Ride or Die?
Metropolitan Bikeshare Systems and Pollution

Sean Furuta

University of California, Berkeley

ECON 195B

Fall 2020

Abstract

This paper tries to establish causality between particulate matter 2.5 concentration and bikeshare quantity demanded, by applying the instrumental variable framework, on a panel of United States cities with bikeshares from 2017 through 2019. The results are statistically significant, showing evidence of increasing bikeshare rides per 1000 residents as pollution concentration increase. These findings are robust to test of endogeneity and instrument strength. From the results, I estimate that a microgram per cubic meter increase in PM2.5 increase rides per 1000 resident by 0.215. This translates to 215 rides for a city population of 1 million.

Keywords

Bikeshare, Scheme, System, Instrumental Variable Regression, Particulate Matter 2.5

Acknowledgements

I thank Professor Graham, Professor Eichengreen, Matthew Tauzer, and Stephanie Bonds for the continuous support and feedbacks as my advisors. All of the remaining errors are my responsibility.

1 Introduction

As an automotive centric society, the United States and its major cities suffer from many problems ranging from traffic congestion and lack of parking. It also contributes to macro environmental problems like climate change and air pollution by burning fossil fuels which produce greenhouse gases. According to the EPA, almost 27% of greenhouse gas emissions originate from automobile tailpipes (Hamilton and Wichman, 2015). Although automobiles in conjunction with a well-connected highway system are effective in long distance travel, they fail miserably in short distances or “last mile” travel within dense cities. “This ‘last-mile’ problem is thought to deter transit use among riders with auto access, even when high quality transit service is provided for the majority of the trip distance” (Zellener et al., 2016).

Local governments have enacted policies to reduce these problems like highway tolls, carpool lanes, and robust public transport system. Also, consumers in major cities around the world have embraced ride sharing schemes like Uber or Lyft to get access to transportation without owning an automobile. The combination of public and private options has led to major cities building bikesharing systems. For example, the city of San Francisco partnered with Ford and Lyft to create a bikesharing system that spans a large portion of the Bay area. Bikeshares do not require fossil fuels to transport individuals across the city and can lower street congestion by taking up less space than a full-size automobile. Bikes smaller size allows them to take paths through cities that may not be available to conventional automobiles making them more versatile in cities. According to the Center of Disease Control, 38% of United States adult population is affected by obesity which makes bikeshares a healthy alternative to automobile transport. Thus, bikeshares can help solve some of the critical pain points of an automotive society and with growing city population in the U.S., city government must be able to shift transportation options away from personal automobiles.

Also, recent technological innovations have made adoption of bikeshare systems more fluid. The most impactful of these technologies are smartphones and their associated application that enable the many sensors on the bikes. Smaller and lighter batteries have also enabled bikeshare systems to incorporate e-bikes to allow for quicker motorized trips. Further technological advancements like dockless bikes will only improve capacity for these bikeshare systems across the world and attract more consumers away from automobile transportation for commutes.

Despite many of these technological advancements being available in the U.S., American cities are relatively slow to bikeshare adoption compared to Europe and China. Some cities to note are London, Paris, Copenhagen, Amsterdam, and Goyang (Nair et al., 2013). More extensive research on implementation and consumer behavior has been conducted in these regions. Two articles that closely relate to my paper are Li and Kamargianni [2018] which explores the relationship between pollution and bikeshare demand in the Chinese city of Taiyuan and Woodcock et al. [2014], which explores health effects on riders. There are concerns surrounding exposure to pollution while utilizing open air transport systems like bikeshare. Both papers found significant effects on consumer due to pollution in the air which I will later discuss in my literature review. Paradoxically, the pollution and congestion problem bikeshares were intended to solve may be a deterrent to using them. However, such interaction between pollution and bikeshare demand has not been thoroughly explored on U.S. cities. Thus, my hypothesis is as pollution concentration increases in cities with bikeshare systems, riders will be discouraged from using the system to protect themselves from pollutant exposure, decreasing demand for bikeshare rides.

Alternatively, as pollution concentration increases, bikeshare demand may increase as a result of environmentally conscious consumers taking into account the need to reduce emissions and choose to use bikeshares in place of automotive transportation.

In the next section, I will outline the relevant literature as it relates to bikeshare and my choice of methodology. In section 3 I outline the choice of data and data sources used. I describe my empirical strategy and its implications in section 4 and in section 5, I discuss my results and its impacts. Section 6 contains my robustness check of my empirical strategy and finally, I conclude my paper in section 7 by summarizing my results.

2 Literature Review

My paper will focus on the effects of pollution on the quantity demanded for bikeshare rides. The growing concern around pollution and last mile transportation has given rise to bikeshare systems popping up in nearly 800 major cities around the world (Fishman, 1). Bikeshares serve as a green alternative to personal automobiles and can efficiently navigate the dense cityscape. However, "the majority of scheme users are substituting from sustainable modes of transport rather than the car" (Fishman et al, 1). Thus, these subscribed users may already be quite environmentally conscious or interested in using cheaper forms of public transportation to get around. The former captures one hypothesis that as pollution worsens in cities with bikeshares, environmentally conscious citizens may choose to use bikeshares to lower their carbon footprint, thus increasing bikeshare demand as pollution increases. Alternatively, potential bikeshare users may be discouraged from using bikeshare systems over enclosed auto transport like trains, buses, or cars due to an increased level of pollution and potential health risks associated. Woodcock et al found that Particulate

Matter 2.5 exposure from average bikeshare rides in London were similar to that of the London underground. Overall, the health benefits for bikeshare systems including less injuries, more exercise, etc. outweigh harms like pollution exposure (Woodcock et al, 2014). My research will explore how the negative health effects of pollution exposure drives decision making around bikeshares.

Such analysis has been conducted in the Chinese city of Taiyuan, which operates one of the most in demand bike-sharing schemes in China, and it was discovered that pollution had negative effects on bikeshare demand (Li and Kamargianni, 2018). My research will focus on American cities and their bikeshare systems, and thus, preferences amongst consumers may be different from Chinese consumers. In China, pollution levels are higher in metropolitan areas than compared to US cities and thus, an increased awareness of potential negative health effects may alter Chinese consumers decisions. Li and Kamargianni used stated preference surveys to analyze consumer transportation choices and constructed a nested logit model. In contrast, my research will used realized quantity demanded data and a 2-stage instrumental variable regression to capture the causal effect of increase pollution concentration on quantity demanded.

Another key aspect in bikeshare system is convince and consumers consider many aspects when choosing to use a bikeshare. For example the use "'personal credit' rather than money if they do not return public bikes within the free use hours) and universal cards (integrating bikesharing systems into other urban transit systems through the use of a rechargeable smart card that can cover a range of payments and trips) can significantly raise bikesharing daily use and turnover rate" (Zhao et al., 2014). As a result, I choose to use fixed entity effects for each city and its respective bikeshare system to capture the unique characteristics of each system. In future research, one should explore the impacts of each of these characteristics on rider demand, but that is outside the scope of this paper.

Another key factor in bikeshare demand and overall transportation demand is weather effects. Martinez [2017] conducted a simple OLS regression on ride counts in the New York City's Citi bikeshare system against several weather factors including precipitation, temperature, snowfall, etc. from 2013 to 2015. Obviously, Martinez's coefficients may be bias due to factors covarying with error terms in the OLS, but they may inform us about general trends. In summary, ride counts decreased with precipitation and snowfall and increased with temperature (Martinez, 2017). Martinez also used simple ride counts to quantify demand which makes it difficult to compare to other cities and does not control for changing populations. However, other research has found that extremely high temperature decreases overall ridership and trip duration which is to be expected given possible health repercussions due to extremely heat like heat stroke (Gebhart and Noland, 2013). Also, rain can have varying effects on quantity of rides demanded depending on its proximity to other forms of transportation like subway stations (Gebhart and Noland, 2014). Gebhart and Noland used log ride counts to quantify demand to control for outliers, but I will use ride count per 1000 residents in a city to control for population variation across cities. Few research papers compare several different cities and weather effects across each city like my research. I will also be including other weather factors that are not included in past literature like humidity, pressure, and dew point. In conjunction with weather effects, most literature relating to bikeshare demand includes seasonal dummies and also day of the week dummies to capture seasonal and weekday commute effects in the regression. Thus, I integrate fixed time effects in my instrumental variable regression as well as weekday dummies.

Finally, my 2 stage least square instrumental variable regression is based on Angrist and Imbens research to estimate average casual effect of variable treatment. The instrument must be valid or does not directly affect the potential outcomes and must monotone with respect to the treatment and outcome (Angrist and Imbens, 1995). Since my endogenous variable is pollution, I selected wind speed to be my instrument given that increased wind speed clear out particulate matter in a location, decreasing concentration of pollutants, and similar analysis has been conducted by Schwartz et al. [2017] in mortality rates in relation to pollution concentration. Similar to Card [1993], I use an interaction between wind speeds and day of week for my instrument to isolate variation in wind speed that do not direct effect ride quantity by multiplying a variable that does not vary with this direct effect.

3 Data

The empirical analysis is based on a city-by-day panel dataset built from multiple publicly available sources, from 2017-2019. Each city requires rideshare companies to publish ridership data including variables such as rides duration, start and end time, and rider type (casual or subscriber). The dataset is comprised of 5664 city-day-usertype observations from San Francisco, Los Angeles, Boston, and Washington DC. User type describes wether the users are subscriber or casual.

3.1 Quantity Demanded

By using ridership data from each rideshare organization for each city, I was able to aggregate daily ride count for the past 3 years. However, rides differ in duration as well as product type because some rides are for one time use while others are subscription based.

3.2 Treatment - Pollution

I acquired pollution data through the EPA's outdoor air quality database. I used a category of pollutant called Particulate Matter 2.5($PM_{2.5}$) which are tiny particles smaller than two and half microns and is produced from burning fuels in cars, buses, trains, etc. These particles can also be produced from burning wood in forest fires. In each city, there are multiple sensor sites dispersed across the city and I took an average for over each day-site observation to get an averaged $PM_{2.5}$ concentration (ug/m^3).

3.3 Weather Control

It is clear by looking at Figure 1 that the ride counts data is very seasonal and this may be a result of changes in temperature, humidity, pressure, precipitation and my instrument, wind speed. To account for this, I pulled daily weather data for each city from OpenWeatherMaps and merged it will the ridership data for each city-day observation.

3.4 Population

I was able to acquire projected total population for each city in 2018 from the United States Census Bureau. These projections are based on the 2010 national census and historical growth in each city. I was unable to find projections for 2017 and 2019 and in future analysis, I would use these population data points to properly calculate the ride count per 1000 residents.

3.5 Summary Statistics

Table 1 presents the means, standard deviations, quantiles, minimum, and maximum of the variables used in the empirical analysis.

Table 1: Summary Statistics Table for All Cities

Variables	Obs.	Mean	Std. Dev.	Min	25%	50%	75%	Max
Count	2764	6489.196093	4199.35859	21	3062.75	5997	9612.25	19113
Daily Mean PM2.5 Concentration	2764	8.084515	8.06006	-2.9	4.4	6.8	10.1	172.6
Temperature	2764	57.865449	15.834352	5.9	46.9	59.6	70.125	94.2
Dew Point	2764	45.48987	17.356372	-12.3	34.575	48.6	57.8	76.9
Humidity	2764	66.47424	15.30103	9.7	55.9	67.5	78	99.9
Wind Speed	2764	9.868017	4.065388	1.2	7	9.3	12.2	37.9
Pressure	2764	29.975796	0.277593	27.5	29.9	30	30.1	30.7
Precipitation	2764	0.10415	0.31466	0	0	0	0.02	4
Count per 1000	2764	9.051215	6.097057	0.030238	4.070916	8.224996	13.745631	27.210991
Year	2764	2017.81259	0.762143	2017	2017	2018	2018	2019

4 Empirical Strategy

4.1 Model

The proposed empirical strategy aims to capture the effects of pollution on the quantity of rides per 1000 residents.

$$Rides_{i,t} = \rho * pol_{i,t} + X'_{i,t} * \omega + \alpha_i + \theta_t + \varepsilon_{i,t}$$

$$pol_{i,t} = \beta * windspeed_{i,t} * dayofweek_t + X'_{i,t} * \delta + \alpha_i + \theta_t + \eta_{i,t}$$

Where quantity $Rides_{i,t}$ is the number of rides per 1000 residents in city i on day *t*. The two equations above describe and 2 stage least squares instrumental variable regression with wind speed as the instrument for pollution concentration. $pol_{i,t}$ is the variable for pollution concentration and ρ is the approximate change from an unit increase in $PM_{2.5}$ concentration. $X'_{i,t}$ are the controls for temperature, humidity, pressure, and precipitation, accounting for variation in weather. $windspeed_{i,t} * dayofweek_t$ is the instrumental variable for pollution concentration and β is the instrumental variable estimator. δ are the coefficients associated with the weather controls in the second equation. Also, a city dummy, α_i, is included to serve as a fixed entity effect to account for the different intercepts of each bikeshare's demand curve and a seasonal dummy, θ_t, is included to account for time variation and seasonal effects. Finally, $\varepsilon_{i,t}$ term is the city-day error in the first equation and $\eta_{i,t}$ is the city-day error in the second equation.

4.2 Identifying assumption

As I mentioned in my model section, I will be using instrumental variable framework to find an estimator for the treatment of pollution. There are two fundamental assumption in instrumental variable regression:

1) The instrument must be uncorrelated with the error term, $\varepsilon_{i,t}$
2) The instrument must be correlated with the casual term, $pol_{i,t}$

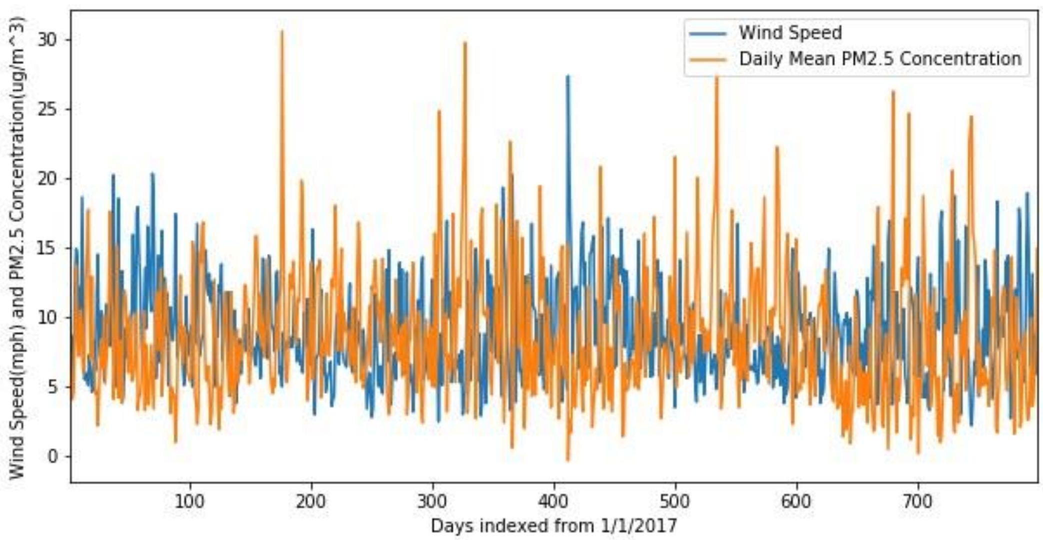
Wind speed-day of week interaction satisfies the first condition because it is most likely unrelated to ride counts because riders are not taking into consideration general wind patterns on a particular day when deciding to ride, so it cannot explain any of the variation in the error term. Although a rider might consider wind speeds in their decision, the direct effect of wind speeds on the decision to ride or not does not vary by day of the week. After all the direct effects of wind speed and day of the week are controlled for, we do not expect the interaction to explain any variability in rides demanded and its error term. As a result, both wind speed and day of the week indicators can be used as an instrument for rides demanded.

The second condition is satisfied because wind speeds are inversely related to pollution concentration. Higher wind speeds prevent the particulate matter from staying in one place, decreasing concentration. In the simple case which wind speed is an indicator and we have weekend indicator (W, S) for the instrument, we get the following equation for β:

$$\beta = E[pol|W = 1, S = 1] - E[pol|W = 1, S = 0] - \{E[pol \mid W = 0, S = 1] - E[pol|W = 0, S = 0]\}$$

This equation can be interpreted as the difference-in-difference of windy days on weekdays and weekends. Weekends will have a lower level of pollution due to traffic and windy days should have a lower concentration of pollution. Thus, the difference of pollution on windy days on a weekday and weekend, the first two terms, is negative and the difference of pollution on a non-windy days on a weekday and weekend, is also negative but slightly different due wind being more effective in clearing pollution concentration when congestion levels differ. Thus, β should be non-zero and our instrument and pollution must be highly correlated. This can be seen in Figure 2 from the peaks and trough of each line illustrate the negative correlation.

Figure 2: Wind Speed and Pollution Comparison over Time in DC



5 Results

Table 2 provides the estimated coefficients of the daily mean $PM_{2.5}$. concentration and weather controls. Each column utilized different combination of fixed effects including day of week, year-month, and city.

Column 1 has none of the three controls. Columns 2, 3, and 4 include city, year-month and day of the week, respectively. Column 5 includes city and year-month dummies. Column 6 includes city and day of week dummies. Column 7 includes day of week and year-month dummy. Finally, column 8 controls for all fixed effects.

Column 8 is the preferred specification because it uses all the controls. These results are inconsistent with my hypothesis of the effects of pollution on ride demand. In all of the columns ρ is estimated to be positive, meaning a unit increase in pollution concentration will result in an increase in rides. Looking at column 8, ρ is 0.215, meaning a unit increase in pollution concentration will approximately increase rides per million residents by 0.215.

I found these results odd, so I conducted on OLS regression and its results are in Table 3. In stark contrast to the IV regression's coefficient for pollution concentration is -0.0106, meaning as pollution concentration increases, rides per 1000 residents decreases. The OLS coefficient is bias when the factor covaries with the error term, so in order to control for such covariance that may be caused by an omitted variable, economist use IV regression to acquire the true coefficient. In the simple univariate case, we have the following formula:

$$\rho = \rho_{ols} + \frac{cov(\varepsilon_{i,t}, pol)}{var(pol)}$$

Plugging in our formula for pol, we see that the second term becomes covariance between ε and η. In this case, we see that the OLS coefficient is higher than the IV regression, so the covariant term between the error of the regression and pollution concentration must be negative. Thus, a possible omitted variable might covary negatively with pollution concentration and explain a positive increase in rides per 1000 resident. One possible candidate for such an effect is demand for automobile. Shocks in automobile demand could explain changes in pollution and effect demand for bikeshares. For example, these shocks would negatively correlate with unobserved bikeshare demand effects in the error term satisfying the first requirement and it would also explain an increase in rides because commuters may be choosing to forgo their cars to use bikeshares. I even conducted a similar IV regression on each individual city and found similar results, strongly contradicting my hypothesis.

6 Robustness Check

I conducted to robustness checks, a test for endogeneity and instrument strength, to ensure the stability of my regression analysis.

6.1 Endogeneity

In order to check if my casual variable pollution concentration is truly endogenous, I test the null hypothesis of exogeneity through the Durbin and Wu-Hausmen scores. In both cases, each had a p-value less than 0.05 and are therefore statistically significant. Thus, I reject the null hypothesis that pollution concentration is exogenous with respects to rides per 1000 residents.

6.2 Instrument Strength

In order to check the strength of wind speed as an instrumental variable for pollution concentration, I conduct a Wald test at a 5% threshold and compare it to my F-statistic. The results show that my F-statistic 638.3 and my Wald test threshold is 15.4. Since my F-statistics is significantly higher than the Wald test threshold, I can confidently say that wind speed is a strong instrument for pollution concentration. However, my partial r-squared statistics was low at 0.25.

6.3 Limitations

One explanation for this result may be that by not adjusting the population for year 2017 and 2019, I do not control for change in population and in major cities, population will most likely grow year over year. A high population and population density may result in higher pollution concentration and more rides. Also, wind speeds can affect rides per 1000 residents through the different weather controls and thus, the instrumental variable regression cannot properly control for pollution concentration. One way to solve this specific issue is to remove the weather controls that are affected by wind speed. Finally, day of the week may vary with the direct effect wind speed has on ride demand which would make the instrument invalid.

7 Conclusion

This paper provides evidence for positive effects of particulate matter 2.5 concentration on bikeshare quantity demanded. Rather than increased pollution concentration deterring individuals from utilizing the bikeshare service in order to protect their health, increased pollution concentration may encourage them to act and reduce emission by using bikeshares to decrease overall pollution that they may see in their cities. This research also contributes to the ongoing transportation literature relating to last mile commutes and possible alternatives to automobile transportation.

I implement a 2-stage instrumental variable regression with pollution concentration as my endogenous variable and wind speed in miles per hour and day of week indicators to establish a causal relationship between pollution concentration and bikeshare rides per 1000 resident in a city. To provide supporting evidence for my analysis, I conduct 2 robustness checks for endogeneity and instrument strength and each bolster my argument by providing statistically significant results. I also compared the IV regression results to that of an OLS to explain possible discrepancies. From the results, I estimate that a microgram per cubic meter increase in $PM_{2.5}$ increase rides per 1000 resident by 0.215. This translates to 215 rides for a city population of 1 million.

These results should encourage wider adoption of bikeshare systems in cities across the United States. By creating bikeshare infrastructure, cities can allow commuters to have a green alternative to transportation as pollution in cities increases. In conjunction with Woodcock et al., this research shows that as pollution concentration increases more rides are taken, improving those riders' overall health despite exposure to pollution.

8 References

Angrist, Joshua, and Guido Imbens. "Identification and Estimation of Local Average Treatment Effects." Econometrica, Feb. 1995, doi:10.3386/t0118.

Card, David. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." Aspects of Labour Economics, 1993, doi:10.3386/w4483.

Fishman, Elliot, Simon Washington & Narelle Haworth (2013) Bike Share: A Synthesis of the Literature, Transport Reviews, 33:2, 148-165, DOI: 10.1080/01441647.2013.775612

Fishman, Elliot (2016) Bikeshare: A Review of Recent Literature, Transport Reviews, 36:1, 92-113, DOI: 10.1080/01441647.2015.1033036

Gebhart, Kyle, and Robert B. Noland. The impact of weather conditions on capital bikeshare trips. No. 13-0563. 2013.

Gebhart, Kyle, and Robert B. Noland. The impact of weather conditions on bikeshare trips in Washington, DC. Transportation 41, 1205–1225 (2014) doi:10.1007/s11116-014-9540-7

Hamilton, Timothy and Wichman, Casey, Bicycle Infrastructure and Traffic Congestion: Evidence from DC's Capital Bikeshare (August 20, 2015). Resources for the Future Discussion Paper 15-20. Available at SSRN: https://ssrn.com/abstract=2649978

Li, Weibo, and Maria Kamargianni. “Air Pollution and Seasonality Effects on Mode Choice in China.” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2634, no. 1, 1 Jan. 2017, pp. 101–109., doi:10.3141/2634-15.

Mitchell, W.J., Borroni-Bird, C.E., Burns, L.D., Reinventing the Automobile-Personal Urban

Mobility for the 21st Century, The MIT Press, March 2010.

Martinez, Mark (2017) "The Impact Weather Has on NYC Citi Bike Share Company Activity," *Journal of Environmental and Resource Economics at Colby*: Vol. 4: Issue 1, Article 12.

Schwartz J, Bind MA, Koutrakis P. 2017. Estimating causal effects of local air pollution on daily deaths: effect of low levels. Environ Health Perspect 125:23–29

Woodcock, J., et al. “Health Effects of the London Bicycle Sharing System: Health Impact Modelling Study.” *Bmj*, vol. 348, no. feb13 1, 2014, doi:10.1136/bmj.g425.

Zhao, Jinbao, et al. “Ridership and Effectiveness of Bikesharing: The Effects of Urban Features and System Characteristics on Daily Use and Turnover Rate of Public Bikes in China.” *Transport Policy*, vol. 35, Oct. 2014, pp. 253–264., doi:10.1016/j.tranpol.2014.06.008.

9 Appendix

9.1 Data – Variable construction and sources

Cities were chosen based on data availability online and if the bikeshare system in each city had been operating during the time period between 2017-2019.

9.1.1 Quantity Demanded

Each dataset was similarly structured which made aggregation quite easy. Each row of the dataset was a particular trip with various additional information outlined in the data section. I grouped each trip by start date and counted each trip to obtain date-count observation for each city from 2017-2019. Some of the bikeshare systems had data from years prior to 2017, but I choose not to use them because the system was not well developed.

9.1.2 Pollution Concentration

The EPA pollution concentration dataset included observations from multiple sensor station in each city and at multiple different times in the day. Each of these observations were again grouped by date and given equal weight when averaged and these averaged daily pollution concentrations were used in the IV regression.

9.1.3 Weather Control

OpenWeatherMaps provides online datasets for daily averaged weather observations for many cities across the United States. I collected data for each city and day in the ridership dataset. Data was collected from a single sensor station. Further research should attempt to aggregate daily weather data across several stations in a city or stations close to bikeshare docks.

9.2 Figures and Tables

Figure 1: Counts of rides per day in San Francisco from July 2017 to December 2018

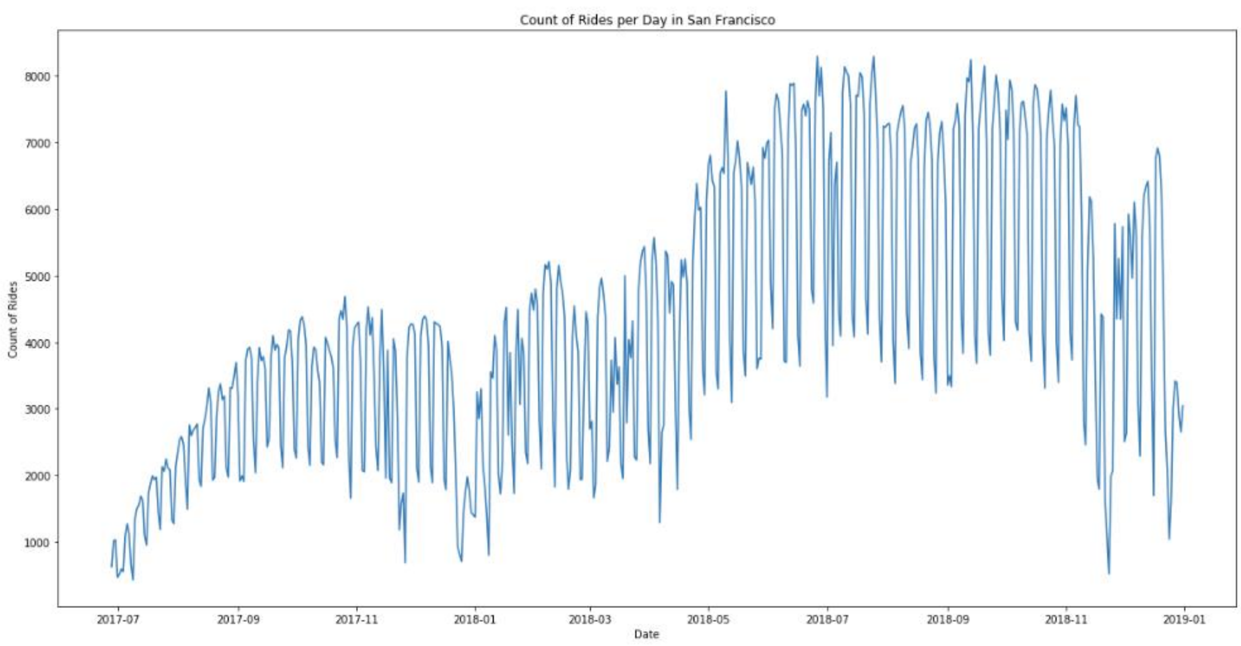


Table 2: Pollution IV Regression Results

VARIABLES	(1) countper1000	(2) countper1000	(3) countper1000	(4) countper1000	(5) countper1000	(6) countper1000	(7) countper1000	(8) countper1000
dailymeanpm25concentration	1.279**	1.131**	1.425*	0.290*	1.229*	0.199*	0.338*	0.215*
	(0.612)	(0.517)	(0.799)	(0.206)	(0.639)	(0.150)	(0.219)	(0.157)
temperature	0.298	0.280*	0.200	-0.0219	0.179	0.0243	-0.0390	-0.00101
	(0.212)	(0.155)	(0.193)	(0.0721)	(0.133)	(0.0463)	(0.0564)	(0.0355)
dewpoint	-0.260	-0.239	-0.208	0.141	-0.179	0.0801	0.148*	0.0833*
	(0.260)	(0.189)	(0.275)	(0.0881)	(0.180)	(0.0557)	(0.0779)	(0.0467)
precipitation	4.069***	2.095**	4.586**	1.882***	2.565*	0.425	1.903***	0.430
	(1.439)	(1.020)	(2.048)	(0.489)	(1.425)	(0.301)	(0.572)	(0.359)
humidity	0.0818	0.0820	0.0666	-0.0905**	0.0531	-0.0630**	-0.0872**	-0.0635***
	(0.113)	(0.0863)	(0.120)	(0.0384)	(0.0812)	(0.0255)	(0.0343)	(0.0211)
pressure	5.129***	2.039**	5.492***	4.731***	2.094**	0.939***	4.622***	0.833***
	(0.557)	(0.890)	(0.858)	(0.202)	(1.064)	(0.273)	(0.264)	(0.288)
windspeed	0.735**	0.606**	0.724*	0.160	0.583*	0.0657	0.157	0.0524
	(0.356)	(0.300)	(0.416)	(0.120)	(0.335)	(0.0876)	(0.115)	(0.0827)
2.ucod	5.285***	5.335***	5.267***	5.283***	5.319***	5.346***	5.285***	5.345***
	(0.271)	(0.232)	(0.290)	(0.101)	(0.245)	(0.0769)	(0.103)	(0.0767)
Constant	-180.8***	-84.39**	-189.3***	-143.3***	-82.89**	-29.54***	-140.6***	-26.34**
	(28.68)	(36.64)	(40.42)	(9.853)	(41.99)	(10.91)	(11.68)	(10.93)
City dummy	No	Yes	No	No	Yes	Yes	No	Yes
Year Month dummy	No	No	Yes	No	Yes	No	Yes	Yes
Day of Week dummy	No	No	No	Yes	No	Yes	Yes	Yes
Observations	5,664	5,664	5,664	5,664	5,664	5,664	5,664	5,664
R-squared				0.301		0.598	0.272	0.601

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Pollution OLS Regression Result

VARIABLES	(1) countper1000	(2) countper1000	(3) countper1000	(4) countper1000	(5) countper1000	(6) countper1000	(7) countper1000	(8) countper1000
dailymeanpm25concentration	-0.0293***	-0.0127***	-0.0220***	-0.0301***	-0.0101***	-0.0134***	-0.0226***	-0.0106***
	(0.00439)	(0.00378)	(0.00397)	(0.00433)	(0.00346)	(0.00374)	(0.00389)	(0.00339)
temperature	-0.120***	-0.0266	-0.112***	-0.124***	-0.0331	-0.0323	-0.116***	-0.0393*
	(0.0242)	(0.0218)	(0.0243)	(0.0241)	(0.0223)	(0.0218)	(0.0242)	(0.0222)
dewpoint	0.265***	0.145***	0.261***	0.269***	0.135***	0.151***	0.264***	0.140***
	(0.0252)	(0.0229)	(0.0255)	(0.0251)	(0.0237)	(0.0229)	(0.0254)	(0.0237)
precipitation	1.149***	0.0138	0.983***	1.171***	-0.0771	0.0403	1.006***	-0.0494
	(0.142)	(0.120)	(0.134)	(0.138)	(0.121)	(0.115)	(0.131)	(0.116)
humidity	-0.144***	-0.0927***	-0.136***	-0.145***	-0.0864***	-0.0952***	-0.137***	-0.0887***
	(0.0118)	(0.0105)	(0.0119)	(0.0117)	(0.0108)	(0.0105)	(0.0119)	(0.0108)
pressure	4.564***	0.624***	4.294***	4.602***	0.487***	0.685***	4.330***	0.549***
	(0.191)	(0.181)	(0.198)	(0.189)	(0.178)	(0.178)	(0.196)	(0.176)
windspeed	-0.0219**	-0.0551***	-0.0258**	-0.0255**	-0.0626***	-0.0573***	-0.0301***	-0.0654***
	(0.0101)	(0.00941)	(0.0104)	(0.01000)	(0.00956)	(0.00921)	(0.0103)	(0.00934)
2.ucod	5.281***	5.349***	5.290***	5.282***	5.351***	5.349***	5.291***	5.351***
	(0.0789)	(0.0664)	(0.0765)	(0.0782)	(0.0650)	(0.0655)	(0.0758)	(0.0642)
Constant	-130.2***	-15.50***	-123.6***	-131.3***	-12.16**	-17.08***	-124.6***	-13.75**
	(6.251)	(5.651)	(6.413)	(6.208)	(5.583)	(5.563)	(6.372)	(5.507)
City dummy	No	Yes	No	No	Yes	Yes	No	Yes
Year Month dummy	No	No	Yes	No	Yes	No	Yes	Yes
Day of Week dummy	No	No	No	Yes	No	Yes	Yes	Yes
Observations	5,664	5,664	5,664	5,664	5,664	5,664	5,664	5,664
R-squared	0.565	0.706	0.594	0.574	0.719	0.714	0.603	0.726

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, *
$p<0.1$

Table 4: Boston and Washington DC Pollution IV Regression Results

City	Boston				Washington DC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000
dailymeanpm25concentration	2.097**	1.964**	-0.0341	-0.276*	1.378**	1.610**	-0.0642	-0.0559*
	(1.064)	(0.933)	(0.210)	(0.154)	(0.620)	(0.754)	(0.309)	(0.328)
temperature	0.771	0.457	-0.207**	-0.131**	0.480**	0.389*	0.112	0.00493
	(0.510)	(0.281)	(0.105)	(0.0579)	(0.214)	(0.235)	(0.0968)	(0.0937)
dewpoint	-0.924	-0.741*	0.331**	0.240***	-0.396*	-0.411	0.0189	0.135
	(0.644)	(0.433)	(0.131)	(0.0804)	(0.233)	(0.299)	(0.106)	(0.122)
precipitation	3.541**	3.171**	0.456	-0.365	2.806**	3.355**	-0.131	0.0454
	(1.651)	(1.560)	(0.343)	(0.285)	(1.338)	(1.576)	(0.653)	(0.673)
humidity	0.404	0.303	-0.169***	-0.123***	0.0460	0.0144	-0.0607*	-0.114***
	(0.293)	(0.187)	(0.0595)	(0.0347)	(0.0809)	(0.0940)	(0.0346)	(0.0362)
pressure	0.923	-0.782	0.707***	0.418**	0.545	-1.363	0.326	0.361
	(0.754)	(0.816)	(0.205)	(0.212)	(0.735)	(1.092)	(0.291)	(0.435)
windspeed	0.429*	0.357*	-0.0742	-0.112***	0.630*	0.671*	-0.143	-0.170
	(0.254)	(0.198)	(0.0511)	(0.0343)	(0.334)	(0.383)	(0.167)	(0.167)
2.ucod	4.405***	4.409***	4.433***	4.436***	8.286***	8.286***	8.286***	8.286***
	(0.320)	(0.271)	(0.0869)	(0.0800)	(0.278)	(0.298)	(0.109)	(0.105)
Constant	-73.63*	-9.554	-10.15	-5.052	-44.47	15.56	-8.957	-5.241
	(39.76)	(21.05)	(9.065)	(6.240)	(28.76)	(27.11)	(12.31)	(9.948)
Year Month dummy	No	Yes	No	Yes	No	Yes	No	Yes
Day of Week dummy	No	No	Yes	Yes	No	No	Yes	Yes
Observations	1608	1608	1608	1608	1410	1410	1410	1410
R-squared	0.656	0.711	0.677	0.726	0.719	0.714	0.787	0.801

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Los Angeles and San Francisco Pollution IV Regression Results

City	Los Angeles				San Francisco			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000	countper1000
dailymeanpm25concentration	0.000635	0.000358	0.00222	0.00144	0.00791	-0.182*	0.00949	0.00749
	(0.00476)	(0.00465)	(0.00397)	(0.00425)	(0.0692)	(0.102)	(0.0367)	(0.0246)
temperature	0.000867	-0.000590	0.000978	-0.000549	-0.0249	-0.160	0.00255	-0.0209
	(0.000918)	(0.00116)	(0.000959)	(0.00120)	(0.0694)	(0.122)	(0.0596)	(0.0543)
dewpoint	-2.71e-05	0.000260	-0.000652	-0.000284	0.0397	-0.0312	0.00694	0.0354
	(0.00190)	(0.00264)	(0.00174)	(0.00252)	(0.0654)	(0.125)	(0.0604)	(0.0569)
precipitation	-0.0263	-0.0130	-0.0205	-0.00961	-0.593	-0.689	-0.811**	-0.268
	(0.0463)	(0.0499)	(0.0455)	(0.0516)	(0.478)	(0.587)	(0.332)	(0.246)
humidity	7.23e-05	-0.000345	0.000338	-0.000114	-0.0228	-0.0621	-0.00717	-0.0220
	(0.000859)	(0.00115)	(0.000844)	(0.00115)	(0.0319)	(0.0540)	(0.0282)	(0.0258)
pressure	-0.0142	0.0106	-0.00143	0.0213	-0.319	-1.057	-0.275	0.799
	(0.0445)	(0.0516)	(0.0419)	(0.0520)	(1.127)	(1.316)	(0.737)	(0.500)
windspeed	-0.000468	-2.54e-05	-0.000166	9.49e-05	0.0492	-0.242*	0.0572	0.00102
	(0.00126)	(0.00125)	(0.00127)	(0.00130)	(0.0750)	(0.129)	(0.0415)	(0.0337)
2.ucod	0.0139***	0.0142***	0.0141***	0.0142***	3.483***	3.483***	3.483***	3.483***
	(0.00488)	(0.00477)	(0.00494)	(0.00479)	(0.101)	(0.162)	(0.0933)	(0.0791)
Constant	0.388	-0.247	0.00274	-0.566	11.02	48.02	8.700	-24.39
	(1.317)	(1.523)	(1.243)	(1.539)	(37.75)	(46.95)	(24.25)	(16.71)
Year Month dummy	No	Yes	No	Yes	No	Yes	No	Yes
Day of Week dummy	No	No	Yes	Yes	No	No	Yes	Yes
Observations	1130	1130	1130	1130	1380	1380	1380	1380
R-squared	0.063	0.097	0.035	0.088	0.529		0.598	0.711

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1