5-15-2020

The Impact of a California Clean Air Regulation on Infant Birth Weight: The Case for a Check Engine Light

Kieran DeMasi
Skidmore College, kdemasi@skidmore.edu

Economics Thesis

The Impact of a California Clean Air Regulation on Infant Birth Weight
The Case for a Check Engine Light

By: Kieran DeMasi
Edited by: Deniz Aytac

Professor Rodrigo Schneider
April 28, 2020

ABSTRACT

This paper aims to analyze the effectiveness of California's Clean Air regulations. Using historical data on infant birthweights from the NCHS, a Difference in Differences (DID) model will demonstrate changes in birthweight after the enforcement of a 1988 regulation. The regulation analyzed affected new cars, mandating that manufacturers install on board diagnostics (OBD) systems on all new vehicles in the state. In the case of the DID, California acts as the treatment, and the neighboring states will act as controls. This is due in part to limitations in the data with regional variables. The analysis also aims to measure the effectiveness of the policies on the basis of environmental indicators as well. This data was made publicly available by the EPA. These variables measure levels of nitrogen dioxide (NO_2) and carbon monoxide (CO) levels. The findings of this research conclude that the enactment of this regulation may have had positive short-term effects on air quality as well as infant birthweight. There exist further opportunities to delve into this research looking at either a longer period or using more specific regional variables than made publicly available in the aforementioned data sets.

INTRODUCTION

For years, social scientists and natural scientists alike have been connecting infant birth weight to environmental factors. The research shows that yet another of the many negative externalities of pollution is a negative effect on birth weight. In particular, air pollution is especially detrimental to fetal growth due to the pervasiveness of the pollutants as well as the limited ways in which people can escape it. In response to poor air quality and concerns of health issues, California has enacted several policies to combat the rise of pollution.

California's air pollution issues began at the beginning of the 20th century. The state faced a massive population boom, and with the increased population density so too grew the number of cars on the roads. In the midst of the second world war, Los Angeles encountered smog for the first time and easily dismissed the issue for a gas attack. However, as the war ended, and the issue persisted the region began to search for the main cause. After shutting down several factories, it became clear that the main cause of air pollution within the region did not come from one centralized entity but instead from the multitude of cars on the road. This experience although especially visible in Southern California, is not unique to the region.

Motor vehicles are found to be the main source of many pollutants within cities, especially small particulate matter, nitrogen dioxide (NO2), and carbon monoxide (CO). As much as 90% of the CO present in cities can come from motor vehicle exhaust. These chemicals are especially detrimental to developing lungs due to their harmful effects. Carbon Monoxide bonds with hemoglobin in the blood, limiting the body's metabolization of oxygen. Nitrogen Dioxide, on the other hand, is a reactive gas that has been found to irritate the lungs and increase the likelihood of respiratory complications (EPA, January 1993). The purpose of this paper is to further research on this topic. While the link between infant health and air pollution is reportedly substantial, how

effective were the policies in limiting air pollution and sparing those with the lowest tolerance? This paper will research the effectiveness of a 1988 California policy which was focused on limiting the effect of these pollutants. The enactment of this policy occurred in a very interesting time. During the 70s and 80s California saw its largest decrease in Smog severity. The policy in question required all vehicles model 1988 and later to have an on board diagnostic systems (a common example would be the check engine light) to help ensure emission standards were met.

LITERATURE REVIEW

In the United States of America, California has consistently been at the forefront of both air quality issues as well as regulations. "Since 1940, over 60 percent of the air pollution in [California] has been associated with cars and trucks. . . Back in 1940, Californians owned about 2.8 million motor vehicles. By 2000, the much larger population drove 23.4 million cars and trucks. Total vehicle miles traveled in those six decades jumped from 24 billion to 280 billion" (Carle 2006). California's air pollution is globally unique for several reasons. Whereas many developing nations' cities in similar situations face air quality issues, their air pollutants can originate from a plethora of sources ranging from manufacturing to energy production. California's pollutants originate primarily from Automobiles. This issue is both extremely interesting and posed a very challenging issue to legislators. London, England, which faced issues with smog in the early 1950s due to Coal plants, could simply pass new legislation and make the conversion to cleaner fuel sources (Carle 2006). Comparatively, California struggled to put limits on individual auto vehicle owners and instead opted to target new sales for vehicles sold in state.

Although the smog issue peaked in the 1970s when significant progress began to steadily decrease Stage 1 smog alerts from 121 in 1977 to none in 2000, California's air remains to be some

of the worst in the nation. (Carle 2006, xvii) The pressing issue of Air pollution in the region even surfaced in the beginning of the 2020 presidential election, when candidate Kamala Harris brought up the improvement of the air in the region but stated that despite the drastic improvement people should not be complacent and instead strive for stricter regulations and cleaner air. The persistence of this issue can also be seen in the region's 87 continuous days breaking federal smog standards in 2018 (Christensen 2019). It is due not only to California's unique situation, but also the continuation of this pressing public health issue, that the interest for this paper originates.

Extensive literature has been written on the effects of certain pollutants on infant health. In Currie and Neidell (2005), the authors delve into the subject focusing on California in the 1990s. This paper is especially interesting due to its research on lower levels of air pollution than previous literature. Where most previous papers focused on cities blanketed in smog in developing nations, Currie and Neidell focus on the impact in California. Through their findings, Currie and Neidell's findings are contrary to many of the preestablished literature. They state that many air pollutants had little to no statistically significant effect on infant health. The exception that they find is carbon monoxide, which plays a statistically significant role in infant deaths. In the end, Currie and Niedell calculate that the reductions in carbon monoxide over the decade saved over 1,000 infant lives.

Matthew Niedell continues the literature on this topic by focusing further on the impact of California's seasonal pollution on childhood asthma. Niedell argues that this data is especially interesting since individuals with asthma and, in particular, those with developing lungs, are especially susceptible to environmental factors such as air pollution. Niedell uses zip codes and months as an important differentiator, because zip code is a precise measure of area, and seasonal pollution in California follows a strong pattern (Niedell). Pollutants that follow this trend range

from various chemical compounds such as NO and CO. However, it has been argued that the particulate matter larger than ten micrometers (PM10) does not follow any regional patterns (Ritz et al., 2000). Particulate matter has also been shown to have a statistically significant effect on birth weight and the rate of premature births (Ritz et al. 2000, Bell et al. 2000). This classification of pollution is an overarching term that defines a wide variety of chemical compositions. In a study on the effects of different forms of PM on birthweight, Bell et al. sought to differentiate the effects based on the types and origins of the particulate matter. They concluded that "Road dust and related constituents such as silicon and aluminum were associated with lower birth weight, as were the motor-vehicle related species such as elemental carbon and zinc and the oil-combustion-associated elements vanadium and nickel" (Bell et al., 2000).

On top of the clear evidence that the most harmful Particulate Matter originates from motor vehicles, there exists substantial evidence within the literature that an increase in vehicles will see a proportional rise in pollutants. In an analysis of a toll-free highway system in China, researchers Fu & Gu (2017) saw a substantial increase in pollution. The authors constructed a difference in differences model (DID) using policies put in place by the Chinese Government. To celebrate the lunar new year, the government eliminated all tolls within a state. This region was used as the treatment, and the surrounding states were used as controls. When combined with historical pollution data over for the holiday season, the authors were able to create a model to measure the effects of the policy. In this way they used a DID to measure an increase in pollution by 20%. By making cars a more convenient mode of transportation a drastic rise in pollution occurred. Similarly, in India, researchers Greenstone and Hanna (2014) found that cities that introduced regulations requiring catalytic converters saw a statistically significant decrease in PM and SO2 while also observing a statistically insignificant decrease in NO. The previously established

proportion of Air pollution in California that is attributable to automobiles as well as these results from areas where automobiles may not be the largest form of pollution suggest that regulations that target cars would be effective in reducing the key pollutants that developing and weakened lungs are susceptible to.

Measuring the effectiveness of regulations using difference in differences and has become increasingly standard. DID is a form of analysis used when dealing with variables that regulations can create a distinct and strong impact on. In terms of air quality regulations, DIDs have frequently been used to measure both the intended and unintended effects of Air Quality regulations. Li, X et al. (2020) use DID to measure the effects of air regulations on employment in polluting industries. Wang K. et al (2019) on the other hand measured the total effect of the air quality regulations on the amount of pollutants within the air itself.

Previous Literature has mostly avoided combing the DID analysis and measuring the effects of regulations on Infant health. For instance, while many papers discuss the effects of air pollution on infant birth weight, very few discuss whether positive change can be seen through regulatory means. Instead we have seen a large amount of works focused on the regulation's effects on infant mortality. A prime example of this is Michael Greenstone and Rema Hanna's paper *Environmental Regulations, Air and Water Pollution, and Infant Mortality in India* (2014). Although Greenstone and Hanna use the bulk of the paper to assess the ability of environmental regulations in developing countries with weak institutions, they set a precedent to using the DID model to assess regulation effectiveness on Infant mortality. Greenstone and Hanna primarily analyze three cornerstone policies to India's pollution regulations. This includes Air pollution regulations that were strongly associated with nitrogen dioxide, particulate matter, and sulfur dioxide.

We have also seen a number of works looking at the effect of pollution on Infant birthweight. Bell et al. (2000) linked heavy metals present in exhaust to lower infant birthweight. Despite the presence of this literature, there is an absence of literature using DIDs to measure the effects of air regulations on infant birthweight.

This paper will focus primarily on the effects of the 1988 air regulation that focused on limiting the pollutants in automobile exhaust. As such the paper will focus primarily on measuring the direct effects of the regulation on air quality and on Infant health. Greenstone and Hanna argue that infant health is an especially powerful measure of the effectiveness of environmental regulations because it is more responsive to pollution changes that adult health. And infant mortality is an especially appealing measure, because the first year of life is especially vulnerable so losses of life may enlarge the data set. Greenstone and Hanna then proceed to test this variable through a DID controlling for metropolitan areas. This is done because the enforcement of the policies tested within the study were done at a local level. What is most interesting about this paper is that the effects of the regulations are first measured through total decrease of pollutants before measuring the effect of this decrease on infant health. This is important because, as the paper discusses, there exists a lot of noise within the data set on infant mortality. However, the effects of the regulations can be seen much more clearly when comparing the effects on infant health and the effects on air quality.

Although many papers measure the effectiveness of pollution regulations in measured effects, the previous literature that associates the presence of pollution with infant health argues that the link is important. Although it is important and easy to calculate an immediate real-world change in the density of pollution, the public health issues can remain. Greenstone and Hana argue that the connection highlights the need for further regulatory expansion while also showing the

effects of the policy. Greenstone and Hana attribute a decrease in the infant mortality rate of 0.64 per 1,000 live births to the introduction of regulations on catalytic converters. Through the implementation of a simple regulation, a measurable change occurred in the rate of infant survivability. Although infants are more reactive towards detrimental environmental factors, we can assume that this decrease in air pollution was not only helpful to infants but all people with conditions that would make them more affected by air pollution (i.e. those with asthma and the elderly). We can also assume that to a lesser extent that all people enjoyed positive health benefits, and although it did not drastically increase their immediate life expectancy, it most likely decreased their likely hood of respiratory and cardiovascular issues in the future (Wang et al 2020).

There are however severe limitations within the regulations that were created. In their interpretation of their findings, economists Shihe Fu and Yizhen Gu (2017) discussed further the implications from toll reductions positive effects on pollution. Fu and Gu found that the elimination of tolls had a statistically significant positive relation on air pollution. They then used this as a call to raise tolls to include the environmental externalities of highway usage. Another piece of information that can be extrapolated from their findings is that without the presence of disincentives, drivers are willing to create congested high-polluting traffic for the sake of convenience. Such an example can be seen in the case of Los Angeles. Although the city has a massive public transit system, it has repeatedly seen ridership decrease in favor of automobile ownership (Tinco 2018). This may instead highlight an issue with the enactment of California's regulations. Historically, California has worked closely with car manufacturers to standardize emission controls and promote the manufacturing of zero-emission vehicles (ZEVs). This puts the majority of the cost of pollution controls on the manufacturer to upgrade their products, and in

fact, many of the older, more polluting models of cars are grandfathered into these strengthening regulations.

In one of their policy assessment reports, RAND Corporation discussed the importance of the idea behind California's Voluntary Accelerated Vehicle Retirement Program (VAVR) citing that although California had enacted many policies to reduce the creation of NO the main contributing pollutant in the creation of Ozone, the regulations had little affect because "in the South Coast in 1998, Light-duty Vehicles (LDVs) at least 15 years old accounted for only 11 percent of total vehicle miles driven by LDVs but 39 percent of total LDV emissions" (Dixon and Garber, 2000) LDVs which are vehicles weighing less than 8.5 tons, are primarily non-commercial vehicles. These cars face a much lower rate of replacement slowing down and limiting the effects of regulations to improve their environmental performance.

California's regulations have almost completely targeted the amount each individual vehicle pollutes while ignoring the findings of Fu & Gu (2017). While these regulations decrease the amount of CO, NO, and PM that each vehicle produces, it ignores a growing trend in Southern California and Los Angeles specifically where vehicle ownership has continued to rise while ridership on public transportation has been continuously decreasing (Tinoco 2018). For these reasons although the regulations may be extremely effective for the automobiles affected, their effects may not be seen until many years later or may be completely ignored due to an increase in total automobile usage.

Because automobiles are the primary source of CO in cities and a main contributor to the creation of NO and PM, regulations effecting automobiles would be one of the most effective ways to limit the concentration of these pollutants in major metropolitan areas. Carbon monoxide has been found to have a statistically significant detrimental effect on infant and fetal health. The same

is true for particulate matter that originates from cars and other vehicles. Nitrogen oxide which is one of the largest contributors to the production of Ozone and Smog has severe effects on vulnerable and compromised respiratory systems. Thus, by studying the effectiveness of these policies which target automobile pollution on the improvement of infant health we can see the effect of a decrease in detrimental pollutants. Following Greenstone and Hanna (2014) this change will be measured through a Difference in Differences using California as a treatment and the surrounding states of Nevada, Utah, and Arizona as controls. The reason for selecting these neighboring states as controls is twofold. First, they did not enact the policy being analyzed, and second, the data being used lacked more specific regional indicators to protect the individuals' privacy. Due to the intent and goal set by the creators of the policies and despite the regulations' limitations it is expected that there will be a decrease in pollution and increase in overall air quality in the treatment areas. Following this change it is anticipated that there will be an increase in infant health and decrease in mortality due to respiratory complications.

METHODOLOGY

Methodology – Infant Data

To test the above hypothesis, several regressions will occur. the National Center for Health Statistics (NCHS) has made several fascinating datasets available with the help of the CDC. The data sets used in this research pertain to infant mortality and infant birth statistics. A third data set exists which links Infant mortality information to the affiliated birth statistics, but this has not been used due to limited years and a change in formatting during the research period.

Methodology - Infant Mortality

Infant mortality data includes information pertaining to the infant such as weight, age, birth date, and the state in which the birth and death occurred. Due to privacy regulations more specific data such on region such as county, town or city of the birth or death were unavailable. Likewise, the specific causes of death were not listed. Therefore, for the sake of this research, we are forced to look at absolute changes in the total infant mortality as compared to changes in infant deaths due to respiratory issues and complications. The data from the control and treatment groups was extracted and isolated from this larger data set and the years pertaining to the research period were merged. Because the policy analyzed in this research was enacted in 1988, and the due to the data's own limitations, the research looks at the years 1986 to 1991. Since the areas most heavily afflicted by the air pollution in the treatment were in the south of California, a combination of the state's more southern neighbors made up the control group. These neighbors include Nevada, Utah, and Arizona. A DID was then done to asses whether a change occurred within the total number of infant deaths within this period. In an attempt to control for various factors, the race of the infant, and the age of the mother were included. The

variable race is also a dummy variable where a value of 1 indicates that the child is white. The

age of the mother on the other hand is a running variable and increases with each year. Due to a

long list of societal issues, people of color are far more likely to live in areas with high inequality

and lower access to vital infant support in the forms of healthcare and childcare. Similarly, the

age of the mother was included in the regression to account for any complications which may

occur from either extremely young mothers or older women giving birth. In equation (1) "i"

represents each individual in year "t".

(1) $\text{Total\_Deaths}_{it} = \beta 0 + D1\ \text{Treatment}_i + D2\ \text{Post}_t + D3\ \text{Interaction}_{it} + D4\ \text{race}_{it} + \beta 5$

$\text{mothers\_age}_{it} + \varepsilon_{it}$

Methodology – Infant Birthweight

Due to the much larger quantity of data available within the Infant Birthweight dataset,

there were many more variables which could be used as controls. The dependent variable tested

was the infant's birthweight *dbirwt*. A mean for variable *dbirwt* was created for each value of

the independent variables. For example, an average birthweight was calculated for all white male

babies born in the month of May in 1987. The independent variables were included, and an

aggregated average was created in an attempt to clean up some of the noise within the data. Due

to its sheer mass, the dependent variable exhibits a lot of noise. This is increased further due to

how many factors can have an impact on infant birthweight. The variable infant sex was included

to account for the difference in average birthweight in the different genders. This variable is a

dummy variable where a value of one indicates that the infant is female. The race of the child is

again taken into account to control for societal inequalities which may affect the infant's birthweight. Mother's age is taken into account again for similar reasons as previously. Mothers education is taken into account as a variable to control for how wealth disparities may affect infant birthweight. The birth month and county of residence is also taken into account and are absorbed within the regression. Here, equation (2) uses "a" to represent each aggregate value in time "t".

$$(2)\ Dbirwt_{ta} = \beta_0 + D_1\ Post_t + D_2\ Treatment_a + D_3\ Interaction_{ta} + D_4\ race_{ta} + \beta_5\ mothers\_age_{ta}$$
$$+ \beta_6\ mothers\_edu_{ta} + D_7\ sex_{ta} + \varepsilon_{ta}$$

Methodology – Infant Mortality Ratio

The two previous datasets are combined and integrated to obtain a new variable. The variable *deathratio* is calculated by the ratio of infants that die each year compared to the total number of births. This variable is used so that the more densely populated California can be more easily compared to the control group made up of less densely populated Nevada, Arizona, and Utah. This ratio is then compared through a simplified DID which includes the Treatment Post and Interaction. Due to the nature of this calculation, and the fact that the ratio is taken from the number of observations in one set over the number of observations in the other, many previously used independent variables are not applicable here. As such equation three uses subscript "st" to show the ratio of this relationship in state "s" and time "t".

$$(3)\ Deathratio_{st} = \beta_0 + D_1\ Post_t + D_2\ Treatment_s + D_3\ Interaction_{st} + \varepsilon_{st}$$

Methodology- Environmental Data

To take a step back, it is vital that we consider the policy not only on its secondary effects of infant and public health, but also on the more direct effects that it had on the environment. Thus, to assess whether this policy enacted any real change on the main pollutants given off by automobiles, data sets from the EPA which tested for air quality throughout the states was used to test for both CO and NO2 concentrations.

Methodology – NO$_2$ Changes

The EPA data set for which evaluated the concentration of NO2 was self-reported and collected from each participating testing site. These locations varied across states but were primarily focused around metropolitan areas. So much so that the data was made available by the EPA for each individual site or the aggregate of each reporting metropolitan area. In the case of NO2, the dependent variable tested in this research is the daily max 1-hour NO2 concentration. This data was collected by the maximum amount of NO2 within each 1-hour timeframe. This value was then averaged per month to eliminate some noise within the data and allow for some absent dates within the dataset. A DID was then run to assess the impact of the policy on the concentration of NO2 within California. Equation (4) shows how the regression for the NO2 levels was taken from each testing location "l" in time "t".

$$(4)\ NO2_{tl} = \beta_0 + D_1\ Post_t + D_2\ Treatment_l + D_3\ Interaction_{tl} + \varepsilon_{tl}$$

Methodology – CO Changes

The EPA dataset for the concentration of CO was also self-reported and collected from individual testing sites. The dependent variable tested for CO in this research was the daily max

8 hour CO concentration. The data was collected in this fashion by the government to offer more protection, and more accurate assessments of high levels of CO. The data was collected by the maximum amount of CO within an 8 hour timeframe. This value was then averaged per month to eliminate some noise within the data and allow for some absent dates within the dataset. A DID was then run to assess the impact of the policy on the concentration of CO within California when compared to the control group.

$$(5)\ CO_{lt} = \beta_0 + D_1\ Post_t + D_2\ Treatment_l + D_3\ Interaction_{lt} + \varepsilon$$

Equation (5) shows how the regression for the NO2 levels was taken from each testing location "l" in time "t". In both equations (4) and (5) we continue to use the same control and treatment groups as in the previous regressions on Infant health. This will not only help with the discussion where changes in infant health can be compared to changes in environmental health in similar areas, but also due to data limitations as well. During the time period being researched the released data was sporadic and limited. Some states had large quantities of testing sites which tested each day, while others lacked strict testing regimen. Utah, for instance, had an entire year within this testing data with no CO data.

RESULTS

Table 1: DID of Infant Mortality

cdeath_year	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
o.Treatment	0.000	.	.	.	.	.	
Post	-6.696	5.520	-1.21	0.225	-17.515	4.123	
Interaction	-70.422	6.123	-11.50	0.000	-82.424	-58.420	***
infantrace	-1.013	2.155	-0.47	0.638	-5.236	3.210	
dmage	0.438	0.158	2.77	0.006	0.128	0.748	***
Constant	3724.215	4.514	825.05	0.000	3715.368	3733.063	***

Mean dependent var		3722.233	SD dependent var			1594.717
R-squared		0.990	Number of obs			27652.000

F-test		213.614	Prob > F			0.000	
Akaike crit. (AIC)		358452.093	Bayesian crit. (BIC)			358493.230	

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

In Table 1, which corresponds to the regression of equation (1), we see a significant value for mothers age and for the variable Interaction. The Coefficients for these variables suggest that by implementing these policies, the treatment saw a decrease of 70.422 infant deaths if all else stayed the same. As predicted, we also see that an increase in mothers age by one year increases the infant's likelihood of dying by 0.43. This is consistent with the various health complications that can arise in infants as woman grow older (March of Dimes 2020). The regression above has an unusually high R squared value showing that there are issues with the regression analysis. As is, the regression can explain for 99% of all variance within the dependent variable. It is extremely likely that there are missing variables within the regression. Currently the variables which are accounted for are being to heavily weighted, creating biased results. Taking this into account it is hard to trust the results of this regression, however the Interaction variable was significant with negative coefficient. This would imply that after the enactment of the regulation there was a net decrease in infant deaths within California.

Table 2: DID on Infant Birthweight*

dbirwt	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
o.Treatment	0.000	.	.	.	.	.	
Post	-14.316	2.785	-5.14	0.000	-19.775	-8.857	***
Interaction	21.098	4.640	4.55	0.000	12.004	30.191	***
dmage	8.188	0.160	51.09	0.000	7.874	8.502	***
dmeduc	0.023	0.069	0.33	0.738	-0.112	0.158	
isex	-183.552	3.898	-47.09	0.000	-191.191	-175.913	***
Mean dependent var		3296.191	SD dependent var			606.420	
R-squared		0.025	Number of obs			408485.000	
F-test		330.364	Prob > F			0.000	
Akaike crit. (AIC)		6383760.081	Bayesian crit. (BIC)			6383945.725	

*** p<0.01, ** p<0.05, * p<0.1
**Coefficients for months have been removed.

Table 2 which corresponds to equation (2) shows the results of a DID on infant birthweight. The above regression analysis is testing the birthweight of children. The R-squared for this regression is extremely low and only accounts for 2.5% of the variability within the data. This can be explained by the many different elements that can affect infant birthweight that were left out of the regression. In Table 2 we see that variables *Post, Interaction, dmage,* and *isex* are all highly statistically significant. Interestingly, *Post* has a coefficient of -14.316 which suggests that across the time period, after 1988 birthweight dropped by 14.316 grams. We also see that *Interaction* shows that during the time of this policy California saw an average increase in birthweight by 21 grams. The 8.188 coefficient on *dmage* suggests that for each additional year older the mother was the infant's weight increased by 8.188 grams if all else was constant. The Infant's sex was also as predicted highly significant and negative. This shows that the average birthweight of a female was 183.552 grams less than a male infant. Despite the high significance of the variables, we cannot explain enough about the variation of the dependent variable to be sure of these results. The missing variables may heavily influence the results, and the results might currently be suffering from omitted variable bias. Despite this, we are again seeing a significant result for the Interaction variable. This variable also has a positive coefficient implying that enacting the policy had a positive effect on infant birthweight.

Table 3: DID on Infant Death Ratio

deathratio	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Treatment	0.131	0.000	1833.52	0.000	0.131	0.131	***
Post	-0.006	0.000	-58.32	0.000	-0.006	-0.006	***
Interaction	-0.013	0.000	-86.30	0.000	-0.013	-0.012	***
Constant	0.017	0.000	324.21	0.000	0.017	0.017	***
Mean dependent var		0.081	SD dependent var			0.067	
R-squared		0.934	Number of obs			303402.000	
F-test		1430050.254	Prob > F			0.000	
Akaike crit. (AIC)		-1606253.280	Bayesian crit. (BIC)			-1606210.789	

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

Table three deals with the regression of equation (3). In the above regression analysis, we see a very high R squared value while having very few variables. Despite only using the treatment and post independent variables as well as their interaction we see that this regression accounts for 93% of variation within the dependent variable. This would imply that a bias might exist within the data which is putting too much weight on the existing variables. Despite this there exists a negative and significant relation between the interaction variable and the ratio of infant deaths to total births. The coefficient of this variable is -0.013 suggesting that enactment of this regulation may have lowered infant deaths by 1.3% within the state. Though the results may be biased they suggest that the enactment of this policy may have lessened infant mortality within California.

Table 4: DID on NO2 levels

dailymax1hourno2 co~n	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]
Post	2.925	2.903	1.01	0.314	-2.767	8.616
Treatment	-1.337	1.193	-1.12	0.262	-3.675	1.001
Interaction	-2.716	3.017	-0.90	0.368	-8.630	3.199
Constant	41.614	1.147	36.27	0.000	39.365	43.863
Mean dependent var		40.442	SD dependent var			24.445
R-squared		0.001	Number of obs			7212.000
F-test		1.221	Prob > F			0.300
Akaike crit. (AIC)		66575.194	Bayesian crit. (BIC)			66602.728

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

Table 5: DID on CO Levels

dailymax8hourcoc on~n	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Post	0.026	0.090	0.29	0.770	-0.150	0.203	
Treatment	-0.688	0.041	-16.86	0.000	-0.768	-0.608	***
Interaction	-0.018	0.104	-0.18	0.859	-0.222	0.185	
Constant	2.518	0.036	70.82	0.000	2.449	2.588	***
Mean dependent var		1.998	SD dependent var			1.458	
R-squared		0.041	Number of obs			7926.000	
F-test		113.122	Prob > F			0.000	
Akaike crit. (AIC)		28141.902	Bayesian crit. (BIC)			28169.814	

Table 4 shows the regression of equation (4). The linear regression analysis is testing the effects of the policy on NO2 quantities. This analysis has an extremely low R-squared value and only accounts for 0.1% of the variability within the data. This could be caused by the limitations in the data which will be discussed further later, or in different factors which may impact NO2 which will be discussed later as well. The findings of this regression (which due to the low R-squared statistic) should be taken with a grain of salt, suggests that the policy may have had a negative effect on NO2 quantity, but that this relationship was not significant. The Coefficient on *Post* suggests that NO2 rose in quantity during the research period by 2.95 parts per million (ppm). Meanwhile *Treatment* suggests that California has had consistently lower NO2 levels than the controls with 1.337 ppm less than the neighboring states. The *Interaction* suggests that through the enactment of this policy, California reduced NO2 emissions by 2.716 ppm. This can not be said definitively, however, due to the low R squared value.

The regression results for equation (5), which explains the impact of the policy on CO levels is found in Table 5. This linear regression analysis has an extremely low R-squared value and only accounts for 4% of the variability within the data. This could be caused by the limitations in the data which will be discussed in the following section. The findings of this regression (which due to the low R-squared statistic) should be taken with a grain of salt, suggests that the policy may have had a negative effect on CO quantity, but that this relationship was not significant. We see in Table 5 a positive coefficient on *Post* suggesting that CO rose by 0.02 ppm over the course of the research period. *Treatment* suggests that California had 0.688 ppm less CO than the control group, while *Interaction* suggests that California decreased CO pollution by another 0.01ppm over the course of the study.

Though individually these findings should be taken with a grain of salt, together they seem to paint the same picture. Over the course of the period, despite worsening overall air quality, the findings suggest that California increased its air quality. We also see findings that suggest that during this period infant birthweight increased in California despite an overall decrease. We also see signs that infant mortality decreased within the treatment group. It is unknown however how much this policy in particular created these results.

DISCUSSION

Infant Data Limitations

The findings of this research were severely limited due to a variety of reasons. First and foremost, there were restrictions within the data, which in turn restricted the quality of this research. The birthweight data did not show smaller geographic regions and for privacy purposes had erased many variables that could have further eliminated noise within the data set or could have explained more about the data itself. Due to this the amount of valuable independent variables that could have been used within the regression were severely limited. The already very noisy data was restricted further by the number of missing variables.

The mortality data was no different. This data was unable to be properly differentiate from different causes of death. This forced the regression to treat infants that died due to accidents the same as it treated children who died due to respiratory complications. This created a lot of unnecessary noise and may have severely biased. The ideal situation would have been to clean the data to only include infants who died due to respiratory issues. This data set also was severely limiting in its geographic descriptors. Forcing all deaths across the states to be handled

the same and eliminating any possibility of taking into account regional constants. Ideally, the dataset would have included county or city level data instead of just the state where the child was born, and where the child died.

Within the research a ratio of births to infant mortality was taken, due to limitations within the death data it was not feasible to properly merge the two infant data datasets. Thus, all independent variables were lost. It would have been ideal, if a Dummy variable could have been made for the infant's survival through the first 12 months. In this way, the observations could have remained distinct. As it was, merging the datasets was impossible without double counting the infants who died. That is why a ratio was used.

EPA Data Limitations

Though the EPA data was much more organized with less noise than the Infant data, there were still many complications that occurred while working with it. The data had large inconsistencies. Many testing sites would cease reporting for months if not years at a time, only to return and report data inconsistently. Utah for instance did not report any data for NO2 for the entirety of 1991. There are also complications with how the data is reported. NO2 1-hour max, which is the variable for NO2 concentrations is the value of the maximum concentration of NO2 during a one-hour period at some point during the responding day. The hour in which the sample was collected was not reported and could theoretically be at any time of day. That means that there may be inconsistencies in the data due to the time of day in which the NO2 sample was collected, and that it may not reflect the true NO2 concentration of the day. For instance, if the sample were to be collected during rush hour an extremely higher concentration of NO2 would be found when compared to a sample taken in the middle of the night when no one is driving.

CO on the other hand is tested by the maximum concentration over an 8-hour period. This means that the value is more likely to show the true maximum concentration for the day.

Policy limitations

The policy was only enforced on new car purchases. Since the average life span of a car is around 10 years, the change if any is observed should be small. The National Highway Traffic Safety Administration reported in 2006 that almost 10% of passenger cars on the road were around than 20 years old (NHTSA 2006). At the same time many other efforts were taken to fight air pollution within this time frame. The different regulations as well as the adoption of new measures for testing could affect the data. Unfortunately, this problem was combined with the limited available datasets released by the CDC. After 1991 all geographic data was erased from publicly available sources. After this time, differentiating between the control and the treatment becomes fundamentally impossible beyond this testing sample without special access to this data. Therefore, it would also be incredibly hard to try and differentiate between the different polices and how much each one was affecting the air quality. Because the other policies were also specifically tailored to mitigating pollution levels and used the same process of implementation where they only affected new automobiles, not considering their presence within this study could have drastically biased the results of this research. At the same time however, we see that the drop in NO2 and CO were not significant in the *Interaction* variable of table 4 or table 5. This could also indicate the potential for missing sources of these pollutants which were not accounted for within the study.

Sample limitations

There exists an entirely different set of limitations which must also be taken into consideration. There is the possibility that the control and treatment groups are fundamentally different. Not only is there the possibility that due to the lower population of the surrounding states that they are less afflicted by air pollution, but also the possibility that they are this push for cleaner air is a tied to a cultural phenomenon that is wholly Californian in nature. There are however many reasons to doubt this possibility, for instance, California's environmental protection acts have been historically used as the basis of many similar acts in both state and federal governments. The second limitation that exists within this sample is the possible spill over affect. Because of the close proximity to California, there is a possibility that some spillover occurred with this policy and the automobiles affected by it.

IMPLICATIONS

Due to the above limitations within the data and research, it is hard to tell whether the observed results of this study can be applied solely to the policy analyzed, or whether they should be applied to California's push to strengthen clean air regulations in general. Though the findings suggest that the policy had minimal affects short term, the literature suggests that a greater affect would be created as the policy continued. This can be seen most strongly in Dixon and Garber (2001). The authors discuss that the very nature of California's many air pollution regulations is what has taken so long for the air quality to improve. They also suggest that it is a failure in the government to not impose a life cycle on the vehicles. Due to this loophole, Dixon and Garber found that vehicles that were over 15 years old accounted for only 11% of all total vehicle miles driven in 1998 but created almost 39% of all the pollutants given off by light-duty vehicles. This policy was created due to the fact that vehicles with system issues have a higher

propensity to pollute, however it inherently ignores issues which may prevent individuals from taking their vehicles to the mechanic in the first place, namely, cost. This policy was instead centered around the belief that while individuals purchase new cars the older cars which were unaffected by the policy will be phased out. However, following this line of thought, individuals buying new cars as opposed to used cars are more likely to be able to afford to take their cars to the mechanic's for regular inspections in the first place, whereas individuals who cannot afford new cars and instead favor the second hand market may instead ignore system issues until they are unavoidable due to being unaware of the issue in the first place. Had the policy been enacted with this in mind and instead created a buy back or retirement system for these older cars like was later put in place, it could have had a much stronger effect within the short term.

Though there is the possibility that this policy did not enact as much change as those prior, there are several substantial reasons for its enactment. Although automobiles are the largest contributor to the air pollution within the region, giving individuals not only power but knowledge about their vehicle makes it easier for them to minimize their person pollution. When the problem is so spread out giving some personal onus to the consumer must be considered a viable move to help cut emissions.

Though I believe in the background of the policy, I also believe that there is room for further policy expansion and potential to improve air quality and living standards further. As we have seen in Fu & Gu (2017), drivers are willing to contribute higher pollution levels for the sake of convenience. We have also seen that there remains a disturbing trend in California of increasing automobile ownership (Tinoco 2018). I would urge policy makers that due to the very nature of this pollution problem that they must make efforts to mitigate this trend. The issue is not only the quality of air filtration on the vehicle but the quantity of vehicles on the road. It is

for this reason that I urge policy makers to either make alternatives to car ownership viable

through subsidies or through the application of a Pigouvian tax on automobiles.

REFERENCES

Bell ML, Belanger K, Ebisu K, et al. Prenatal exposure to fine particulate matter and birth weight: variations by particulate constituents and sources. *Epidemiology*. 2010;21(6):884–891. doi:10.1097/EDE.0b013e3181f2f405

Carle, D. (2006). INTRODUCTION: The Nature of California Air. In *Introduction to Air in California* (pp. Xiii-Xxi). University of California Press. Retrieved February 18, 2020, from www.jstor.org/stable/10.1525/j.ctt1pns9c.4

Christensen, J. (2019, September 13). Fact-check: Democratic debate from Houston. Retrieved from https://www.10news.com/news/national/fact-check-democratic-debate-from-houston

Currie, Janet, and Matthew Neidell. "Air Pollution and Infant Health: What Can We Learn From Californias Recent Experience." *Quarterly Journal of Economics,* 120 (2005). https://doi.org/10.3386/w10251.

Dixon, Lloyd, and Steven Garber. Fighting Air Pollution in Southern California by Scrapping Old Vehicles. Santa Monica, CA: RAND Corporation, 2001. https://www.rand.org/pubs/monograph_reports/MR1256.html. Also available in print form.

Fu, S., & Gu, Y. (2017). Highway toll and air pollution: Evidence from Chinese cities. *Journal of Environmental Economics and Management*, *83*, 32–49. doi: 10.1016/j.jeem.2016.11.007

Greenstone, M., & Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in india. *The American Economic Review, 104*(10), 3038-3072. doi:http://dx.doi.org.lib-proxy01.skidmore.edu:2048/10.1257/aer.104.10.3038

Li, X., Wang, Y., Zhou, H., & Shi, L. (2020). Has China's war on pollution reduced employment? Quasi-experimental evidence from the Clean Air Action. Journal of Environmental Management, 260. https://doi-org.lib-proxy01.skidmore.edu/10.1016/j.jenvman.2019.109851

March of Dimes. (n.d.). Pregnancy after age 35. Retrieved April 20, 2020, from https://www.marchofdimes.org/complications/pregnancy-after-age-35.aspx#

Neidell, Matthew J. "Air Pollution, Health, and Socio-Economic Status: the Effect of Outdoor Air Quality on Childhood Asthma." *Journal of Health Economics* 23, no. 6 (2004): 1209–36. https://doi.org/10.1016/j.jhealeco.2004.05.002.

Ritz, B., Yu, F., Chapa, G., & Fruin, S. (2000). Effect of Air Pollution on Preterm Birth among Children Born in Southern California between 1989 and 1993. *Epidemiology, 11*(5), 502-511. Retrieved February 11, 2020, from www.jstor.org/stable/3703990

Tinoco, Matt. "Transit Ridership Is Falling Because Angelenos Keep Buying Cars, UCLA

Report Says." Curbed LA. Curbed, LA, January 31, 2018.
https://la.curbed.com/2018/1/31/16950224/metro-ridership-decline-stats-car-ownership-study.

U.S. Environmental Protection Agency. The Benefits and Costs of the Clean Air Act, 1990-2010. Report to the U.S. Congress, November 1999.

Vehicle Survivability and Travel Mileage Schedules. (2006). U.S. Department of Transportation, National Highway Traffic Safety Administration.

Wang, K., Yin, H., & Chen, Y. (2019). The effect of environmental regulation on air quality: A study of new ambient air quality standards in China. Journal of Cleaner Production, 215, 268–279. https://doi-org.lib-proxy01.skidmore.edu/10.1016/j.jclepro.2019.01.061

Wang, Z., Peng, J., Liu, P., Duan, Y., Huang, S., Wen, Y., Liao, Y., Li, H., Yan, S., Cheng, J., & Yin, P. (2020). Association between short-term exposure to air pollution and ischemic stroke onset: a time-stratified case-crossover analysis using a distributed lag nonlinear model in Shenzhen, China. Environmental Health: A Global Access Science Source, 19(1), 1–12. https://doi-org.lib-proxy01.skidmore.edu/10.1186/s12940-019-0557-4