Full of Hot Air? Replicating Estimates of Regulatory Effects on Air Quality

Jack (Quan Cheng) Xie
Topics in Econometrics—Uppsala University
Professor Adrian Adermon

5 June 2022

1 Introduction

Given the large health and environmental tolls of air pollution, how to reduce harmful emissions such as ground level ozone is a key regulatory concern. However, not only is empirical estimation of regulatory effects on air quality in the United States complicated by wide industrial and political heterogeneity, to model complex interactions between pollution and weather systems is a notoriously difficult statistical and meteorological challenge.

Using econometric methods, Maximilian Auffhammer and Ryan Kellogg (2011), henceforth AK, seek to identify the casual effects of gasoline content regulation on ground level ozone concentrations in the United States. They examine the effect of three regulations in particular: federal Reid vapor pressure (RVP) regulation, federal reformulated gasoline (RFG), and California reformulated gasoline (CARB).

Federal RVP regulations limited the intensity of emissions from gasoline sold during the summer months without restrictions on the chemical composition of emissions. The policy was implemented in two phases. Phase I was enforced between 1989 to 1991 and limited emissions RVP to either 10.5, 9.5 or 9.0 psi across different counties. Phase II began in 1992 and enforces a summer RVP limit of either 9.0 and 7.8 psi.

Federal RFG was enforced since 1995 in areas of severe nonattainment of EPA air quality standards. RFG regulation targets the chemical content of emissions, and has a stricter content and performance criteria than both phases of RVP.

CARB, California's own reformulated gasoline program that began enforcement in 1996, expanded on federal RFG regulations and is the most stringent regulatory treatment examined.

Using panel data from weather monitors tracking ground-level ozone concentrations, AK estimate the effects using both difference-in-differences (DD) and regression discontinuity design (RD) approaches. They find that federal gasoline standards allowing flexible compliance did not improve air quality, while chemical content regulation was able to improve air quality. Gasoline refiners are market agents most affected by such regulations, and AK argue that the flexibility of non-content specific federal emission standards allow refiners to reduce the cost of compliance to regulation without actually reducing emissions.

I replicate AK's main empirical results and explore the robustness of their process and estimates. While their reasoning of the differential effectiveness of air quality regulations are credible, their implementation of econometric techniques on studying air quality regulation outcomes raises many issues of the internal and external validity.

2 Data Preparation

AK use data on ozone concentrations from the Environmental Protection Agency's (EPA) Air Quality Standards Database for 1989-2003. The panel data tracks daily maximum ozone concentration levels recorded by pollution monitors across 49 states. They also combine these air quality observations with weather data from the National Climatic Data Center's Cooperative Station Data (NOAA 2008).

AK perform several important data processing steps in preparation for their DD estimation, some of which appear discretionary. First, they remove data points from less than nine hours of observations between 9 AM and 9 PM according to EPA standards. Second, they keep data only on monitors with more than 75% observed days in each season. AK does not motivate the use of the 75% threshold.

Third, to rule out potential spill-over effects and contaminated measurements, AK drop data from counties neighboring other counties under more stringent regulation. There may be a minor error in AK's Stata code when merging the treated-neighbor dataset.[1] I correct this merging procedure and check the robustness of the estimates with the adjusted data sample.

Since RVP and some aspects of RFG are enforced only between summer months from June 1st and August 31st, AK keep only observations during these months for their DD estimation. Lastly, they remove observations without weather data from NOAA.

AK also construct complex weather, region, and time covariates for their estimations. Weather variables include cubic temperature polynomials, quadratics in rain and snow, lagged temperature variables. Monitors are assigned to one of four US census regions. Time variables include day of week and day of year. They also create many variables interactions, including weather with day of week, and linear and quadratic regulation-region trends.

AK reports yearly tabulations of their processed DD data sample with 1,144,025 total observations, which I replicate sample used in Table 1. Our total and yearly observations are the same. However, due to changes in the order of data processing in my replication, the composition of the treatment groups tabulations differ slightly. This difference may be indicative of the sensitivity of the resulting data to relatively minor changes in the complex cleaning process.

The replicated yearly tabulations by treatment groups are also summarized in Table 1. It is important to note that the baseline group are all monitors in counties with a 9.0 psi RVP limit. This means that no units in the data are untreated with any regulation. In fact, monitors in counties subject to a the baseline 9.0 psi RVP follow a more stringent standard than those in the RVP Phase I treatment group, which follow RVP limits of only 9.5 or 10.0 psi. Assuming the 9.0 psi baseline treatment has an effect in the same direction as the other treatments, the DD model would underestimate the effects of the non-baseline treatments on the outcome variables (Fricke 2017).

In the next section I discuss the empirical strategies of AK's DD implementation and my robustness checks.

3 Difference-in-Differences (DD) Strategies

In this section I summarize AK's main DD strategies. Then I discuss the adjustments I make on the AK's model specifications to test the the robustness of their implementation. Lastly, I introduce the staggered DD estimation strategy I use to test the validity of AK's identifying assumption for casual DD estimates. The staggered DD estimates period effects on air quality before and after the treatment begins. Testing if the pre-treatment effects are statistically zero verifies that the parallel trends assumptions are

[1] After merging by county using `AER20090377_NeighborData.dta`, AK keeps only `not matched` counties from the `master` data. The correct process should be to drop the treated neighbors in the `matched` counties and the `not matched` observations from the `using` data.

Table 1—Replicated Summary Statistics on Monitor and Regulations for Summer Ozone Season

	Daily		Counts of active monitors			Counts of treatments			
	Observations	Counties	Total monitors	Urban	Rural	RVPI[a]	RVPII[b]	RFG95	CARB
1989	63,076	418	720	153	244	326			
1990	66,108	436	751	157	268	343			
1991	69,164	451	782	151	297	369			
1992	69,848	452	789	155	300		129		
1993	72,606	469	815	167	301		133		
1994	74,440	473	835	163	316		135		
1995	77,007	477	865	170	330		103	107	
1996	76,462	471	854	165	330		66	99	48
1997	78,283	478	873	166	336		64	101	47
1998	79,544	487	889	165	344		67	97	47
1999	80,750	485	899	168	344		67	97	47
2000	82,466	489	915	178	346		71	90	46
2001	83,781	490	929	178	355		69	89	45
2002	85,230	495	943	177	361		69	87	47
2003	85,260	498	945	180	362		67	87	48
Total	1,144,025	7,069	12,804	2,493	4,834	1,038	1,040	854	375

Notes: Data constructed by removing (i) observations with less than 9 hours between 9 AM and 9 PM; (ii) monitors with less than 75% of observed summer days; (iii) observations missing weather or income; (iv) counties with more stringently regulated neighbors; (v) observations outside of summer months. AK remove data in the following order: (i) (ii) (iv) (iii), and (v).
[a]The RVPI Column lists number of counties with a 9.5 or 10.5 psi RVP requirement.
[b]The RVPII Column lists number of counties with a below 7.8 psi RVP requirement.

not violated.

AK estimate a basic DD model with a log-linear specification,

$$\ln(y_{it}) = \boldsymbol{\alpha} \cdot \mathbf{Treat}_{ct} + \mu_i + \eta_{ry} + \varepsilon_{it}, \tag{1}$$

for monitor i in county c, where y_{it} is the ozone dependent variable, $\mathbf{Treat}_{ct}$ is the vector of treatment indicators at date t in county c, μ_i is the monitor fixed effect and η_{ry} is the region-year fixed effect. The identification assumption for the causal estimate of $\boldsymbol{\alpha}$ is that $E[\mathbf{Treat}_{ct} \cdot \varepsilon_{it} | \mu_i, \eta_{ry}] = 0$.

AK also estimates equation (2) which adds various control variables discussed in the previous section.

$$\begin{aligned} \ln(y_{it}) =& \boldsymbol{\alpha} \cdot \mathbf{Treat}_{ct} + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + \boldsymbol{\gamma}_r \cdot \mathbf{D}_t + \delta \cdot \mathbf{I}_{ct} \\ &+ \boldsymbol{\theta} \cdot \mathbf{Trend}_{rct} + \mu_i + \eta_{ry} + \varepsilon_{it}. \end{aligned} \tag{2}$$

The variables $\mathbf{W}_{it}$ control for monitor specific weather shocks with lags and interactions, $\mathbf{D}_t$ denotes day-of-week and day-of-year, $\mathbf{I}_{ct}$ denotes county-level income, and $\mathbf{Trend}_{rct}$ controls for linear time trends in each county and census region. Similar to equation (1), then identifying assumption for causal estimate of $\boldsymbol{\alpha}$ in equation (2) is that the unobserved factors are not correlated with treatment, i.e. $E[\mathbf{Treat}_{ct} \cdot \varepsilon_{it} | \mathbf{W}_{it}, \mathbf{D}_t, \mathbf{I}_{ct}, \mathbf{Trend}_{rct}, \mu_i, \eta_{ry}] = 0$.

AK estimates equations (1) and (2) with both the full sample of monitors and a restricted sample of monitors with observation in each year of the 15-years time span in the data. They cluster by state-year to allow for correlation within state-year cells, including within-year serial correlation and within-state cross-sectional correlation.

Additionally, AK's Stata code differences the dependent, treatment, and certain control variables with their monitor-specific means. This procedure is strange since AK does not mention this transformation in their paper and do not implemented it in the RD estimation. The main motivation of this step is likely to improve the computational efficiency and accuracy of the estimates. I check the robustness of the results without this mean-difference step. The conceptual empirical implications is apparently simple[2], but as shown in this next section this step significantly impacts the DD results.

Another interpretation of the identifying assumptions for equations (1) and (2) is the parallel trends assumption, which requires that the level difference between the treated and baseline group outcomes are constant across time in the absence of treatment. Although the parallel trends cannot be verified in the post-treatment periods due to the unobservable counter-factuals, the pre-treatment trends can be checked to verify that the treated and baseline groups are suitable for a DD estimation. AK plots the residuals of ozone concentrations removing weather shocks, which can serve as an "eyeball test" of parallel pre-trends. I attempt to verify the pre-trend requirement more explicitly by estimating equation (3), a staggered extension of AK's DD specification, and testing whether the pre-treatment coefficients are jointly zero.

$$\ln y_{it} = \sum_{j=-\underline{j}}^{\bar{j}} \boldsymbol{\alpha}_j \cdot \mathbf{Z}_{ct}^j + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + \boldsymbol{\gamma}_r \cdot \mathbf{D}_t + \delta \cdot \mathbf{I}_{ct} \tag{3}$$

$$+ \boldsymbol{\theta} \cdot \mathbf{Trend}_{rct} + \mu_i + \eta_{ry} + \varepsilon_{it},$$

In equation (3), t is the time period in months, $\underline{j}$ and $\bar{j}$ are the number of leads months before treatment and lag months after treatment, $\mathbf{Z}_{ct}^j$ are indicators for $t - j = T_c^0$ in the treated county c, and T_c^0 is the initial treatment month in county c. All variables index by t are monthly panel means of the original daily data. If the parallel trends hold in the pre-treatment periods, we should see that lead period estimates should be statistically zero, i.e. $\alpha_{-\underline{j}} = \alpha_{-\underline{j}+1} = ... = \alpha_{-1} = 0$. In the next section I report the results of the replicated and robustness check DD estimations.

4 Difference-in-Differences Results

Table 2 columns (1) to (4) report my replicated results of the original estimates, which are columns (1) and (5) of the full sample estimates in AK's Table 2 and the estimates of the restricted sample of monitors in AK's Table 4. In columns (5) to (8) I replicate the same results without performing monitor-mean differences to check the sensitivity of the results to this panel-level transformation. Lastly, in columns (9) to (12) I estimate the DD treatment effects without mean differences and correcting the treated neighbor merge error, which adds an additional 60,986 observations in the full sample and 11,989 observations in the restricted sample. This coding correction also tests the robustness of AK's results to sampling.

In columns (1) and (2), I report my replicated results with the full monitor sample. Column (1) uses the simple DD specification from equation (1) and column (2) uses the specification with the full set of weather, time, and geographic controls from equation (2). All estimates in Table 2 are compared against a baseline of 9.0 psi summer RVP limit. This is notable since if the 9.0 psi RVP baseline reduces ozone concentrations, the effect of the other treatments on reducing ozone levels will be underestimated from an untreated outcome (Fricke 2017). Perhaps unsurprisingly, AK finds positive though statistically insignificant effects of RVP Phase I (RVP I) on log ozone, the 9.0 psi RVP baseline is a more stringent limit than the 9.5 and 10.5 psi RVP I treatment. For the 7.8 psi RVP Phase II (RVP II), AK finds negative and insignificant effects. Federal RFG has an estimated effect of -0.029 reduction in log ozone

[2]With the monitor mean-difference transformation, AK actually estimates the equation $\ln(y_{it}) - \bar{y}_i = \boldsymbol{\alpha} \cdot (\mathbf{Treat}_{ct} - \bar{\mathbf{Treat}}_i) + \boldsymbol{\zeta}_{irct} - \bar{\boldsymbol{\zeta}}_{\boldsymbol{i}} + \varepsilon_{it}$, where $\boldsymbol{\zeta}_{irct}$ is the combined covariates. This is the same as estimating the original equations with a monitor-specific constant, $\ln(y_{it}) = \boldsymbol{\alpha} \cdot \mathbf{Treat}_{ct} + \boldsymbol{\zeta}_{irct} + \varepsilon_{it} + K_i$, where $K_i = \bar{y}_i - \boldsymbol{\alpha}\bar{\mathbf{Treat}}_i - \bar{\boldsymbol{\zeta}}_{\boldsymbol{i}}$ is a constant for monitor i across time.

at the 0.05 significance level under the simple DD estimation in column (1). However, it loses statistical significance under the full controls in column (2). CARB gasoline has a comparative effect of between -0.095 and -0.065 in the simple and full control DD specifications, which are respectively significant at the 1% and 5% level.

Columns (3) and (4) estimates the same specifications as in columns (1) and (2) on a restricted sample of monitors with data spanning all 15 years of observation between 1989 to 2003. The results are similar for RVP II and RFG. Point estimates for the effect of CARB on reducing ozone increase to -0.148 and -0.123 for the simple and full control DD models, and both significant at the 1% level. Point estimates for RVP Phase I are still insignificant, though the signs change from positive to negative.

In columns (5) to (8), I repeat the same estimations from columns (1) to (4) without applying the monitor-mean difference. This dramatically reduces the statistical significance of the estimates, though intuition suggests that the monitor-mean difference should only affect the monitor fixed effect estimates. The standard error increases dramatically for all treatment point estimates. The effects for CARB drop close to zero except for the full-covariate restricted-sample estimate in column (8), which is significant at the 5% level with a similar point estimate with its mean-difference counterpart in column (4). Point estimates for RVP I from the simple DD specification in columns (5) and (7) become large but very imprecise. Since the DD results appear highly sensitive to the mean-differencing procedure, AK should have provided their empirical motivation of the mean-difference process.

In columns (9) to (12) I repeat the same estimations as in columns (5) to (8), correcting for AK's potential coding error when eliminating more strictly treated neighbors. This adjustment increases the full sample from 1,144,025 to 1,205,011 observations and the 15-year monitors sample from 455,084 to 467,073 observations. The results are similar to those in columns (5) to (8), which suggests that the results are not sensitive to this particular sample adjustment.

Table 2—Difference-in-Differences: Replication and Robustness Checks Results

	Full sample		All year monitors		Full sample		All year monitors		Full sample		All year monitors	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
RVPI (9.5, 10.5 psi)	0.016	0.005	-0.009	-0.014	-0.743	0.018	-0.927	-0.012	-0.721	0.019	-0.868	0.011
	(0.016)	(0.017)	(0.015)	(0.018)	(0.406)	(0.026)	(0.517)	(0.023)	(0.400)	(0.024)	(0.505)	(0.018)
RVPII (7.8 psi)	-0.007	-0.012	-0.009	-0.022	-0.008	0.015	-0.018	-0.003	0.002	0.014	-0.015	-0.002
	(0.008)	(0.011)	(0.009)	(0.013)	(0.014)	(0.018)	(0.014)	(0.021)	(0.014)	(0.018)	(0.014)	(0.021)
Federal RFG	-0.029**	-0.018	-0.031**	-0.030*	0.042*	-0.031	0.029	-0.048	0.054***	-0.026	0.038*	-0.044
	(0.009)	(0.012)	(0.010)	(0.014)	(0.017)	(0.023)	(0.017)	(0.025)	(0.016)	(0.022)	(0.017)	(0.025)
CARB Gasoline	-0.095***	-0.065**	-0.148***	-0.123***	-0.002	-0.013	0.003	-0.118**	0.004	-0.012	0.005	-0.117**
	(0.013)	(0.020)	(0.014)	(0.025)	(0.023)	(0.034)	(0.023)	(0.040)	(0.023)	(0.033)	(0.022)	(0.040)
Income		-0.299		-0.407		-0.715***		-0.594***		-0.705***		-0.596***
		(0.234)		(0.245)		(0.099)		(0.088)		(0.096)		(0.088)
Monitor FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Full controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Difference	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Corrected Merge	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Observations	1,144,025	1,144,025	455,084	455,084	1,144,025	1,144,025	455,084	455,084	1,205,011	1,205,011	467,073	467,073

Notes: This table shows coefficients estimated from OLS regressions of the indicated dependent variable on log daily maximum ozone concentrations. Standard errors clustered by state-year are in parentheses. All regulatory effects are relative to the 9.0 psi RVP baseline. The sample only contains data from summer months, between June 1st and August 31st. Columns 1 and 2 replicate AK equation (1). Columns 3 and 4 replicate AK equation (2). Columns 5, 6, 7, and 8 replicate AK without monitor-mean differenceing. Columns 9, 10, 11, and 12 replicate AK correcting the treated neighbor merge error and without monitor-mean differencing. Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

I also conduct a robustness check of parallel pre-trends by examining the pre-treatment lead estimates in a staggered DD estimation on monthly summer data. I exclude RVP I from the staggered implementation since there is no data is available before 1989, when RVP I regulations began. The results of the staggered implementation are illustrated in Figure 1 and the full set of estimates are reported in Table A1 of the appendix. The appendix table also reports the F-tests results with null hypotheses of jointly zero pre-treatment estimates.

The first row of Figure 1 shows the period estimates with a window of three months before and after treatment for RVP II, RFG, and CARB. The second and third rows use six-month and twelve-month windows before and after treatment. Point estimates plots are connected with blue lines and the 95% confidence interval bounds[3] are plotted with the red lines.

For RVP II and RFG, the pre-trend confidence intervals contain zero in most cases, except for a few periods in the 12-month window specifications. Other than the 12-month window estimates, the p-values of jointly zero test indicate pre-trend effects cannot be rejected at the five-percent level except in one specification, RFG's 1-month window estimates with full controls in column (8) of Table A1. Ignoring 12-month window estimates, these pre-trends estimates provide some support for the parallel pre-trends requirement for the causal identifying assumptions of RVP II and RFG, though the post-treatment effects are also non-significant.

For CARB, the pre-trends estimate significantly depart from zero in all window specifications. Jointly zero pre-treatment effects can be rejected at less than the one-percent level for all staggered DD estimates. This provides evidence that the pre-trends are not parallel for CARB and its baseline, which casts doubt on the identifying assumption for estimating causal of effects of CARB using a DD approach.

From the post-treatment lag coefficients in Figure 1, effects similar to AK's original DD specification are observed. RVP II and RFG have zero post-treatment coefficients in almost all cases, while CARB has significant negative coefficients. However, since the pre-trends coefficients are not zero for CARB, its RD estimates are likely not appropriate for causal interpretations.

5 Regression Discontinuity (RD) Designs

AK also implement a regression discontinuity design to allow for a more flexible model and to capture the spatial heterogeneity of the regulatory effects. They estimate the regression

$$\ln(y_{it}) = \boldsymbol{\alpha}_i \cdot \mathbf{Treat}_{ct} + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + f_i(Date_i) + \mu_i + \varepsilon_{it}, \tag{4}$$

for each monitor i where more than 75% of days within the monitor's date range is observed across dates t, in seasons where more than 75% of days are observed. The vector $\mathbf{Treat}_{ct}$ is derived from the treatment indicators, using a 30 day linear phase-in period where each treatment variable increases from 0 to 1 within the 30 day period leading up to the treatment. The treatment variables in $\mathbf{Treat}_{ct}$ are RVP II, RFG, CARB, and combined RVP II and RFG. RVP I is not included in the RD implementation because there are no observations in the data before the treatment.

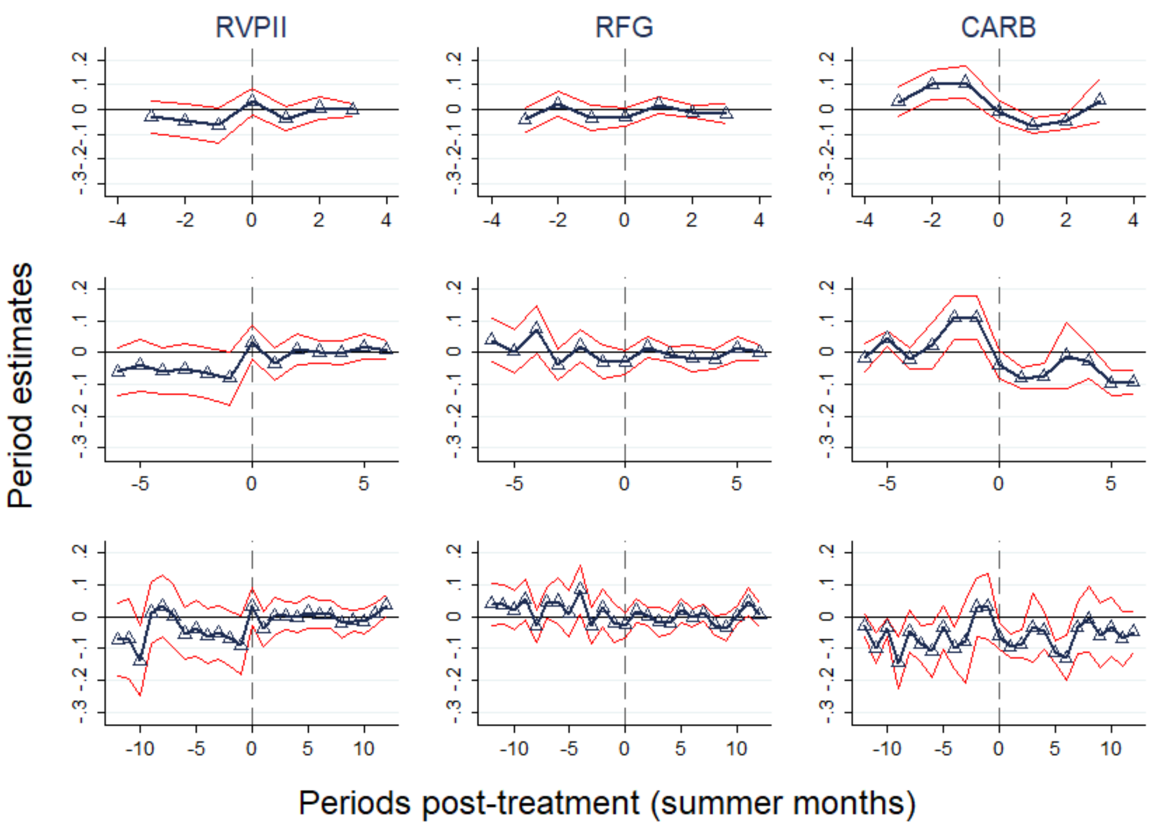
Ak uses a similar set of W_{it} weather and time covariates from the DD implementation. Region covariates are not used since the model is estimated for each monitor, and there are no geographic variation within each monitor's observations.

AK models the running variable $Date_i$ with an eighth order Chebychev polynomial. The n-th order polynomial terms $P_n(Date_i)$ are generated with a recurrence relation, where $T_0(x) = 0, T_1(x) = x$, and $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$ for $n \geq 2$.

AK's RD implementation is highly complex due to the numerousness of monitor-specific estimates, the treatment linear phase-in, and the use of the Chebychev polynomial. The monitor-specific approach makes general interpretations of the treatment effects difficult since AK suggests no method of aggregating the estimates for each treatment other than taking the simple mean of the point estimates, which ignore the estimates' standard errors. The linear phase-in treatment conflicts with the main motivation

[3] The 95% confidence intervals are calculated by multiplying 1.96 with the standard errors, and adding and subtracting from the point estimates.

Figure 1—Staggered Difference-in-Differences Estimates and Trends



Notes: The first row contains 3-month leads and lag estimates for all treatments from the staggered DD equation (3). The second and third row contain the 6 and 12-month estimates, respectively. The 95% confidence intervals are plotted with red lines. Point estimates and F-test results are reported in appendix Table A1.

of estimating RD regressions, which is to exploit discontinuities in outcomes around cutoffs in treatment running variables. The use of the high-ordered global polynomial model also raises issues of overfitting (Hausman and Rapson, 2018).

My replication of the AK's RD results are reported in the next section. To check the robustness of AK's complex RD implementation, I also propose and estimate a more standard RD implementation with equation (5), which groups the monitors by treatment.

$$\ln(y_{it}) = \boldsymbol{\gamma} \cdot \text{Treat}_{ct} + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + \delta_1 t + \varepsilon_{it}. \tag{5}$$

In equation (5), t is the normalized daily post-treatment period where $t = 0$ is the time of treatment, and Treat_{ct} is a scalar treatment variable with a 30-day linear phase-in. The phase-in is necessary to create any significant estimates with daily data, but I discuss the validity of this transformation in the results section.

I estimate this treatment-specific RD model with a three-month bandwidth before and after treatment to estimate the local discontinuity effect at the regulation start dates. AK's RD global polynomial estimates use the full 15-year sample. This large bandwidth combined with the high-ordered global polynomial estimation likely allow AK's RD specification to overfit the data, producing non-meaningful estimates (Hausman and Rapson, 2018).

I also estimate the RD with first to third-order polynomials of t allowing for differential terms before and after treatment. The regression equation for each of the n-th order polynomial specifications are given by

$$\ln(y_{it}) = \boldsymbol{\gamma} \cdot \text{Treat}_{ct} + \boldsymbol{\beta} \cdot \mathbf{W}_{it} + \sum_{m=1}^{n} \left(\delta_j t^m + \sigma_j \text{Treat}_{ct} t^m \right) + \varepsilon_{it}. \tag{6}$$

Though this specification may not be optimal for this setting, I believe RD estimates on each treatment group instead of on each monitor allow clearer interpretations of the RD treatment effects and their validity.

6 Regression Discontinuity Results

Tables A2 to A5 in the appendix report the replicated results of AK's RD specification for each monitor. Figure 2 shows the frequency distribution and Epanechnikov kernel-smoothed distribution of the replicated monitor-specific RD estimates with the full sample of locations. The mean effects for RVP II, RFG, and CARB on log ozone are $-0.0032, -0.453$, and -0.0608 respectively, compared to AK's reported mean RD treatment effects of $-0.0001, -0.021$, and -0.060.

The mean results for RFG and CARB deviate significantly from AK's DD estimates. AK argue that the monitor-specific RVP results are confounded by NO_x regulations and that the CARB results are spatially heterogeneous. Though AK's monitors-specific findings provide insight about the heterogeneity of the treatment effects, the large number of estimates from Tables A2 to A5 are difficult to interpret. To my knowledge, the means of the point estimates are not readily useful in interpreting the causal effect of the treatments. To report the estimate for any specific monitor as a representative of the entire treatment group is also unconvincing.

To evaluate the robustness of the RD approach in this setting, I also estimate a more traditional treatment-grouped RD specification according to equations (5) and (6) with first, second, and third polynomial orders. The results are reported in Table 3.

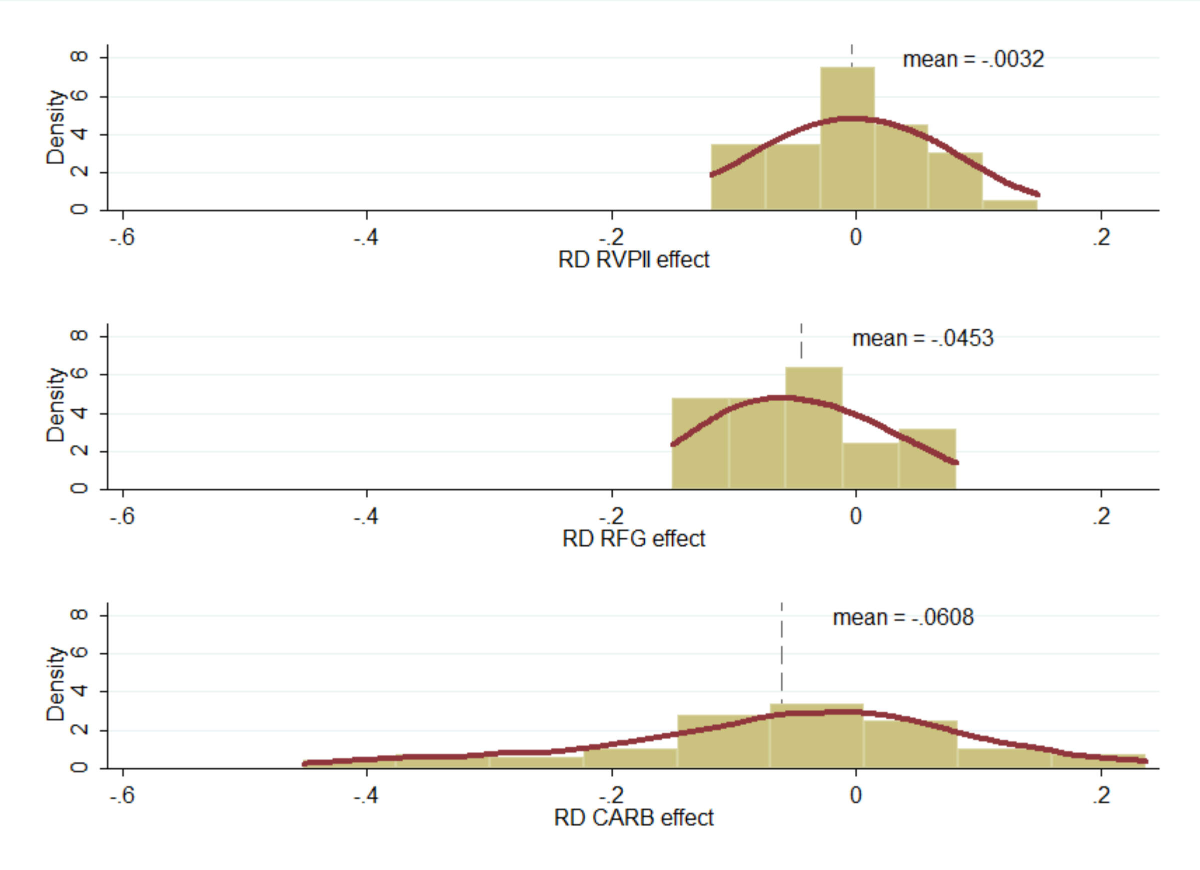
Columns (1) to (4) of Table 3 report the estimated RD treatment effects for all monitors in counties subject to RVP II in the data. The estimated treatments lose statistical power compared to AK's DD estimates. The RVP treatment effects vary significantly with changes in polynomial order. Estimated RVP effects on log ozone range from -0.045 to -0.137, and are not statistically significant at the 10% level.

Columns (5) to (8) report the RFG RD treatment effects, which range from -0.128 to -0.142. The simple linear-trend results in column (5) estimated from equation (5) is significant at the 5% level. The differential polynomial treatments estimates with equation (6) specifications are significant at the 10% level.

Columns (9) to (12) report the CARB treatment effects, which differ dramatically from AK's mean monitor-specific results. The point estimates range from 0.395 to 0.410 and are significant at the 5% level. These are quite large effects in the opposite sign the mean of AK's monitor-specific results.

The treatment-grouped specification from equations (5) and (6) adhere more closely to the standard RD strategy than AK's parameterization in (4). The standard specification results are less likely to be overfitted and perhaps are conceptually clearer. Although the standard RD specifications likely do not optimize the statistical power of the estimations, the non-significant results may indicate that a discontinuity approach is inappropriate to estimate regulatory effects on air quality. Since time and cost intensive planning and investments are likely required to adhere to new regulations, refiners likely do not adopt to new regulations sharply at the cutoff.

Figure 2—Monitor Regression Discontinuity Kernel Density



Notes: The bars plots show the frequency of the estimated treatment effects from equation (4). The red lines show the smoothed cross-monitor distribution of the estimates using a Epanechnikov kernel with a bandwidth of 0.05.

Table 3—Regression Discontinuity Results: Replication and Robustness Checks

	RVP Phase II				RFG Federal				CARB Gasoline			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Estimate	-0.045	-0.106	-0.106	-0.137	-0.140**	-0.130*	-0.128*	-0.142*	0.395**	0.457*	0.438*	0.410*
	(0.128)	(0.124)	(0.127)	(0.120)	(0.020)	(0.036)	(0.038)	(0.042)	(0.056)	(0.107)	(0.110)	(0.112)
Polynomial Order	1	1	2	3	1	1	2	3	1	1	2	3
Differential Slope	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,863	7,863	7,863	7,863	4,766	4,766	4,766	4,766	16,053	16,053	16,053	16,053

Notes: This table shows the coefficients of OLS regressions of the indicated regression discontinuity treatment effect on log ozone. The first rows under each treatment header are coefficients estimated from equation (5), and the following coefficients are estimated from equation (6). Standard errors are clustered on year-season. Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

7 Conclusion

In this paper, I replicate and conduct robustness checks on AK's DD and RD estimates of the effect of gasoline-emissions regulation on reducing ground-level ozone concentrations. The specific regulations examined are federal Reid vapor pressure (RVP) standards, federal reformulated gasoline (RFG), and California reformulated gasoline (CARB).

I find that though DD estimates of CARB effects are significant, the identifying assumption required for causal interpretations is likely not satisfied. AK point to the spatial heterogeneity of treatment effects in motivating the RD implementation, but this exact heterogeneity undermines the common-trends assumption for the DD estimations. An alternative estimation strategy could be to create a synthetic control, not relying any specific baseline monitors to provide a valid parallel trend.

The use of the 9.0 psi RVP baseline in the estimations is also problematic, since all treatment effects are estimated relative to already treated units. AK's conclusions that RVP I and II regulations are ineffective may be misleading. It is quite possible that RVP I and II are effective at reducing ozone concentrations, but that the treat effects are not statistically different from that of the 9.0 psi RVP baseline.

I also find that AK's complex RD implementation is likely not suitable in this setting. Best practices for RD strategy warn that such high-ordered parameterizations and large bandwidths allow the model to overfit the data instead of extracting meaningful estimates.

The use of linear phase-in also complicates RD interpretation. Without this transformation, the discontinuous effects at the daily cut-offs are too small to be estimated. Standard RD specifications are perhaps also not suitable for these estimations since meaningful discontinuities likely do not exist around the cutoff date. AK's monitor-specific approach also obscures the general interpretation of the results, and should probably estimate more general effects of the treatments.

While AK's arguments about regulatory effects on air quality are conceptually credible, my examination of the internal validity of their empirical results raise considerable issues. More advancement in econometrics techniques and better meteorological modeling are likely needed to address the estimation challenges posed by large spatial heterogeneity and complex weather interactions.

References

- ▷ **Auffhammer, Maximilian, and Ryan Kellogg.** (2011). "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality." *American Economic Review 101, no. 6 (October 2011): 2687–2722.* http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.6.2687

- ▷ **Fricke, Hans.** (2017). Identification based on Difference-in-Differences Approaches with Multiple Treatments. *Oxford Bulletin of Economics and Statistics.* http://onlinelibrary.wiley.com/doi/10.1111

- ▷ **Hausman, Catherine and Rapson, David S.** (2018) Regression Discontinuity in Time: Considerations for Empirical Applications. *Annual Review of Resource Economics 10, no. 1 (October 2018): 533-552.* https://doi.org/10.1146/annurev-resource-121517-033306

Appendix

Table A1—Staggered Difference-in-Differences Estimates

	RVP Phase II						RFG Federal						CARB Gasoline					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
$\alpha_{\hat{t}=-12}$					-0.082	-0.073					-0.071	0.040					-0.080***	-0.028
					(0.072)	(0.058)					(0.099)	(0.034)					(0.006)	(0.017)
$\alpha_{\hat{t}=-11}$					-0.093	-0.069					-0.031	0.038					-0.039**	-0.099**
					(0.057)	(0.063)					(0.095)	(0.031)					(0.012)	(0.024)
$\alpha_{\hat{t}=-10}$					-0.184**	-0.138*					-0.046	0.021					0.080***	-0.035*
					(0.062)	(0.056)					(0.096)	(0.031)					(0.003)	(0.015)
$\alpha_{\hat{t}=-9}$					-0.052	0.012					-0.096	0.055					-0.091**	-0.145**
					(0.045)	(0.049)					(0.145)	(0.032)					(0.029)	(0.041)
$\alpha_{\hat{t}=-8}$					0.027	0.033					-0.154	-0.029					-0.116**	-0.045
					(0.050)	(0.049)					(0.139)	(0.026)					(0.032)	(0.035)
$\alpha_{\hat{t}=-7}$					-0.014	0.001					-0.091	0.044					-0.086*	-0.085*
					(0.047)	(0.051)					(0.143)	(0.025)					(0.035)	(0.030)
$\alpha_{\hat{t}=-6}$			-0.033	-0.060	-0.039	-0.054			-0.007	0.038	-0.038	0.048			-0.040	-0.017	-0.116**	-0.108*
			(0.035)	(0.038)	(0.041)	(0.042)			(0.038)	(0.035)	(0.058)	(0.037)			(0.066)	(0.023)	(0.033)	(0.043)
$\alpha_{\hat{t}=-5}$			-0.022	-0.040	-0.028	-0.038			-0.006	0.003	-0.037	0.009			-0.068	0.045**	-0.143***	-0.033
			(0.041)	(0.041)	(0.048)	(0.045)			(0.046)	(0.035)	(0.064)	(0.037)			(0.066)	(0.013)	(0.033)	(0.035)
$\alpha_{\hat{t}=-4}$			-0.029	-0.059	-0.035	-0.062			0.057	0.073	0.026	0.083*			-0.037	-0.020	-0.113**	-0.098*
			(0.043)	(0.038)	(0.049)	(0.044)			(0.042)	(0.038)	(0.060)	(0.040)			(0.066)	(0.017)	(0.033)	(0.033)
$\alpha_{\hat{t}=-3}$	-0.059	-0.031	-0.068	-0.051	-0.075	-0.052	0.013	-0.041	0.014	-0.039	-0.006	-0.027	-0.180***	0.031	-0.218**	0.022	-0.299***	-0.078
	(0.037)	(0.033)	(0.040)	(0.040)	(0.043)	(0.043)	(0.048)	(0.024)	(0.049)	(0.025)	(0.059)	(0.029)	(0.008)	(0.031)	(0.057)	(0.037)	(0.028)	(0.066)
$\alpha_{\hat{t}=-2}$	-0.018	-0.046	-0.027	-0.065	-0.034	-0.069	0.024	0.022	0.025	0.021	0.005	0.029	-0.079***	0.099**	-0.118	0.110**	-0.198***	0.030
	(0.041)	(0.035)	(0.045)	(0.041)	(0.047)	(0.043)	(0.051)	(0.025)	(0.051)	(0.026)	(0.062)	(0.030)	(0.006)	(0.031)	(0.056)	(0.034)	(0.026)	(0.047)
$\alpha_{\hat{t}=-1}$	-0.023	-0.063	-0.032	-0.081	-0.039	-0.091	-0.017	-0.036	-0.017	-0.029	-0.036	-0.019	-0.055***	0.108**	-0.094	0.108**	-0.174***	0.031
	(0.037)	(0.036)	(0.042)	(0.043)	(0.045)	(0.047)	(0.054)	(0.026)	(0.055)	(0.027)	(0.065)	(0.031)	(0.006)	(0.033)	(0.056)	(0.035)	(0.026)	(0.053)
$\alpha_{t=0}$	-0.020	0.031	-0.011	0.032	-0.011	0.033	-0.061**	-0.032	-0.061**	-0.032	-0.074*	-0.027	-0.071	-0.010	-0.079**	-0.039	-0.066	-0.061*
	(0.035)	(0.027)	(0.035)	(0.028)	(0.037)	(0.031)	(0.023)	(0.019)	(0.023)	(0.019)	(0.032)	(0.020)	(0.038)	(0.022)	(0.026)	(0.022)	(0.034)	(0.021)
$\alpha_{t=1}$	-0.014	-0.037	-0.006	-0.036	-0.005	-0.038	0.034	0.019	0.034	0.016	0.022	0.018	-0.061	-0.066***	-0.070*	-0.080***	-0.056	-0.093***
	(0.025)	(0.024)	(0.024)	(0.025)	(0.026)	(0.028)	(0.023)	(0.017)	(0.023)	(0.017)	(0.032)	(0.019)	(0.040)	(0.015)	(0.025)	(0.016)	(0.035)	(0.019)
$\alpha_{t=2}$	0.068*	0.007	0.076**	0.009	0.077**	0.002	-0.009	-0.011	-0.009	-0.006	-0.021	0.001	-0.035	-0.047**	-0.044	-0.074**	-0.031	-0.085**
	(0.027)	(0.023)	(0.027)	(0.025)	(0.029)	(0.030)	(0.024)	(0.013)	(0.024)	(0.013)	(0.033)	(0.015)	(0.042)	(0.015)	(0.027)	(0.021)	(0.036)	(0.022)
$\alpha_{t=3}$	-0.015	-0.001	0.005	0.002	0.007	0.003	-0.015	-0.019	-0.013	-0.020	-0.024	-0.017	-0.064**	0.036	-0.073	-0.010	-0.054	-0.034
	(0.017)	(0.013)	(0.027)	(0.018)	(0.028)	(0.023)	(0.015)	(0.020)	(0.022)	(0.021)	(0.030)	(0.023)	(0.022)	(0.044)	(0.060)	(0.054)	(0.038)	(0.055)
$\alpha_{t=4}$			0.016	-0.001	0.018	-0.003			-0.015	-0.020	-0.027	-0.021			0.009	-0.028	0.028	-0.044
			(0.023)	(0.020)	(0.025)	(0.024)			(0.023)	(0.016)	(0.031)	(0.017)			(0.052)	(0.027)	(0.031)	(0.030)
$\alpha_{t=5}$			0.058*	0.019	0.060	0.013			0.021	0.013	0.009	0.020			-0.012	-0.096***	0.007	-0.112***
			(0.029)	(0.020)	(0.031)	(0.026)			(0.024)	(0.019)	(0.032)	(0.019)			(0.043)	(0.021)	(0.024)	(0.018)
$\alpha_{t=6}$			0.019	0.007	0.026	0.008			-0.011	-0.001	-0.016	-0.003			-0.171*	-0.093***	-0.137**	-0.129**
			(0.013)	(0.015)	(0.021)	(0.022)			(0.023)	(0.013)	(0.026)	(0.015)			(0.061)	(0.019)	(0.045)	(0.035)
$\alpha_{t=7}$					-0.004	0.007					0.034	0.011					0.018	-0.034
					(0.025)	(0.022)					(0.022)	(0.013)					(0.029)	(0.042)
$\alpha_{t=8}$					0.035	-0.020					-0.043	-0.026					0.098*	-0.007
					(0.022)	(0.023)					(0.028)	(0.016)					(0.042)	(0.053)
$\alpha_{t=9}$					-0.037*	-0.013					-0.093**	-0.035					-0.016	-0.058
					(0.019)	(0.016)					(0.032)	(0.022)					(0.033)	(0.051)
$\alpha_{t=10}$					0.015	-0.015					0.047*	0.003					-0.002	-0.034
					(0.024)	(0.020)					(0.021)	(0.015)					(0.042)	(0.047)
$\alpha_{t=11}$					0.041*	0.009					0.060	0.047*					-0.034	-0.068
					(0.019)	(0.018)					(0.031)	(0.023)					(0.035)	(0.044)
$\alpha_{t=12}$					0.003	0.036*					-0.011	0.007					0.027*	-0.048
					(0.015)	(0.017)					(0.028)	(0.019)					(0.010)	(0.033)
income		-0.125		-0.125		-0.122		-0.292***		-0.294***		-0.299***		0.090		0.092		0.091
		(0.181)		(0.180)		(0.179)		(0.072)		(0.073)		(0.073)		(0.203)		(0.204)		(0.203)
Window	±3	±3	±6	±6	±12	±12	±3	±3	±6	±6	±12	±12	±3	±3	±6	±6	±12	±12
Monitor FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Full controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean differenced	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Corrected merge	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	2.4105	1.3001	1.5102	1.2153	4.8075	1.5771	.6023000000000001	2.7839	2.0891	2.1038	2.6686	2.5814	255.1545	5.444500000000001	455.6928	9.8362	348.4603	106.9087
P-value	.0668	.2743	.1738	.2977	0	.0963	.614	.0411	.0544	.0527	.002	.0028	0	.0108	0	.0002	0	0
Observations	16,706	16,706	16,706	16,706	16,706	16,706	10,732	10,732	10,732	10,732	10,732	10,732	7,152	7,152	7,152	7,152	7,152	7,152

Notes: Effects shown are the estimated coefficients on the treatment dummies of the staggered difference-in-difference specification (3). Standard errors are clustered on state-year. Estimated effects are relative to the baseline of a 9.0 psi RVP standard. Sample uses monthly data from summer months June 1 to August 31. Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

Table A2—Monitor-Specific Regression Discontinuity Estimates for RVP II

Index	State	County	Monitor ID	RD RVPII effect	RVPII standard error	RVPII t-score	RVPII p-value
1	CO	Adams	3001	.057**	.028	2.08	.042
2	CO	Arapahoe	2	.1***	.031	3.252	.002
3	CO	Denver	14	.078*	.041	1.931	.058
4	CO	Jefferson	2	.043	.032	1.33	.188
5	FL	Broward	2003	-.034	.031	-1.078	.286
6	FL	Broward	8002	-.111***	.032	-3.462	.001
7	FL	Duval	77	.018	.026	.712	.479
8	FL	Hillsborough	81	.044	.035	1.27	.209
9	FL	Hillsborough	1035	.008	.023	.328	.744
10	FL	Miami-Dade	21	-.118	.03	-3.956	
11	FL	Miami-Dade	27	-.079***	.027	-2.903	.005
12	FL	Miami-Dade	29	.015	.035	.438	.663
13	FL	Miami-Dade	30	-.102**	.044	-2.318	.024
14	FL	Palm Beach	2004	-.115	.111	-1.043	.301
15	FL	Pinellas	4	-.019	.027	-.692	.492
16	FL	Pinellas	18	-.044	.038	-1.148	.256
17	FL	Pinellas	5002	-.07**	.034	-2.023	.048
18	IL	Madison	8	.022	.032	.702	.486
19	IL	Madison	1009	.013	.036	.36	.72
20	IL	Madison	2007	.025	.028	.901	.371
21	IL	Madison	3007	.003	.027	.101	.92
22	IL	Saint Clair	10	.069**	.032	2.129	.037
23	LA	Beauregard	2	.006	.054	.116	.908
24	LA	Calcasieu	2	-.015	.056	-.267	.79
25	LA	East Baton Rouge	3	.014	.039	.362	.719
26	LA	East Baton Rouge	1001	-.017	.036	-.462	.646
27	LA	Grant	1	.003	.035	.099	.922
28	LA	Jefferson	1001	.149***	.042	3.55	.001
29	LA	Orleans	12	-.066	.047	-1.401	.166
30	LA	Pointe Coupee	1	-.029	.032	-.931	.356
31	LA	St. Bernard	2	-.086*	.048	-1.799	.077
32	LA	St. James	2	-.049	.068	-.719	.475
33	LA	St. Mary	3	.047	.053	.884	.38
34	LA	West Baton Rouge	1	-.06	.039	-1.55	.126
35	NV	Washoe	1005	.033	.032	1.041	.302
36	TN	Davidson	11	.064	.045	1.421	.161
37	TN	Davidson	26	-.027	.05	-.545	.588
38	TX	Bexar	32	-.015	.057	-.259	.796
39	TX	El Paso	37	.087***	.033	2.676	.01
40	TX	Gregg	1	.014	.031	.444	.658
41	TX	Jefferson	9	-.099	.066	-1.497	.14
42	TX	Jefferson	11	.084	.065	1.291	.202
43	TX	Nueces	25	-.013	.045	-.275	.784
44	TX	Nueces	26	-.027	.053	-.515	.609
45	TX	Travis	14	.05	.037	1.348	.183

Notes: Effects shown are the monitor-specific estimated coefficients on the treatment dummies of the regression discontinuity specification (4). Standard errors are clustered on year-season. Estimated effects RVP phase II (less than or equal to 7.8 psi) are relative to the baseline of a 9.0 psi RVP standard. Sample uses data from all seasons of 1989-2003.
Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

Table A3—Monitor-Specific Regression Discontinuity Estimates for RFG

Index	State	County	Monitor ID	RD RFG effect	RFG standard error	RFG t-score	RFG p-value
1	DE	New Castle	1003	-.118**	.059	-2.012	.049
2	IL	Cook	50	-.006	.05	-.125	.901
3	IL	Cook	64	-.063	.047	-1.345	.184
4	IL	Cook	7002	.066**	.03	2.241	.029
5	IL	DuPage	6001	.03	.027	1.118	.268
6	IL	Kane	5	.082***	.029	2.853	.006
7	IL	Lake	1002	-.008	.037	-.222	.825
8	IL	Lake	3001	-.02	.036	-.543	.59
9	IL	Will	1008	.07**	.03	2.333	.023
10	MA	Hampden	8	-.026	.042	-.624	.535
11	MA	Hampshire	4002	-.035	.027	-1.311	.195
12	NJ	Atlantic	5	-.138	.026	-5.396	
13	NJ	Camden	3	-.048	.034	-1.412	.163
14	NJ	Camden	1001	-.15	.031	-4.865	
15	NJ	Cumberland	7	-.119	.025	-4.703	
16	NJ	Gloucester	2	-.105***	.029	-3.557	.001
17	NJ	Hudson	6	-.032	.029	-1.097	.278
18	NJ	Hunterdon	1	-.034	.033	-1.014	.315
19	NJ	Mercer	5	-.112***	.041	-2.714	.009
20	NJ	Monmouth	5	-.057*	.029	-1.979	.053
21	NJ	Morris	3001	-.075**	.033	-2.236	.029
22	NY	Dutchess	7	-.038	.03	-1.262	.212
23	NY	Essex	2	-.083***	.023	-3.635	.001
24	NY	Essex	3	-.073***	.023	-3.14	.003
25	NY	New York	63	-.094*	.049	-1.917	.06
26	NY	Suffolk	2	-.081**	.032	-2.529	.014
27	PA	Philadelphia	24	.043	.028	1.5	.139

Notes: Effects shown are the monitor-specific estimated coefficients on the treatment dummies of the regression discontinuity specification (4). Standard errors are clustered on year-season. Estimated effects RFG are relative to the baseline of a 9.0 psi RVP standard. Sample uses data from all seasons of 1989-2003.
Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

Table A4—Monitor-Specific Regression Discontinuity Estimates for CARB

Index	State	County	Monitor ID	RD RVPII effect	RD RFG effect	RD CARB effect	RVPII standard error	RFG standard error	CARB standard error	RVPII t-score	RFG t-score	CARB t-score	RVPII p-value	RFG p-value	CARB p-value
1	CA	Alameda	3	.11	.	.136**	.028		.054	3.913		2.506			.015
2	CA	Alameda	5	.023	.	.019	.053		.075	.444		.248	.659		.805
3	CA	Alameda	1001	.029	.	.067	.033		.063	.876		1.055	.385		.296
4	CA	Butte	2	-.024	.	.078	.039		.097	-.62		.796	.538		.43
5	CA	Contra Costa	2	.004	.	.092*	.033		.047	.126		1.959	.9		.055
6	CA	Contra Costa	1002	.039	.	.061	.031		.06	1.278		1.016	.206		.314
7	CA	Contra Costa	3001	.126***	.	.237	.037		.044	3.385		5.4	.001		
8	CA	Fresno	7	.107***	.	.193***	.033		.066	3.206		2.934	.002		.005
9	CA	Fresno	4001	.083*	.	.19**	.048		.091	1.737		2.085	.088		.041
10	CA	Kern	7	-.068	.	-.116**	.057		.058	-1.191		-2.021	.238		.048
11	CA	Kern	8	.019	.	.003	.032		.059	.602		.052	.549		.959
12	CA	Kern	232	.098**	.	.067	.047		.063	2.078		1.065	.042		.291
13	CA	Kern	5001	.055**	.	.102	.026		.062	2.122		1.631	.038		.108
14	CA	Kern	6001	.008	.	.061	.034		.065	.248		.933	.805		.355
15	CA	Lake	3001	-.138	.	.006	.086		.092	-1.613		.061	.112		.952
16	CA	Los Angeles	2	.216	.157***	-.146*	.051	.045	.083	4.195	3.465	-1.76		.001	.084
17	CA	Los Angeles	16	.226	.144***	-.129	.044	.042	.088	5.111	3.404	-1.473		.001	.146
18	CA	Los Angeles	113	.031	-.032	-.326***	.049	.07	.111	.635	-.449	-2.938	.528	.655	.005
19	CA	Los Angeles	1002	.188***	.23***	-.327**	.063	.064	.133	2.995	3.567	-2.463	.004	.001	.017
20	CA	Los Angeles	1103	.255***	.051	-.452	.073	.07	.117	3.481	.722	-3.856	.001	.473	
21	CA	Los Angeles	1201	.054	-.032	-.108	.053	.063	.109	1.033	-.517	-.995	.306	.607	.324
22	CA	Los Angeles	1301	.077	.043	-.017	.059	.083	.154	1.3	.511	-.111	.199	.611	.912
23	CA	Los Angeles	1601	.09*	-.055	-.379	.046	.064	.087	1.97	-.858	-4.373	.054	.394	
24	CA	Los Angeles	1701	.126***	.099*	-.437***	.045	.058	.127	2.808	1.72	-3.435	.007	.091	.001
25	CA	Los Angeles	2005	.112**	.094*	-.277	.05	.053	.072	2.228	1.773	-3.838	.03	.081	
26	CA	Los Angeles	4002	.033	.125	-.091	.056	.083	.128	.595	1.493	-.709	.554	.141	.481
27	CA	Los Angeles	5001	.07*	-.024	-.124	.04	.053	.079	1.759	-.443	-1.566	.084	.659	.123
28	CA	Los Angeles	6002	.011	-.194***	-.285	.048	.054	.064	.232	-3.618	-4.476	.818	.001	
29	CA	Marin	1	.081**	.	.174***	.037		.059	2.226		2.928	.03		.005
30	CA	Monterey	2	.001	.	-.011	.025		.044	.049		-.244	.961		.808
31	CA	Napa	3	.015	.	0	.035		.054	.427		-.004	.671		.997
32	CA	Orange	1	.122**	.176***	-.07	.06	.066	.118	2.039	2.665	-.598	.046	.01	.552
33	CA	Orange	2001	.002	.079	.125	.044	.048	.078	.046	1.651	1.61	.963	.104	.113
34	CA	Orange	5001	.088*	.018	-.211**	.05	.049	.081	1.769	.361	-2.612	.082	.719	.011
35	CA	Riverside	2002	.04	0	-.366	.032	.031	.091	1.246	-.001	-4.002	.218	.999	
36	CA	Riverside	5001	.058**	.04	.037	.027	.037	.037	2.14	1.085	.986	.036	.282	.328
37	CA	Riverside	6001	.126***	.041	.01	.043	.036	.073	2.958	1.113	.131	.004	.27	.896
38	CA	Riverside	8001	.147***	-.045	-.213	.042	.035	.046	3.458	-1.279	-4.658	.001	.206	
39	CA	Riverside	9001	.071	.026	-.027	.054	.046	.068	1.313	.567	-.401	.194	.573	.69
40	CA	Sacramento	2	.069	.	.02	.066		.085	1.049		.237	.298		.814
41	CA	Sacramento	6	.102**	.	-.193**	.047		.092	2.193		-2.107	.032		.039
42	CA	Sacramento	10	.026	.	-.064	.048		.084	.537		-.767	.594		.446
43	CA	San Benito	2	.051**	.	.07*	.023		.037	2.157		1.886	.035		.064
44	CA	San Benito	3	.028	.	-.007	.026		.04	1.071		-.171	.288		.865
45	CA	San Bernardino	1	.022	-.047	-.013	.049	.054	.083	.444	-.854	-.151	.659	.397	.88
46	CA	San Bernardino	5	.162***	.113**	-.082	.047	.045	.064	3.476	2.509	-1.281	.001	.015	.205
47	CA	San Bernardino	12	.1**	.023	-.046	.048	.053	.048	2.085	.437	-.959	.041	.664	.341
48	CA	San Bernardino	1004	.331	.2***	-.314**	.058	.071	.134	5.684	2.828	-2.35		.006	.022
49	CA	San Bernardino	2002	.173	.044	-.206***	.037	.046	.063	4.614	.961	-3.291		.34	.002
50	CA	San Bernardino	4003	.1**	.049	-.074	.041	.039	.06	2.44	1.251	-1.227	.018	.216	.225
51	CA	San Bernardino	9004	.156	.061*	-.024	.035	.031	.03	4.478	1.971	-.81		.053	.421
52	CA	San Diego	1	.028	-.072	-.23***	.032	.055	.078	.87	-1.316	-2.941	.388	.193	.005
53	CA	San Diego	3	.063**	.087**	-.138**	.025	.039	.056	2.533	2.206	-2.482	.014	.031	.016
54	CA	San Diego	5	.048	.002	-.231***	.034	.062	.08	1.389	.033	-2.873	.17	.974	.006
55	CA	San Diego	6	-.014	-.065	-.07	.031	.061	.076	-.446	-1.069	-.917	.657	.29	.363
56	CA	San Diego	1001	-.022	-.037	-.213***	.026	.054	.065	-.823	-.691	-3.294	.414	.492	.002
57	CA	San Diego	1002	.023	-.026	-.13***	.026	.047	.048	.897	-.564	-2.719	.373	.575	.009
58	CA	San Diego	1006	.071***	.051	-.122*	.023	.037	.063	3.142	1.392	-1.941	.003	.169	.057
59	CA	San Diego	1007	-.07**	-.096	-.103	.033	.062	.067	-2.103	-1.548	-1.533	.04	.127	.13
60	CA	San Francisco	5	.053	.	-.017	.053		.079	.992		-.211	.325		.833
61	CA	San Joaquin	1002	.001	.	.063	.038		.08	.026		.78	.98		.439
62	CA	San Luis Obispo	2001	.046**	.	-.054	.022		.05	2.096		-1.083	.04		.283
63	CA	San Luis Obispo	2002	.089*	.	-.144*	.046		.075	1.948		-1.929	.056		.059
64	CA	San Luis Obispo	3001	.03	.	.043	.022		.04	1.398		1.076	.167		.286
65	CA	San Luis Obispo	8001	.048**	.	-.022	.02		.042	2.367		-.524	.021		.602
66	CA	San Mateo	1001	.014	.	-.15	.046		.107	.298		-1.402	.767		.166
67	CA	Santa Barbara	8	-.01	.	-.046	.024		.046	-.435		-1.004	.665		.32
68	CA	Santa Barbara	10	.024	.	-.201**	.033		.08	.736		-2.514	.465		.015
69	CA	Santa Barbara	1013	-.023	.	-.02	.025		.048	-.952		-.427	.345		.671
70	CA	Santa Barbara	1014	-.004	.	-.03	.03		.064	-.125		-.465	.901		.644
71	CA	Santa Barbara	1018	.025	.	.059	.02		.053	1.246		1.1	.218		.276
72	CA	Santa Barbara	1021	-.007	.	-.033	.023		.044	-.323		-.766	.748		.447
73	CA	Santa Barbara	1025	.047**	.	-.08	.02		.057	2.306		-1.416	.025		.162
74	CA	Santa Barbara	2004	-.005	.	-.138***	.023		.047	-.203		-2.942	.84		.005
75	CA	Santa Barbara	3001	.017	.	-.06	.019		.046	.876		-1.309	.384		.196
76	CA	Santa Barbara	4003	-.03	.	.067	.034		.083	-.894		.805	.375		.424
77	CA	Santa Clara	4	.158***	.	.034	.052		.082	3.024		.42	.004		.676
78	CA	Santa Clara	1001	.073*	.	.017	.041		.08	1.779		.217	.08		.829
79	CA	Santa Cruz	3	.011	.	.01	.025		.047	.444		.223	.658		.824
80	CA	Shasta	3003	.008	.	-.046	.021		.044	.379		-1.057	.706		.295
81	CA	Solano	4	.036	.	.144***	.036		.051	.983		2.844	.33		.006
82	CA	Sonoma	3	.054	.	.088	.033		.058	1.604		1.525	.114		.133
83	CA	Stanislaus	5	.062	.	.101*	.038		.057	1.609		1.776	.113		.081
84	CA	Tulare	6	.089***	.	.209***	.029		.064	3.088		3.285	.003		.002
85	CA	Tulare	2002	.017	.	.006	.037		.061	.455		.097	.651		.923
86	CA	Ventura	4	-.063	-.038	-.048	.043	.047	.063	-1.462	-.805	-.764	.149	.424	.448
87	CA	Ventura	5	.059**	-.014	-.141***	.028	.028	.046	2.082	-.482	-3.045	.042	.631	.003
88	CA	Ventura	2002	.072	-.012	-.354	.053	.049	.086	1.368	-.241	-4.116	.176	.811	
89	CA	Ventura	2003	-.067**	-.05	-.006	.033	.045	.073	-2.029	-1.116	-.077	.047	.269	.939
90	CA	Ventura	3001	-.046	-.09	-.099	.033	.057	.1	-1.398	-1.566	-.988	.167	.123	.327

Notes: Effects shown are the monitor-specific estimated coefficients on the treatment dummies of the regression discontinuity specification (4). Standard errors are clustered on year-season. Estimated effects CARB are relative to the baseline of a 9.0 psi RVP standard. Sample uses data from all seasons of 1989-2003.
Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.

Table A5—Monitor-Specific Regression Discontinuity Estimates for combined RVP and RFG

Index	State	County	Monitor ID	RD RVPII effect	RD RFG effect	RVPII standard error	RFG standard error	RVPII t-score	RFG t-score	RVPII p-value	RFG p-value
1	AZ	Maricopa	19	-.168	-.087*	.04	.045	-4.156	-1.95		.056
2	AZ	Maricopa	1004	-.081	-.079**	.05	.033	-1.618	-2.37	.111	.021
3	DC	District of Columbia	25	.123*	.059	.071	.06	1.74	.977	.088	.333
4	MD	Baltimore	3001	.002	-.055	.045	.042	.036	-1.314	.971	.194
5	MD	Harford	1001	.057*	-.039	.032	.025	1.802	-1.533	.077	.131
6	TX	Galveston	1002	-.216***	-.003	.077	.066	-2.79	-.049	.007	.961
7	TX	Harris	46	-.093	-.003	.057	.044	-1.638	-.067	.107	.947
8	TX	Harris	47	-.041	.016	.055	.042	-.739	.377	.463	.708
9	TX	Harris	62	-.109*	.022	.056	.051	-1.933	.433	.058	.666
10	TX	Harris	1035	-.032	.02	.061	.048	-.53	.418	.598	.678
11	TX	Tarrant	1002	.048	.208	.041	.032	1.184	6.483	.242	
12	TX	Tarrant	2003	-.079	.1***	.05	.034	-1.572	2.943	.122	.005
13	VA	Fairfax	1004	.063	-.016	.057	.053	1.107	-.308	.273	.76
14	VA	Fairfax	5001	-.018	.007	.06	.052	-.306	.129	.761	.898

Notes: Effects shown are the monitor-specific estimated coefficients on the treatment dummies of the regression discontinuity specification (4). Standard errors are clustered on year-season Estimated effects of RVP phase II (less than or equal to 7.8 psi) and federal RFG are relative to the baseline of a 9.0 psi RVP standard. Sample uses data from all seasons of 1989-2003.
Stars indicate the following p-values: *** p<0.01, ** p<0.05, * p<0.10.