Safia Read

May 5, 2024

QTM 345

QTM 345 Final Paper

I. Introduction

This study was conducted by Maximillian Auffhammer and Ryan Kellog and published in 2011. It aimed to study the effect of U.S. gasoline regulation policies on ozone concentrations. Ozone is a pollutant that is formed by the reaction of volatile organic chemicals (VOCs) and nitrogen oxides (NO_x). Exposure to high ozone concentrations can lead to asthma, susceptibility to pneumonia and bronchitis, and increased rates of early mortality. VOCs can be produced as vehicle emissions, industrial byproducts, or through biogenic factors. Therefore, national and statewide legislation has been enacted to try to curb ozone emissions. However gasoline regulations have tradeoffs, namely the effectiveness of the measure against the cost of monitoring, and effects that regulation has on the gasoline market. This study aims to discover the effectiveness of historical regulations on ozone concentrations so they can begin to interrogate if the tradeoffs are worth it.

This paper uses the Difference in Differences designs and Regression Discontinuity designs. These methods can be employed because of spatial and temporal differences in the application of gasoline regulations. The paper evaluates three main regulations. The Reid Vapor Pressure laws (RVP) were implemented in 1989, becoming stricter in 1992 and limited the presence of light components in gasoline, primarily butane, in the summer (June-August). The first phase varied by state and the strictest requirement was 9.0 psi. All contiguous states had a summer limit of 9.0 psi, but Southern states that did not attain EPA ozone standards had a limit of 7.8 psi. In 1995, Federal Reformulated Gasoline (RFG) laws were implemented, becoming stricter in 2000. RFG limits were only required in areas designated to be in severe nonattainment of EPA standards, but other regions could opt in. RFG laws limited toxic air pollutants (TAP, five types of VOCs) and the amount of NO_x that could be in gasoline. They also looked at a state regulation, the California Air Resources Board law (CARB, 1996), which required a seasonal 7.0 psi RVP limit and year-round content criteria that limit concentrations of olefins (6 percent by volume) and aromatic hydrocarbons (25 percent by volume) as well as an 80 percent reduction in sulfur content to limit NOx emissions. Both the DD and RD designs could compare before and after the implementation of a regulation. The majority of the data is from EPA Ozone monitoring stations and supplemented with NOAA weather station data in order to control for weather conditions. There are also some other datasets used in order to control for income, neighboring counties, and ?

The results of the DD analysis show that RFG and CARB regulations significantly reduced ozone concentrations. RFG is shown to reduce ozone concentrations by about -1.5 to -3% while CARB has a greater effect of -6 to -9%. RVP Phase 1 did not even produce a negative intercept, and RVP Phase 2 was mostly negative but not significantly. When broken out by urban, suburban, and rural, suburban areas experienced the greatest decrease in ozone pollution and rural areas the least. This makes sense considering the density of traffic and cars. They also did a DD analysis on just the ozone monitors that were operating over the entire analysis period (1989-2003). These results were more negative than their first analysis, which they attribute to the new ozone monitors being added to areas of high pollution. The RD analysis was monitor-specific, with RVP conducted on 46 monitors, RFG on 41 monitors, and CARB on 48 monitors. The mean effect of RVP on the monitors was −0.0001 and of RFG was 0.021.

However, the RFG results are impacted by an independent policy on the eastern seaboard that installed NOx control equipment at electric power plants. Controlling for that policy, the effect changes to 0.016, showing neither RVP or RFG significantly decrease ozone concentrations. The ATE of CARB was −0.060, but the effects were spatially heterogeneous. Coastal monitors had almost no reduction, but inland monitors experienced reductions of up to 35.2%.

These results are useful because they show that the most useful regulations (CARB) are the most specific and strict regulations. Less detailed regulations such as RVP let companies choose what they wish to regulate, which are the cheapest, least potent VOCs, do not even have a limited effect in most cases and are not worth the market disruptions. On the other hand, regulations like CARB save lives, the study calculates up to 660 lives in California each year, and at a cost a third of the EPA's statistical value of a life. Another important takeaway from the results is the magnitude of effects depend on whether they are urban, suburban, or rural as well as coastal or inland. Therefore, gasoline regulations may have a limited effect outside of suburban, inland areas, and other policies, such as reducing the industrial output of VOCs may have a larger effect in other communities.

II. Data

The majority of the data comes from the EPA's Air Quality Standards database. The database has hourly readings from ozone monitors across the country. The researchers used that data to get the daily maximum concentration and calculate the eight hour maximum (dividing the day into 8 hour periods and taking the maximum of each). These variables are what the EPA regulations are based off of. There were 1,144,025 days of data, the amount of monitors grew over time from 720 to 945, and 20% of the monitors were in urban areas.

The study drops monitoring days that have less than 9 hours recorded between 9 am and 9 pm and years that have less than 25% of days in the summer recorded, following EPA standards. This could be problematic if certain areas of high ozone concentration are the ones most often not recording, perhaps because they are in underserved areas. Additionally, if low concentration area monitors go long periods without being repaired because they are low priority, that could be a concern. The study also drops the data of counties that border counties with stricter regulations because sometimes the more regulated gas is distributed there by companies and the study wants to avoid no interference. They use a neighbor dataset and a list of counties for this. A concern would be how the monitors are geographically distributed.

The study also incorporates daily minimum and maximum temperature, rainfall, and snowfall data from the National Climatic Data Center's Cooperative Station. This is because ozone production is higher in higher temperatures and lower during days with precipitation. Weather data allows the study to control for these confounders. To match the weather station data to the ozone monitor and to fill in the gaps in the missing data, the researchers developed an algorithm which they say has 95% accuracy. To control for confounders, they also use income data and the CumulativeNOxInstallations.dta.

	treat_rvpI	treat_rvpII	treat_rfg	treat_CARB
year				
1989.0	371.0	0.0	0.0	0.0
1990.0	381.0	0.0	0.0	0.0
1991.0	395.0	0.0	0.0	0.0
1992.0	0.0	132.0	0.0	0.0
1993.0	0.0	140.0	0.0	0.0
1994.0	0.0	140.0	0.0	0.0
1995.0	0.0	111.0	111.0	0.0
1996.0	0.0	71.0	106.0	48.0
1997.0	0.0	76.0	107.0	48.0
1998.0	0.0	82.0	108.0	49.0
1999.0	0.0	87.0	107.0	49.0
2000.0	0.0	97.0	106.0	49.0
2001.0	0.0	97.0	107.0	47.0
2002.0	0.0	100.0	108.0	49.0
2003.0	0.0	100.0	108.0	50.0

Table 1. Count of each treatment type per year

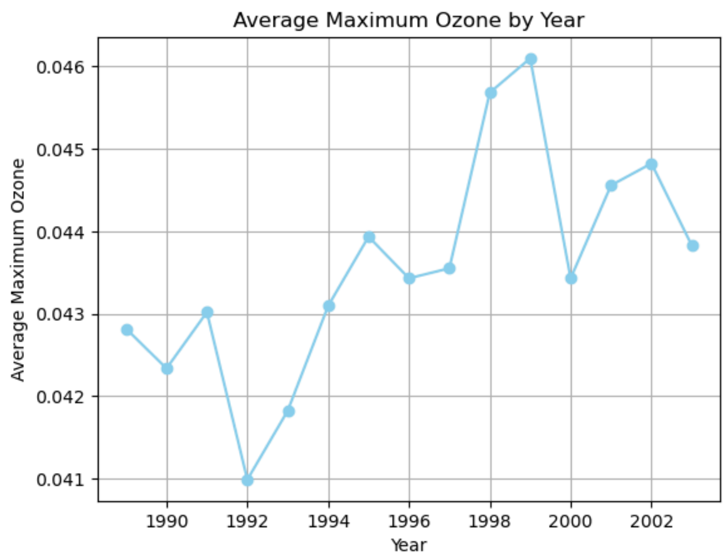

Figure 2. Average Maximum Ozone Concentration Over Time

III. Methods

The paper uses a Difference in Differences design to estimate the natural log of the monitor's daily maximum reading and its maximum eight hour reading during the summer months. It uses an additively separable model where α is a four-element vector of parameters whose estimation is of primary interest, Treat_ct is a vector of four variables indicating whether the county c in which monitor i is located is subject to one of the four possible regulatory treatments at time t (excluding the baseline RVP standard of 9.0 psi), μ_i are fixed effects due to the unobserved factors that cause spatial differences in ozone concentrations, η_ry are fixed effects to control for year-to-year shocks, and ε_it is the error term.

$$\ln(y_{it}) = \boldsymbol{\alpha} \cdot \mathbf{Treat}_{ct} + \mu_i + \eta_{ry} + \varepsilon_{it}.$$

Figure 3. DD Equation 1

$$\ln(y_{it}) = \boldsymbol{\alpha} \cdot \mathbf{Treat}_{ct} + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + \gamma_r \cdot \mathbf{D}_t + \delta \cdot I_{ct} + \boldsymbol{\theta} \cdot \mathbf{Trend}_{rct} + \mu_i + \eta_{ry} + \varepsilon_{it}.$$

Figure 4. DD Equation 2

The second part of the paper uses a Regression Discontinuity design. The model is additively separable. Its outcome variable is a monitor-specific ozone concentration, and it uses monitor specific treatment effects, weather coefficients, and time trends (an eighth-order Chebychev polynomial). The running variable is time. This design is more flexible than the DD design because it is specialized to the location of each monitor and allows for comparison of how location affects outcomes. On the other hand, the results are harder to condense and present than the DD design.

$$\ln(y_{it}) = \boldsymbol{\alpha}_i \cdot \mathbf{Treat}_{ct} + \boldsymbol{\beta}_i \cdot \mathbf{W}_{it} + f_i(Date_t) + \mu_i + \varepsilon_{it}.$$

Figure 4. RD Equation

The unbiasedness assumption of the paper is predicated on there being no confounders. The design addresses this by controlling for weather, income, and two different fixed effects. They also demonstrated that controlling for these variables does not greatly influence the outcome. However, there could still be some factors the study doesn’t account for. Another source of ozone pollution is industrial manufacturing. While the fixed effect controls for the different starting ozone concentrations, factories opening or changing their regulations could change the treatment effect. Another factor that the authors did not consider is race. Even wealthy black areas have higher levels of pollution than middle-class white areas. However, if the other confounders did not make a large impact, perhaps these would not either. The arrow of time assumption is also part of the unbiasedness proof. This is also mostly plausible, with the exception that because the policies were announced before they were implemented, and implemented in phases, there could be some treatment effect as companies transitioned to making the new gasoline.

The consistency assumption states that there are no hidden treatment effects: the observed treatment group and assigned treatment group are the same. The paper makes this assumption because the treatment is a law, and therefore required. It assumes that the penalties of breaking the law will deter defiers. However, it is possible that some areas did not implement the laws. The nature of environmental laws is that there are many levels of regulation; local, state, and federal, and the factors that affect VOC and NOx production could be affected by a variety of legislation. So, there are potentially some treatment effects that could not be accounted for.

The asymptotic normality assumption is that the distribution will converge to a normal distribution and follow the Central Limit Theorem. This assumption is plausible because of Figure 6, which shows the distribution of ozone concentrations. Because the dataset's values are roughly normally distributed, the researchers can reasonably assume that the population is normally distributed. However, it is worth noting that this is not a random sample.

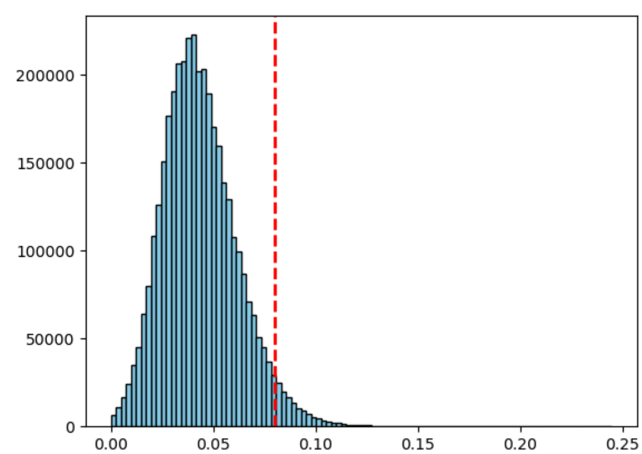

Figure 6. Distribution of Max 8hr Ozone Concentrations, with the EPA cutoff for severe nonattainment in red.

IV. Results

Since there were two experiment designs, I chose to replicate the first one, the Difference in Differences. My estimations were different from the studies, however this is probably down to coding errors. I also got the most negative slope and lowest p-values for the CARB policy. CARB was the only policy that produced consistently negative results. However, my CARB coefficients were much smaller than the study's output: around -2-3%. RVP I, RVP II, and RFG fluctuated in whether the slope was negative or positive and had much higher p values.

I had trouble replicating the results of the paper exactly. Part of this was due to the nature of the replication code provided. A lot of code I saw was in Stata, and I am not very familiar with Stata or Python. Also, they left a lot of variable manipulation and cleaning in their replication code rather than providing a clean version of the dataset. I was able to get regression outputs for Table 2, but wasn't able to organize them well. I set up the code for the Table 4 results and the Figure 5 graph, but kept running into minor mistakes in the code that prevented them from running properly.

	coef	std err	z	P>|z|	[0.025	0.975]
treat_rvpII	3.485e-09	0.003	1.11e-06	1.000	-0.006	0.006
treat_rfg	2.873e-09	0.003	9.91e-07	1.000	-0.006	0.006
treat_rvpI	4.954e-09	0.005	1.06e-06	1.000	-0.009	0.009
treat_CARB	-3.519e-09	0.008	-4.24e-07	1.000	-0.016	0.016
_RY	-2.975e-11	3.38e-05	-8.79e-07	1.000	-6.63e-05	6.63e-05
_RW	2.626e-08	0.014	1.92e-06	1.000	-0.027	0.027
_RD	-1.048e-10	0.000	-4.43e-07	1.000	-0.000	0.000
treat_rvpIID	-0.0008	0.006	-0.134	0.893	-0.012	0.011
treat_rfgD	0.0013	0.006	0.228	0.820	-0.010	0.012
treat_rvpID	0.0098	0.007	1.364	0.173	-0.004	0.024
treat_CARBD	-0.0311	0.011	-2.906	0.004	-0.052	-0.010
_RYD	0.0003	0.000	1.579	0.114	-8.44e-05	0.001
_RWD	0.0005	0.014	0.033	0.974	-0.026	0.027
_RDD	6.639e-05	0.000	0.281	0.779	-0.000	0.001

Figure 7. Regression results for max_ozone. treat_CARBD has a more negative slope and higher pvalue than treat_rvpID, treat_rvpIID, or treat_rfgD.

V. Discussion

The drawbacks of using DD in this setting was that the time trends and confounding variables had to be nationalized. This also meant that without the specificity of the individual RD designs, confounders such as the northeastern coast NOx policy could be overlooked. This setting is also difficult to get all of the assumptions to hold up, particularly consistency. However, individual RDs for over 100 monitors are difficult to compare and analyze. The authors only provided specific information about an average instance and the distributions. I found the distributions informative, but would have liked to know more about the spatial distribution within the RD sample distribution. I think the analysis drew from a large sample size, but could have made its results more robust by bootstrapping. This way the DD study would have a sampling distribution of regression coefficients instead of singular results.

The CARB regulation results could also have more interrogation. CARB was the only state-level policy chosen, which asks whether state and federal policy effectiveness can be compared. Are they enforced by the same institutions? I also wonder if a stricter state policy is indicative of stronger enforcement of environmental laws in general. Itt is also possible they needed this law because ozone concentrations were much worse than in other places in the U.S. These questions are not to dispute that the policy had an effect, which is most clear in the Los Angeles RD results, but rather to question their comparison to other locations.

This study evaluated many policies and designs in order to draw conclusions about the effectiveness of gasoline regulations on ozone pollution. While I appreciate the thoughtfulness of interrogating so many avenues of inquiry, if they had narrowed their focus to one design or policy, they may have been able to spend more time discussing the intricacies of those results and justifying their assumptions. Environmental policy has significant economic and health consequences, and so evaluating their effect is necessary to inform whether similar policies should be used in the future. Despite some of the assumptions potentially not being met, CARB having an effect is convincing and points that this is the path for future legislation. It would be interesting to see a follow up policy of recent ozone policy and see whether gasoline regulation is still the primary avenue of regulation.