average reduction in the AQI, this does not necessarily mean that Uber improves the air quality for all treated counties. In this section we explore if our results hold across

results show that the only particle with fewer days of bad air quality is O_3. Table A.7 shows the impact of Uber on CO, PM_{10}, and SO_2; we find no significant estimate for any of these contaminants. Finally, table A.8 presents results for the 90th percentile value of O_3, NO_2, and PM_{25}; as expected, the only significant effect relates to O_3.

US census regions. [23] The map in figure 9 shows the point estimates and standard errors of running our preferred specification for each region.

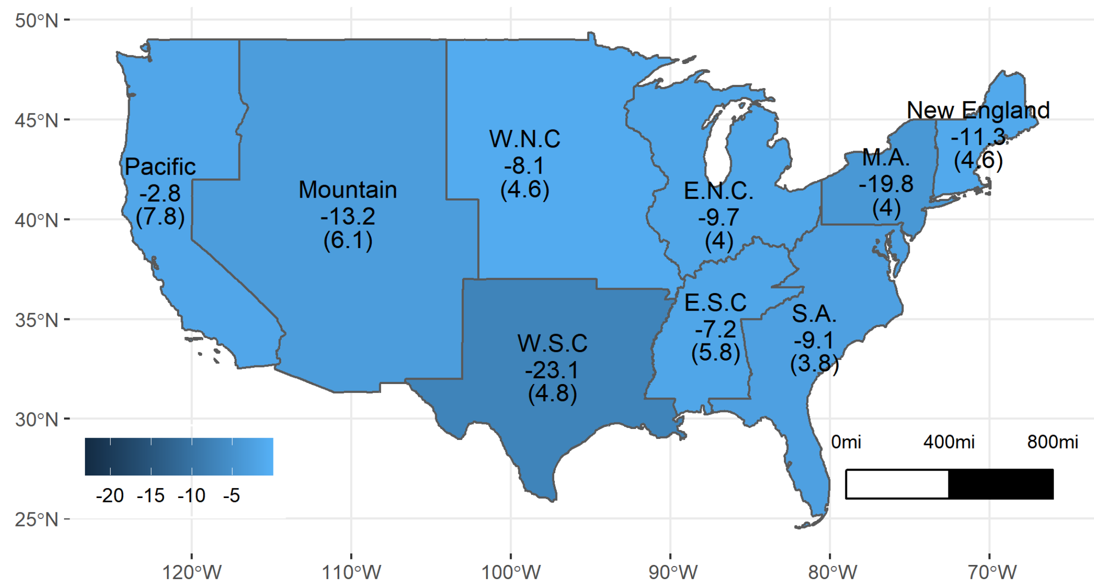

Figure 9: Effect of Uber on the maximum value of the air quality index across census regions

Notes: This map portrays point estimates and standard errors (in parenthesis) of a Callaway and Sant'Anna's difference-in-differences (CS-DD) design on the impact of Uber on the maximum-air quality index (AQI) across all nine US census regions. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The treated and control groups are urban counties with and without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level.

Results show reductions in the AQI across all census regions. Specifically, we find five percent significant declines in New England, Mid-Atlantic, South Atlantic, East North Central, West North Central, and Mountain, ten percent for West North Central, and statistically insignificant effects for the Pacific and East South Central Regions. Thus, even though we find no qualitative differences in the impact of Uber on air quality, we encounter differences in the intensity and statistical significance of point estimates. For instance, the effect of Uber appears to be considerably higher in the Mid-Atlantic and West South Central states, suggesting that we need further research on the regional or city-specific mechanisms behind the relationship between ride-hailing technologies and

[23] Specifically, the US Census Buro divides the country into nine regions; New England, Mid-Atlantic (MA), South-Atlantic (SA), East North Central (ENC), East South Central (ESC), West North Central (WNC), West South Central (WSC), Mountains, and Pacific.

air quality. Yet, finding consistently negative estimates across census regions reduce concerns that high reductions in a subset of urban agglomerations drive our results.[24]

4.5.2. Unobserved confounders

Even though we cannot reject the common trends assumption with our main sample, this does not necessarily imply that the common trends assumption holds after treatment. For instance, unobservable covariates can potentially bias point estimates if they systematically correlate with the introduction date of Uber. This section examines the robustness of our results to three well-known sources of air pollution changes in the US; changes in the fleet of fossil-fuel power plants, forest fires, and the violation of NAAQS.[25] For this, we exclude from the treatment and control groups all periods after a county reported changes in its power plants fleet, had a forest fire of more than 2,000 acres, or violated the North American Air Quality Standards (NAAQS). Furthermore, we also show a specification excluding neighboring demarcations because of the spread-out behavior of most air pollutants. Table 4 shows the results of each robustness exercise.

Excluding counties with changes in the power fleet, forest fires, or violations of NAAQS leads to qualitatively similar results to the main specification. Furthermore, table A.9 of the appendix presents the same estimates for the count of days with AQI values larger than 100 units. Overall, they mimic the estimates from table 4.

[24]Figure 6 of the appendix portrays the same exercise with the number of days of bad air quality. Point estimates remain negative although less significant than for the maximum AQI because there is more limited variation in the number of bad air quality days than in the raw AQI.

[25]The Energy Information Administration (EIA) provides detailed information on the status of power generating units (EIA, 2021). We extract forest-fire data from the US Department of Agriculture (USDA, 2021) and the United States Geological Service (USGS, 2021). For NAAQS violations, we use the EPA's Green Book (EPA, 2021b).

Table 4: Effect of Uber on the air quality index (AQI) for counties with no power fleet changes, forest fires larger than 2,000 acres, or violations of NAAQS

	Power fleet changes		Forest fires		O_3 violations of NAAQS		PM_{25} violations of NAAQS	
	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties
	-18.33^{*}	-5.62^{*}	-19.63^{*}	-34.90^{*}	-7.17^{*}	-8.93	-12.87^{***}	-16.86^{***}
	(8.60)	(2.48)	(9.01)	(16.67)	(3.42)	(4.63)	(3.67)	(4.00)
N.Counties	697	693	700	687	698	698	698	696
N.Groups	8	8	8	8	8	8	8	8
N.Periods	18	18	18	18	18	18	18	18

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber on the maximum value of the air quality index (AQI). Treated and control counties are those with and without Uber at time t. We provide results for three different samples: Power fleet changes exclude all counties reporting a change in their fleet of fossil-fuel power plants; Forest fires exclude all counties that reported a forest fire larger than 2,000 acres within our observation period; and NAAQS violations excludes all counties that violated NAAQS. The CS-DD model controls for county and year fixed effects and cluster standard errors at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

4.5.3. Conditional parallel trends

This section departs from the unconditional parallel trends assumption we kept throughout the study and shows results conditioning the trends on average pre-treatment income, population density, and temperature. Conditioning on observables is relevant if we believe that covariate specific time-trends are modifying the value of the AQI. In our case, first-treated counties are, on average, more densely populated and wealthier than latter-treated demarcations. If the inter-temporal path of the AQI depends on either income or population density, conditioning the parallel trends on these covariates is potentially better. Table 5 shows the point estimates of the conditional trends models for the maximum AQI and the number of unhealthy air quality days.

Reassuringly, conditioning the parallel trends on income and temperature leads to point estimates statistically equivalent to the preferred specification. Conditioning on population density, on the other hand, does decrease the size of the coefficient for the maximum value of the AQI. However, the qualitative effect is still negative and statistically different from zero.

Table 5: Effect of Uber on the air quality index (AQI) with conditional parallel trends

	Conditional on income		Conditional on Pop. Density		Conditional on Temperature	
Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days	
	-10.51^{***}	-2.89^{***}	-4.68^{*}	-1.85^{**}	-13.05^{***}	-3.05^{***}
	(2.66)	(0.62)	(2.24)	(0.68)	(2.84)	(0.56)
N.Counties	694	694	700	700	680	680
N.Groups	8	8	8	8	8	8
N.Periods	18	18	18	18	14	14

Notes: This table contains the results of the Callaway and Sant'Anna's difference-in-differences (CS-DD) estimates of the impact of Uber on the maximum value of the AQI and the number of days of unhealthy air quality, i.e., days with AQI values greater than one-hundred units. We provide results conditioning parallel trends on three different variables; income, population density, and average temperature. The treated and control groups are counties with and without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

4.5.4. Two ways fixed effects

In this section, we look at our estimates robustness to the more typical TWFE-DD design. Using TWFE-DD to estimate the effect of Uber on air quality has the advantage that it allows us to incorporate time-varying covariates.

Equation 5 shows the econometric specification of the TWFE-DD model. In it, AQI_{cy} is the value of the AQI (or the number of days of bad air quality) for county c at year y. Uber$_{cy}$ is an indicator dummy equal to one if Uber operates in county c at time y. β contains the point estimate of interest, i.e., the effect of Uber on air quality. Ψ_{cy} controls for exogenous shocks to the air quality index like forest fires, violations of NAAQS, and power plants openings and closures. W_{cy} further adds weather controls in the form of temperature, relative humidity, wind speed, and atmospheric pressure. Finally, λ_c and ω_y are county and year fixed effects.

$$AQI_{cy} = \beta_1 Uber_{cy} + \delta\Psi_{cy} + \gamma W_{cy} + \lambda_c + \omega_y + \varepsilon_{ct} \tag{5}$$

Table 6 contains the results of the TWFE-DD model across three specifications: (1) only controls for county and year fixed effects; (2) adds the matrix of exogenous shocks to the model; and (3) includes weather covariates. It is comforting to see that the effect of Uber on the AQI is not statistically different across specifications. Furthermore, point estimates are also quantitatively equivalent to the CS-DD model. In the raw specification, Uber decreases the maximum value of the AQI by 9.06 units and the number of bad air quality days by 3.26.

Table 6: Two-way fixed effects difference-in-differences (TWFE-DD) estimates on the effect of Uber on the Air quality index (AQI)

	Maximum Air Quality Index			Bad Air Quality Days		
	(1)	(2)	(3)	(1)	(2)	(3)
ATT	-9.06^{***}	-7.70^{***}	-8.13^{***}	-3.26^{***}	-1.74^{**}	-1.93^{***}
	(1.97)	(1.96)	(1.84)	(0.59)	(0.59)	(0.44)
Power Plant Closure		-0.30	-0.88		0.92	0.70
		(1.75)	(1.64)		(0.79)	(0.66)
Power Plant Opening		15.98^{*}	13.20		5.52^{**}	7.00^{***}
		(8.08)	(7.72)		(2.05)	(0.97)
Forest Fire		-0.09	-1.30		0.98^{*}	0.67
		(1.71)	(1.80)		(0.40)	(0.36)
NAAQS Violations		-6.91^{***}	-4.68^{*}		-8.18^{***}	-7.39^{***}
		(2.00)	(1.90)		(0.95)	(0.84)
Temperature			0.29^{*}			0.07^{*}
			(0.12)			(0.03)
No.Obs	11,192	11,192	8,012	11,192	11,192	8,012

Notes: This table contains the results of the Two-way fixed effects difference-in-differences (TWFE-DD) estimates of the impact of Uber on the maximum value of the AQI and the number of days of unhealthy air quality, i.e., days with AQI values greater than one-hundred units. Column (1) only controls for county and year fixed effects. Column (2) adds the matrix of exogenous shocks to the model, i.e., forest fires, violations of NAAQS, and changes in the composition of the power plant's fleet. And column (3) includes weather covariates in the form of temperature, relative humidity, wind speed, and atmospheric pressure. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Another interesting result is the decrease in the AQI after a county violates NAAQS, confirming previous literature on the effectiveness of policies related to the America's Clean Air Act (Currie and Walker, 2019). Additionally, we see a five and seven percent significant increase in the second and third specifications after opening new power facilities.

5. Conclusion

Because of the complexity of the transportation system and the intricate relationship between air pollutants in the lower atmosphere, it is challenging to assess the air quality effects of Uber. In this study, we estimate this effect by looking at Uber's impact on the AQI, a well-known proxy for average air quality that incorporates information on the concentration of all six criteria pollutants measured by the EPA.

We infer causality by leveraging the spatio-temporal variation in the introduction date of Uber with CS-DD designs. In layman's terms, we compare the value of the AQI on treated and non-treated demarcations before and after Uber started operations.

Our findings show that Uber improves air quality mainly through its effect on ozone (although we also find suggestive evidence for a decrease in PM_{25}). We perform several robustness tests to check the stability of our results across different control samples, regions, and specifications where we exclude all counties reporting forest fires, power plant closures, and violations of NAAQS. Additionally, our estimates are also robust to TWFE-DD designs and conditioning the parallel trends with income, population density, and temperature.

These results stand contrary to current political claims on the adverse effects of Uber on air pollution. The reason for this discrepancy is that intuitively, one could think that air pollution increases with traffic. However, this is not necessary because of three main reasons. First, more Uber cars do not imply higher emissions when these new vehicles substitute high emitting taxis or private cars. Second, even if Uber replaces some share of public transportation rides, the pollution question remains open if the effect of this reduction on air pollution is less significant than the influence of substituting old taxis and private vehicles with newer Uber cars. Third, even if Uber increases the concentration of combustion contaminants like nitrogen dioxides, the inverse relationship between these particles and atmospheric ozone in urban agglomerations could improve air quality in regions with high ozone levels.

Our findings provide a holistic picture of the air quality effects of Uber with several relevant policy implications for the regulation of these disruptive technologies. We recommend that researchers and policymakers be careful when using the findings of transportation studies on the effects of ride-hailing technologies as evidence of their pernicious effects on air quality. In the end, its impact on the transportation system is only one piece of a more complex puzzle. Future studies could expand this work by looking at the air pollution effects of introducing electric vehicles for ride-hailing services, the impact of these technologies on greenhouse gas emissions, and heterogeneous effects across different urban centers.

References

Bill de Blassio (2019) Curbing city traffic and air pollution: Mayor de Blasio on reforming how Uber and Lyft operate. URL: `https://www.nydailynews.com/opinion/ny-oped-fhv-20190816-mdkgwdn7fbhoxjovdhqlhyi3wa-story.html`.

Callaway B. and Sant'Anna P. (2020) Difference-in-Differences with multiple time periods. Journal of Econometrics.

Clewlow R. R. and Mishra G. S. (2017) Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. Institute of Transportation Studies, University of California.Technical report, Davis, Research Report UCD-ITS-RR-17-07.

Currie J. and Walker R. (2019) What do economists have to say about the Clean Air Act 50 years after the establishment of the Environmental Protection Agency? Journal of Economic Perspectives, 33(4): 3–26.

De Chaisemartin C. and d'Haultfoeuille X. (2020) Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9): 2964–96.

EIA (2021) Preliminary Monthly Electric Generator Inventory (based on Form EIA-860M as a supplement to Form EIA-860). URL: `https://www.eia.gov/electricity/data/eia860m/`.

EPA (2021a) Air Data: Air Quality Data Collected at Outdoor Monitors Across the US). URL: `https://www.epa.gov/outdoor-air-quality-data`.

——— (2021b) Non-attainment Areas for Criteria Pollutants (Green Book). URL: `https://www.epa.gov/green-book`.

Feigon S. and Murphy C. (2016) Shared mobility and the transformation of public transit.Technical Report Project J-11, Task 21, American Public Transportation Association.

Goodman-Bacon A. (2021) Difference-in-differences with variation in treatment timing. Journal of Econometrics.

Hampshire R., Simek C., Fabusuyi T., Di X., and Chen X. (2017) Measuring the impact of an unanticipated disruption of Uber/Lyft in Austin, TX. Lyft in Austin, TX (May 31, 2017).

Holland S. P., Mansur E. T., Muller N. Z., and Yates A. J. (2016) Are there environmental benefits from driving electric vehicles? The importance of local factors. American Economic Review, 106(12): 3700–3729.

Keating D. (2019) Uber Adding To Air Pollution In Europe. URL: `https://www.forbes.com/sites/davekeating/2019/11/20/uber-adding-to-air-pollution-in-europereport/?sh=7f60dd1d5041`.

Lalive R., Luechinger S., and Schmutzler A. (2018) Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management, 92: 744–764.

Rayle L., Shaheen S., Chan N., Dai D., and Cervero R. (2014) App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in San Francisco University of California Transportation Center (UCTC). University of California, Berkeley, United States.

Sarmiento L., Wägner N., and Zaklan A. (2021) Effectiveness, Spillovers, and Well-Being Effects of Driving Restriction Policies. RFF Discussion Paper.

Schaller B. (2017) The Growth of App-Based Ride Services and Traffic, Travel and the Future of New York City.

Sun L. and Abraham S. (2020) Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics.

Uber (2021) Use Uber in cities around the world. URL: `https://www.uber.com/global/en/cities/`.

U.S. Census Bureau (2015-2019) Means of Transportation to Work American Community Survey 5-year estimates. URL: `https://censusreporter.org`.

——— (2020) About Micropolitan and Metropolitan Statistical Areas. URL: `https://www.census.gov/programs-surveys/metro-micro/about.html`.

USDA (2021) National Interagency Fire Occurrence 1992-2015. URL: `https://catalog.data.gov/ne/dataset/national-interagency-fire-occurrence-1992-2015-feature-layer-54440`.

USGS (2021) Historic Fire Data. URL: `https://rmgsc.cr.usgs.gov/outgoing/GeoMAC/historic_fire_data/`.

Ward J. W., Michalek J. J., and Samaras C. (2021a) Air Pollution, Greenhouse Gas, and Traffic Externality Benefits and Costs of Shifting Private Vehicle Travel to Ridesourcing Services. Environmental Science & Technology.

Ward J. W., Michalek J. J., Samaras C., Azevedo I. L., Henao A., Rames C., and Wenzel T. (2021b) The impact of Uber and Lyft on vehicle ownership, fuel economy, and transit across US cities. Iscience, 24(1): 101933.

Yan X., Levine J., and Zhao X. (2019) Integrating ridesourcing services with public transit: An evaluation of traveler responses combining revealed and stated preference data. Transportation Research Part C: Emerging Technologies, 105: 683–696.

Acknowledgments

We truly appreciate all the constructive comments and suggestions from seminar participants at RFF-CMCC European Institute on Economics and the Environment. We are very grateful for constructive comments and feedback from Jacopo Bonan. Also, we are gratefully acknowledged the financial support from the H2020 ERC Starting Grant funded by European Commission (#853487).

A. Appendix

For Online Publication

A.1. Introduction

Table A.1: Effect of Uber on the share of public transport and total number of available private cars

	Share of public transit commuters			Total number of available private cars		
	Never and still not treated	Never treated	Still not treated	Never and still not treated	Never treated	Still not treated
	0.08	0.12*	0.08	27,522.08***	24,685.04***	30,867.30***
	(0.05)	(0.06)	(0.05)	(5714.24)	(6605.90)	(4481.16)
N.Counties	700	700	564	538	538	479
N.Groups	8	8	7	7	7	6
N.Periods	18	18	17	8	8	7
Parallel trends Wald Test (P-value)	1	1	1	1	1	1

Notes: This table contains the results of the Callaway and Sant'Anna (2020)'s difference-in-differences design (CS-DD) on the impact of Uber on the share of workers using public transit for their daily commute (left) and the total count of available private cars per county (right). Both variables come from the American Community Survey. Treated counties are those where Uber started operations between 2010 and 2017. Control counties are those without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

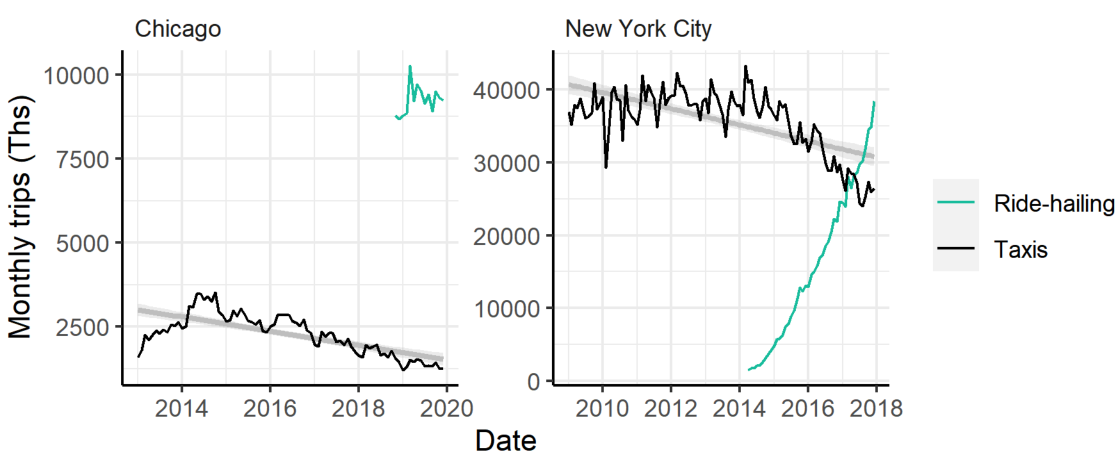

Figure 1: Number of monthly trips in Chicago and New York City

A.2. Empirical design

Goodman-Bacon (2021) shows that the TWFE-DD coefficient is a weighted average of all possible group-period DD estimators across three groups; earlier vs. later treated, later vs. earlier treated, and treated vs. untreated. Figure 2 depicts the decomposition proposed by Goodman-Bacon (2021) on the effect of Uber on the maximum value of the air quality index. The vertical axis shows the estimates for each 2x2 DD and its corresponding weights on the horizontal axis.[26] The solid horizontal lines depict the TWFE-DD estimates of each comparison and the dotted lines the weighted average TWFE-DD estimate.

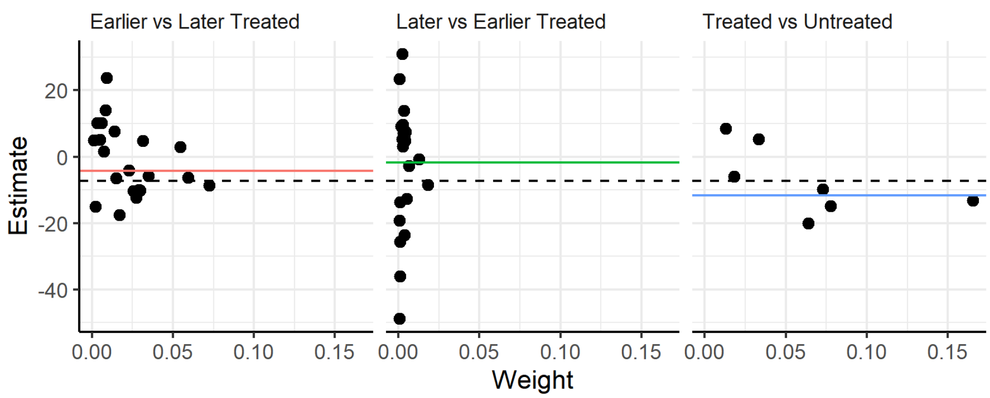

Figure 2: Goodman-Bacon decomposition on the effects of Uber on the air quality index

The figure shows that TWFE-DD the comparison of later vs. earlier and earlier vs. later treated counties can potentially bias the TWFE-DD estimates. For instance, in retrospect, we can see that the positive point estimates from later vs. earlier treated arise because already-treated counties experience substantial decreases in pollutant levels years after Uber. Using their post-treatment outcomes as a control group for stations treated in later years underestimates the true Uber impact on the Later Treated.

[26]Each estimate weight comes from the size of the treatment group for each 2x2 DD comparison.

A.3. Data

AQI Basics for Ozone and Particle Pollution

Daily AQI Color	Levels of Concern	Values of Index	Description of Air Quality
Green	Good	0 to 50	Air quality is satisfactory, and air pollution poses little or no risk.
Yellow	Moderate	51 to 100	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.
Orange	Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is less likely to be affected.
Red	Unhealthy	151 to 200	Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects.
Purple	Very Unhealthy	201 to 300	Health alert: The risk of health effects is increased for everyone.
Maroon	Hazardous	301 and higher	Health warning of emergency conditions: everyone is more likely to be affected.

Figure 3: Levels of concern for the general population based on the value of the air quality index (AQI)

Notes: `https://www.airnow.gov/aqi/aqi-basics/`

Table A.2: Descriptive statistics of air quality index (AQI) across treatment and control groups

	2010	2011	2012	2013	2014	2015	2016	2017	Avg. Treated	Control
Full sample										
Maximum AQI	137	145	167	148	142	134	137	125	142	120
90th Perc. AQI	68	72	82	82	76	71	70	65	73	67
%Unhealthy Days	3	4	6	6	5	3	3	2	4	3
Pre-treatment										
Maximum AQI	140	156	177	160	149	137	140	125	148	-
90th Perc. AQI	68	78	89	89	80	73	72	66	77	-
%Unhealthy Days	3	5	7	7	6	4	4	2	5	-
Post-treatment										
Maximum AQI	132	128	147	116	118	120	113	114	124	-
90th Perc. AQI	68	62	70	62	62	58	55	55	62	-
%Unhealthy Days	2	2	3	1	2	1	1	1	2	-
N.obs	144	1,009	1,007	999	3,807	1,073	977	1,039	1,257	1,499
N.Counties	8	58	57	56	219	64	59	64	73	140
Pre-treatment Periods	10	11	12	13	14	15	16	17	-	-
Post-treatment Periods	8	7	6	5	4	3	2	1	-	-

Notes: This table shows the average of the maximum and ninety percentile values of the AQI, as well as the share of days when counties report episodes of unhealthy air quality for one control and eight treatment groups. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. Unhealthy episodes refer to days with AQI values beyond 100 units. Each treatment group contains all counties where Uber began operations in that particular year; e.g., the 2010 treatment group only includes counties treated in 2010. The control group refers to all counties without Uber as of 2017. Pre-treatment and Post-treatment values relate to the average in treated counties before and after the introduction date of Uber.

Table A.3: Descriptive Socio-demographic characteristics across treatment and control groups

	2010	2011	2012	2013	2014	2015	2016	2017	Avg. Treated	Control
Full sample										
Income per capita (Ths.)	85	65	53	51	47	44	42	45	54	42
Population density	2,994	5,934	1,204	625	494	294	204	210	1,495	247
Pub. Trans. Index	11	13	3	1	1	1	1	1	4	1
Pre-treatment										
Income per capita (Ths.)	85	65	53	51	47	44	42	45	54	-
Population density	2,994	5,893	1,215	625	495	296	205	211	1,492	-
Pub. Trans. Index	11	13	3	1	1	1	1	1	4	-
Post-treatment										
Income per capita (Ths.)	85	66	53	51	47	44	42	45	54	-
Population density	2,994	6,000	1,182	624	492	286	195	198	1,496	-
Pub. Trans. Index	11	13	3	1	1	1	1	1	4	-

Notes: This table shows the income per capita, population density, and public transportation index for one control and eight treatment groups. The public transportation index indicates the percentage of persons commuting by public transit (U.S. Census Bureau, 2015-2019). Each treatment group contains all counties where Uber began operations in that particular year; e.g., the 2010 treatment group only includes counties treated in 2010. The control group refers to all counties without Uber as of 2017. Pre-treatment and Post-treatment values relate to the average in treated counties before and after the introduction of Uber.

A.4. Results

Table A.4: Effects of Uber on the air quality index (AQI) for additional variables

	Never and still not treated			Never treated			Still not treated		
	Median AQI	90th percentile AQI	% of bad days	Median AQI	90th percentile AQI	% of bad days	Median AQI	90th percentile AQI	% of bad days
	-0.9^{*}	-4.2^{***}	-0.8^{***}	-1.6^{***}	-5.8^{***}	-0.9^{***}	-0.2	-2.5^{***}	-0.7^{***}
	(0.4)	(0.7)	(0.1)	(0.6)	(1.0)	(0.1)	(0.4)	(0.7)	(0.1)
N.Counties	700	700	700	700	700	700	564	564	564
N.Groups	8	8	8	8	8	8	7	7	7
N.Periods	18	18	18	18	18	18	17	17	17
W.test	1	1	1	1	1	1	1	1	1

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences design (CS-DD) estimates of the impact of Uber on the median and 90th percentile value of the AQI as well as on the share of days with unhealthy air quality levels, i.e., days with AQI values greater than one-hundred units. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t. The "never treated" group only includes counties without Uber as of 2017. And the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.5: Effects of Uber on the air quality index (AQI) (Treatment trigger restricted to the introduction date of Uber X)

	Never and still not treated		Never treated		Still not treated	
	Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days	Maximum AQI	Unhealthy Days
	-10.00^{***}	-1.71^{***}	-12.69^{***}	-2.41^{***}	-5.00	-1.17^{**}
	(2.67)	(0.44)	(3.61)	(0.56)	(2.73)	(0.44)
N.Counties	702	702	702	702	564	564
N.Groups	8	8	8	8	7	7
N.Periods	18	18	18	18	17	17
Parallel trends Wald Test (P-value)	1	1	1	1	1	1

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber-X on the maximum value of the AQI and the number of days of unhealthy air quality, i.e., days with AQI values greater than one-hundred units. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber-X at time t, the "never treated" group only includes counties without Uber-X as of 2017, and the "still not treated" group only contains eventually treated counties without Uber-X at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

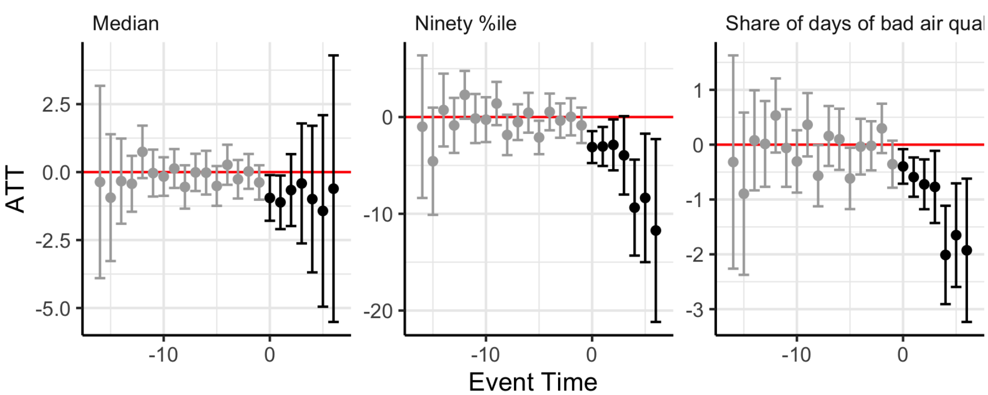

Figure 4: Dynamic estimates for the effect of Uber on the air quality index (AQI)

Notes: This figure portrays event-time point estimates on the impact of Uber on the median and 90th percentile value of the AQI as well as on the share of days with unhealthy air quality levels, i.e., days with AQI values greater than one-hundred units. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The vertical lines are 95% confidence intervals. The treated group contains all counties where Uber started operations between 2010 and 2017. The control group refers to all counties without Uber at time t. The CS-DD model comes from the methodology outlined in Callaway and Sant'Anna (2020), and controls for county and year fixed effects. Standard errors are clustered at the county level.

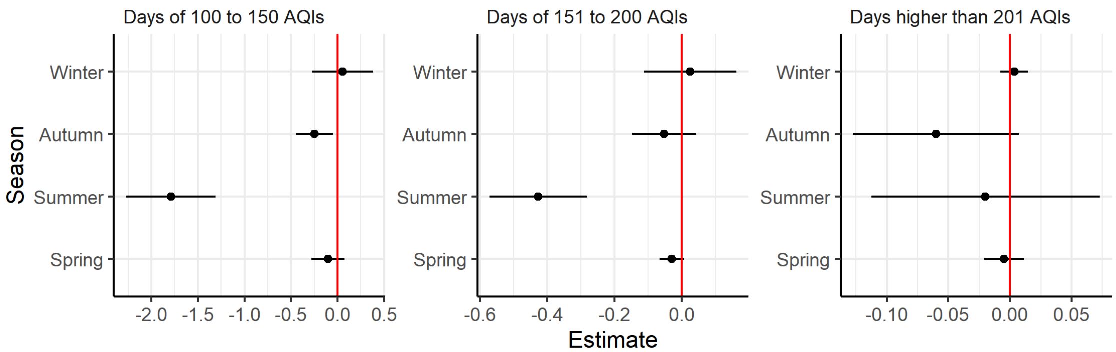

Figure 5: Seasonal effect of Uber on the number of days of bad air quality by risk level

Notes: This figure portrays the seasonal results of the Callaway and Sant'Anna's difference-in-differences (CS-DD) design on the impact of Uber on the number days of unhealthy air quality episodes. The air quality index (AQI) standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. Unhealthy air quality episodes refer to days with AQI values beyond one hundred. The treated group contains all counties where Uber started operations between 2010 and 2017. The control group refers to all counties without Uber as of 2017. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level.

Table A.6: Effect of Uber on the number of days of bad air quality for selected contaminants

	Never and still not treated			Never treated			Still not treated		
	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}
	0.02	-2.40^{***}	-0.37	0.01	-2.61^{***}	-0.61	0.01	-2.40^{***}	0.13
	(0.03)	(0.40)	(0.31)	(0.04)	(0.46)	(0.36)	(0.03)	(0.37)	(0.35)
N.Counties	249	565	533	249	565	533	164	485	423
N.Groups	8	8	8	8	8	8	7	7	7
N.Periods	18	18	18	18	18	18	17	17	17
Parallel trends Wald Test (P-value)	1	1	1	1	1	1	1	1	1

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber on the number of days of bad air quality, i.e., days with an air quality index (AQI) value higher than 100 units. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t, the "never treated" group only includes counties without Uber as of 2017, and the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.7: Effect of Uber on the maximum air quality index (AQI) for selected contaminants

	Never and still not treated			Never treated			Still not treated		
	CO	PM_{10}	SO_2	CO	PM_{10}	SO_2	CO	PM_{10}	SO_2
	−4.52	−31.50	−9.11	−4.68	−32.18	−13.28	−2.13	−27.34	−4.65
	(7.94)	(24.66)	(7.10)	(8.26)	(25.81)	(7.65)	(5.29)	(27.49)	(6.70)
N.Counties	190	366	312	190	366	312	45	160	147
N.Groups	7	8	8	7	8	8	6	7	7
N.Periods	18	18	18	18	18	18	17	17	17
Parallel trends Wald Test (P-value)	0	1	1	0	1	1	1	1	1

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber on the maximum value of the AQI for selected contaminants. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t, the "never treated" group only includes counties without Uber as of 2017, and the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.8: Effect of Uber on the 90th percentile value of the air quality index (AQI) for selected contaminants

	Never and still not treated			Never treated			Still not treated		
	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}	NO_2	O_3	PM_{25}
	−3.12	-3.83^{***}	−0.20	−3.85	-4.73^{***}	−0.64	−1.00	-2.97^{***}	0.95
	(1.93)	(0.81)	(0.69)	(2.00)	(1.15)	(0.88)	(1.94)	(0.80)	(0.75)
N.Obs	249	565	533	249	565	533	164	485	423
N.Groups	8	8	8	8	8	8	7	7	7
N.Periods	18	18	18	18	18	18	17	17	17

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber on the 90th percentile value of the AQI for selected contaminants. We provide results for three different control groups. The "never and still not treated" group encompasses all counties without Uber at time t, the "never treated" group only includes counties without Uber as of 2017, and the "still not treated" group only contains eventually treated counties without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.9: Effect of Uber on the number of bad air quality days for counties with no power fleet changes, forest fires larger than 2,000 acres, or violations of NAAQS

	Power fleet changes		Forest fires		O_3 violations of NAAQS		PM_{25} violations of NAAQS	
	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties	Rep. County	Rep. and Neighboring Counties
	-3.05^{***}	-1.23^{**}	-3.26^{***}	-5.00^{***}	-1.11^{*}	-1.07	-2.20^{***}	-2.64^{***}
	(0.71)	(0.47)	(0.63)	(1.13)	(0.51)	(0.58)	(0.49)	(0.55)
N.Counties	697	693	700	688	698	698	698	696
N.Groups	8	8	8	8	8	8	8	8
N.Periods	18	18	18	18	18	18	18	18

Notes: This table contains the results of Callaway and Sant'Anna (2020)'s difference-in-differences (CS-DD) estimates of the impact of Uber on the number of days of bad air quality, i.e., days with an air quality index (AQI) value higher than 100 units. Treated and control counties are those with and without Uber at time t. We provide results for three different samples: Power fleet changes exclude all counties reporting a change in their fleet of fossil-fuel power plants; Forest fires exclude all counties that reported a forest fire larger than 2,000 acres within our observation period; and NAAQS violations excludes all counties that violated North American Air Quality Standards (NAAQS). The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level. Significance levels denoted by $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

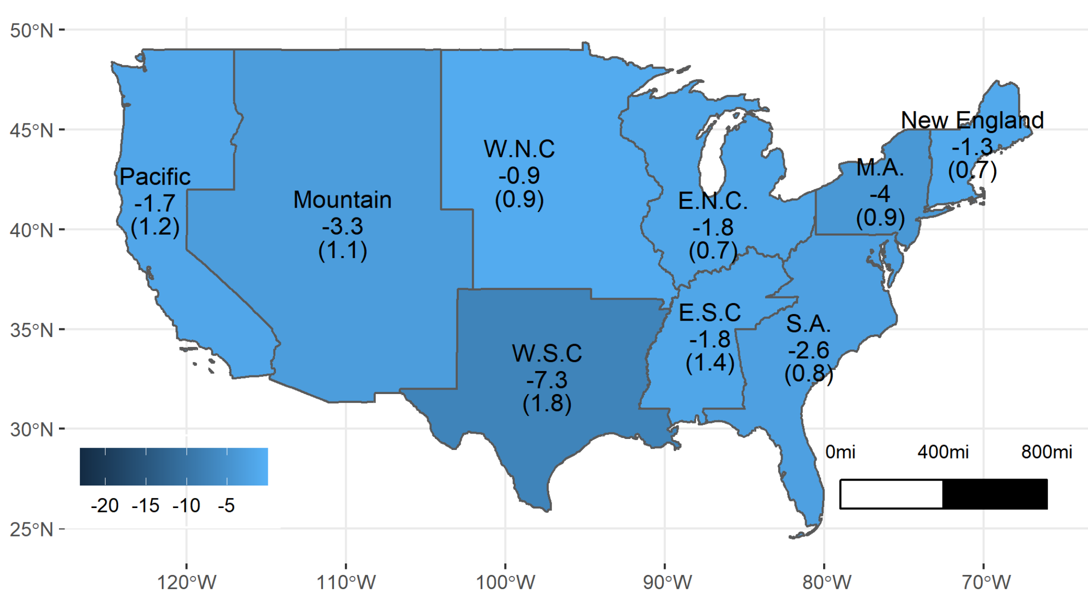

Figure 6: Effect of Uber on the number of bad air quality episodes across census regions

Notes: This map portrays point estimates and standard errors in parenthesis of a Callaway and Sant'Anna's difference-in-differences (CS-DD) design on the impact of Uber on the number of days of bad air quality, i.e., days with maximum air quality index (AQI) values higher than 100 units. The AQI standardizes the concentration of criteria contaminants into a single scale running between 0 and 500 units. The treated and control groups contain all counties with and without Uber at time t. The CS-DD model controls for county and year fixed effects. Standard errors are clustered at the county level.