Uber's Effect On the Quality of Air

Tyler Kim

November 27, 2022

Abstract

Does Uber affect air quality in cities that it operates in? Using data on the dates that Uber entered the cities it services along with longitudinal data on air quality throughout the United States, this paper estimates the impact of Uber's effect on air quality. Through a difference-in-differences estimation, my analysis finds that the entrance of Uber worsens air quality by 3.49 AQI points on average. This increase is statistically significant at the 0.05 level. My analysis also estimates that Uber's negative impact on air quality increases in magnitude over time. In aggregate, these findings suggest that Uber's services have a detrimental effect on air quality in the cities that it operates in.

Keywords: Uber, Air Quality, Pollution, Difference-in-Difference

1 Introduction

Since its inception in 2010, Uber has radically transformed the landscape of for-hire transportation. Through its peer-to-peer matching algorithm, Uber allows anyone that owns a car to essentially serve as an independent taxi service, servicing demand for transportation within its mobile app. For riders, this means that on-demand transportation is available through a few taps of their phone screen. In contrast to the incumbent taxi industry, Uber also allows riders to view their trip route, pay electronically, and agree to a trip price upfront.

Uber and other ride-sharing apps have transformed the way we travel. In 2019, ride-sharing companies provided 91.1 million rides in Massachusetts alone. This equates to an average of 250,000 rides per day, in just one US state! While we can appreciate the massive human coordination service that Uber provides, what are the hidden costs of such a massive amount of travel? According to the Washington Department of Ecology, burning gasoline and diesel fuel creates harmful byproducts like nitrogen dioxide, carbon monoxide, hydrocarbons, benzene, and formaldehyde. In addition, vehicles emit carbon dioxide, the most common human-caused greenhouse gas. The more trips Uber services, the more gasoline is needed to fulfill those trips. This potential for environmental damage inspires my research, which investigates the relationship between Uber and air quality. Namely, this paper asks the question, "What is Uber's effect on the quality of air?"

My research has broad implications. Since the 1800s, the public has grown increasingly aware of the interrelationship between their health and the infrastructure of their cities[1]. Today, with climate change being at the front of the public consciousness, individual people are mindful of how their consumption patterns are contributing to human-inflicted pollution. Investigations into the externalities of Uber can allow consumers to adjust their behavior to account for costs beyond just monetary value. Increased public awareness of the full spectrum of the costs of Uber could encourage Uber to mitigate these externalities. One example of such behavior is their Green Future program, which encourages the use of electric vehicles for Uber drivers by offering increased earnings, special ride options for electric vehicles, and more.

[1] See *John Snow, Cholera, the Broad Street Pump; Waterborne Diseases Then and Now*, Tulchinsky, 2018

Beyond public awareness, investigating the externalities of Uber also is extremely useful for informing legislation. While Uber may have avoided legislation in its infancy, as the service matures, legislation will emerge to address the full spectrum of its effects. One example is California Senate Bill 1014, which imposes rules to reduce the ride-hailing industry's greenhouse gas emissions. This paper seeks to establish what the causal effect of Uber is on air quality, and investigations like these will hopefully inform the legislation directed towards Uber and other emergent technologies.

The question of Uber's impact on air quality does not have an immediately obvious answer. On one hand, it is reasonable to say that Uber offers a decreased time cost relative to walking or public transportation, thereby increasing demand for automotive transportation that was previously met through more environmentally friendly means. Additionally, it's estimated that about 40% of the miles logged by ride-hailing vehicles are deadhead miles[2]. These deadhead miles mean that traveling from point A to point B in an Uber creates **strictly more** emissions than the same trip taken in a personal car, all else being equal. On the other hand, it is also reasonable to model the Uber driver supply as a competitive market. The marginal cost will be lower for a driver of a fuel-efficient car by the nature of it using less gas, and vice versa. More fuel-efficient cars will enter the market of drivers, and less fuel-efficient cars will be competed out of the market as a result of unsustainably high marginal costs. The end result is a fleet of fuel-efficient Uber drivers that replaces trips taken by less economical vehicles.

Importantly, both of the above scenarios are hypotheticals. It is critical to investigate the causal relationship between Uber and air quality so that we can move out of the realm of theory and into empiricism.

One difficulty in evaluating the air quality effects of Uber is that pollution is a complex phenomenon that has various influences. According to the World Heath Organization, common sources of air pollution include household combustion devices, motor vehicles, industrial facilities, and forest fires. For our purposes, this means that Uber makes up a subset of a subset of the human-related drivers of air quality. Additionally, there are myriad pollutants that affect air quality, and only

[2]Deadhead miles refers to the miles driven that are not transporting a passenger. Examples include traveling to pick up a passenger, or cruising while waiting to be assigned a ride request

some of those are emitted by cars.

This paper uses the Environmental Protection Agency's (EPA) air quality index (AQI) statistic. The AQI measures the concentrations of five major pollutants and maps these concentrations to a number on a scale of 0-500. Higher values of AQI correspond to higher concentrations of measured pollutants. Higher values of AQI also correspond directly to increased population risk, with moderately low AQI values posing health risks to elderly populations or those with other health issues, and extremely high AQI values pose a health threat to all populations. Figure 1 breaks down the thresholds for AQI and population risk.

In order to mitigate the effect of outlier events, such as wildfires, on our analysis, we use the annual median AQI as our outcome variable of interest (as opposed to maximum AQI or average AQI, for instance). Throughout the rest of this paper, any reference of AQI will be referring to median AQI unless otherwise noted.

I use difference-in-difference estimation to determine the causal effect of Uber on AQI. This method compares the difference in the AQI before and after Uber entered a city, with the difference in AQI for the same time period in a city where Uber did not enter. We consider cities without Uber in a given year as the control group, and those with Uber in a given year as the treatment group.

Specifically, I use Callaway and Sant'Anna's difference-in-differences (CS-DID) design. Recent literature (Baker, Larcker, Wang, 2021; Goodman-Bacon, 2021) has shown evidence that standard two-way fixed effects difference-in-differences (TWFE-DD) setups generate biased estimates of average treatment effects on the treated (ATT) when treatment timing is staggered. Because Uber enters different cities at different times, my empirical design contains staggered treatment timing, and using TWFE-DD could lead to biased estimates of ATT. By batching my observations into cohorts, or groups, based on when the observation recieved treatment, CS-DID allows me to estimate ATT's in the presence of staggered treatment times.

My results show that the introduction of Uber worsens air quality. The average ATT across all groups and periods is estimated to be 3.5, and this estimation is statistically significant at the 5% level. This corresponds to an 9.7% increase in AQI relative to the average AQI of the control group. Furthermore, my analysis indicates that the ATT strictly increases in the years after treatment, suggesting that the impact of Uber's entry negatively impacts air quality more severely over time. Uber's ATT 3 years after treatment has an associated 95% confidence interval of $[.323, 5.707]$. This indicates that the true treatment effect on the treated ranges lies within a range of a minuscule effect (.323) and a moderate effect (5.707) with 95% confidence. However, the ATT 7 years after treatment has a 95% confidence interval of $[3.968, 22.726]$. The upper bound of the 95% confidence interval has nearly quadrupled in just four years. An upper bound of 22.726 is not insignificant. For reference, the average median annual AQI for the control group in Figure 6.2 is 35.72, which is considered a good AQI that poses little or no pollution risk. An AQI increase of 22.726 would make the air quality moderately concerning, and the pollution poses a risk to those who are unusually sensitive to air pollution. In a relatively short period of time, Uber's effect on air quality has the potential to lead to health risks for groups who are highly sensitive to air pollution.

Uber's effect on air quality has been treated by Sarmiento and Kim, however, their study considers the effects of Uber's entry on maximum annual AQI, and studies the effects at the county level, as opposed to the city level. I also control for the population when estimating the ATT, while their study did not. Interestingly, our results conflict – their study concludes that Uber improves the air quality, while mine finds the opposite. Other studies have determined positive relationships between ride-sharing and public transportation use in the US(Zhang and Zhang, 2018), as well as positive (yet location-dependent) effects of public transportation on air quality (Beaudoin, Farzin and Lawell, 2015).

My research contributes to the ongoing policy debate of public health and environmental externalities of Uber by determining a statistically significant negative causal effect of Uber's entrance on AQI. Practically, this research should urge policymakers in cities without Uber to be cautious in allowing their services to enter their city, and policymakers in cities with Uber to consider how they can use their legislative power to mitigate the negative effects of Uber on air quality. My research also demonstrates increasingly severe negative impacts on AQI over time, indicating that policymakers should remain vigilant in monitoring air quality effects even after proposed legislation runs its course, as Uber's impact on air quality is demonstrated to be dynamic. The rest of this paper is structured as follows: Section 2 presents the data sources used in this paper, Section 3 outlines the empirical design used in my CS-DID specification, Section 4 provides the results of the research, and Section 5 concludes my paper and gives a brief commentary on potential extensions of the research conducted in this paper.

2 Data

The data on AQI is provided by the EPA. This data set collects annual AQI data for all Core-Based Statistical Areas (CBSA) in the US. The concept of a CBSA was created and defined by the United States Office of Management and Budget based on the concept of a core area containing a substantial population nucleus, with adjacent communities having a high degree of social and economic integration with that core. While there are many different specifications of AQI that the EPA collects, we filter out all but the median AQI. We consider data from the years 2013-2021 in our analysis. We have non-treated units as our control, allowing us to consider the period starting in 2013 and still have a sizeable control group. We end the period of consideration at 2021 to avoid capturing the incomplete AQI readings of the current year 2022.

By using AQI to measure the effect of Uber on air quality, we are given an abstraction that is representative of all the harmful pollutants, without being overly specific. Moreover, the EPA provides very clear guidelines on how AQI corresponds to population health risk (Figure 1), which allows us to easily interpret our results in terms of public welfare. Consider a thought experiment to appreciate the significance of the interpretability of AQI. Assume that instead of AQI, we used the absolute concentration of traffic-related pollutants such as nitrogen dioxide (NO_2). We would conduct our empirical test, and obtain data on the causal effect of Uber on NO_2 concentration. What exactly does this concentration mean? For instance, when high concentrations of NO_2 interact with ground-level ozone (O_3), the NO_2 degrades the O_3 back into regular oxygen (O_2).

In this hypothetical setup, the results of our empirical test would be inconclusive for our research question, since the effect of NO_2 on air quality would be inconclusive! The AQI is a great choice for an outcome variable of interest because of its direct translation to health risk.

The AQI data is very high quality, with no present null values. However, one quirk of the data is that CBSAs are determined in a way that occasionally combines multiple cities into one CBSA. For instance, Boston's CBSA is encoded as Boston-Cambridge-Newton. For this paper, we select only the first city in a CBSA that comprises multiple cities. We do this because cities are listed in order of decreasing size, so choosing the first city in the CBSA allows us to let the biggest city represent the CBSA. While I considered splitting each CBSA into its constituent cities and assigning all cities the AQI value of the overall CBSA, I decided against this approach to avoid data duplication issues.

The Uber entry data is taken from a study by Jonathan Hall[3] that was downloaded from the NBER. This data contains entry dates of when Uber entered cities, and includes the type of Uber that launched (UberX, UberBlack, etc.) as well as notes that validate the entry date by linking to news articles announcing the launch. Our data includes Uber entry dates that range from year 2013-2017.

This data set is also of high quality, and the notes that validate the entry dates add a measure of confidence that the entry dates are accurate. Missing values are not a prevalent issue, with the entry for a record's city having the most null values. These null values only made up 2% of all observations, which is a sufficiently small amount that I was comfortable discarding these observations, as they would not affect the validity of the analysis. Beyond dropping null values, I also cleaned the data by truncating the state initials from the "City/State" column, so the new "City" column matches the form of the AQI data.

Figure 3 shows a geographical visualization of the locations where Uber entered. At first glance, the distribution of treated cities appears to be concentrated in the East part of the US. Taken in isolation, this could be indicative that Uber has chosen cities to treat in a biased fashion. However,

[3] "Is Uber a substitute or complement for public transit?" Hall, Palsson, Price, 2018

a heat map of the population in the US shown in figure 2 reveals that the distribution of treated cities tracks very closely with the overall population distribution in the US. Cities that receive treatment are typically highly population-dense areas. To account for the bias inherent in Uber's launch strategy, we include population as a regressor in our final regression.

Finally, I use population data from the US Census to allow us to include city populations as a covariate in our regression. This data is high quality as well, with no present null values. We take population data from the time periods under consideration (2013-2021), and transform the dataset to have each observation represent a unique (City, Year, Population) record. In terms of data cleaning, the city names were encoded with the type of municipality at the end (so Boston was encoded as "Boston city"). We simply remove this trailing municipality from the column. To visualize the of the data, Figure 4 is a histogram depicting the distribution of population for the cities under consideration for my analysis.

After performing the initial data cleaning on the three datasets, I then join the AQI data with the population data on common (City, Year) pairs, and join the Uber entry data with the merged data on common (City) values. The final output is a table where each observation represents a unique entry of (City, Year, Uber Entry Date, Median AQI, Population). For use in the CS-DID model, we transform the Uber Entry Date to a "treated" column which takes on the value:

$$treated(entry\ year, entered) = \left\{ \begin{array}{ll} entry\ year, & \text{if } entered = True \\ 0, & \text{if } entered = False \end{array} \right\} \tag{1}$$

where *entry year* is the year that Uber entered a city, and *entered* is a boolean that is True if Uber has entered a city in a given year for the observation in question. We consider the entry year to be:

$$entry\ year(entry\ date = MM/DD/YYYY) = \left\{ \begin{array}{ll} YYYY, & \text{if } MM <= 06 \\ YYYY + 1, & \text{if } MM > 06 \end{array} \right\} \tag{2}$$

In plain English, we consider Uber's entry year to be cut off in June. Any entry dates that occur after June are considered to have their entry year in the next year. We make this consideration to provide a distinction between Uber entering a city in January and December, for instance.

In general, my data is high quality because the data is mostly provided by government agencies. As a result, little data is lost when we are joining all the tables together. This can be seen in the treatment group data; the summary statistics in Table 6.1 show that there are 2011 observations in the treatment group. Table 6.2 gives the summary statistics for the control group.

3 Empirical Design

Uber first introduced its services under the name UberCab in San Francisco, initially providing ride-sharing with luxury vehicle transport only (what is now called Uber Black). Beginning in 2012, Uber shifted to the business model it employs today, allowing private car owners (of any make/model) to provide transportation services. The company began national expansion to large metro areas like Boston, New York City, and Chicago in 2011. By 2014, Uber had established operations in 100 cities, and a year later, the company was valued at $50 billion after raising a $1 billion Series F round.

By nature of it being a for-profit company, Uber's entrance into cities was assuredly a highly strategic decision. One could imagine that Uber may have viewed cities with higher wealth levels, higher population densities, and established transportation networks as particularly attractive locations to enter. When choosing an empirical strategy, we must account for these potential sources of biases. Therefore, we choose difference-in-difference estimation, which allows us to assume that the unobserved difference between the treatment group and control group remains constant.

While Uber's entrance has been previously blocked in various cities across the United States, as it stands today, Uber is not banned in any US cities. Furthermore, these temporary blockings were, in all cases, due to legal complications such as driver registration and employment classification laws. We do not expect that the blocking of Uber was related to any systemic predictors of air quality, and can safely assume that Uber's temporary blocking does not introduce bias into our model.

The canonical difference-in-difference setup has two periods: pre-treatment and post-treatment. In this setup, all treated units receive treatment at the same time. Because Uber entered different cities at different times, our data has staggered treatment times, and there are more than two periods. Recent literature (Borusyak and Jaravel, 2017; de Chaisemartin and D'Haultfoeuille, 2020; Goodman-Bacon, 2021) has shown evidence that standard two-way fixed effects difference-in-differences (TWFE-DD) are not guaranteed to recover interpretable causal coefficients when applied to settings with staggered treatment times. To combat this problem, we use the difference-in-difference design outlined by Callaway and Sant'Anna (2020) using the csdid package in Stata. CS-DID overcomes the problem of staggered treatment times by estimating ATTs for each cohort (group) of units treated at the same time. The parameter of interest is

$$ATT(g,t) = E[Y_t^1 - Y_t^0 | G_g = 1]$$

where Y_t^1 represents the outcome of the treated units at time t, Y_t^0 represents the outcome of untreated units at time t, and the $G_g = 1$ conditions the ATT on the group parameter g. Thus, $ATT(g,t)$ represents the ATT for cohort g at time t. We estimate the ATT by replacing expectations with the sample equivalents, namely,

$$\widehat{ATT}(g,t) = \frac{1}{N_g} \sum_{i:G_i=g} [Y_{it} - Y_{i,g-1}] - \frac{1}{N_G} \sum_{i:G_i \in \mathcal{G}} [Y_{it} - Y_{i,g-1}]$$

where $\mathcal{G}$ represents the control group. In our case, we have never treated and not yet treated cities as our control group.

Similar to TWFE-DD methods, the CS-DID specification relies on the common trends assumption. CS-DID tests this assumption by ensuring that there are no statistically significant ATTs measured before the treatment period across all groups. My study ensures this in two ways: graphically and through a hypothesis test. The results of the graphical common trends test can be seen in Figure 5. The blue bars/dots correspond to the pre-treatment periods. While it appears that there are no statistically significant ATTs in the pre-treatment period, the size of the error bars makes it difficult to be certain that statistical significance is not achieved. To be precise, I conduct a hypothesis test of the same pre-treatment common trends assumption. In this test, the null hypothesis is that $ATT(g, t) = 0$ for all groups and time periods before treatment. The p-value of this test is .5807, and therefore we fail to reject the null hypothesis that there are no pre-treatment ATTs that are statistically different than zero. Thus, our design does not show evidence of violating the common trends assumption.

4 Results

4.1 Estimated ATT Across All Groups and Periods

Table 6.3 displays the most simple ATT estimation: an average of the ATTs across all groups and time periods. The estimated ATT is 3.49, and this value is statistically significant at the 0.05 level. This means that on average, across all treatment groups and time periods, Uber's entrance increased the median AQI by 3.49. It is important to recall that AQI is a measure of the concentration of pollutants in the air. Therefore, positive values of ATT represent increases in the AQI, which correspond to **worse** air quality. Using the average median AQI for the control group in Table 6.2, the increase in median AQI caused by Uber represents a 9.7% increase relative to the average median AQI of the control group. In other words, we expect that, on average, introducing Uber into a city that previously did not have Uber will lead to a 9.7% increase in median AQI[4].

[4] We can say this because we have normalized the causal effect of Uber's entry by the average median AQI in cities without Uber. Put a different way, the expected increase in AQI after Uber's entry is 3.49. The average median AQI in cities without Uber is 35.72 (per Table 6.2). The expected increase in AQI after Uber's entry is 9.7% of the average median AQI in cities that do not have Uber, and therefore we expect an increase of 9.7% on the average median AQI if we introduce Uber into cities that previously did not have Uber.

4.2 Estimated ATT Per Period Across All Groups

Table 6.4 displays the estimated ATT for each period, across each treatment group. Recall that our estimand of interest was $ATT(g,t) = E[Y_t^1 - Y_t^0|G_g = 1$. The results in Table 6.4 consider each time period t between 2013-2021 and average the ATT in that time period across all groups.

ATTs in the initial time periods until the year 2016 are not statistically significant. After 2016, the ATT strictly increases with each successive year, with the highest ATT occurring in the year 2021. In the same period, the ATTs become monotonically more statistically significant, with the estimated ATTs in years 2017-2018 being statistically significant at the 0.05 level, and in years 2019-2021, the ATTs are statistically significant at the 0.01 level. The combination of these two facts indicates that Uber's negative impact on air quality becomes more significant over time, irrespective of when treatment was implemented in a given city.

4.3 Estimated ATT Per Group Across All Periods

Figure 6.5 shows the estimated ATT for each treatment group across all time periods. Plots of ATT by period for each group are located in the appendix. In order from group 2013-2017, the plots are shown in Figure 6, Figure 7, Figure 8, Figure 9, and Figure 10

The result that immediately jumps out is the ATT associated with group 2013. While the result is only statistically significant at the 0.10 level, the ATT is 20.216 – a considerable increase in AQI when you consider that the thresholds of AQI health risk (Figure 1) are divided into increments of 50. In general, the earlier a city received treatment, the higher the group's ATT was. Group 2014 has an ATT of 5.75 and is statistically significant at the 0.05 level, while Group 2015 has an ATT of 1.36 and is statistically significant at the 0.01 level. Groups 2016 and 2017 have small ATT values and are not statistically significant. Taken in tandem with the group-independent results from section 4.2, these results are further indications that Uber's detrimental effect on air quality becomes more severe over time.

4.4 Estimated ATT Relative to First Treated Across All Groups

Finally, Table 6.6 computes the ATT for the time periods relative to the first treatment date. "TmN" refers to the ATT N years **before** the first treatment, and "TpN" refers to the ATT N years **after** the first treatment.

Consistent with our test of the common trends assumption, "Pre_avg", and all "Tm"s have ATTs not statistically different from zero[5]. Consistent with sections 4.2 and 4.3, we can observe that ATT increases as more time has elapsed since the first treatment. It's important to note that "Tp8" should be treated as a bit of an anomaly since it only captures the interactions of Group 2013 and period 2021[6]. The ATT 1 year after treatment (Tp1) is not statistically significant from zero. Tp2-Tp7 all have ATTs that are strictly increasing on this domain. The Tp2 ATT is statistically significant at the 0.10 level, Tp3 is statistically significant at the 0.05 level, and Tp4-7 are all statistically significant at the 0.01 level.

5 Conclusion

Because of the myriad factors that contribute to pollution and poor air quality, it is difficult to theorize about the effect of Uber on air quality. I choose the AQI as a proxy for air quality, as it combines various pollutants that contribute to the quality of air into a numerical scale that has direct interpretations in terms of health risks. To estimate causal effects, I utilize the CS-DID method to account for biases arising from Uber's staggered treatment implementation that would be unaccounted for in other DID designs. This empirical method compares the difference AQI in treated cities and untreated cities, controlling for population, city fixed-effects, and year fixed-effects.

My results show that not only does Uber negatively impact air quality, but these negative impacts

[5] Remember, the CS-DID common trends assumption was tested by verifying that pre-treatment ATTs were not statistically significant.

[6] Because this is the only group/period pair with 8 years elapsed between them

become more severe over time. The harm to air quality increases over time not only in terms of the magnitude of the effect on AQI but also in terms of statistical significance. These results are consistent across multiple granularities: groups, periods, and relative time elapsed since treatment all exhibit the aforementioned trends.

These findings have the potential to influence consumption patterns of environmentally aware Uber riders, but also will be highly useful to inform legislation that can systemically mitigate the negative externalities of Uber in a way that transcends the capabilities of individuals motivated by a shared cause.

My findings provide an estimation of the causal effects of Uber on air quality, with important consumer behavior and policy implications. Future studies might take the growing prevalence of electric vehicles into account when analyzing Uber's effect on air quality.

6 Appendix

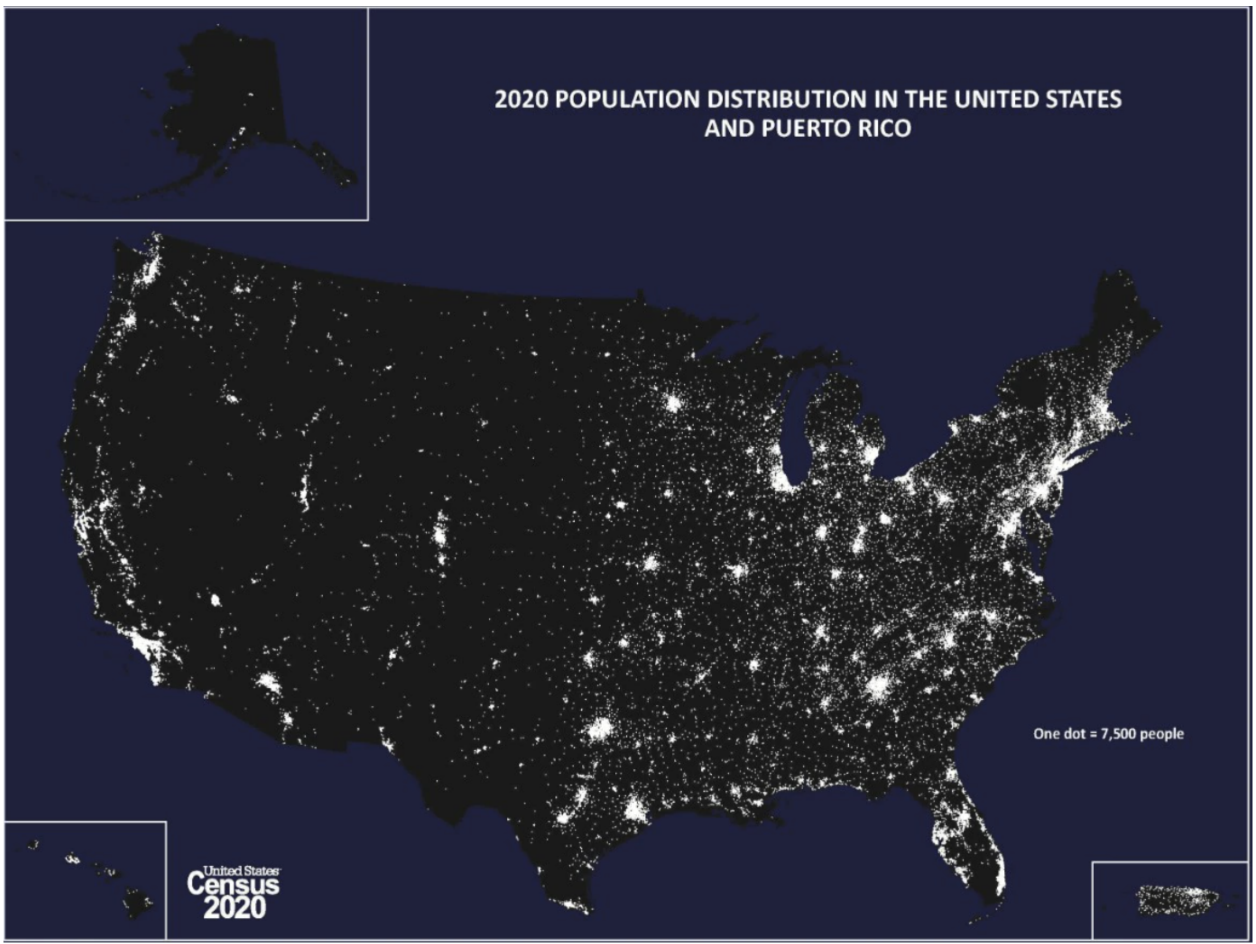
Figure 1: AQI Mapping to Risk Level

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.

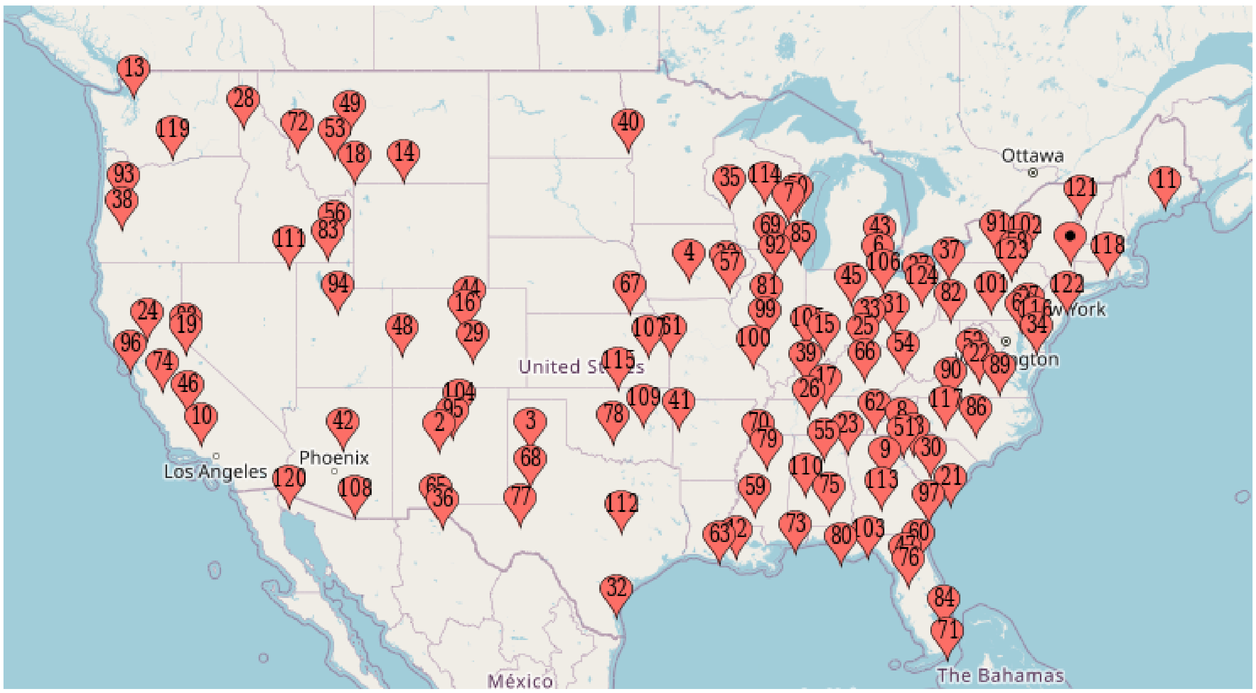
Source: Environmental Protection Agency

Figure 2: Heatmap of Population Distributions Across the US



Source: United States Census Bureau

Figure 3: Geographical Map of Uber's Entry



Note: I generated this plot by manually entering the cities that Uber had entered. The numbers on the pin marks should be disregarded

Figure 4: Population Distribution in Data

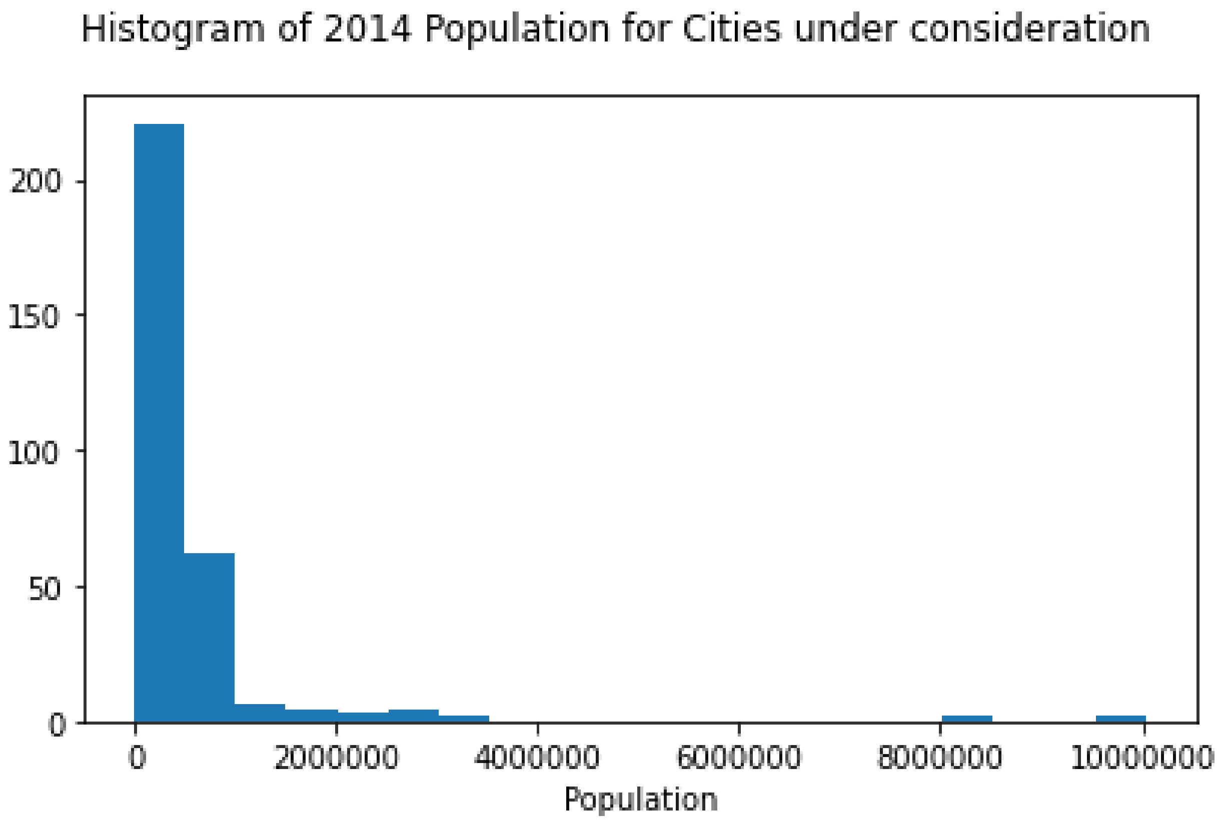


Figure 5: Graphical pre-treatment test for common trends assumption

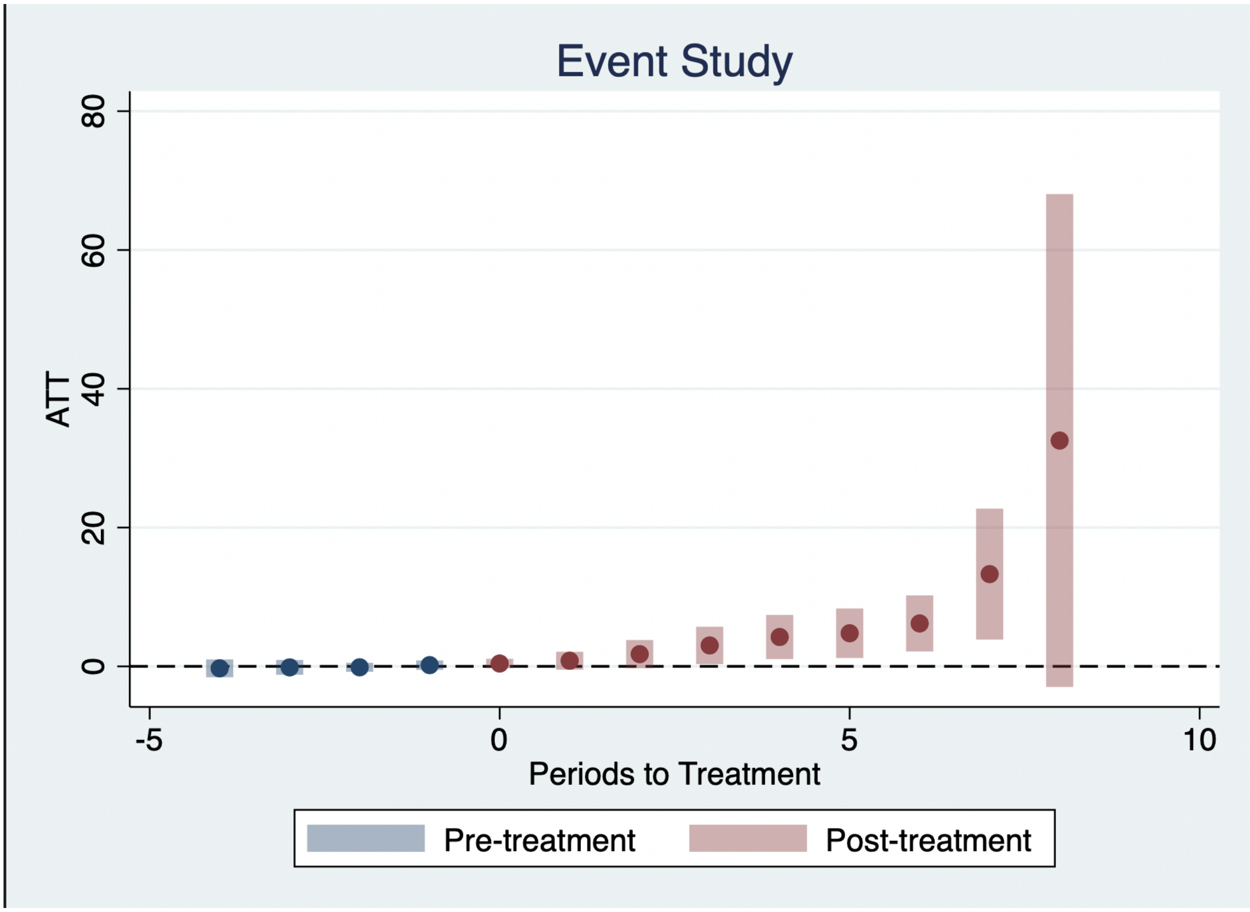


Figure 6: Plot of ATT's for Group 2013

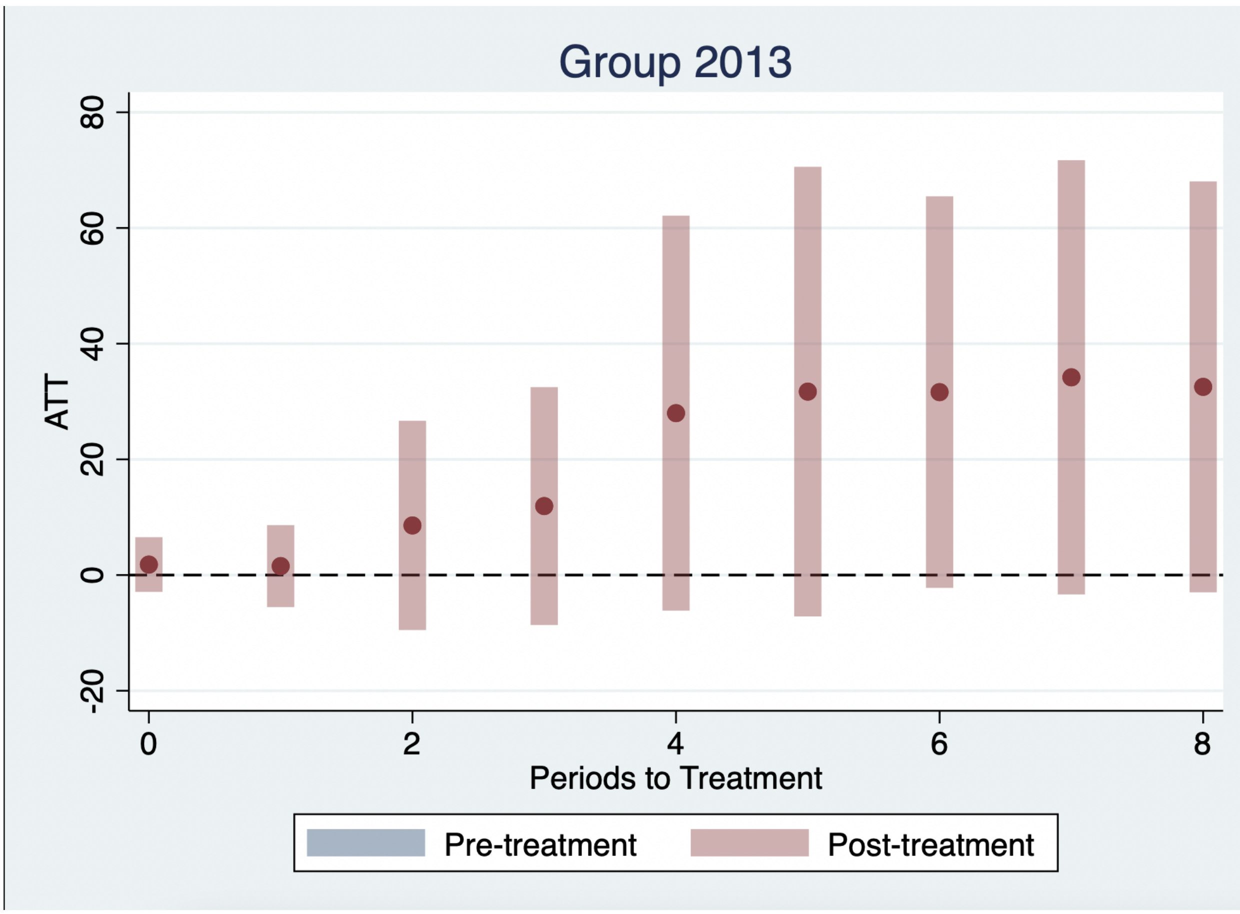


Figure 7: Plot of ATT's for Group 2014

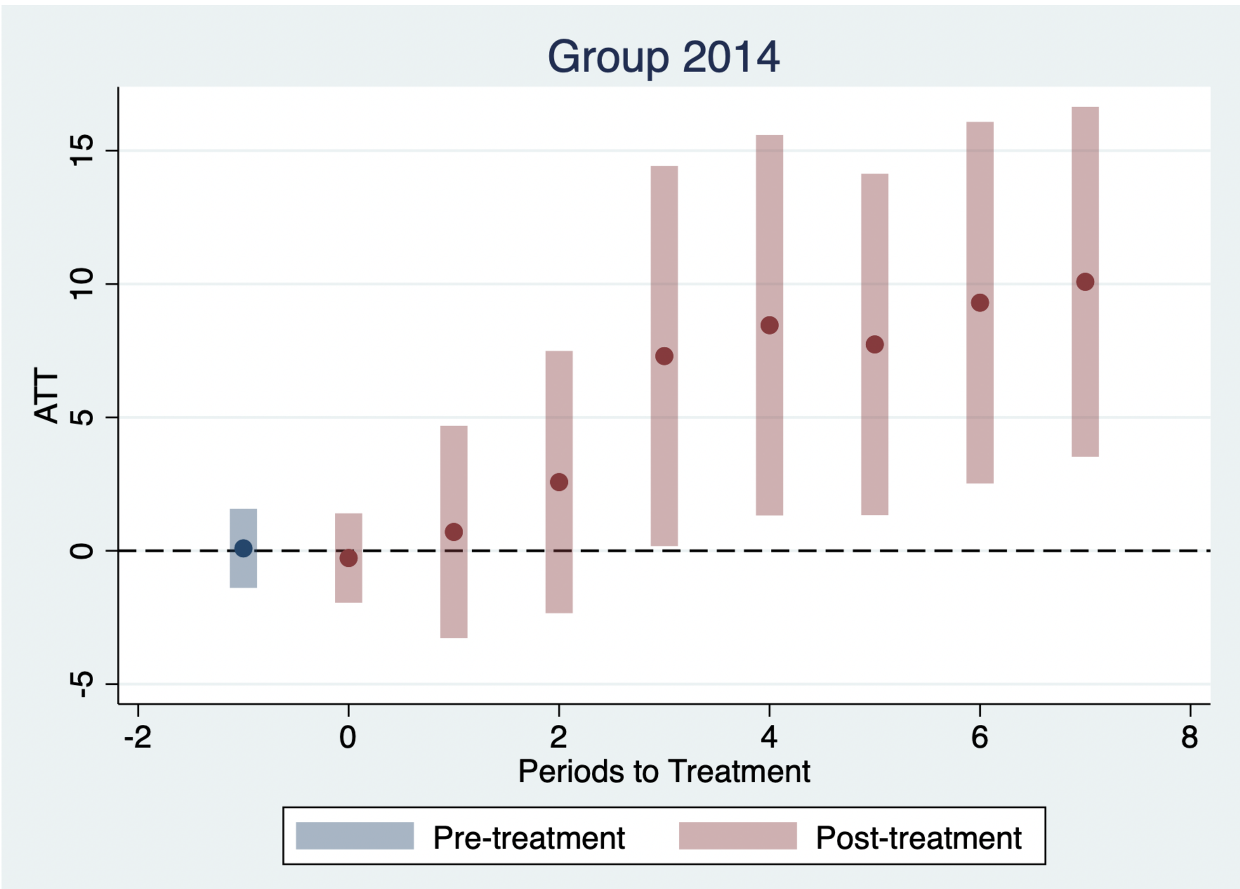


Figure 8: Plot of ATT's for Group 2015

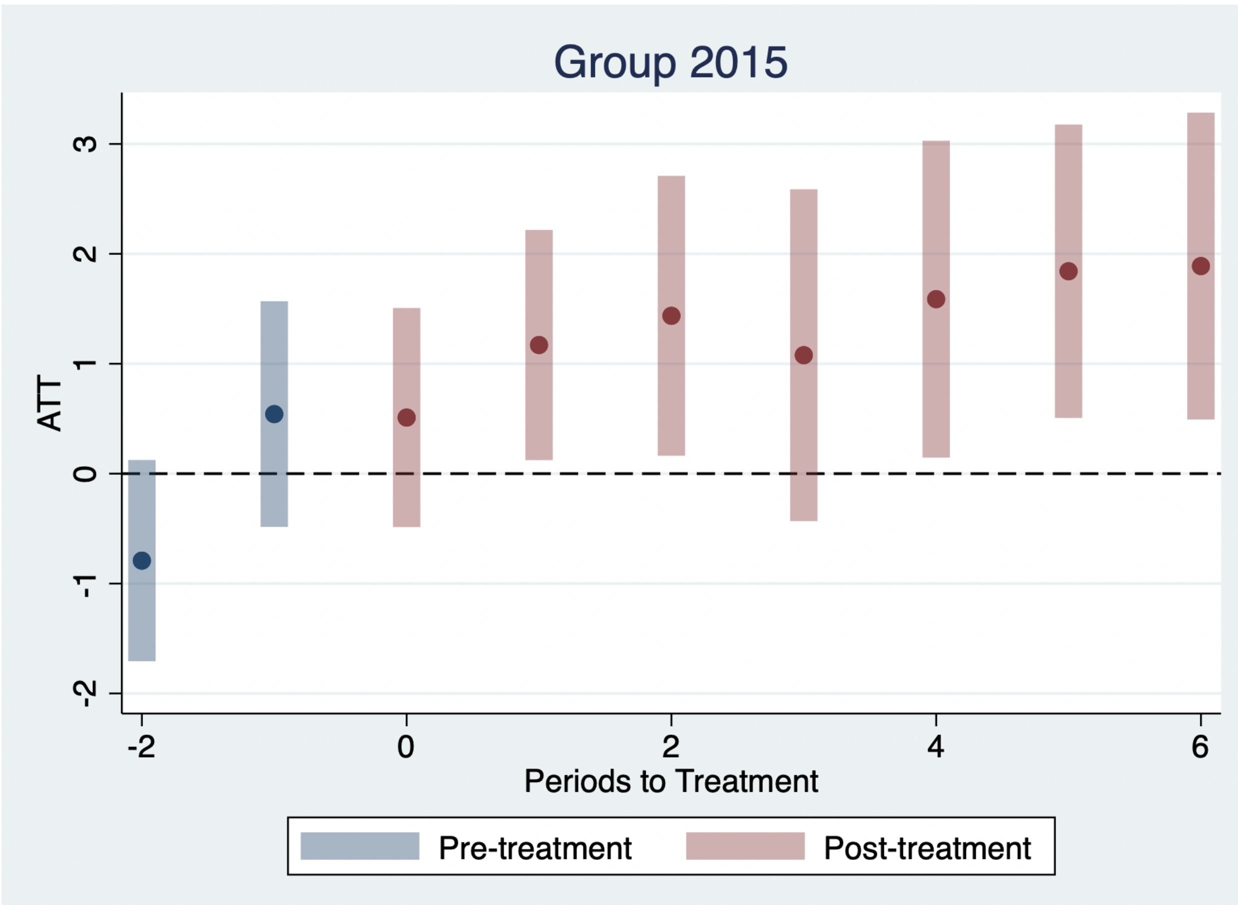


Figure 9: Plot of ATT's for Group 2016

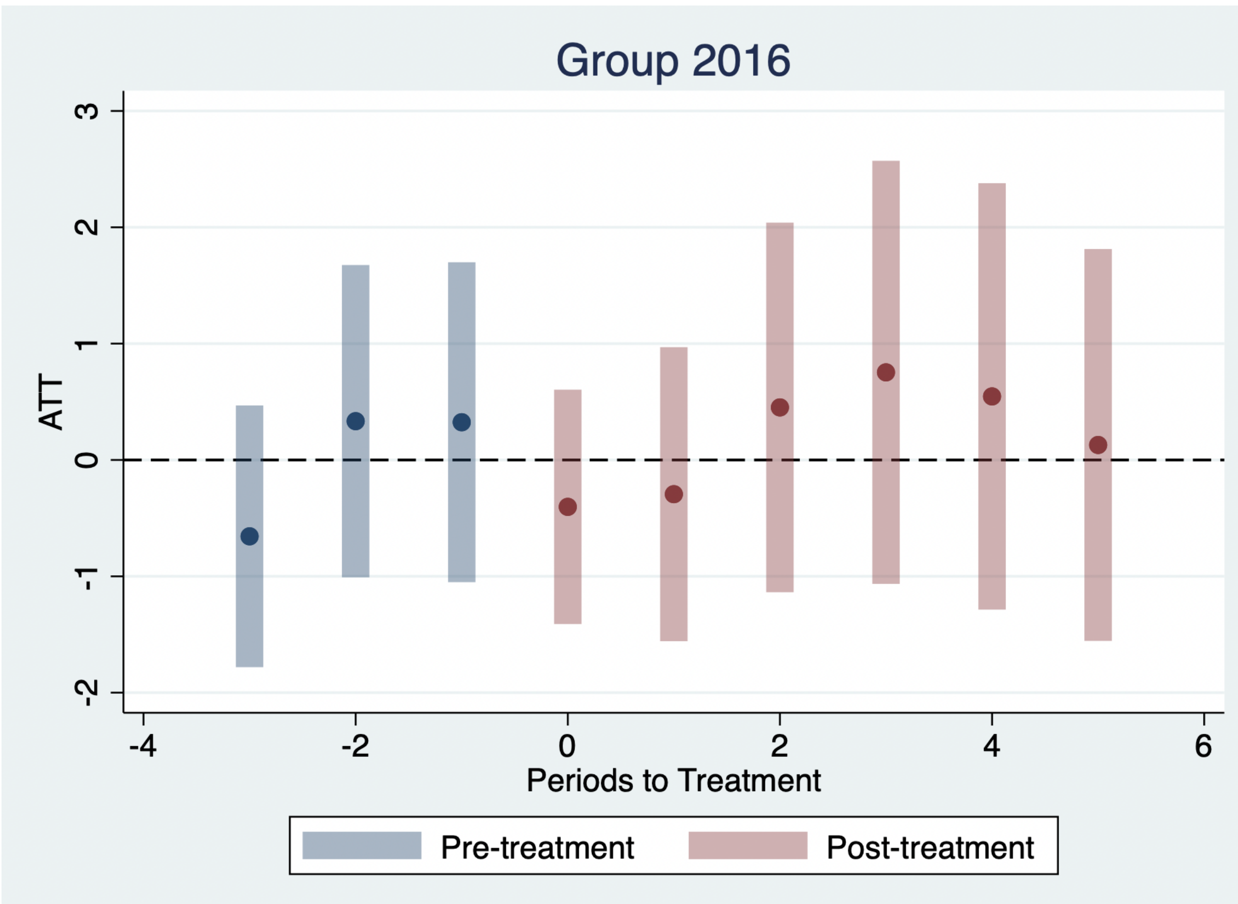


Figure 10: Plot of ATT's for Group 2017

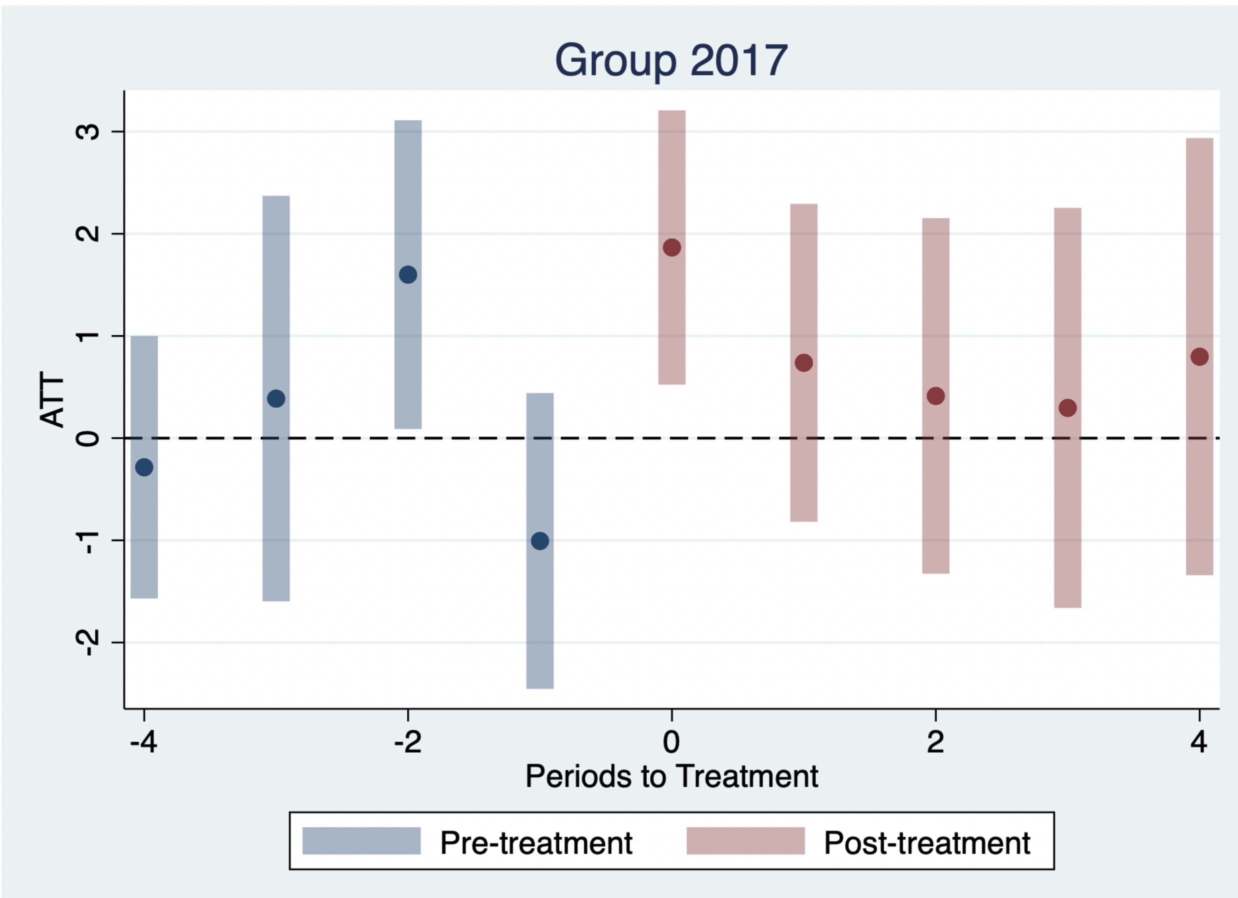


Table 6.1: Summary Statistics for Treatment Group

	N	Mean	SD	Min	Max
median_aqi	2011	42.86	11.13	10	123.00
population	2011	447,983.37	1,003,318.96	1108	10,033,449.00

Table 6.2: Summary Statistics for Control Group

	N	Mean	SD	Min	Max
median_aqi	2623	35.72	9.85	3	77.00
population	2623	65,361.34	91,041.41	3817	936,178.00

Table 6.3: Estimated ATT of Uber on median annual AQI for all groups across all periods

	(1)
ATT	3.496**
	(1.365)
N	4,634

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.4: Estimated ATT of Uber on median annual AQI for each period, across all groups

	(1)
T2013	1.818
	(2.413)
T2014	-0.030
	(1.056)
T2015	1.028
	(1.344)
T2016	1.849
	(1.240)
T2017	3.891**
	(1.718)
T2018	4.142**
	(1.865)
T2019	4.166***
	(1.569)
T2020	4.749***
	(1.656)
T2021	4.920***
	(1.566)
N	4,634

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$T2013$ is the ATT across all cohorts in the calendar year 2013, etc.

Table 6.5: Estimated ATT of Uber on median annual AQI for each group across all periods

	(1)
G2013	20.216*
	(11.662)
G2014	5.748**
	(2.592)
G2015	1.359***
	(0.526)
G2016	0.193
	(0.601)
G2017	0.822
	(0.656)
N	4,634

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$G2013$ is the ATT across all periods for the 2013 cohort, etc.

Table 6.6: Estimated ATT relative to the period first treated across all groups. TmN is the ATT N years before treatment, TpN is the ATT N years after treatment.

	(1)
Pre_avg	-0.097
	(0.198)
Post_avg	7.447**
	(3.221)
Tm4	-0.285
	(0.655)
Tm3	-0.144
	(0.544)
Tm2	-0.135
	(0.335)
Tm1	0.176
	(0.348)
Tp0	0.411
	(0.347)
Tp1	0.813
	(0.658)
Tp2	1.762*
	(1.037)
Tp3	3.015**
	(1.374)
Tp4	4.231***
	(1.625)
Tp5	4.771***
	(1.823)
Tp6	6.188***
	(2.061)
Tp7	13.297***
	(4.811)
Tp8	32.538*
	(18.119)
N	4,634

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Andrew Baker, David Larcker, and Charles Wang. How Much Should We Trust Staggered Difference-In-Differences Estimates? 2021.

Justin Beaudoin, Y. Hossein Farzin, and Cynthia Lin Lawell. Public transit investment and sustainable transportation: A review of studies of transit's impact on traffic congestion and air quality. 2015.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. 2021.

Jonathan Hall, Craig Palsson, and Joseph Price. Is Uber a substitute or complement for public transit? 2018.

Jonathan Roth, Pedro Sant'Anna, Alyssa Bilinski, and John Poe. What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. 2022.

Luis Sarmiento and Yeong Jae Kim. The Air Quality Effects of Uber. 2021.

Yuanyuan Zhang and Yuming Zhang. Exploring the Relationship between Ridesharing and Public Transit Use in the United States. 2018.

Responses in Red

1. It’s a little hard to know at this point in the Intro what your research is.

Added new paragraph that motivates the research and explicitly states the research question

2. Seems a bit in the weeds for the Intro. But you could put this discussion in the Data section.

Moved to the section that discusses the AQI data

3. Use the confidence intervals that accompany your estimates to tell us about the (with 95% confidence) max and min treatment effects implied by your results. Then perhaps one of those is surprising to the reader (based on prior literature) and you can point that out.

Included 95% confidence intervals for the 3 & 7 years after treatment

4. Seems the related lit discussion is just missing?

Added

5. This doesn’t look right. Compare it to what is in the lecture slides or the “CS-DID” paper.

Response below

6. It’s great that you’re allowing for heterogenous treatment effects across cities and time (and elapsed time since treatment). But as the CS-DID paper makes clear you aren’t going to be able to estimate each such effect, only certain weighted averages. You don’t need to go into the details of all that in the paper but the discussion around here looked confused.

Per points 5 and 6, I updated the paper to reflect equation 8 of *What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature* (Hall, Sant’Anna, Bilinski, Poe 2022)

```
----------------------------------------------------------------------------
> ----------------------------------------------------------------------------
> ----------------
      name:  <unnamed>
       log:  /Users/tylerkim/Documents/14.33/data/summary.log
  log type:  text
 opened on:  29 Nov 2022, 00:53:08

.
. use uberx_data.dta

.
. rename POPESTIMATE2014 population

.
. encode city, g(ncity)

.
. drop city

.
. drop if year == 2022
(320 observations deleted)

.
. rename ncity city

.
. est clear

.
. estpost summarize median_aqi population if treated != 0

             |  e(count)   e(sum_w)    e(mean)     e(Var)      e(sd)     e(mi
> n)     e(max)     e(sum)
-------------+--------------------------------------------------------------
> -------------------------
  median_aqi |      2011       2011   42.85579   123.8051   11.12677
> 10        123      86183
  population |      2011       2011   447983.4   1.01e+12    1003319       11
> 08   1.00e+07   9.01e+08
```

```
.
. esttab using "./tables/table1.tex", replace cells("count mean(fmt(%20.2fc))
>  sd(fmt(%20.2fc)) min max(fmt(%20.2fc))") nonumber noobs booktabs collabels
> ("N" "Mean" "SD"
>  "Min" "Max")
(output written to ./tables/table1.tex)

.
. est clear

.
. estpost summarize median_aqi population if treated == 0

             |  e(count)   e(sum_w)    e(mean)     e(Var)      e(sd)     e(mi
> n)     e(max)     e(sum)
-------------+----------------------------------------------------------------
> -------------------------
  median_aqi |      2623       2623   35.72207   96.94828   9.846232
>  3         77      93699
  population |      2623       2623   65361.34   8.29e+09   91041.41       38
> 17     936178   1.71e+08

.
. esttab using "./tables/table2.tex", replace cells("count mean(fmt(%20.2fc))
>  sd(fmt(%20.2fc)) min max(fmt(%20.2fc))") nonumber noobs booktabs collabels
> ("N" "Mean" "SD"
>  "Min" "Max")
(output written to ./tables/table2.tex)

.
. est clear

.
. csdid median_aqi population, ivar(city) time(year) gvar(treated) notyet met
> hod(reg)
Panel is not balanced
Will use observations with Pair balanced (observed at t0 and t1)
.............................................
Difference-in-difference with Multiple Time Periods
```

```
                                                      Number of obs = 4,58
> 3
Outcome model  : regression adjustment
Treatment model: none
------------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+----------------------------------------------------------------
> -
g2013        |
 t_2012_2013 |   1.818255   2.413148     0.75   0.451    -2.911427    6.54793
> 8
 t_2012_2014 |   1.550475   3.616256     0.43   0.668    -5.537256    8.63820
> 7
 t_2012_2015 |   8.582716   9.229193     0.93   0.352     -9.50617     26.671
> 6
 t_2012_2016 |   11.92983   10.49032     1.14   0.255    -8.630819    32.4904
> 9
 t_2012_2017 |   27.99454    17.4195     1.61   0.108    -6.147047    62.1361
> 2
 t_2012_2018 |   31.71555   19.82694     1.60   0.110    -7.144533    70.5756
> 3
 t_2012_2019 |   31.62938   17.26838     1.83   0.067    -2.216027    65.4747
> 9
 t_2012_2020 |   34.18776   19.15129     1.79   0.074    -3.348076     71.723
> 6
 t_2012_2021 |   32.53821   18.11897     1.80   0.073     -2.97432    68.0507
> 4
-------------+----------------------------------------------------------------
> -
g2014        |
 t_2012_2013 |    .090687   .7562915     0.12   0.905    -1.391617    1.57299
> 1
 t_2013_2014 |  -.2729004    .855502    -0.32   0.750    -1.949654    1.40385
> 3
 t_2013_2015 |   .7057611   2.030337     0.35   0.728    -3.273627    4.68514
> 9
 t_2013_2016 |   2.575181   2.508342     1.03   0.305     -2.34108    7.49144
> 2
 t_2013_2017 |   7.299622   3.635956     2.01   0.045     .1732798    14.4259
> 6
 t_2013_2018 |   8.456183   3.638873     2.32   0.020     1.324123    15.5882
> 4
 t_2013_2019 |   7.735981   3.265471     2.37   0.018     1.335776    14.1361
> 9
 t_2013_2020 |   9.300885   3.458749     2.69   0.007      2.52186    16.0799
> 1
 t_2013_2021 |   10.08277   3.346827     3.01   0.003     3.523105    16.6424
```

```
> 3
-------------+----------------------------------------------------------------
> -
g2015        |
 t_2012_2013 |  -.7912159   .4670545    -1.69   0.090    -1.706626    .124194
> 1
 t_2013_2014 |    .541999    .523912     1.03   0.301    -.4848497    1.56884
> 8
 t_2014_2015 |   .5104483   .5084904     1.00   0.315    -.4861744    1.50707
> 1
 t_2014_2016 |   1.170451   .5343528     2.19   0.028     .1231389    2.21776
> 3
 t_2014_2017 |   1.436385   .6496292     2.21   0.027     .1631349    2.70963
> 5
 t_2014_2018 |   1.077911    .770233     1.40   0.162    -.4317179     2.5875
> 4
 t_2014_2019 |   1.587946   .7354369     2.16   0.031     .1465157    3.02937
> 5
 t_2014_2020 |   1.841839   .6810126     2.70   0.007     .5070791    3.17659
> 9
 t_2014_2021 |   1.888759   .7123223     2.65   0.008     .4926332    3.28488
> 5
-------------+----------------------------------------------------------------
> -
g2016        |
 t_2012_2013 |  -.6561179   .5739282    -1.14   0.253    -1.780997    .468760
> 7
 t_2013_2014 |   .3336037   .6847651     0.49   0.626    -1.008511    1.67571
> 9
 t_2014_2015 |    .324579   .7013664     0.46   0.644    -1.050074    1.69923
> 2
 t_2015_2016 |  -.4028631   .5139444    -0.78   0.433    -1.410176    .604449
> 5
 t_2015_2017 |  -.2941151   .6448689    -0.46   0.648    -1.558035    .969804
> 7
 t_2015_2018 |   .4518047   .8105071     0.56   0.577     -1.13676    2.04036
> 9
 t_2015_2019 |   .7533638   .9276697     0.81   0.417    -1.064835    2.57156
> 3
 t_2015_2020 |   .5466409   .9350625     0.58   0.559    -1.286048     2.3793
> 3
 t_2015_2021 |   .1288349   .8592613     0.15   0.881    -1.555286    1.81295
> 6
-------------+----------------------------------------------------------------
> -
g2017        |
 t_2012_2013 |  -.2849763   .6552211    -0.43   0.664    -1.569186    .999233
> 5
 t_2013_2014 |   .3871202   1.012556     0.38   0.702    -1.597454    2.37169
```

```
> 4
 t_2014_2015 |   1.599966   .7708362     2.08   0.038     .0891552    3.11077
> 7
 t_2015_2016 |  -1.005993   .7385355    -1.36   0.173    -2.453496    .441509
> 6
 t_2016_2017 |   1.865914   .6845872     2.73   0.006     .5241473     3.2076
> 8
 t_2016_2018 |   .7371248   .7939142     0.93   0.353    -.8189183    2.29316
> 8
 t_2016_2019 |   .4129377   .8878153     0.47   0.642    -1.327148    2.15302
> 4
 t_2016_2020 |   .2958839   .9986705     0.30   0.767    -1.661474    2.25324
> 2
 t_2016_2021 |   .7970169   1.091332     0.73   0.465    -1.341955    2.93598
> 9
-----------------------------------------------------------------------------
> -
Control: Not yet Treated

See Callaway and Sant'Anna (2021) for details

.
. // csdid_plot, title("Group 2013") name(g2013) group(2013)
. // csdid_plot, title("Group 2014") name(g2014) group(2014)
. // csdid_plot, title("Group 2015") name(g2015) group(2015)
. // csdid_plot, title("Group 2016") name(g2016) group(2016)
. // csdid_plot, title("Group 2017") name(g2017) group(2017)
.
. estat pretrend
Pretrend Test. H0 All Pre-treatment are equal to 0
chi2(10) =     8.4945
p-value  =     0.5807

.
. est clear

.
```

```
. estat simple, estore(simp)
Average Treatment Effect on Treated
------------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+----------------------------------------------------------------
> -
         ATT |   3.496479   1.365452     2.56   0.010     .8202416    6.17271
> 7
------------------------------------------------------------------------------
> -

.
. esttab simp using "./regression/reg1.tex", replace b(3) se(3) star(* 0.10 *
> * 0.05 *** 0.01)
(output written to ./regression/reg1.tex)

.
. estat group, estore(group)
ATT by group
------------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+----------------------------------------------------------------
> -
    GAverage |   3.054364   1.107451     2.76   0.006     .8838004    5.22492
> 8
       G2013 |    20.2163   11.66155     1.73   0.083    -2.639922    43.0725
> 3
       G2014 |   5.747554   2.591886     2.22   0.027     .6675505    10.8275
> 6
       G2015 |   1.359106   .5263608     2.58   0.010     .3274573    2.39075
> 4
       G2016 |   .1934551   .6014674     0.32   0.748    -.9853992    1.37230
> 9
       G2017 |   .8217754   .6556734     1.25   0.210    -.4633209    2.10687
> 2
------------------------------------------------------------------------------
> -
```

```
.
. esttab group using "./regression/reg4.tex", replace b(3) se(3) star(* 0.10
> ** 0.05 *** 0.01)
(output written to ./regression/reg4.tex)

.
. est clear

.
. estat calendar, estore(calendar)
ATT by Calendar Period
----------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+--------------------------------------------------------------
> -
    CAverage |   2.948084   1.157153     2.55   0.011     .6801055    5.21606
> 3
       T2013 |   1.818255   2.413148     0.75   0.451    -2.911427    6.54793
> 8
       T2014 |  -.0297837   1.056047    -0.03   0.978    -2.099597     2.0400
> 3
       T2015 |   1.028081   1.344472     0.76   0.444    -1.607036    3.66319
> 8
       T2016 |     1.8486   1.240309     1.49   0.136    -.5823617    4.27956
> 1
       T2017 |   3.890843   1.717729     2.27   0.024     .5241568     7.2575
> 3
       T2018 |   4.142033   1.864734     2.22   0.026     .4872211    7.79684
> 5
       T2019 |   4.165509   1.569486     2.65   0.008     1.089374    7.24164
> 4
       T2020 |    4.74929   1.656244     2.87   0.004     1.503111    7.99546
> 9
       T2021 |    4.91993   1.566144     3.14   0.002     1.850344    7.98951
> 5
----------------------------------------------------------------------------
> -
```

```
.
. esttab calendar using "./regression/reg2.tex", replace b(3) se(3) booktabs
> star(* 0.10 ** 0.05 *** 0.01)
(output written to ./regression/reg2.tex)

.
.
. est clear

.
. estat event, estore(event)
ATT by Periods Before and After treatment
Event Study:Dynamic effects
------------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+----------------------------------------------------------------
> -
     Pre_avg |  -.0971149    .197541    -0.49   0.623    -.4842882    .290058
> 3
    Post_avg |   7.447422   3.221124     2.31   0.021     1.134134    13.7607
> 1
         Tm4 |  -.2849763   .6552211    -0.43   0.664    -1.569186    .999233
> 5
         Tm3 |  -.1443407   .5437681    -0.27   0.791    -1.210107    .921425
> 1
         Tm2 |  -.1348977   .3354505    -0.40   0.688    -.7923687    .522573
> 3
         Tm1 |    .175755   .3475825     0.51   0.613    -.5054943    .857004
> 2
         Tp0 |    .411009   .3468013     1.19   0.236    -.2687091    1.09072
> 7
         Tp1 |   .8132869   .6581812     1.24   0.217    -.4767245    2.10329
> 8
         Tp2 |   1.762168   1.037048     1.70   0.089    -.2704077    3.79474
> 4
         Tp3 |   3.014832    1.37351     2.19   0.028     .3228018    5.70686
> 2
         Tp4 |   4.231413   1.625482     2.60   0.009     1.045526      7.417
> 3
         Tp5 |   4.771354   1.822684     2.62   0.009     1.198959    8.34374
> 8
         Tp6 |   6.187763   2.060824     3.00   0.003     2.148622     10.226
> 9
         Tp7 |   13.29676   4.810917     2.76   0.006      3.86754    22.7259
> 9
         Tp8 |   32.53821   18.11897     1.80   0.073     -2.97432    68.0507
> 4
```

```
--------------------------------------------------------------------------------
> -

.
. estat event
ATT by Periods Before and After treatment
Event Study:Dynamic effects
--------------------------------------------------------------------------------
> -
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval
> ]
-------------+------------------------------------------------------------------
> -
     Pre_avg |  -.0971149    .197541    -0.49   0.623    -.4842882    .290058
> 3
    Post_avg |   7.447422   3.221124     2.31   0.021     1.134134    13.7607
> 1
         Tm4 |  -.2849763   .6552211    -0.43   0.664    -1.569186    .999233
> 5
         Tm3 |  -.1443407   .5437681    -0.27   0.791    -1.210107    .921425
> 1
         Tm2 |  -.1348977   .3354505    -0.40   0.688    -.7923687    .522573
> 3
         Tm1 |    .175755   .3475825     0.51   0.613    -.5054943    .857004
> 2
         Tp0 |    .411009   .3468013     1.19   0.236    -.2687091    1.09072
> 7
         Tp1 |   .8132869   .6581812     1.24   0.217    -.4767245    2.10329
> 8
         Tp2 |   1.762168   1.037048     1.70   0.089    -.2704077    3.79474
> 4
         Tp3 |   3.014832    1.37351     2.19   0.028     .3228018    5.70686
> 2
         Tp4 |   4.231413   1.625482     2.60   0.009     1.045526      7.417
> 3
         Tp5 |   4.771354   1.822684     2.62   0.009     1.198959    8.34374
> 8
         Tp6 |   6.187763   2.060824     3.00   0.003     2.148622     10.226
> 9
         Tp7 |   13.29676   4.810917     2.76   0.006      3.86754    22.7259
> 9
         Tp8 |   32.53821   18.11897     1.80   0.073     -2.97432    68.0507
> 4
--------------------------------------------------------------------------------
> -
```

```
.
. csdid_plot, title("Event Study") name(event)

.
. esttab event using "./regression/reg3.tex", replace b(3) se(3) star(* 0.10
> ** 0.05 *** 0.01)
(output written to ./regression/reg3.tex)

.
. log close
      name:  <unnamed>
       log:  /Users/tylerkim/Documents/14.33/data/summary.log
  log type:  text
 closed on:  29 Nov 2022, 00:53:14
-----------------------------------------------------------------------------
> -----------------------------------------------------------------------------
> ----------------
```

```
clear all
capture log close
set more off
cd /Users/tylerkim/Documents/14.33/data

log using summary.log, replace

use uberx_data.dta

rename POPESTIMATE2014 population

encode city, g(ncity)

drop city

drop if year == 2022

rename ncity city

est clear

estpost summarize median_aqi population if treated != 0

esttab using "./tables/table1.tex", replace cells("count mean(fmt(%20.2fc)) sd(fmt(%20.2fc)) min max(fmt(%20.2fc))") nonumber noobs booktabs collabels("N" "Mean" "SD" "Min" "Max")

est clear

estpost summarize median_aqi population if treated == 0

esttab using "./tables/table2.tex", replace cells("count mean(fmt(%20.2fc)) sd(fmt(%20.2fc)) min max(fmt(%20.2fc))") nonumber noobs booktabs collabels("N" "Mean" "SD" "Min" "Max")

est clear

csdid median_aqi population, ivar(city) time(year) gvar(treated) notyet method(reg)

// csdid_plot, title("Group 2013") name(g2013) group(2013)
// csdid_plot, title("Group 2014") name(g2014) group(2014)
// csdid_plot, title("Group 2015") name(g2015) group(2015)
// csdid_plot, title("Group 2016") name(g2016) group(2016)
// csdid_plot, title("Group 2017") name(g2017) group(2017)

estat pretrend

```

```
esttab simp using "./regression/reg1.tex", replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01)

estat group, estore(group)

esttab group using "./regression/reg4.tex", replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01)

est clear

estat calendar, estore(calendar)

esttab calendar using "./regression/reg2.tex", replace b(3) se(3) booktabs star(* 0.10 ** 0.05 *** 0.01)


est clear

estat event, estore(event)

estat event

csdid_plot, title("Event Study") name(event)

esttab event using "./regression/reg3.tex", replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01)

log close

clear all

```