8022 Replication: "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality"

Sarah Hay

April 2025

I used data posted by the author: Yes

Substantive addition: County-year asthma mortality rates, 1980-2014.

In "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality," Auffhammer and Kellogg (2011) investigate whether five gasoline content regulations successfully reduced ground-level ozone concentrations between 1989 and 2003. They build a panel dataset tracking daily observations from the Environmental Protection Agency (EPA)'s national network of air quality monitors, and they combine it with weather data and income data. The authors use a difference-in-differences (DD) analysis and regression discontinuity (RD) analysis to identify the effects of the policy. I successfully replicate their DD analysis and conduct two sensitivity analyses. I also extend their work by using synthetic control methods to evaluate if the policies studied are associated with lower asthma mortality rates through the policies' effects on ozone levels.

In this paper, I present the policy background in Section 1. Sections 2 and 3 discuss the data and methods, respectively. In section 4, I present the DD replication results and sensitivity analyses. Section 5 discusses the results of the synthetic control analyses, and Section 6 presents additional discussion of the replication findings and current DD literature.

1 Policy Background

In 1963, Congress passed the Clean Air Act (CAA), the first legislation authorizing the government to mitigate air pollution (EPA 2024a). The purpose of the Clean Air Act is "to protect and enhance the quality of the Nation's air resources so as to promote the public health and welfare and the productive capacity of its population," 42 U.S.C. § 7401(b) (2023). Congress notably amended the CAA in 1970, 1977, and 1990 to grant EPA additional authority to fulfill its mission (CRS 2022). The 1970 CAA Amendments granted EPA the authority to set National Ambient Air Quality Standards (NAAQS) for six criteria pollutants: sulfur dioxide, particulate matter, nitrogen dioxide, carbon monoxide, lead, and ozone (CRS 2022). These pollutants were identified because they "endanger public health or welfare" (CRS 2022). Ozone causes respiratory health problems, such as coughing and trouble breathing, and it can contribute to new or worsening respiratory conditions like asthma (EPA 2025b).

Auffhammer and Kellogg's paper focuses on federal and state policies that intend to minimize ground-level ozone by regulating gasoline content. Ground-level ozone is created when nitrogen oxides (NOx) react with volatile organic compounds (VOCs) found in vehicle emissions (EPA 2025a). These reactions are most likely to occur on "hot sunny days," but ozone can still form during the cooler months (EPA 2025a).

EPA issues gasoline standards to limit fuel emissions that contribute to ozone formation (EPA 2024c). The agency has used two types of fuel regulations over time: standards limiting Reid vapor pressure (RVP), and reformulated gasoline (RFG) requirements (EPA 2024c). Both of these policies require "specially formulated gasoline that evaporates less at higher temperatures than regular gasoline" (EPA 2024c). RVP fuels are designed to minimize RVP, which is a common measure of gasoline volatility (EPA 2024b), whereas RFG fuel is "blended to burn more cleanly than conventional gasoline" by limiting certain compounds that are directly associated with ozone formation (EPA 2025c).

Due to California's historically poor air quality (and a provision in the CAA that effectively allows California to set more stringent environmental standards than the federal government), the California Air Resources Board (CARB) issued its own fuel regulations that took effect in 1996 (CARB 2025a). The CARB standards have a more stringent limit on RVP than the federal standards (Auffhammer & Kellogg 2011a), and they place more specific limitations on the levels of certain compounds in fuels than the federal RFG standards (CARB 2025b).

These policies do not apply equally across the country. The more stringent federal policies only apply to counties that are not in attainment of EPA's NAAQS standards for ground-level ozone.

Although California's CARB policy is even more restrictive than the federal policy, it applies to all California counties equally, regardless of that county's NAAQS attainment status. Additionally, the policies have overlapping implementation dates and continue to overlap during the second half of the study period.

For their paper, the authors identify five policies to assess:

1. "Summer RVP of 9.0 psi (some counties, 1989-1991; most ozone attainment counties, 1992 onward);
2. Summer RVP of 9.5 psi or 10.5 psi (many counties, 1989-1991);
3. Summer RVP of 7.8 psi or below (southern ozone nonattainment counties and northern opt-in counties, 1992 onward);
4. Federal RFG (severe ozone nonattainment and opt-in counties, 1995 onward); and
5. CARB (all California counties, 1996 onward)" (Auffhammer & Kellogg, 2011a).

2 Data

Auffhammer and Kellogg posted an extensive replication package online with complete data and Stata code (Auffhammer & Kellogg, 2011b). Given the extensive data processing the authors did originally, I used the authors' replication data. In this section, I discuss the authors' data and the data I have added as a substantive addition.

2.1 Authors' Data

The unit of analysis is the ozone monitor-day (hereafter, "monitor" and "monitor-day"). For the dependent variable, the authors retrieved "ambient air concentrations of ozone from the EPA's Air Quality Standards Database" from air quality monitors located across the nation (Auffhammer & Kellogg, 2011a, p. 2694). The authors then calculated daily ozone maximum and daily eight hour maximum concentrations for each monitor-day; these two metrics are used by EPA for determining NAAQS attainment (Auffhammer & Kellogg, 2011a). The authors took additional steps to clean the air quality data based on EPA data standards and qualitative interviews with EPA staff, which

I matched in my replication. In the analysis, the authors only keep daily observations from June 1 and August 31 of each year because summer months are when ozone risk is highest (Auffhammer & Kellogg, 2011a).

The authors control for weather conditions because conditions can affect ozone development. They retrieved weather data from the National Oceanic and Atmospheric Administration's database that provides "daily minimum and maximum temperatures, rain, and snowfall" for over 20,000 locations nationwide (Auffhammer & Kellogg, 2011a, pp. 2695-2696). The authors used an algorithm to match monitors to data from nearby weather stations (Auffhammer & Kellogg, 2011a). Each monitor-day has unique weather observations.

To control for income, the authors use data from the Bureau of Economic Analysis's "county-level total annual personal income" (Auffhammer & Kellogg, 2011, 2700). The income data are at the county-year level and are matched to each observation by county FIPS code and year.

To clean the data for use in analysis, I replicated the authors' Stata code in the R programming software. Table 1, in the appendix, compares summary statistics that I generated against the authors' published results. My results yield an average of 7 additional monitor-days per year than the authors', for a total of 104 additional observations across the 15 year study period. These additional results constitute 0.01 percent of the total number of observations. Excluding them did not affect the results in the subsequent analyses.

2.2 Data Extension

My primary substantive addition to the authors' data is county-year asthma mortality rates from the Institute for Health Metrics and Evaluation (2017). I use this data in a synthetic control analysis to assess whether the reductions in ozone concentrations from some of these gasoline regulations resulted in decreased mortality from asthma. I discuss these data and results further in Section 5.

I also added data on county population and county land area to control for population, population density, and to calculate per capita income. For county population data, I used pre-processed Census data from Brown University's American Communities Project (n.d.) for years 1980, 1990, 2000, and 2010. I linearly imputed values for years in between the decennial censuses. To calculate per capita income, I divided the annual income variable from the Bureau of Economic Analysis data the authors used by the annual county population. Population and per capita income are at the county-year level.

For county land area data, I used Census data from the year 2000 from the Missouri Census Data Center's Geographic Correspondence Engine (2010). To calculate county population density, I divided the population by county land area. Population density is at the county-year level.

3 Replication Methods

My analysis only replicates the authors' DD estimation. In this section, I highlight the primary endogeneity problem and discuss the estimating equations I use in my replication.

3.1 Endogeneity Problem

The primary endogeneity problem in this paper is selection bias, as a result of policy design. The federal gasoline policies (the treatments) are not applied randomly: they apply only to areas that are in nonattainment of the ozone NAAQS. Said differently, the policies generally only apply to areas that already have higher ozone levels. As a result, the treatment and comparison groups are systematically different. The DD results will be biased unless we account for selection bias.

3.2 Estimating Equations

Using a twoway fixed effects model to estimate DD aims to eliminate selection bias in our data. "Twoway fixed effects" refers to having fixed effects for both unit and time. The monitor fixed effects allow us to compare each monitor to itself over time, which controls for unobserved qualities about a given monitor's location that are constant over time. We assume that there are no time-variant changes at the monitor level that correlate with the error term. By comparing each monitor to itself, we account for selection bias, although this alone may not account for non-parallel trends. The authors also use census region-year (hereafter, "region" and "region-year") fixed effects to control for uniform regional shocks that occur at a specific point in time.

The basic specification for the replication is

$$ln(ozone_{it}) = \alpha + \beta\{policies_{it}\} + \mu monitor_i + \theta year_t * region + \varepsilon_{it}$$

where $\{policies_{it}\}$ is a vector of four binary policy variables for every monitor i and time t, β is a vector of coefficients for each policy, $monitor_i$ is a unit-level fixed effect for every air pollution monitor, and $year_t * region$ is a time fixed effect interacted with census region.

Throughout the paper, the authors introduce further controls. The extended model is

$$ln(ozone_{it}) = \alpha + \beta\{policies_{it}\} + \mu\{monitor_i\} + \theta\{year_t * region\} + \gamma W_{it} + \eta D_t + \lambda I_{it} + \nu T_{it} + \varepsilon_{it}$$

where the additional values are vectors of controls for monitor-day weather and weather time trends (W); "dummy variables for day-of-week and day-of-year" (D); county-year income (I); and county-level linear time trends (T) that relate to policy adoption (Auffhammer & Kellogg, 2011a, pp. 2699-2700).

3.3 Identifying Assumptions

In DD, we assume that any policy change results in a level difference—that is, that the policy does not affect the rate of change, but rather the value of the outcome—and that without treatment, the treatment and control groups would have evolved similarly (the "parallel trends" assumption).

Based on the way the way the policy is implemented, it is likely that the level difference assumption is met. The authors note that policy implementation is almost immediate because manufacturers supply the new fuels at the same time, so all cars have access to the new fuels at the same time (Auffhammer & Kellogg, 2011a, p. 2700).

Auffhammer & Kellogg's data do not allow us to precisely evaluate the parallel trends assumption because the untreated comparison group is unclear as a result of the data structure. The policies are coded as a vector of four binary variables, with the omitted category being the "baseline" policy of 9.0 psi. However, the authors themselves say that this policy applies to "most ozone attainment counties" starting in 1991 (Auffhammer & Kellogg, 2011a, p. 2697), implying that it does not apply to all of them. Because this policy is omitted as the reference group, it is the default status for any county where another policy does not apply instead. This makes it difficult to precisely estimate the effects of the policies on treated counties.

4 Findings

In this section, I compare my replication findings to the authors' primary DD results as well as their secondary results, which are segmented by urban, suburban, and rural location. I also present sensitivity analyses of the authors' original results.

4.1 Primary Reuslts

Tables 2A, 2B, and 2C replicate the authors' main results table. Each table shows my result under the heading "Replication," the authors' original result under "Original," and the calculated percent change between my result and the authors' result. Tables 2A and 2B show findings using the dependent variable log(daily maximum ozone concentration), whereas Table 2C shows findings using the dependent variable log(daily maximum 8-hour ozone concentration). The different dependent variables reflect two different metrics EPA uses to assess NAAQS compliance.

I successfully replicate the authors' results for Models 1-4 (Tables 2A and 2B) and Models 6-7 (Table 2C). While my results in Models 5 and 8 are slightly different than the authors', they are in the same direction as the original results and are substantively similar.

Model 1 and Model 6 use the basic estimating equation that only includes unit and time fixed effects. The results are consistent with what we would expect from the policies: the coefficient on RVP I is slightly positive (recall that RVP I has a higher psi ceiling than the baseline policy) and statistically insignificant, while the coefficient on RVP II is slightly negative and statistically insignificant. The coefficients on both Federal RFG and CARB are significant at the 1 percent level for both dependent variables. The federal RFG policy is associated with a small decline in logged ozone levels, whereas the CARB policy is associated with a large decline. In Model 1 and Model 6, the CARB policy is associated with an approximately 9 percent decrease in logged ozone levels compared to the baseline policy.

The other models add additional controls. The coefficients change most when controlling for income and controlling for linear regulation-region time trends. The results are similar throughout the models: the CARB policy is consistently associated with a large and statistically significant decrease in logged ozone levels, the federal RFG policy with a small but statistically significant decrease in logged ozone levels, and the RVP policies are associated with statistically insignificant results.

4.2 Secondary Results: Urban, Suburban, and Rural

The authors then run separate analyses for urban monitors (Table 3A), suburban monitors (Table 3B), and rural monitors (Table 3C). For each type of location, they ran one model with weather controls, income controls, and day of week and day of year fixed effects, and a second model with linear regulation-region trends. Across all three geographic types, the largest effect sizes were found in the suburban analyses. The largest coefficient for the CARB policy occurs in the suburban analysis (Table 3B, Model 3). In that analysis, the CARB policy is associated with an approximately 11 percent decrease in logged ozone concentration and is significant at the 1 percent level.

4.3 Sensitivity Analyses

As part of my analysis, I tested whether the authors' models were sensitive to controls for income and weather. In the original paper, the authors never described their rationale for controlling for income. I considered that income may function as a proxy for population. Population could be related to ozone levels and policy implementation because highly populated areas could have more cars, which could lead to higher ozone levels and thus lead to policy assignment. High population levels are also more likely to be found in urban areas, which themselves are more likely to have higher incomes.

Table 4 shows sensitivity analyses for controlling for population. Column 1 shows the basic model (found in Table 2A, Model 1 and reproduced in Table 4) with only monitor and year fixed effects. Column 2 shows the basic model with added income controls. Column 3 shows the basic model with added linear population controls, whereas column 4 shows the basic model with quadratic population controls to account for the possibility that population does not have a linear relationship with logged ozone. Column 5 shows results for controlling for income, population, and quadratic population.

The CARB policy, controlling for income, is associated with an approximately 8 percent decrease in ozone compared to the baseline policy. That same policy, controlling for quadratic population, is also associated with an approximately 8 percent decrease in ozone. When controlling for both quadratic population and income in Column 5, the results are substantively the same as when we only control for quadratic population. Additionally, the federal RFG policy coefficient in this analysis is only statistically significant (at the 10 percent level) when controlling for quadratic population, which indicates that income may not be the most appropriate control. There is a more compelling theoretical case for population controls than for income. Based on this analysis, I believe that quadratic population would have been a better control than income to use throughout the models.

Table 5 shows the results of phasing in the weather controls. Model 1 is the basic model. Model 2 includes maximum daily temperature polynomials, Model 3 adds minimum daily temperature polynomials, Model 4 adds rainfall and snowfall polynomials, and Model 5 adds weather-day of week and weather-day of year interactions. This sensitivity analysis demonstrates that the authors' original analysis is not sensitive to any specific weather controls. The results in Model 3, which only controls for maximum and minimum daily temperature polynomials, are substantially the same as the results from Model 5, which adds precipitation polynomials and interaction terms.

5 Extension: Synthetic Control Method Analysis of CARB Policies on Asthma Mortality in California

In their conclusion, Auffhammer & Kellogg suggest that "CARB (and RFG) may convey health benefits through reductions in emissions of toxic air pollutants" (2011a, p. 2720). To extend their paper, I use synthetic control methods (SCM) to identify the effects of the CARB policy on asthma mortality rates. I selected asthma mortality as my dependent variable primarily due to data availability, but theory does support this choice of dependent variable.

One concern with using asthma mortality as the dependent variable is that it would be a lagging indicator. There are two components underlying this concern: first, that reductions in ozone levels would not be immediate following policy implementation, and second, that changes in asthma mortality would not be immediate following the reduction in ozone levels. The authors address the first point in their discussion of regression discontinuity (RD). They note that are able to use RD with a tight time frame because "imposition of a gasoline standard affects all cars simultaneously," and that we can expect to see a change shortly after the policy goes into effect (Auffhammer &

Kellogg, 2011a, p. 2700). As such, I feel confident that changes in ozone levels will be sensitive to implementation over a short time period.

Regarding the second concern, the environmental health literature indicates that "short-term exposures" to ozone are associated with asthma mortality, even after controlling for other pollutants (Liu et al., 2019). When the literature discusses lags in disease presentation following increased ozone levels, lags are discussed in terms of days, not years. For example, Zanobetti et al. (2003) used a model with a 40-day lag to study how air pollution affects respiratory mortality. The use of a relatively short lag period may indicate that changes in air pollution in one time period may not affect mortality in much later time periods. Because the data in my analysis reflect a long time horizon, I feel comfortable moving forward without lags.

5.1 Methods

I conduct two SCM analyses: one using logged ozone levels as the dependent variable to prove that this method produces comparable results to the DD, and the second using state-year average asthma mortality. In both analyses, the independent variable is the CARB policy, through the use of California as the treatment group. California at the state level is the appropriate treatment group because the CARB policy is applied equally across the state and because this policy is consistently associated with the largest effect size in the DD.

SCM uses statistical weighting to construct a "synthetic" comparison group based on the pre-treatment characteristics of the treatment group. In this case, California is the treatment group because of its unique CARB policy. I identified the two ozone concentration measures, population density, average maximum and minimum temperature, average rainfall, and per capita income as the characteristics on which to match for the first model evaluating ozone levels. For the model assessing the policy's impact on asthma mortality, I also include asthma mortality as a predictor variable.

The computer assesses all of those characteristics for the rest of the "donor pool"—the other states in the dataset—and assigns each state a weight that will be multiplied by the values of all of that state's predictor variables and added into, essentially, a weighted average that is supposed to match California's values.

In a SCM analysis, we hope to see that the pretreatment plots for California and "Synthetic

California" are perfectly aligned, and that the plot for California diverges from Synthetic California post-treatment.

The donor pool for both analyses excludes California, Arizona, and South Dakota for analytical reasons. California is the treatment, so it is not in the donor pool. I excluded Arizona from the donor pool because Arizona implements the CARB policies for part of each year (Auffhammer & Kellogg, 2011a, 2694), which means that it is not a valid control for California. South Dakota only had observations for 4 out of the 15 years in this period, so I excluded it to balance the panel. The authors' dataset had no observations from Alaska, Hawaii, Idaho, and Wyoming, so those states are not represented in the donor pool. The rest of the states, plus DC, are included in the donor pool.

5.2 Ozone Findings

Tables 6 and 7 show the weighted composition of Synthetic California and the pretreatment values of California, Synthetic California, and the overall donor pool means, respectively. For the ozone analysis, Synthetic California is derived from a weighted average of Utah, Maryland, and Colorado. In Table 7, we see that Synthetic California is more similar to real California than the donor pool for pretreatment values of both ozone measurements, minimum temperature, average rainfall, and population density. Synthetic California's means of the other variables—maximum temperature and per capita income—are somewhat less similar than the donor pool.

Despite some dissimilarities in the pretreatment values comparison, Figures 1-3 show that this method generally mirrors the results from the DD analysis. Figure 1 compares the overall trend in logged ozone for California and Synthetic California. While the trends are not identical in the pretreatment period, they are very similar. Following the implementation of the CARB policy in 1996, the trends diverge, and the California line consistently shows lower logged ozone values than Synthetic California. Figure 2, which plots the difference between California and Synthetic California (a "gap plot"), shows the same result by plotting the difference of California and Synthetic California. Following treatment in 1996, California's logged ozone levels are meaningfully below the midline, which would represent no difference between the state and its synthetic counterpart.

Because SCM does not offer the traditional significance tests used in other models, one way to demonstrate significance is through placebo testing. To run a placebo test, we undergo the same method described above while assigning the "treated" status to different units. We then plot the difference between each treated observation and its synthetic control on the same plot, alongside the true treated observation, to see if the true treated observation is a notable outlier.

Figure 3 is the placebo plot that compares California's gap plot to gap plots generated using every other state in the donor pool as the "treated" group. California's line, shown in black, is one of the lowest lines after treatment. When using the other states as the "treated" group, no other state is as consistently low following the treatment as California. Considering that parts of these other states were subject to the federal policies studied in the DD above and are not truly untreated, it is notable that California's gap is as different as it is. These results mirror the DD analysis and increase our confidence in using this method to evaluate the effects of the CARB policy on asthma mortality.

5.3 Asthma Mortality Findings

Tables 8 and 9 show the weighted composition of Synthetic California and the pretreatment values comparison using asthma mortality as the dependent variable, with the CARB policy again being the independent variable through California being the treated state. In order to maximize pretreatment similarity, I also included pretreatment asthma mortality rates in the set of criteria to match. For the asthma analysis, Synthetic California is composed of Utah, Maryland, Nevada, and DC.

California and Synthetic California match well on pretreatment values for both ozone measures, rainfall, and asthma mortality. Synthetic California's pretreatment values for maximum and minimum temperature variables, population density, and per capita income are all less similar to actual California than the donor pool mean. This table does not instill confidence that Synthetic California in this analysis is a good comparison for actual California.

In Figure 4, we see that Synthetic California's path tracks closely with California's path both before and after treatment. Based on Figure 4, there is no meaningful difference between California and Synthetic California with regard to asthma mortality rates. Figure 5 shows the same information as the difference between California and Synthetic California.

In the placebo plot in Figure 6, California's policy does not stand out as an effective tool for reducing asthma mortality. The California line is close to the 0 line throughout the plot, showing that there is no meaningful difference in asthma mortality rates between California and its synthetic comparison.

These results suggest that either SCM was less successful at estimating the effects of the CARB policy on asthma mortality, or that the policy does not have a significant effect on asthma mortality rates. Perhaps asthma mortality is not the right variable of interest: Other important analyses could study overall respiratory health, hospitalizations for respiratory conditions, or mortality from other chronic respiratory diseases. Additionally, matching on some variable representing healthcare access could have been useful. Given more time, the author would hope to pursue further analyses along these lines.

6 Discussion

This paper seems to paint a compelling picture that the CARB gasoline policy and the federal RFG policy are much more effective at reducing ozone pollution than any of the RVP policies. The results are robust to sensitivity analyses, and the CARB policy in particular can be successfully replicated using SCM.

This conclusion is supported by theory: RFG policies, which include CARB, are much stricter than RVP policies ("Date of Switch to Summer-grade Gasoline Approaches," 2013). Additionally, the authors note that the RVP and federal RFG policies give manufacturers more flexibility to meet the standards, whereas CARB limits particular compounds that are most likely to contribute to ozone creation (Auffhammer & Kellogg, 2011a, p. 2719). Because CARB limits specific compounds, the policy is more targeted at limiting ozone creation. SCM did not reveal a meaningful relationship between the CARB policy and reductions in asthma mortality.

Although the authors' results are consistent with theory, their estimation may be biased. Since this paper's publication in 2011, the DD literature on estimating models with overlapping treatment implementation has evolved. In 2021, Andrew Goodman-Bacon published a paper in the Journal of Econometrics calling into question the effectiveness of the twoway fixed effects estimator at providing unbiased estimates when "treatment timing varies across units." In these cases, the estimation provides "a weighted average of all possible simple 2x2 DDs that compare one group that changes treatment status to another group that does not" (Goodman-Bacon, 2021). The assumptions are rather strict: that a “variance-weighted average” of the untreated groups before treatment is 0, and that the “average treatment effects for each timing group do not change over time” (Goodman-Bacon, 2021). When those assumptions are not met, twoway fixed effects is not an unbiased estimator.

First, this dataset does not provide the ability to assess the first assumption. The data begin in 1989, concurrently with the first policy studied, which is used as the baseline throughout. Second, it is unlikely that these policies have an average treatment effect that does not change over time.

Three topics come to mind that could affect the average treatment effect of the policies each year: Weather, manufacturing changes, and electric vehicles. As discussed earlier, weather is a key component to ozone creation. Since weather trends can be inconsistent across years, and with climate change driving hotter temperatures each year, the background conditions for ozone creation are not constant and could affect ozone creation. Second, since most of these policies offer manufacturers flexibility for developing fuels that meet the standards, manufacturing changes could affect whether the fuel actually reduces ozone levels, even if it continues to be in compliance. Finally, the advent of electric vehicles could affect how much gas is used and contributes to ozone creation. In this time period, the California Air Resources Board “adopted the Zero-Emission Vehicle requirement in 1990” that spurred electric vehicle development and deployment in the state (CARB, 2025c).

Overall, it is unclear that Auffhammer and Kellogg’s original paper using twoway fixed effect estimators for DD produces an unbiased estimate of the effects of these policies. Further study should reevaluate the analysis in light of Goodman-Bacon’s developments in the literature.

References

[1] Auffhammer, Maximilian and Ryan Kellogg. "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality." *American Economic Review* 101 (2011a): 2687-2722.

[2] Auffhammer, Maximilian, and Kellogg, Ryan. Replication data for: Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality. Nashville, TN: American Economic Association [publisher], 2011b. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-10-11. https://doi.org/10.3886/E112465V1.

[3] Clean Air Act, 42 U.S.C. § 7401(b) (2023).

[4] California Air Resources Board. "California Reformulated Gasoline." Accessed April 5, 2025a. https://ww2.arb.ca.gov/our-work/programs/fuels-enforcment-program/california-reformulated-gasoline.

[5] California Air Resources Board. "California RFG Fact Sheets." Accessed April 5, 2025b. https://afdc.energy.gov/files/pdfs/3002.pdf.

[6] California Air Resources Board. "Zero-Emission Vehicle Program." Accessed April 27, 2025c. https://ww2.arb.ca.gov/our-work/programs/zero-emission-vehicle-program/about.

[7] Congressional Research Service (CRS). *Clean Air Act: A Summary of the Act and Its Major Requirements* by Richard K. Lattanzio. RL30853, September 13, 2022. https://www.congress.gov/crs-product/RL30853.

[8] "Date of Switch to Summer-grade Gasoline Approaches." U.S. Energy Information Administration, (2013). Accessed April 27, 2025. https://www.eia.gov/todayinenergy/detail.php?id=11031.

[9] Environmental Protection Agency. "Evolution of the Clean Air Act." Last modified November 12, 2024a. https://www.epa.gov/clean-air-act-overview/evolution-clean-air-act.

[10] Environmental Protection Agency. "Gasoline Reid Vapor Pressure." Last modified November 12, 2024b. https://www.epa.gov/gasoline-standards/gasoline-reid-vapor-pressure.

[11] Environmental Protection Agency. “Ground-level Ozone Basics.” Last modified March 11, 2025a. https://www.epa.gov/ground-level-ozone-pollution/ground-level-ozone-basics.

[12] Environmental Protection Agency. “Health Effects of Ozone Pollution.” Last modified March 13, 2025b. https://www.epa.gov/ground-level-ozone-pollution/health-effects-ozone-pollution.

[13] Environmental Protection Agency. “Learn about Gasoline.” Last modified August 22, 2024c. https://www.epa.gov/gasoline-standards/learn-about-gasoline.

[14] Environmental Protection Agency. “Reformulated Gasoline.” Last modified March 26, 2025c. https://www.epa.gov/gasoline-standards/reformulated-gasoline.

[15] Goodman-Bacon, Andrew. “Difference-in-differences with variation in treatment timing.” *Journal of Econometrics* 225, no. 2 (2021).

[16] Institute for Health Metrics and Evaluation (IHME). United States Chronic Respiratory Disease Mortality Rates by County 1980-2014. Seattle, United States of America: Institute for Health Metrics and Evaluation (IHME), 2017.

[17] Liu, Yuewei, Jingju Pan, Hai Zhang, Chunxiang Shi, Guo Li, Zhe Peng, Jixuan Ma, Yun Zhou, and Lan Zhang. “Short-term Exposure to Ambient Air Pollution and Asthma Mortality.” *American Journal of Respiratory and Critical Care Medicine* 200, no. 1 (2019).

[18] Missouri Census Data Center. “Geocorr 2000: Geographic Correspondence Engine.” MCDC Data Applications (2010). Accessed April 19, 2025, from https://mcdc.missouri.edu/applications/geocorr2000.html.

[19] “Racial and Ethnic Diversity, 1980-2010: Data Files.” Brown University American Communities Project (2010). Accessed April 19, 2025, from https://s4.ad.brown.edu/projects/diversity/Data/Download4.htm.

[20] Zanobetti, Antonella, Joel Schwartz, Evi Samoli, Alexandros Gryparis, Giota Touloumi, Janet Peacock, Ross H. Anderson, Alain Le Tertre, Janos Bobros, Martin Celko, Ayana Goren, Bertil Forsberg, Paola Michelozzi, Daniel Rabczenko, Santiago Perez Hoyos, H. Erich Wichmna, and Klea Katsouyanni. “The Temporal Pattern of Respiratory and Heart Disease Mortality in Response to Air Pollution.” *Environmental Health Perspectives* 111, no. 9 (2003): 1188-1193.

A Tables and Figures

A.1 Summary Statistics Table Replication

Table 1: Summary Statistics on Monitor-Day Observations for Summer Ozone Season

		Monitor-Day Observations			Number of Counties		
	Year	Observations, Replication	Observations, Original	Difference	Counties, Replication	Counties, Original	Difference
	1989	63,084	63,076	8	418	418	0
	1990	66,119	66,108	11	436	436	0
	1991	69,174	69,164	10	451	451	0
	1992	69,853	69,848	5	452	452	0
	1993	72,616	72,606	10	469	469	0
	1994	74,449	74,440	9	473	473	0
	1995	77,024	77,007	17	477	477	0
	1996	76,470	76,462	8	471	471	0
	1997	78,286	78,283	3	478	478	0
	1998	79,555	79,544	11	487	487	0
	1999	80,753	80,750	3	485	485	0
	2000	82,468	82,466	2	489	489	0
	2001	83,786	83,781	5	490	490	0
	2002	85,231	85,230	1	495	495	0
	2003	85,261	85,260	1	498	498	0
Total	—	1,144,129	1,144,025	104	7,069	7,069	0
Average	—	76,275	76,268	7	471	471	0

A.2 Primary Results Table Replication: DD Estimation Results

Table 2A: Difference-in-Difference Estimation Results

	Dependent var: ln(daily maximum ozone concentration)								
	Model 1			Model 2			Model 3		
	Replication	Original	Pct. Chg.	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.016	0.016	0%	0.013	0.013	0%	0.015	0.015	0%
	(0.016)	(0.016)	0%	(0.015)	(0.015)	0%	(0.016)	(0.016)	0%
RVP Phase II	-0.007	-0.007	0%	-0.011	-0.011	0%	-0.007	-0.007	0%
	(0.008)	(0.008)	0%	(0.007)	(0.007)	0%	(0.007)	(0.007)	0%
Federal RFG	-0.029***	-0.029***	0%	-0.030***	-0.030***	0%	-0.016**	-0.016**	0%
	(0.009)	(0.009)	0%	(0.007)	(0.007)	0%	(0.008)	(0.008)	0%
CARB Gasoline	-0.095***	-0.095***	0%	-0.090***	-0.090***	0%	-0.077***	-0.077***	0%
	(0.013)	(0.013)	0%	(0.011)	(0.011)	0%	(0.011)	(0.011)	0%
County income ($ billion)	-	-	-	-	-	-	-1.379***	-1.379***	0%
	-	-	-	-	-	-	(0.285)	(0.285)	0%
Monitor FEs	Yes			Yes			Yes		
Region-year FEs	Yes			Yes			Yes		
Region-DOW FEs	No			Yes			Yes		
Region FE-DOY interaction	No			Yes			Yes		
Weather controls	No			Yes			Yes		
Income	No			No			Yes		
Regulation-region trends	No			No			No		
Regulation-region quad trends	No			No			No		
Observations	1,144,129	1,144,025	0.01%	1,144,037	1,144,025	0.001%	1,144,037	1,144,025	0.001%
R-squared (within-monitor)	0.024	0.024	0.00%	0.261	0.261	0.00%	0.261	0.261	0.00%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

Table 2B: Difference-in-Difference Estimation Results

	Dependent var: ln(daily maximum ozone concentration)					
	Model 4			Model 5		
	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.001	0.001	0.0%	0.004	0.004	0.0%
	(0.016)	(0.016)	0.0%	(0.018)	(0.018)	0.0%
RVP Phase II	-0.012	-0.012	0.0%	-0.011	-0.012	-8.3%
	(0.009)	(0.009)	0.0%	(0.012)	(0.011)	9.1%
Federal RFG	-0.036***	-0.036***	0.0%	-0.018	-0.019	-5.3%
	(0.011)	(0.011)	0.0%	(0.012)	(0.012)	0.0%
CARB Gasoline	-0.065***	-0.065***	0.0%	-0.063***	-0.064***	-1.6%
	(0.019)	(0.019)	0.0%	(0.022)	(0.020)	10.0%
County income ($ billion)	-0.226	-0.225	0.4%	-0.303	-0.302	0.3%
	(0.245)	(0.245)	0.0%	(0.233)	(0.236)	-1.3%
Monitor FEs	Yes			Yes		
Region-year FEs	Yes			Yes		
Region-DOW FEs	Yes			Yes		
Region FE-DOY interaction	Yes			Yes		
Weather controls	Yes			Yes		
Income	Yes			Yes		
Regulation-region trends	Yes			Yes		
Regulation-region quad trends	No			Yes		
Observations	1,144,037	1,144,025	0.001%	1,144,037	1,144,025	0.001%
R-squared (within-monitor)	0.264	0.264	0.00%	0.264	0.264	0.00%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

Table 2C: Difference-in-Difference Estimation Results

	Dependent var: ln(daily maximum 8-hour ozone concentration)								
	Model 6			Model 7			Model 8		
	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.017	0.018	-5.6%	0.015	0.015	0.0%	0.005	0.004	25.0%
	(0.017)	(0.017)	0.0%	(0.018)	(0.017)	5.9%	(0.020)	(0.020)	0.0%
RVP Phase II	-0.005	-0.005	0.0%	-0.009	-0.009	0.0%	-0.011	-0.011	0.0%
	(0.008)	(0.008)	0.0%	(0.007)	(0.007)	0.0%	(0.012)	(0.012)	0.0%
Federal RFG	-0.028***	-0.028***	0.0%	-0.028	-0.028	0.0%	-0.021	-0.022	-4.5%
	(0.009)	(0.009)	0.0%	(0.008)***	(0.008)***	0.0%	(0.013)	(0.013)*	0.0%
CARB Gasoline	-0.090***	-0.090***	0.0%	-0.086	-0.086	0.0%	-0.062***	-0.063***	-1.6%
	(0.013)	(0.013)	0.0%	(0.012)***	(0.012)***	0.0%	(0.021)	(0.021)	0.0%
County income ($ billion)	-	-	-	-	-	-	-0.103	-0.102	1.0%
	-	-	-	-	-	-	(0.246)	(0.246)	-0.0%
Monitor FEs	Yes			Yes			Yes		
Region-year FEs	Yes			Yes			Yes		
Region-DOW FEs	No			Yes			Yes		
Region FE-DOY interaction	No			Yes			Yes		
Weather controls	No			Yes			Yes		
Income	No			No			Yes		
Regulation-region trends	No			No			Yes		
Regulation-region quad trends	No			No			Yes		
Observations	1,144,129	1,144,025	0.01%	1,144,037	1,144,025	0.001%	1,144,037	1,144,025	0.001%
R-squared (within-monitor)	0.026	0.026	0.00%	0.255	0.255	0.000%	0.259	0.259	0.000%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

A.3 Secondary Results Table Replication: DD Segmented by Urban, Suburban, and Rural

Table 3A: Difference-in-Difference Estimation Results for Urban Monitors

	Dependent var: ln(daily maximum ozone concentration)					
	Model 1			Model 2		
	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.020	0.020	0.0%	0.007	0.007	0.0%
	(0.021)	(0.021)	0.0%	(0.021)	(0.021)	0.0%
RVP Phase II	0.002	0.002	0.0%	0.002	0.002	0.0%
	(0.014)	(0.014)	0.0%	(0.014)	(0.014)	0.0%
Federal RFG	-0.004	-0.004	0.0%	-0.032**	-0.032**	0.0%
	(0.013)	(0.016)	-18.8%	(0.015)	(0.015)	0.0%
CARB Gasoline	-0.071***	-0.071***	0.0%	-0.072***	-0.072***	0.0%
	(0.016)	(0.016)	0.0%	(0.025)	(0.025)	0.0%
County income ($ billion)	-1.236***	-1.236***	0.0%	0.470	0.470	0.0%
	(0.397)	(0.397)	0.0%	(0.428)	(0.428)	0.0%
Monitor FEs	Yes			Yes		
Region-year FEs	Yes			Yes		
Region-DOW FEs	Yes			Yes		
Region FE-DOY interaction	Yes			Yes		
Weather controls	Yes			Yes		
Income	Yes			Yes		
Regulation-region trends	No			Yes		
Observations	222,982	222,982	0.0%	222,982	222,982	0.0%
R-squared (within-monitor)	0.281	0.281	0.0%	0.285	0.285	0.0%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

Table 3B: Difference-in-Difference Estimation Results for Suburban Monitors

	Dependent var: ln(daily maximum ozone concentration)					
	Model 3			Model 4		
	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.024	0.024	0.0%	-0.001	-0.001	0.0%
	(0.018)	(0.018)	0.0%	(0.016)	(0.016)	0.0%
RVP Phase II	-0.011	-0.011	0.0%	-0.020*	-0.020*	0.0%
	(0.009)	(0.009)	0.0%	(0.011)	(0.011)	0.0%
Federal RFG	-0.025***	-0.025***	0.0%	-0.049***	-0.049***	0.0%
	(0.009)	(0.009)	0.0%	(0.014)	(0.014)	0.0%
CARB Gasoline	-0.113***	-0.113***	0.0%	-0.101***	-0.102***	-1.0%
	(0.014)	(0.014)	0.0%	(0.022)	(0.022)	0.0%
County income ($ billion)	-1.614***	-1.614***	0.0%	-0.728***	-0.727***	0.1%
	(0.269)	(0.268)	0.4%	(0.219)	(0.219)	0.0%
Monitor FEs	Yes			Yes		
Region-year FEs	Yes			Yes		
Region-DOW FEs	Yes			Yes		
Region FE-DOY interaction	Yes			Yes		
Weather controls	Yes			Yes		
Income	Yes			Yes		
Regulation-region trends	No			Yes		
Observations	490,539	490,539	0.0%	490,539	490,539	0.0%
R-squared (within-monitor)	0.275	0.275	0.0%	0.278	0.278	0.0%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

Table 3C: Difference-in-Difference Estimation Results for Rural Monitors

	Dependent var: ln(daily maximum ozone concentration)					
	Model 5			Model 6		
	Rep.	Orig.	Pct. Chg.	Rep.	Orig.	Pct. Chg.
RVP Phase I	0.011	0.011	0.0%	0.0005	0.0005	0.0%
	(0.018)	(0.018)	0.0%	(0.022)	(0.022)	0.0%
RVP Phase II	-0.006	-0.006	0.0%	-0.011	-0.011	0.0%
	(0.009)	(0.009)	0.0%	(0.011)	(0.011)	0.0%
Federal RFG	-0.019	-0.018	5.6%	-0.027**	-0.027**	0.0%
	(0.011)	(0.011)	0.0%	(0.013)	(0.013)	0.0%
CARB Gasoline	-0.040***	-0.040***	0.0%	-0.017	-0.017	0.0%
	(0.013)	(0.013)	0.0%	(0.024)	(0.024)	0.0%
County income ($ billion)	-1.144*	-1.145*	-0.1%	0.051	0.051	0.0%
	(0.629)	(0.629)	0.0%	(0.842)	(0.841)	0.1%
Monitor FEs	Yes			Yes		
Region-year FEs	Yes			Yes		
Region-DOW FEs	Yes			Yes		
Region FE-DOY interaction	Yes			Yes		
Weather controls	Yes			Yes		
Income	Yes			Yes		
Regulation-region trends	No			Yes		
Observations	430,516	430,504	0.003%	430,516	430,504	0.003%
R-squared (within-monitor)	0.243	0.243	0.000%	0.244	0.244	0.000%

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

A.4 Sensitivity Analyses

Table 4: Difference-in-Difference Estimation Results With Population Controls

	Dependent var: ln(daily maximum ozone concentration)					
	(1)		(2)	(3)	(4)	(5)
	Replication	Original				
RVP Phase I	0.016	0.016	0.017	0.022	0.017	0.017
	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)
RVP Phase II	-0.007	-0.007	-0.001	-0.005	-0.005	-0.004
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Federal RFG	-0.029***	-0.029***	-0.012	-0.014	-0.017*	-0.017*
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)
CARB Gasoline	-0.095***	-0.095***	-0.079***	-0.093***	-0.080***	-0.080***
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)
County income ($ billion)	-	-	-0.169***	-	-	-0.047
	-	-	(0.000)	-	-	(0.422)
Population (thousands)	-	-	-	-0.019***	-0.002	-0.002
	-	-	-	(0.005)	(0.006)	(0.006)
Quad. Population (thousands)	-	-	-	-	-0.0002***	-0.0002***
	-	-	-	-	0.0003	-0.000
Monitor FEs	Yes		Yes	Yes	Yes	Yes
Region-year FEs	Yes		Yes	Yes	Yes	Yes
Income	No		Yes	No	No	Yes
Population	No		No	Yes	Yes	Yes
Quadratic Population	No		No	No	Yes	Yes
Observations	1,144,129	1,144,025	1,144,129	1,144,129	1,144,129	1,144,129
R-squared (within-monitor)	0.024	0.024	0.0251	0.0250	0.0254	0.025

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

Table 5: Difference-in-Difference Estimation Results With Phased-In Weather Controls

	Dependent var: ln(daily maximum ozone concentration)					
	(1)		(2)	(3)	(4)	(5)
	Replication	Original				
RVP Phase I	0.016	0.016	0.013	0.014	0.013	0.013
	(0.016)	(0.016)	(0.014)	(0.015)	(0.015)	(0.015)
RVP Phase II	-0.007	-0.007	-0.011	-0.012*	-0.012	-0.011
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Federal RFG	-0.029***	-0.029***	-0.032***	-0.030***	-0.030***	-0.029***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
CARB Gasoline	-0.095***	-0.095***	-0.084***	-0.091***	-0.091***	-0.091***
	(0.013)	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)
Monitor FEs	Yes		Yes	Yes	Yes	Yes
Region-year FEs	Yes		Yes	Yes	Yes	Yes
Max Temperature Polynomial	No		Yes	Yes	Yes	Yes
Min Temperature Polynomial	No		No	Yes	Yes	Yes
Precipitation Polynomials	No		No	No	Yes	Yes
Weather Interactions	No		No	No	No	Yes
Observations	1,144,129	1,144,025	1,144,129	1,144,129	1,144,129	1,144,037
R-squared (within-monitor)	0.024	0.024	0.192	0.220	0.228	0.250

Standard errors clustered by state-year. *** significant at the 1 percent level. ** significant at the 5 percent level. * significant at the 10 percent level. Data restricted to summer months only.

A.5 Extension: Synthetic Control

Table 6: Synthetic Control Weights for Logged Ozone

Weights	States	Weights	States
0	Alabama	0	Montana
n.d.	Alaska	0	Nebraska
exc.	Arizona	0	Nevada
0	Arkansas	0	New Hampshire
0.118	**Colorado**	0	New Jersey
0	Connecticut	0	New Mexico
0	Delaware	0	New York
0	DC	0	North Carolina
0	Florida	0	North Dakota
0	Georgia	0	Ohio
n.d.	Hawaii	0	Oklahoma
n.d.	Idaho	0	Oregon
0	Illinois	0	Pennsylvania
0	Indiana	0	Rhode Island
0	Iowa	0	South Carolina
0	Kansas	exc.	South Dakota
0	Kentucky	0	Tennessee
0	Louisiana	0	Texas
0	Maine	**0.58**	**Utah**
0.302	**Maryland**	0	Vermont
0	Massachusetts	0	Virginia
0	Michigan	0	Washington
0	Minnesota	0	West Virginia
0	Mississippi	0	Wisconsin
0	Missouri	n.d.	Wyoming

n.d. = no data in dataset; exc. = excluded from donor pool. Values are weights assigned to states in creating synthetic California.

Table 7: Balance Table for Logged Ozone

	California		
	Real	Synthetic	Donor Pool Mean
Ozone (8 hour max.)	0.056	0.057	0.050
Ozone (daily max.)	0.067	0.066	0.057
Maximum Temperature (Mean)	85.319	86.799	84.314
Minimum Temperature (Mean)	57.901	61.204	62.322
Rainfall (Mean)	0.488	6.256	12.435
Population Density (Mean)	699.128	898.449	1264.733
Income Per Capita (Mean, $ Thousands)	21.329	19.209	20.644

Figure 1: Trend in Logged Ozone: California and Synthetic California

Figure 2: Logged Ozone Gap between California and Synthetic California

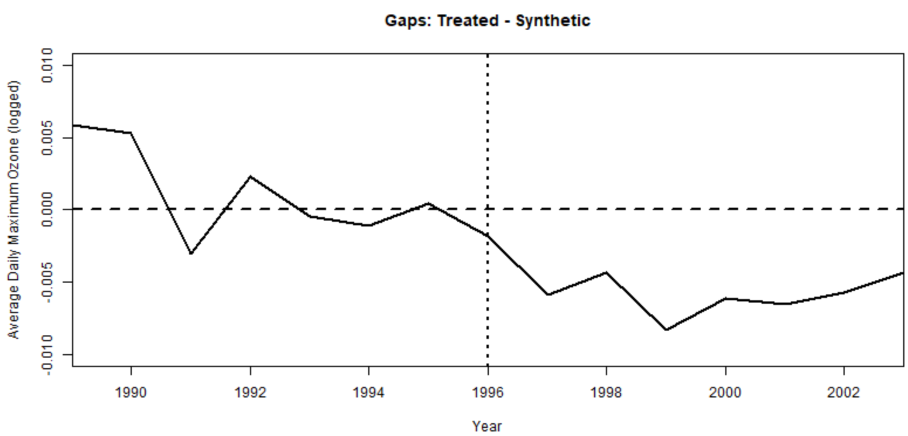

Figure 3: Logged Ozone Gap in California and Placebo Plots for Other States

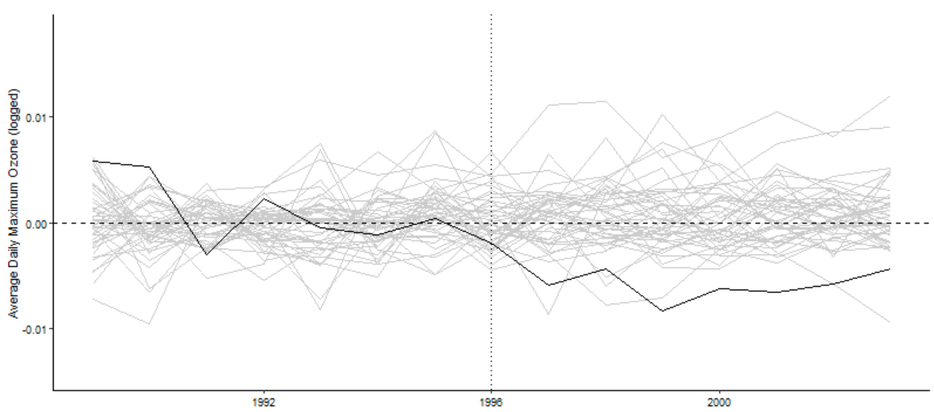

Table 8: Synthetic Control Weights for Annual Average Asthma Mortality

Weights	States	Weights	States
0	Alabama	0	Montana
n.d.	Alaska	0	Nebraska
exc.	Arizona	**0.151**	**Nevada**
0	Arkansas	0	New Hampshire
0	Colorado	0	New Jersey
0	Connecticut	0	New Mexico
0	Delaware	0	New York
0.053	**DC**	0	North Carolina
0	Florida	0	North Dakota
0	Georgia	0	Ohio
n.d.	Hawaii	0	Oklahoma
n.d.	Idaho	0	Oregon
0	Illinois	0	Pennsylvania
0	Indiana	0	Rhode Island
0	Iowa	0	South Carolina
0	Kansas	exc.	South Dakota
0	Kentucky	0	Tennessee
0	Louisiana	0	Texas
0	Maine	**0.469**	**Utah**
0.320	**Maryland**	0	Vermont
0	Massachusetts	0	Virginia
0	Michigan	0	Washington
0	Minnesota	0	West Virginia
0	Mississippi	0	Wisconsin
0	Missouri	n.d.	Wyoming

n.d. = no data in dataset; exc. = excluded from donor pool. Values are weights assigned to states in creating synthetic California.

Table 9: Balance Table for Annual Average Asthma Mortality

	California		
	Real	Synthetic	Donor Pool Mean
Ozone (8 hour max.)	0.056	0.057	0.05
Ozone (daily max.)	0.067	0.066	0.057
Maximum Temperature (Mean)	85.319	87.686	84.314
Minimum Temperature (Mean)	57.901	62.39	62.322
Rainfall (Mean)	0.488	6.251	12.435
Population Density (Mean)	699.128	1320.648	1264.733
Income Per Capita (Mean, $ Thousands)	21.329	20.492	20.644
Asthma Mortality (Mean)	2.442	2.441	2.213

Figure 4: Trend in Asthma Mortality: California and Synthetic California

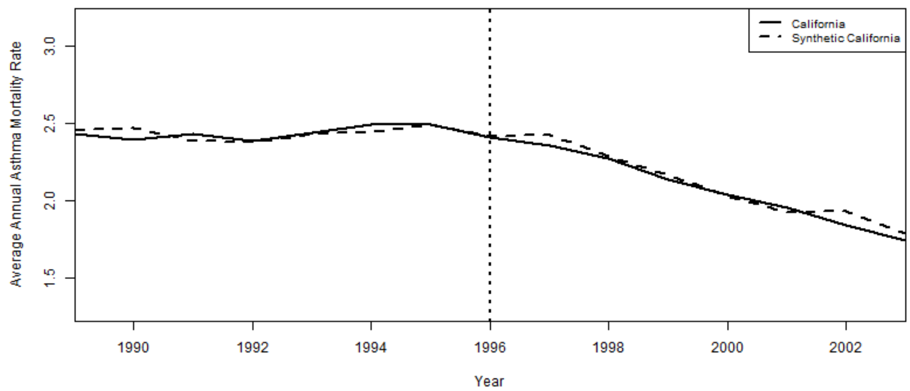

Figure 5: Asthma Mortality Gap between California and Synthetic California

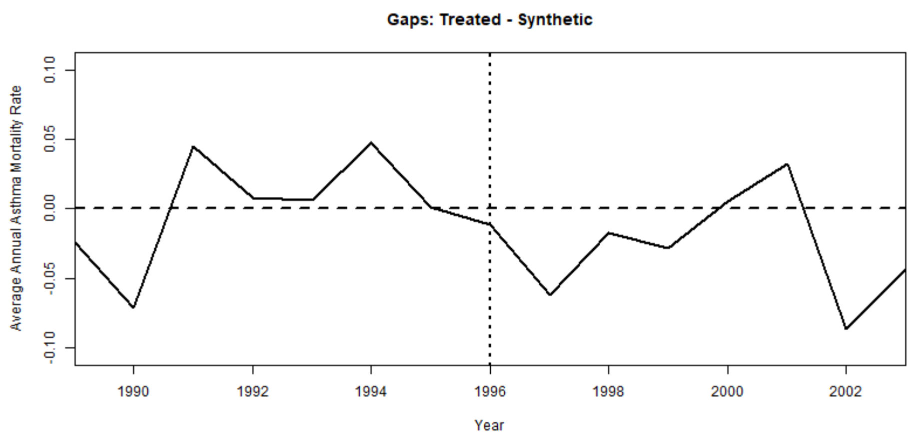

Figure 6: Asthma Mortality Gap in California and Placebo Plots for Other States