Online Appendix

for "Air Filters, Pollution, and Student Achievement"

by Michael Gilraine

A Additional Information on Residential Assistance within 5 miles of Gas Leak

This appendix discusses the assistance that residents received if they resided within five miles of the gas leak. It then presents evidence on whether assistance rates for students' families are likely to be discontinuous across schools on either side of the five-mile boundary.

Residential Assistance: Residents within the five-mile boundary received assistance in the form of reimbursement for any expenses associated with relocation, including accommodation costs and transportation costs to take their child to and from school. To aid residents that did not wish to relocate, SoCalGas offered services to improve air quality inside homes, including installing air filters or scrubbers and weatherizing homes. Given that residents closer to the gas leak were likelier to evacuate, schools nearer to the gas leak presumably have more students who relocated. If moving negatively affects student performance, this implies that we expect there to be a positive relationship between student performance and distance to the leak, which will need to be controlled for. Previewing the empirical design, I will do so by controlling for distance to the leak, although I show robustness to alternative specifications including directly controlling for the percent of individuals who reside within five miles of the leak within a given school's attendance zone in Section IV.D. Section V.A also shows that effect sizes are similar among students who reside and do not reside within the five-mile boundary (according to their ZIP Code of residence).[40]

---

[40]Admission to the local school is guaranteed if one resides within that school's attendance zone. State law, however, mandates intradistrict open enrollment which means that parents can apply for their child to attend another school within the LAUSD. The district must allow the transfer if the school has sufficient capacity, with lotteries determining placements if schools are oversubscribed.

By the time the gas leak was halted, over 8,300 households had been relocated, 3,060 homes had received plug-in air filters, 5,300 homes had air scrubbers installed, and 5,200 homes had been weatherized.[41] To place into context, the author calculates there were 145,000 residents occupying 46,500 residences within five miles of the gas leak.[42] Given this, about twenty percent of eligible households evacuated and thirty percent of eligible households received some form of residential air quality improvement.

Residential Assistance Rates Within Nearby Schools: The high rates of relocation among residents within five miles of the leak are likely to negatively impact student achievement.[43] Given that this will be key to control for, Figure B.8(a) displays the map of school attendance zones in 2015-16 for the twenty-eight schools in the main analysis sample and denotes whether a given school received air filters or not. Given that relocation assistance is provided to individuals living within five miles of the leak, it is clear that schools located close to the five-mile air filter eligibility threshold have attendance zones that cross the five-mile boundary. In addition, while residing within an attendance zone guarantees admission to that school, LAUSD has a (state-mandated) open enrollment policy (subject to capacity constraints) and so schools draw students from throughout Los Angeles. All schools therefore receive a mix of children eligible and ineligible for relocation assistance. (This allows treatment effects to be estimated separately among those who do and do not receive residential assistance – see Section V.A).

Using the 2010 census, Figure B.8(b) shows the percent of the population within each school's attendance zone that resides within five miles of the leak by the distance of the school to the gas leak. Schools near the five-mile boundary have a clear mix of students eligible

---

[41]See https://www.sempra.com/newsroom/press-releases/aliso-canyon-gas-leak-incident-update-february-8-2016.

[42]To do so, I use census block counts from the 2010 census. I exclude individuals north of the gas leak in these counts as they were not impacted (as discussed).

[43]Alternatively, the air quality improvements provided for homes within five miles of the leak could raise achievement. Given the positive relationship between distance to the leak and test scores in the data, however, it appears that the negative effects of relocation outweigh the positive effects of household air quality improvements.

and ineligible for relocation assistance. Of particular interest is that the school furthest from the gas leak that still received air filters, Andasol Avenue Elementary, has less than thirty percent of students eligible for relocation assistance, which is substantially less than other schools that received air filters. If relocation negatively affects student achievement, we might expect this school to perform particularly well since it both received air filters and had few students relocate (and indeed this is what we will observe).

B Appendix Figures and Tables

Figure B.1: Map of Resident Odor Complaints
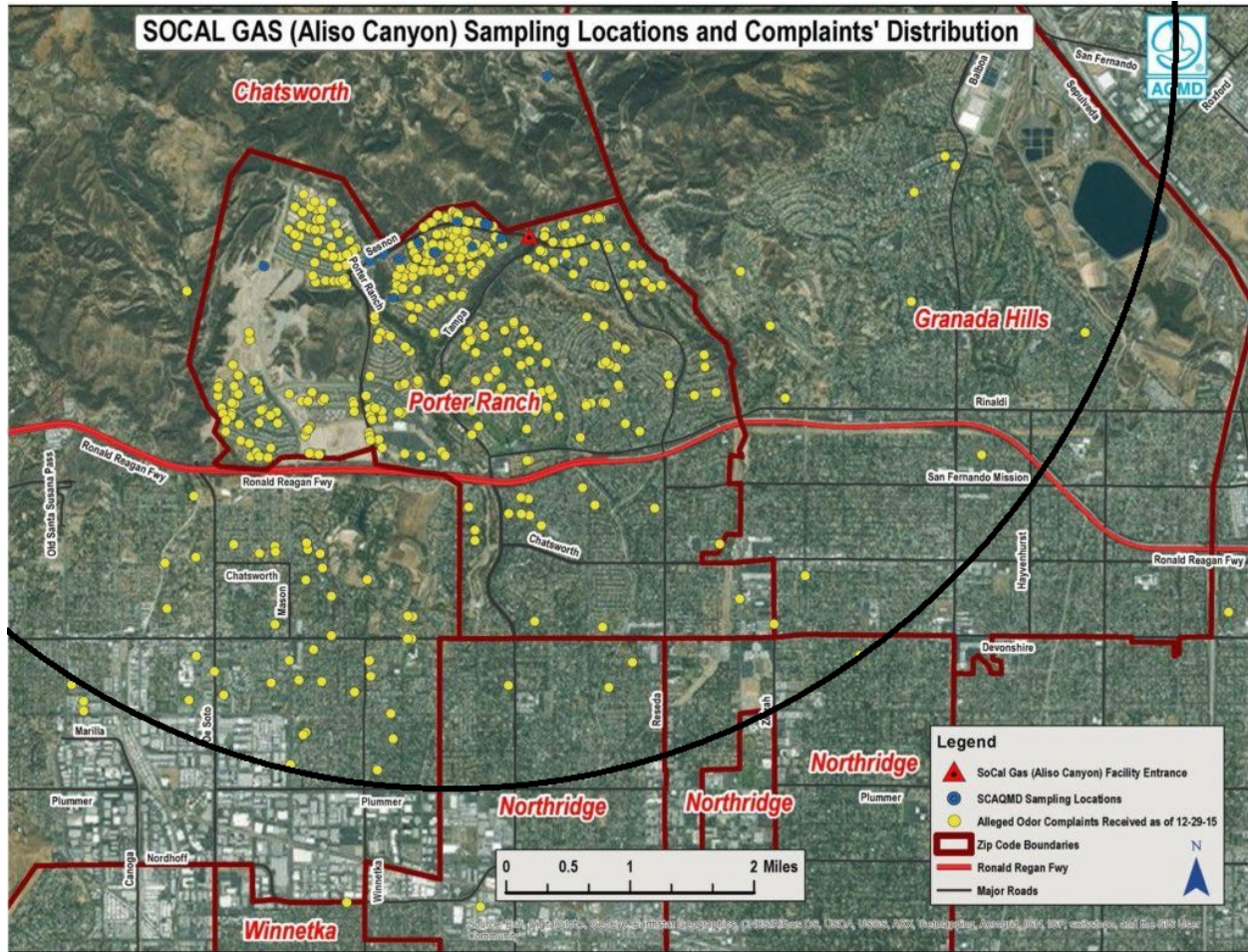
Notes: This figure shows alleged odor complaints reported to SoCalGas as of December 29, 2015. A black circle representing five miles from the gas leak is superimposed over the figure. Zip Code boundaries are also shown. The figure comes from an interoffice correspondence from the Los Angeles Unified School District Office of Environmental Health and Safety and is available at `https://achieve.lausd.net/cms/lib/CA01000043/Centricity/Domain/135/Informative%20re%20Aliso%20Canyon%20Gas%20LeakUpdate%20%2001%2021%2016.pdf`.

Figure B.2: Map of Methane Measurements

Notes: This figure is directed southwestward and shows estimated natural gas presence from a laser-based methane detector used by researchers Rob Jackson, Nathan Phillips, and Bob Ackley to document the extent of the Aliso Canyon gas leak. The vertical height of the bars indicate the amount of natural gas detected. Visually, the natural gas presence appears concentrated north of the Ronald Reagan Freeway which is about two miles south of the Aliso Canyon gas leak.

Figure B.3: RD Validity: Covariates

(a) Lagged Mathematics Score
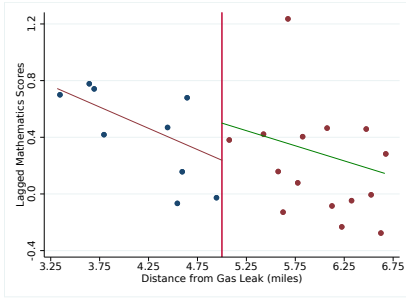
(b) Lagged English Score
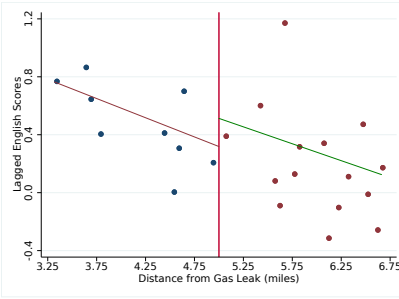
(c) Percent Hispanic
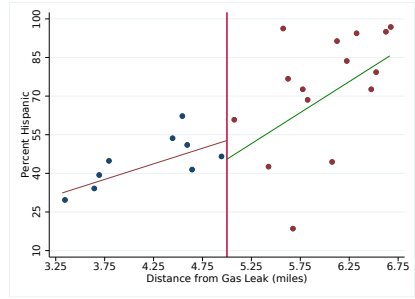
(d) Percent White
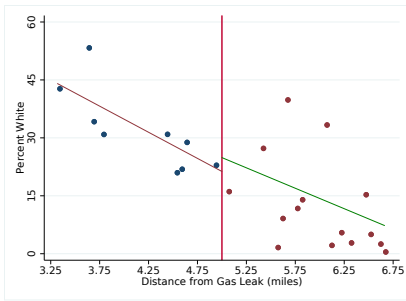
(e) Percent Asian
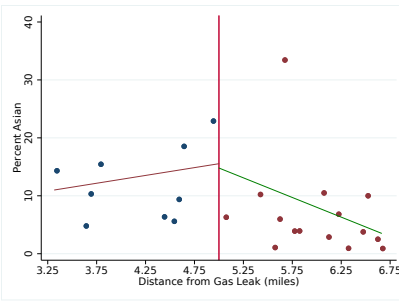
(f) Percent English Learner
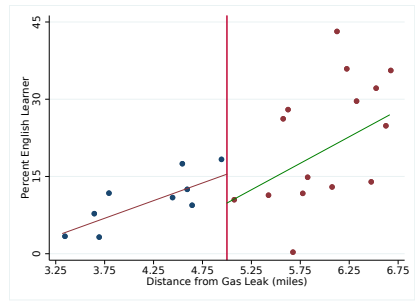
(g) Percent Free or Reduced Lunch
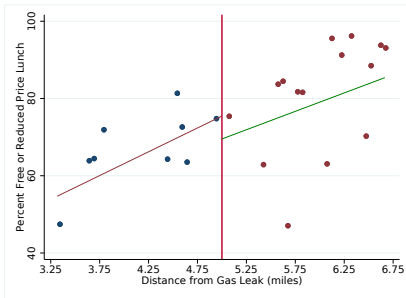
(h) Percent Parent HS Dropout
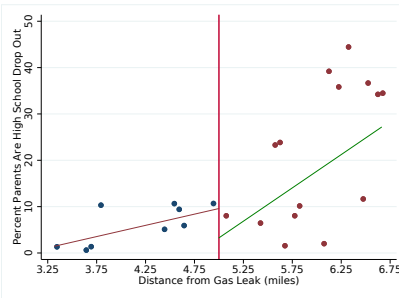
(i) Percent Parent College Graduate
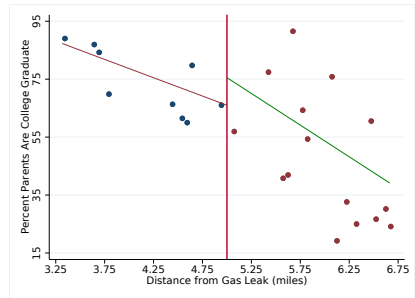
Notes: This figure plots various covariates around the 5-mile boundary in 2015-16 (the year of the gas leak). Sample is restricted to elementary school students with valid current and lagged mathematics scores. Table 6 presents results from a formal test of discontinuity in each of these covariates.

Figure B.4: Robustness: Placebo Years

(a) Mathematics Score
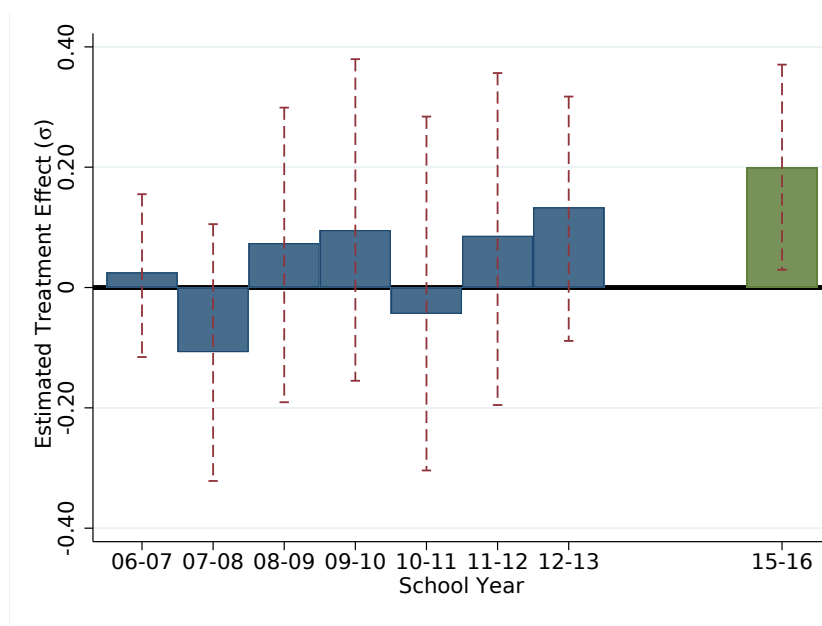
(b) English Score
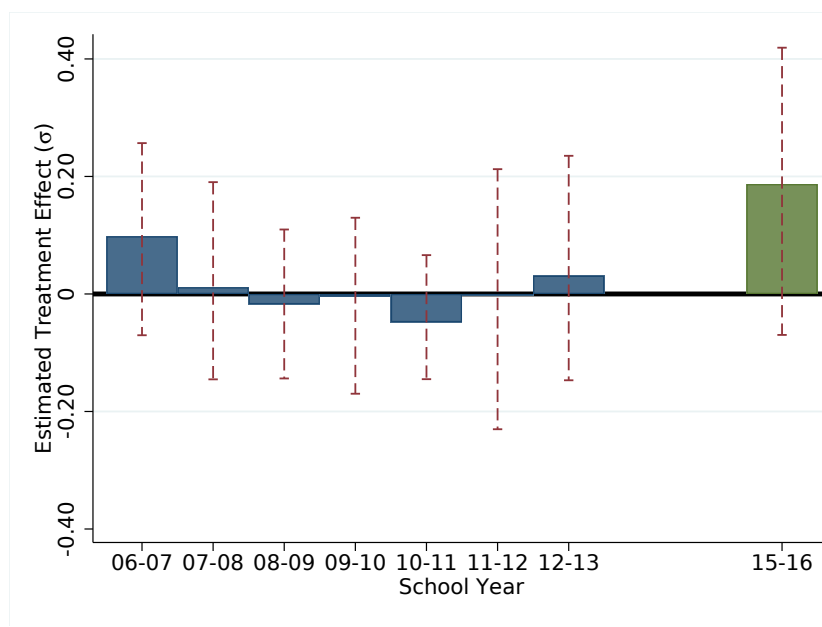
Notes: This figure plots spatial RD estimates from equation (IV.1) by year from 2006-07 through 2015-16. School years 2013-14 and 2014-15 are omitted due to no test scores being available in 2013-14. Controls for lagged test scores, demographics, student switching, and residential ZIP Code fixed effects are used and thus the point estimate for the 2015-16 school year is identical to the one reported in column (2) of Table 5. The dashed whiskers represent 95 percent confidence intervals constructed using the wild cluster bootstrap procedure from Cameron et al. (2008), clustered at the school level.

Figure B.5: Robustness: RD Bandwidth

(a) Mathematics Score
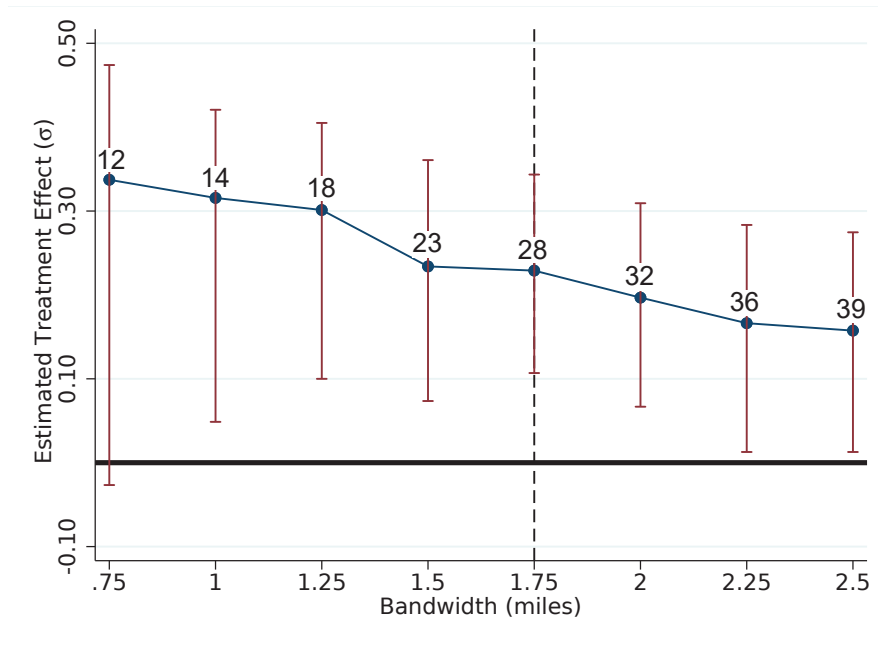
### (b) English Score
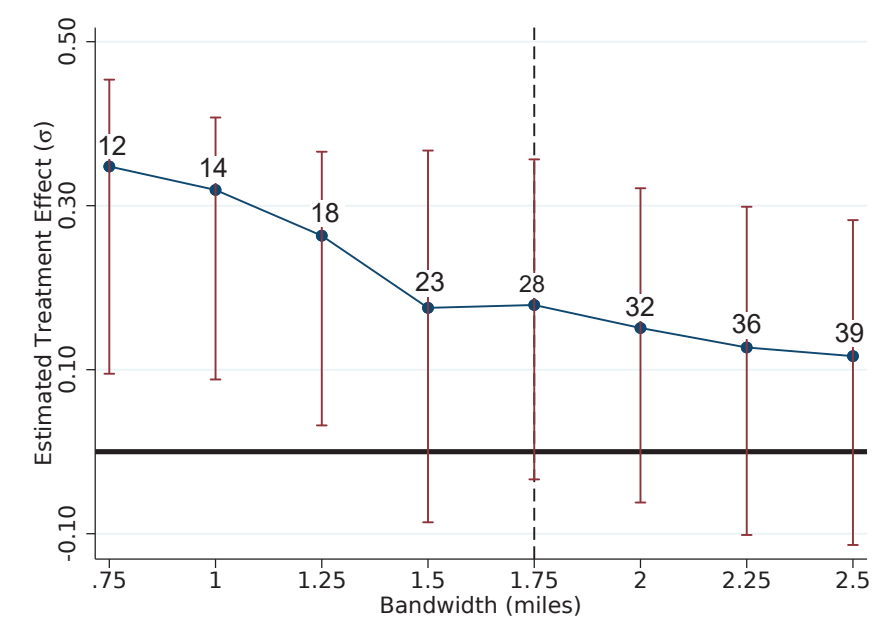
Notes: These figures plot spatial RD estimates using various bandwidths from 0.75 to 2.5 miles. The number of schools for each bandwidth is reported over that bandwidth's 'dot.' Controls for lagged test scores, demographics, student switching, and residential ZIP Code fixed effects are used. The vertical dashed line represents a bandwidth of 1.75 miles which is used for the main analysis (with its point estimate being the same of that reported in column (2) of Table 5). The whiskers represent 95 percent confidence intervals with inference done clustering at the school level using the wild cluster bootstrap procedure from Cameron et al. (2008). Note that the wild clustered bootstrap procedure generates asymmetric confidence intervals.

Figure B.6: Robustness: Diff-in-Disc Bandwidth

### (a) Mathematics Score
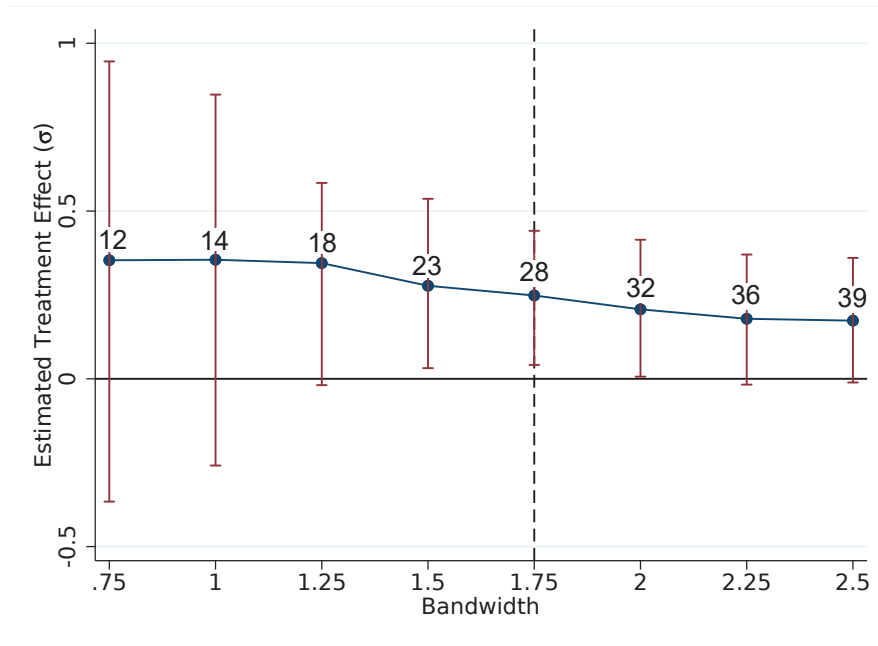
(b) English Score
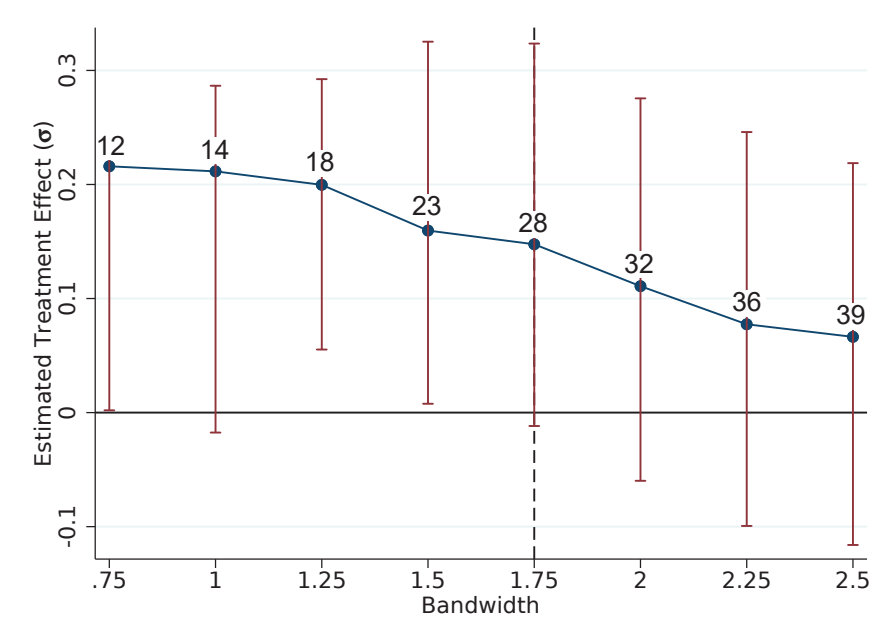
Notes: These figures plot difference-in-discontinuity estimates using various bandwidths from 0.75 to 2.5 miles. The number of schools for each bandwidth is reported over that bandwidth's 'dot.' Controls for student switching, residential ZIP Code, student, and school fixed effects are used. The vertical dashed line represents a bandwidth of 1.75 miles which is used for the main analysis (with its point estimate being the same of that reported in column (3) of Table 7). The whiskers represent 95 percent confidence intervals with inference done clustering at the school level using the wild cluster bootstrap procedure from Cameron et al. (2008). Note that the wild clustered bootstrap procedure generates asymmetric confidence intervals.

Figure B.7: Robustness: Number of Pre-Years Used in Difference-in-Differences

(a) Control Group: All LAUSD Schools
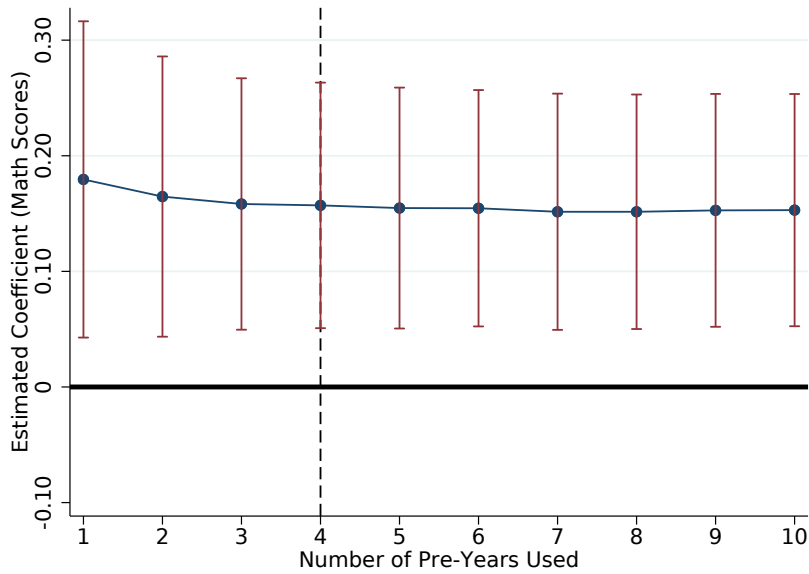
(b) Control Group: 'RD Sample' (Schools within 1.75 Miles of Boundary)
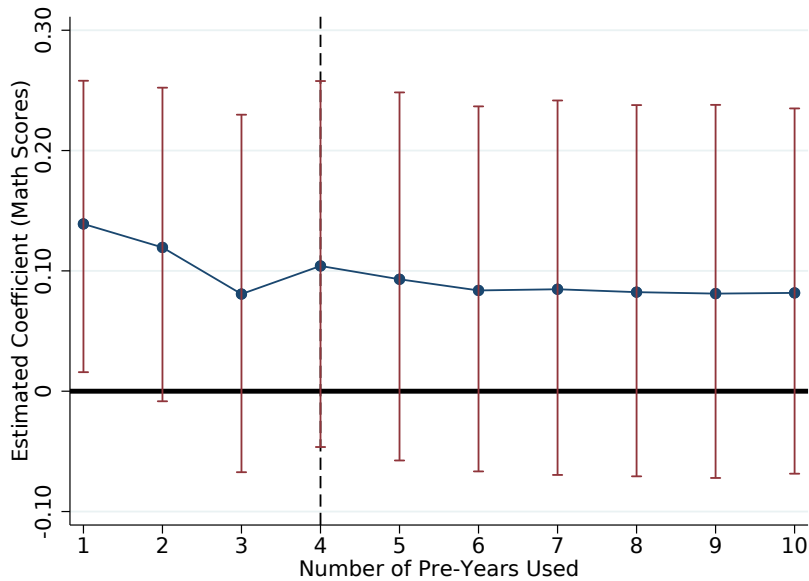

Notes: These figures plots the difference-in-differences estimates from equation (IV.2) when data cover from 1 to 10 years of data in the pre-period. Controls for 'Student Switching' along with student, school, and residential ZIP Code fixed effects are included. The vertical dashed line represents 4 years of pre-data which is used for the main analysis (with its point estimate being the same of that reported in column (2) of Table B.2). Figure B.7(a) includes all LAUSD schools in the control group, while Figure B.7(b) restricts the data to only those schools in the RD sample. The dashed whiskers represent 95 percent confidence intervals constructed using standard errors clustered at the school level for Figure B.7(a) or constructed using the wild cluster bootstrap procedure clustered at the school level from Cameron et al. (2008) in Figure B.7(b).

Figure B.8: Relocation Assistance Assignment by School

(a) Map of School Attendance Zones
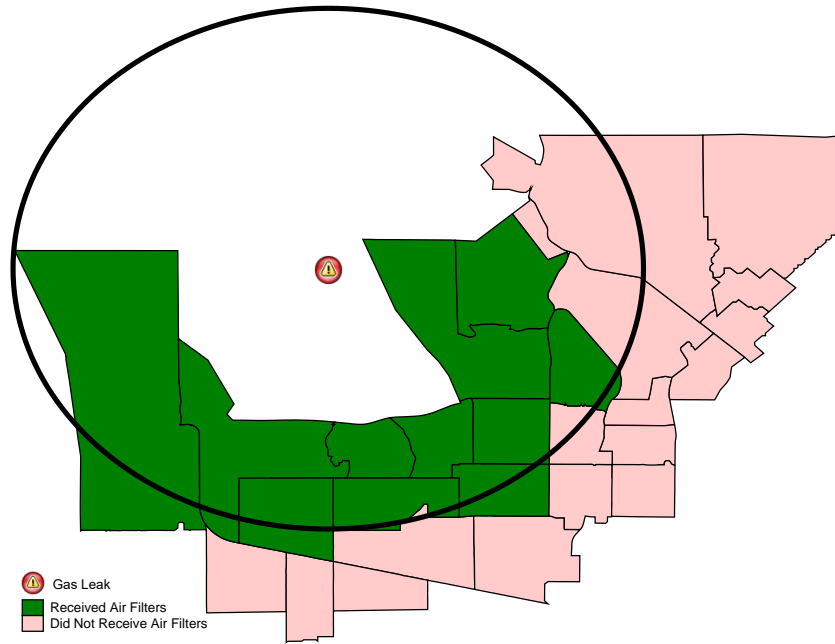
(b) Percent in Attendance Zone Residing Within 5 Miles of Leak
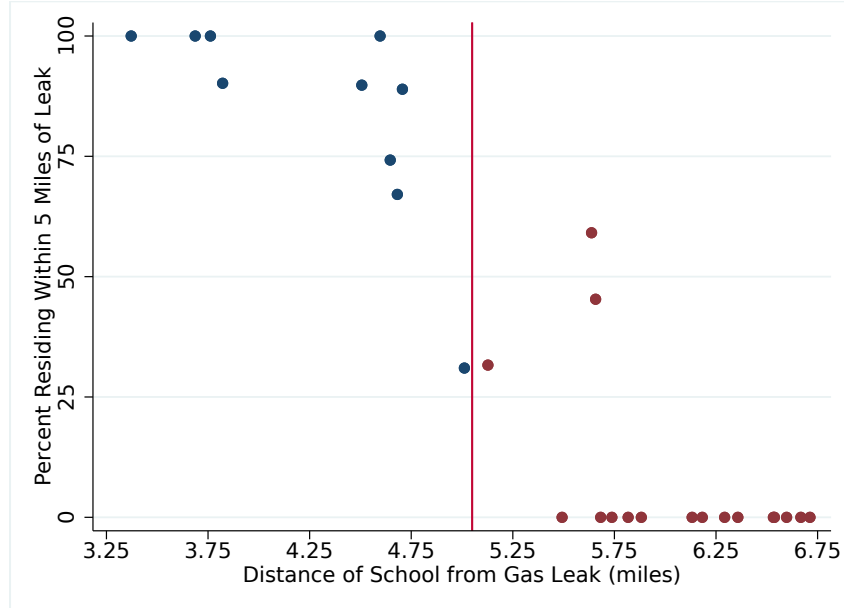
Notes: Figure B.8(a) displays the map of school attendance zone in 2015-16 for the twenty-eight schools in the RD sample and denotes whether a given schools received air filters or not. A circle is superimposed that denotes the five-mile radius around the gas leak in which residents would be eligible for relocation assistance from SoCalGas. One school that did not receive filters is excluded as it is an open enrollment magnet school as so has no attendance zone. Figure B.8(b) then shows the percent of the population within a school attendance zone that resides within five miles of the gas leak by each *school's* distance to the gas leak. Population counts come from the 2010 census and are calculated at the census block level and then are assigned to being within the five-mile boundary and a given attendance zone according to their centroid.

Table B.1: Air Testing Results from Schools within 5 Miles

School Name	Distance to Leak (miles) (1)	Times Tested (2)	Methane (ppbv) (3)	Ethane (ppbv) (4)	Benzene (ppbv) (5)	Toluene (ppbv) (6)	Ethyl-benzene (ppbv) (7)	Xylenes (ppbv) (8)
Castlebay Lane Charter[1]	1.7	15	6,400	120	0.92	14.4	11.37	39.9
Porter Ranch Community[1]	1.8	15	12,940	460	0.73	1.8	0.24	1.7
Robert Frost Middle	3.2	3	3,970	None	0.60	4.7	0.54	3.6
Van Gogh Charter	3.3	3	3,670	None	0.38	3.1	0.29	1.6
Beckford Charter	3.4	5	5,050	None	0.73	1.8	0.55	3.7
El Oro Way Charter	3.7	3	3,220	None	0.45	2.6	0.24	1.2
Darby Avenue Charter	3.8	5	5,950	None	1.10	17.0	2.80	8.2
Germain Academy	3.9	6	4,300	None	1.20	9.9	0.78	5.0
Granada Hills Charter High	4.2	2	3,440	None	0.50	4.3	None	3.1
Nobel Charter Middle	4.4	2	3,240	None	0.47	12.0	None	3.1
Chatsworth Charter High[2]	4.5	2	5,040	None	0.56	7.6	None	2.1
Knollwood Preparatory	4.5	4	3,420	None	1.60	12.0	0.50	3.5
Ernest Lawrence Middle	4.6	2	3,980	None	0.57	9.8	0.67	5.4
Granada Community Charter	4.6	2	3,320	None	0.38	9.8	0.44	2.9
Chatsworth Park Elementary	4.6	2	3,090	None	0.34	8.4	6.20	4.4
Superior Street Elementary	4.7	4	4,280	None	0.84	7.5	1.10	6.3
Topeka Drive Charter	4.7	2	3,120	None	0.64	7.1	None	1.6
Patrick Henry Middle	4.8	2	3,420	None	0.54	13.0	0.29	1.4
Andasol Avenue Elementary	5.0	2	3,050	None	0.37	7.4	None	4.7
Regulatory Limit[3]	-	-	500,000	1,000,000	0.92	80	450	160

[1] Castlebay Lane Charter and Porter Ranch Community were tested every school day from Nov 30-Dec 18 (except Dec 15 as all LAUSD were closed due to a terrorist threat) as well as February 22.

[2] Due to their proximity, Chatsworth Charter High and Stoney Point Continuation were tested as one school.

[3] Regulatory limits for methane and ethane are from the National Institute for Occupational Safety and Health, while the regulatory limits for benzene, toluene, ethylbenzene and xylenes are from the California Office of Environmental Health Hazard Assessment Chronic Reference Exposure Level.

Notes: This table shows the *maximum* reading of six pollutants that were tested Waterstone Environmental as part of the air testing program conducted by the LAUSD in response to the Aliso Canyon gas leak. Pollutant units are in parts per billion by volume (ppbv). The two schools nearest to the leak, Castlebay Lane Charter and Porter Ranch Community, were tested every school day from Nov 30-Dec 18 and then were relocated to new schools outside the five-mile boundary after Dec 18. Every other school within five miles of the gas leak was tested at least twice starting on Jan 19 as part of the program, although several schools were tested more often and Beckford Charter was also tested on Dec 2. Schools appeared to have been tested more than twice for two reasons: (i) schools initially tested on or before Jan 28 were retested in mid-February, and (ii) any school with a benzene reading above 0.92 ppbv were retested soon after. Air testing results are available at https://achieve.lausd.net//site/Default.aspx?PageID=10329.

Table B.2: Difference-in-Differences Estimates

| | Mathematics Score | | | English Score | | |
	Elementary Schools (1)	'Outlier' Excluded (2)	All Schools (3)	Elementary Schools (4)	'Outlier' Excluded (5)	All Schools (6)
Panel A. Control Group: All LAUSD Schools						
Air Filter × Post	0.157***	0.154***	0.104**	0.064	0.062	0.049
[clustered s.e. p-value]	[0.004]	[0.006]	[0.027]	[0.283]	[0.307]	[0.421]
Observations	823,528	822,330	1,227,317	818,361	817,181	1,219,190
Panel B. Control Group: 'RD Sample' (Schools within 1.75 Miles of Boundary)						
Air Filter × Post	0.108	0.104	0.127*	0.070	0.067	0.137**
[wild cluster bootstrap p-value]	[0.148]	[0.189]	[0.080]	[0.271]	[0295]	[0.018]
Observations	46,021	44,823	63,061	45,747	44,567	62,645
Controls						
Student FEs	Yes	Yes	Yes	Yes	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	Yes	Yes
Student Switching	Yes	Yes	Yes	Yes	Yes	Yes
Residential ZIP Code FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates report the effect of air filters on student achievement from the difference-in-differences identification strategy described in equation (IV.2). The outcome variable used is standardized math or English scores and so effect sizes are in terms of standard deviations of the student test score distribution. Panel A includes all LAUSD schools in the control group, while Panel restricts the data to only those schools in the RD sample. Data cover 4 years of pre-data, which given the lack of test scores in 2013-14 extends back five years before the gas leak. Columns (2) and (5) exclude Andasol Avenue Elementary from the regression as it experienced large test score growth in 2015-16 and so may be considered an outlier. Columns (3) and (6) report results including middle schools. 'Student Switching' controls contain an indicator if the student is attending a new school, the distance of their current school from their school last year, and an indicator if the student switched schools within the school year. 'Residential ZIP Code FEs' are fixed effects for the ZIP Code of student residence. All regressions include grade fixed effects and controls for school stability rate, percent of teachers in three experience bins (0-5, 5-10, and 10+), and average class size. P-values clustered at the school level are reported below the point estimates in square brackets. Given there are only twenty-eight school clusters for Panel B, p-values clustered at the school level are calculated using the wild clustered bootstrap procedure from Cameron et al. (2008) in Panel B. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table B.3: Functional Form Robustness: Regression Discontinuity

| | None | Control for Geographic Location: | | | |
| | | Percent Residing Within 5 Miles | Linear | Quadratic | Triangular Kernel |
	(1)	(2)	(3)	(4)	(5)
Panel A. Outcome: Mathematics Scores					
Within 5 Miles	0.084*	0.257*	0.205***	0.277*	0.241**
[wild cluster bootstrap p-value]	[0.064]	[0.079]	[0.007]	[0.061]	[0.011]
Panel B. Outcome: English Scores					
Within 5 Miles	0.141*	0.267*	0.200	0.201	0.197
[wild cluster bootstrap p-value]	[0.026]	[0.071]	[0.107]	[0.302]	[0.123]
Controls					
Lagged Test Scores	Yes	Yes	Yes	Yes	Yes
Student Demographics	Yes	Yes	Yes	Yes	Yes
Student Switching	Yes	Yes	Yes	Yes	Yes
Residential ZIP Code FEs	Yes	Yes	Yes	Yes	Yes
# of Students	4,854	4,854	4,854	4,854	4,854
# of Schools	28	28	28	28	28

Notes: This table shows robustness to the functional form that controls for geographic location by estimating the effect of air filters on student achievement from the regression discontinuity identification strategy as described in equation (IV.1) using various functional forms (including none). The outcome variable used is standardized math or English scores and so effect sizes are in terms of standard deviations of the student test score distribution. Column (1) features no geographic controls and so is identical to column (4) of Table 4. Column (2) then controls for the percent of people in a school's attendance zone that reside within five miles of the leak, while column (3) is identical to column (2) in Table 5 as both control for geographic location using linear distance to the leak. 'Lagged test scores' control for lagged mathematics and English scores interacted with grade dummies. 'Student Demographics' include gender, ethnicity, parental education, free and reduced price lunch status, English learner status, and language spoken at home. Missing indicators are used to control for missing demographics or lagged other-subject scores. 'Student Switching' controls contain an indicator if the student is attending a new school, the distance of their current school from their school last year, and an indicator if the student switched schools within the school year. 'Residential ZIP Code FEs' are fixed effects for the ZIP Code of student residence. All regressions include grade fixed effects and controls for a school's magnet status, affiliated charter status, school stability rate, percent of teachers in three experience bins (0-5, 5-10, and 10+), and average class size. Number of observations are reported for panel A; panel B has four fewer observations as these students lack lagged English scores. Given there are only twenty-eight school clusters, p-values clustered at the school level using the wild clustered bootstrap procedure from Cameron et al. (2008) are reported below the point estimates in square brackets. Conducting inference using the wild clustered bootstrap procedure, ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table B.4: Functional Form Robustness: Difference-in-Discontinuities

	None	Control for Geographic Location: Percent Residing Within 5 Miles	Linear	Quadratic	Triangular Kernel
	(1)	(2)	(3)	(4)	(5)
Panel A. Outcome: Mathematics Scores					
Within 5 Miles	0.108*	0.298**	0.179**	0.278*	0.248**
[wild cluster bootstrap p-value]	[0.075]	[0.013]	[0.044]	[0.056]	[0.021]
Panel B. Outcome: English Scores					
Within 5 Miles	0.078	0.166	0.143	0.205*	0.188**
[wild cluster bootstrap p-value]	[0.205]	[0.140]	[0.136]	[0.085]	[0.025]
Controls					
Student Switching	Yes	Yes	Yes	Yes	Yes
Residential ZIP Code FEs	Yes	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations (student-year)	43,588	43,588	43,588	43,588	43,588
# of Schools	28	28	28	28	28

Notes: This table shows robustness to the functional form that controls for geographic location by estimating the effect of air filters on student achievement from the difference-in-discontinuities identification strategy as described in equation (IV.1) using various functional forms (including none). The outcome variable used is standardized math or English scores and so effect sizes are in terms of standard deviations of the student test score distribution. Column (1) features no geographic controls. Column (2) then controls for the percent of people in a school's attendance zone that reside within five miles of the leak, while column (3) is identical to column (3) in Table 7 as both control for geographic location using linear distance to the leak. 'Lagged test scores' control for lagged mathematics and English scores interacted with grade dummies. 'Student Demographics' include gender, ethnicity, parental education, free and reduced price lunch status, English learner status, and language spoken at home. Missing indicators are used to control for missing demographics or lagged other-subject scores. 'Student Switching' controls contain an indicator if the student is attending a new school, the distance of their current school from their school last year, and an indicator if the student switched schools within the school year. 'Residential ZIP Code FEs' are fixed effects for the ZIP Code of student residence. All regressions include grade fixed effects and controls for a school's magnet status, affiliated charter status, school stability rate, percent of teachers in three experience bins (0-5, 5-10, and 10+), and average class size. Number of observations are reported for panel A; panel B has fewer observations as some students lack lagged English scores. Given there are only twenty-eight school clusters, p-values clustered at the school level using the wild clustered bootstrap procedure from Cameron et al. (2008) are reported below the point estimates in square brackets. Conducting inference using the wild clustered bootstrap procedure, ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.